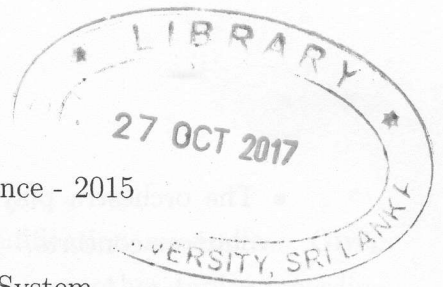


Eastern University, Sri Lanka  
Department of Mathematics  
Special Degree Examination in Computer Science - 2015  
(End of Course Examination)



CSS 01: Advanced Database Management System

- Answer **all questions**. (This paper has 6 questions on 5 pages.)
- At the bottom of the front page of your answer book, write the question numbers in the order you answered.
- Time allowed: **Three Hours**.

Databases and database systems are an essential component of life in modern society. Most of us encounter several activities every day that involve some interaction with a database.

- Describe the three-schema architecture of a database.
  - State clearly what is meant by a distributed database system.
  - State the advantages and disadvantages of a distributed database system comparing to a centralised database system. database management systems.
  - State what is meant by a database **transaction** and explain how execution of transactions in a **single user** system differs from a **multi user** system.
  - Define what is meant by query tree.
    - What is meant by heuristic optimisation and discuss the main heuristics that are applied during query optimisation.
    - Consider the relations:
      - Vehicle (regno, make, colour)
      - Person (eno, name, address)
      - Owner (eno, regno)
- ( $\alpha$ ) For the following query draw the initial query tree:  
SELECT eno, name, regno  
FROM Person, Owner  
WHERE Person.eno = Owner.eno and Person.name = 'Hari';
- ( $\beta$ ) State how to optimise the initial query into final optimised query and draw the optimised query tree.

[100% marks]

Symphonic Band is an orchestra that plays different types of concerts. The orchestras popularity is growing fast and they are starting to have problems to keep track of the musicians that should play in each concert as well as the musical works that are most suitable for the concert. Help the orchestra to create a database model, as a first step to implement a database, so that the orchestra can keep track of both musicians and musical works. The database model must represent the following points:

- The orchestra plays three types of concerts: church concerts, private parties, and outdoor concerts.
- The orchestra plays three types of music: classical, popular, and American folk. The orchestra always plays classical music in their church concerts. The orchestra always plays American folk on private parties. Finally, the orchestra plays a blend of the three types of music when playing outdoors.
- It should be possible to find in the database the music works that are suitable for each type of concert so that the repertoire can be easily planned well in advance.
- For each musical work, the database should store which musical setting (i.e. T instruments) are required to play.
- The database should store information for each coming concert. The information should include the place, date and time of the concert as well as the type of concert and the repertoire that will be played.
- For each musician in the orchestra, the database should store his/her name, the instrument that he/she plays, and in which of the coming concerts he/she will participate.

Draw an EER diagram that captures the above information. Identify any constraints that are not captured by the EER diagram. [1]

3. A Company specializing in the maintenance of computers and related devices stores, in a single table, information about the repairs that its technicians have carried out on computers and other electronic devices.

It records the serial number of the equipment repaired, a one-word description of the equipment, the employee number and name of the technician who carried out (or supervised) the repair (no more than one name will be recorded), the date the repair was begun, brief notes on the nature of the problem, and the technicians pager number.

A partial snapshot of this relation might look like following:

SerialNo	RepairDate	Description	Technician	Name	Notes	Pager
R299822	11-10-2015	Router	P32014	Sakhel	Dropped packets	#932
PS993301	23-05-2015	Power supply	P32014	Sakhel	Fuse replacement	#932
NB39393	23-12-2008	Netbook	P88317	Alice	Keyboard faults	#290
NB99019	01-09-2015	Netbook	P88317	Alice	Cracked screen	#290
HD30022	13-12-2015	BackupHD	P93858	Ksenia	Fuse replaced	#998
DT83298	23-10-2015	Desktop	P88317	Alice	Boot-up failure	#290
DT83298	23-11-2015	Desktop	P32014	Sakhel	Replaced HD	#932
DT40332	03-12-2015	Desktop	P32014	Sakhel	Sticky Keys	#932

For instance, the table records that on the 11<sup>th</sup> of October in 2015, a technician named Sakhel, whose employee number is P32014, and whose pager number is 932, repaired a Router, serial number R299822, which was having a problem with dropped packets.

- (a) Assuming that no device has more than one failure per day, what is the primary key of this table.

- (b) Identify the functional dependencies in this table.
- (c) This table is susceptible to insertion, deletion and modification anomalies. Given an example, based on the table, each kind. Assume that no other tables recording information on equipment or technicians exist. Assume that the data you see in the table is all the data the table holds.
- (d) Bring the data in this table to BCNF, specifying the primary and foreign keys where appropriate, of each table.
- (e) The company's database has a table, with hundreds of thousands of entries, which records details of its customers. Three of columns of this table are as follows:

Street	City	PostCode
234 Elm St	Smallville	SM456
249 Elm St	Smallville	SM456
276 Elm St	Smallville	SM457

The company accesses this database frequently, to send out mailings to its customers. A data analyst has pointed out that the table is not fully normalized, since there is a functional dependency from PostCode to City. (In other words, PostCode functionally determines City). However, he recommends that the table not be split into two BCNF tables. What arguments might he have made to justify his recommendation?

[100% marks]

Indexing is a data structure technique to efficiently retrieve records from the database files based on some attributes on which the indexing has been done.

- (a) Explain the difference between each of the following with an suitable example:
- Primary versus Secondary indexes.
  - Dense versus Sparse indexes.
  - Clustered versus Unclustered indexes.
- (b) List the differences between  $B$ -tree and  $B^+$  tree.
- (c) Draw a sample index structure using  $B$ - tree and  $B^+$ - tree each.
- (d) Consider the relation PARTS with `partNo` as the key field includes records with the following `partNo` values: 23, 65, 37, 60, 46, 92, 48, 71, 56, 59, 18, 21, 10, 74, 78, 15, 16, 20, 24, 28, 39, 43, 47, 50, 69, 75, 8, 49, 33, 38.
- Suppose that the search field values are inserted in the given order in a  $B^+$ -tree of order  $p = 4$  and  $p_{leaf} = 3$ . show how the tree will expand and what the final tree will look like.

[100% marks]

5. Concurrent systems such as Operating systems and Database management systems are generally designed to operate indefinitely.

(a) State what is meant by interleaved concurrency and state two advantages of interleaved concurrency.

(b) Consider the following description of a ticket booking system and answer questions given below:

A company which organises bus service throughout the country has recently started an on-line ticket booking facility. The system can show available buses and for a selected bus can show the seats that are available and the charge for a seat. A customer can select an available seat and book it after making a payment using an electronic card. At particular intervals the system counts the number of available seats in a bus and changes the ticket price according to some formula such that the ticket price would be cheap at the beginning and higher when the number of available seats becomes low. The system allows many users to login to the ticketing system at one time and to book tickets concurrently.

i. State four problems that would arise when transactions are executed concurrently without any concurrency control measures and discuss about each problem giving situations that would probably occur during a ticket booking on the system described above.

ii. Discuss about two possible problems that would arise if there are no recovery measures adopted in the system.

(c) Transactions should possess several properties, often called the ACID properties. State the ACID properties and explain them using examples drawn from the ticket booking system described in part (b).

6. The order of execution of operations from all the transactions executed concurrently is known as a schedule.

(a) State what is meant by conflicting operations in transactions and conflict equivalence.

(b) Describe in detail one approach to testing the correctness of transactions when transactions are executed concurrently.

(c) Consider the schedule for the three transactions  $T_1$ ,  $T_2$  and  $T_3$  given on the next page and answer the following:

i. Draw the precedence graph for the given schedule.

ii. Determine whether the given schedule is conflict-serialisable. If it is conflict-serialisable list all the equivalent serial schedules.

Transaction T <sub>1</sub>	Transaction T <sub>2</sub>	Transaction T <sub>3</sub>
read_item(a) $a \leftarrow a + 1000$ write_item(a)		read_item(c) read_item(a) $c \leftarrow c + a$ write_item(c)
	read_item(a) read_item(c) $a \leftarrow a + c$	
read_item(z) $z \leftarrow z + 10$ write_item(z)		read_item(z) $z \leftarrow z - a$ write_item(z)
	read_item(z) $a \leftarrow a + z$ write_item(a)	

- (d) Schedules of operations of transactions executed concurrently can be categorised as recoverable and unrecoverable. Explain how a schedule becomes unrecoverable.

[100% marks]