THE OPEN UNIVERSITY OF SRI LANKA
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

BACHELOR OF TECHNOLOGY – LEVEL 5
FINAL EXAMINATION 2015/2016

ECX5245 –DATABASE MANAGEMENT SYSTEMS

DATE: December 11, 2016 **TIME: 0930-1230 HRS**

Instructions to candidates

1. This question paper consists of eight questions. Answer **FIVE** questions **ONLY**.
2. This is a **CLOSED BOOK** examination.

Question 1 – Database Programming (20 marks)

Consider the following relational schema. **Enrolled** has one record per student-class pair such that student is enrolled in the class. There is a teacher for each class.

Student (*snum:* int, *sname:* string, *age:* int)

Enrolled (*snum:*int, *cname:*string)

Class (*cname:*string, *meets_at:*time, *room:*string, *tid:*int)

Teacher (*tid:*int, *tname:*string, *deptid:*int)

Write SQL statements for the following:

(a.) Find the names of all teachers who are in the database system but does not have any classes assigned yet.

(5 marks)

(b.) Find the names of students who are enrolled in more than two classes.

(5 marks)

(c.) Write a trigger to ensure that each class cannot have more than 15 students enrolled.

(10 marks)

Question 2 – Object Relational Databases (20 marks)

Consider the following object relational database schema for recording the assignments of employees to projects:

Object types:

proj_t (pno: char(6), pname: varchar2(12), stdate: date, endate: date)
assignment_t (project: ref proj_t, adate: date, hours: number(2))
assignment_tb: table of assignment_t
emp_t (eno: number(8), ename: varchar2(12), assignments: assignment_tb)

Tables:
projects of proj_t(pno primary key);
emp of emp_t(eno primary key) nested table assignments store as emp_assigntb;

The attributes of proj_t are project number (pno), project name (pname), start date (stdate), and end date (endate). The attributes of assignment_t include ref of proj_t (project), date on which the employee was assigned to the project (adate), and number of hours per week (hours). In the type emp_t, the attributes are employee number (eno), name (ename), and the set of project assignments of type assignment_tb which is a table of assignment_t. Some sample data for this database is shown below.

PROJECTS

PNO	PNAME	STDATE	ENDATE
MARS01	MARS PROBE	01-Jan-2007	31-Aug-2008
LUNA05	LUNAR LANDER	01-Jan-2006	31-Dec-2008
SPCE02	SPACE STATION	01-Jun-2006	30-Jun-2008
COME03	COMET PROBE	05-Sep-2007	31-Dec-2007

EMP

ENO	ENAME	ASSIGNMENTS		
		PROJECT	ADATE	HOURS
12347648	Carol SMART	MARS01	01-Jan-2007	10
		LUNA05	10-Jun-2007	10
		SPCE02	20-Mar-2007	15
14527845	Jim TERRIFIC	MARS01	11-Feb-2007	15
		LUNA05	12-Apr-2007	12
		SPCE02	20-Jun-2007	13

(a) Write Oracle OBJECT SQL statement to answer the following query. Use columns of REF type instead of joins to link tables.

For each employee, get the employee number, name, and the total number of hours per week the employee is assigned to work on projects.

(4 marks)

(b) Assuming that the database contains only the given sample data, assign Carol SMART to project number COME03 for 2 hours per week from 5 September 2007.

(5 marks)

(c) Add a member method named projcnt to get the number of projects assigned to an employee after a specified date (for example, the number of projects assigned after 1 June 2007). Write Oracle SQL statements to modify the object type emp_t.

(9 marks)

(d) Using the method defined above, write an Oracle SQL statement to display the eno, ename, and the number of projects assigned to each employee after 1 September 2007. (A date can be input as a string of the form '01-Sep-07'.)

(2 marks)

Question 3 – Formal Query Language (20 marks)

Consider the following schema:

Product (maker, model, type)
PC (model, speed, ram, hd, rd, price)
Laptop (model, speed, ram, hd, screen, price)
Printer (model, color, type, price)

The **Product** relation gives the manufacturer, model number and type (PC, laptop or printer) of various products.

The **PC** relation gives for each model number that is a PC the speed (of the processor, in megahertz), the amount of RAM (in megabytes), the size of the hard disk (in gigabytes), the speed and type of the removable disk (CD or DVD), and the price.

The **Laptop** relation is similar, except that the screen size (in inches) is recorded in place of information about the removal disk.

The **Printer** relation records for each printer model whether the printer produces color output (true, if so), the process type (laser, ink-jet or bubble), and the price.

Write the following queries in **relational algebra**.

(a) Which PC models have a speed of at least 1000?

(3 marks)

(b) Which manufacturers make laptops with a hard disk of at least one giga-byte?

(4 marks)

(c) Find the model number and price of all products (of any type) made by manufacturer B.

(5 marks)

(d) Find those manufacturers that sell Laptops, but not PC's.

(4 marks)

(e) Find those hard-disk sizes that use same hard disk size in two or more PC's.

(4 marks)

Question 4 – Disk and Files (20 marks)

(a) Explain the terms related to disk access; Seek time, Rotational delay, Data transfer time

(6 marks)

(b) Consider a disk with sector size of 512 bytes, 4000 tracks per surface, 20 sectors per track, and five single-sided platters. Each disk block spans two sectors. The disk platters rotate at 6000 rpm (revolutions per minute).

i. What is the average access time for a disk block if the average seek time is 100 msec (0.1 seconds)? Assume ½ revolution for average rotational delay. You can transfer 20 sectors in a single revolution.
Note: Disk Access Time = Seek Time + Rotational Delay + Transfer Time

(5 marks)

ii. What is the capacity of the disk in bytes?

(3 marks)

(c) Which of the three basic file organizations would you choose for a file where the most frequent operations are as follows?

i. Search for records based on a range of field values.
ii. Perform inserts and scans where the order of records does not matter
iii. Search for records based on a particular field value

(3 marks)

(d) Consider a relation stored as a randomly ordered file for which the only index is an unclustered index on a field called *sal.* If you want to retrieve all records with *sal* > 20, is using the index always the best alternative? Explain.

(3 marks)

Question 5 – XML Databases (20 marks)

(a) What are the main components of a FLWOR expression and explain them using a suitable example.

(5 marks)

Consider the following XML document for the questions given below:

```
<Contact>
        <Names>
                <Name type="Legal">
                        <First>Thomas</First>
                        <Middle>Cruise</Middle>
                        <Last>Mapother</Last>
                </Name>
                <Name type="Stage">
                        <First>Tom</First>
                        <Middle></Middle>
                        <Last>Cruise</Last>
                </Name>
        </Names>
        <Addresses>
                <Address type="Primary">
                        <Street>12345 Main Street</Street>
                        <City>San Diego</City>
                        <State>CA</State>
                        <Zip>92130</Zip>
                </Address>
                <Address type="Other">
                        <Street>6200 Cruise Avenue</Street>
                        <City>San Fernando</City>
                        <State>CA</State>
                        <Zip>92126</Zip>
                </Address>
```

```
        </Addresses>
        <Phones>
                <Phone type="Mobile">8085554422</Phone>
                <Phone type="Home">8085553399</Phone>
        </Phones>
</Contact>
```

Note that the above XML document stored in *Contractors (id int, conDetails xml)* table created in MS SQL Server and it contains only single record.

(b) What is the output of the following XPath expression:

```
SELECT conDetails.query('//Phone[@type="Home"]')
FROM Contractors
```

(2 marks)

(c) Display the details of the first address. What would be the XPath expression and its output?

(3 marks)

(d) Write an XQuery to print the name (i.e. first name and last name) of both Contractors.

Output:
```
<name>Thomas Mapother</name>
<name>Tom Cruise</name>
```

(5 marks)

(e) Write an expression to insert the following contractor name after the *"Thomas"*

```
<Name type="Legal">
        <First>Saman</First>
        <Middle>Keshara</Middle>
        <Last>Jayaweera</Last>
</Name>
```

(5 marks)

Question 6 – Indexing and Query Processing (20 marks)

(a) What are the steps in Query Processing? Explain each step.

(4 marks)

(b) What is the justification for using I/O costs as the main measure in relational query optimizer?

(2 marks)

(c) Consider the following schema:

Emp(eid:integer; ename:string, age:integer, pno:integer)

Proj(pno:integer, pname:string, budget:float)

Consider the following query:

```
SELECT Emp.eid
FROM   Emp, Proj
WHERE Emp.pno = Proj.pno
```

i. Write two relational algebra expressions for the above SQL query in which, natural join is used for one and Cartesian product for the other.

(2 marks)

ii. Draw the graphical representation of the relational algebra expressions in (i.).

(2 marks)

iii. What query plan in (ii.) would you suggest to execute on a DBMS and briefly explain the query plan you consider?

(2 marks)

iv. The above SQL query which prints employees have been assigned to the projects. Re-write the same query in SQL without join tables.

(2 marks)

v. What index would you suggest to speed up the query in (vi.)? Assume that DBMS supports only B+ tree indexes.

(2 marks)

vi. Consider the following information.

Employee (Emp) relation consists of 1000 pages with 100 tuples per page. Project (Proj) relation consists of 500 pages with 200 tuples per page.
Assume that 90% of employees have been assigned to projects and all attributes of employee table are equal in size. You are free to assume all null values are placed in together in index file.
Estimate the total cost for the query plan in (v.) ignoring output costs. State any assumptions you made.

(4 marks)

Question 7 – Transactions and Concurrency Control (20 marks)

(a) Briefly explain the properties of a transaction.

(4 marks)

(b) Briefly explain, what a *Serializable Schedule* is?

(1 mark)

(c) Briefly explain the rules in Strict 2 Phase Locking Protocol.

(2 marks)

(d) Consider a database with objects X and Y and assume that there are two transactions T1 and T2. Transaction T1 reads object X and Y and then writes object X. Transaction T2 reads objects X and Y and then writes objects X and Y. Both T1 and T2 commit after all read and write actions of them.

i. Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-read conflict.

(2 marks)

ii. Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a read-write conflict.

(2 marks)

iii. Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-write conflict.

(2 marks)

iv. For each of the three schedules, show that Strict 2PL disallows the conflicts.

(2 marks)

(e) Consider the following sequence of actions, listed in the order they are submitted to the DBMS. The Strict 2PL has been used for concurrency control.

T1:R(P), T2:W(Q), T2:W(P), T3:R(R), T3:R(Q), T1: W(R)

T_1	T_2	T_3
S(P)		
R(P)		
	X(Q)	
	W(Q)	
	X(P)	
	W(P)	
		S(R)
		R(R)
		S(Q)
		R(Q)
X(R)		
W(R)		

Assume that older transaction has higher priority always. The DBMS processes actions in the order shown.

i. Follow wound-wait policy to deal with deadlock in above schedule. Draw schedule again.

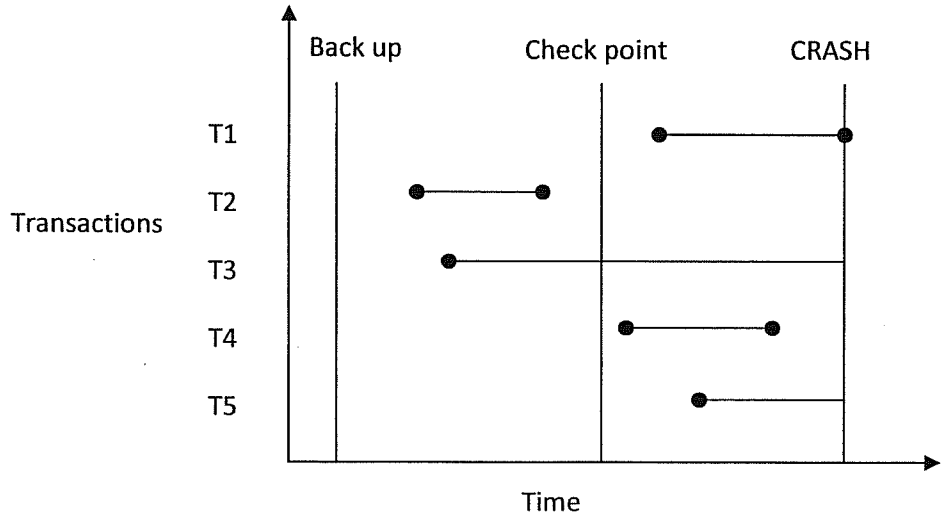
(3 marks)

ii. Follow deadlock detection approach to deal with deadlocks in above schedules. Draw a wait-for graph for the schedule given above to detect the deadlock and briefly explain how to identify the deadlock and remove.

(2 marks)

Question 8 – Crash Recovery (20 marks)

(a.) Consider the diagram below:



What is the desired state that the database should be in after Crash Recovery? Briefly explain.

(5 marks)

(b.) What is meant by Write-Ahead Logging. Why it is necessary to implement this protocol in a DBMS?

(5 marks)

(c.) What are the roles of Analysis, Redo and Undo phases in ARIES?

(5 marks)

(d.) What are the differences between update log records and CLRs?

(5 marks)

-- End of the Question Paper --