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

FINAL EXAMINATION 2008-2009
BACHELOR OF TECHNOLOGY – LEVEL 5

085

ECX5240 – INFORMATION SYSTEMS

DATE: APRIL 07, 2009 **TIME: 0930-1230 HRS**

Answer **only five** questions.

1.

a) What is a constraint in a RDBMS? [2 Marks]

b) Identify and briefly explain different types of constraints in a RDBMS. [3 Marks]

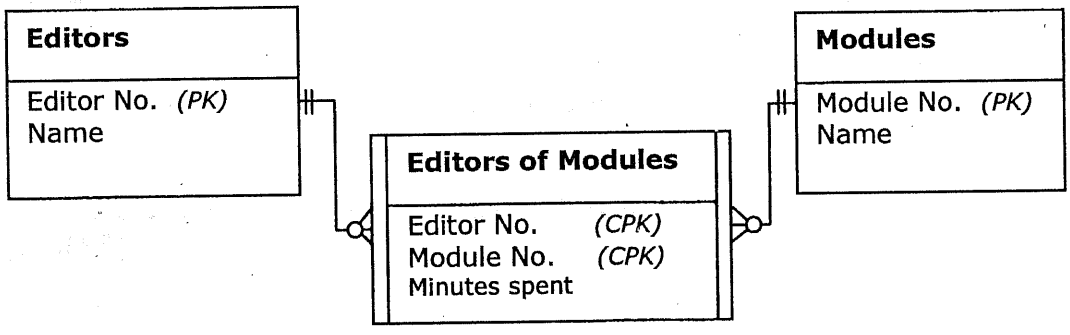
c) Briefly describe the significance of application domain knowledge in identifying constraints in a database. [4 Marks]

d) List two reasons to perform database refactoring. [4 Marks]

e) Explain the database recovery process after a disk failure. Assume the DBMS is using the deferred update. [7 Marks]

2.

Consider the following ER diagram in crow's foot notation to answer this question.



a) Redraw the ER diagram including 'Reviewers' as a new entity. It is not essential to redraw an existing entity if it does not establish a relationship with 'Reviewers'. Business rules applicable for reviewers are as follows,

- Each and every *module* should be reviewed by one and only one reviewer
- Although a *reviewer* doesn't have to be associated with a *module* at all, he or she cannot be associated with more than five modules at any given time.

[4 Marks]

b) Write SQL queries for following statements:

i.) Select the module names edited by Mr Murali. [4 Marks]

ii.) Select the module names edited by three editors and the sum of *minutes spent* is greater than 150. [4 Marks]

iii.) Select editors who have not involved in editing any module. [4 Marks]

c) Draw an appropriate query plan for question 3 (b) iii. [4 Marks]

3.

a) List two uses of relational algebra in a RDBMS. [4 Marks]

b) Relational schemas for three relations in a project information database are as follows.

Project (Project ID, description, type)
Staff (Staff ID, Name, NIC, Grade)
Project_Staff (Project ID, Staff ID)

i) Find the *Project IDs* of projects of type *A* or type *B*? [3 Marks]

ii) Find the *Staff IDs* of the staff members *not* allocated to a project? [3 Marks]

iii) Find the *Staff IDs* of the staff members allocated to all projects? [3 Marks]

c) Briefly describe how you can improve the performance of query execution by examining the query plan. [7 Marks]

4.

a) Briefly explain the ACID properties of a database transaction. [4 Marks]

b) Identify what information is to be recorded, if you are to audit activities in a web enabled database. [4 Marks]

c) Briefly describe two functionalities of a distributed database when compared to a centralised database. [7 Marks]

d) Briefly explain how indices improve the performance of database query execution. [5 Marks]

5.

a) Briefly describe 3 motivation factors for contribute in Free/Open Source Software Development. [6 Marks]

b) Using an illustration distinguish the data transfer method in an ajax based web application and a traditional web application. [5 Marks]

c) Discuss the viability in Sri Lanka (in terms of different needs of computer users, cost of software, legal aspects, user friendliness, access to knowledge and freedom to share) if you are to propose Free/Open Source Software as an alternative to the proprietary software systems. [9 Marks]

6.

a) List 4 advantages of having a DBMS. [4 Marks]

b) Briefly explain the advantages of building a data warehouse or a business intelligent solution for an organisation involved in online sales. [7 Marks]

c) Identify 3 challenges of improving query execution performance in web enabled databases. [3 Marks]

d) Describe how the DBMS's capability of implementing threads within multiple processes does increases the running performance of database operations in contemporary multi-processor and multi-core architectures? [6 Marks]

7.

a) What is data localisation in respect to distributed database systems? [3 Marks]

b) Briefly explain how data localisation can improve the performance of a distributed database. [5 Marks]

c) Using an illustration, briefly describe Two-Phase Commit (2PC) protocol in terms of messages exchanged and log records written. [6 Marks]

d) Database views are materialised to improve the performance. Do you propose to materialise every view in a database? Justify your answer. [6 Marks]

8.

a) How can you use data encryption to protect unauthorized access in a distributed database? [5 Marks]

b) A student is comparing the performance of a centralised database with a parallel database. His centralised database consists of only a single machine. Parallel database consists of a cluster of 6 machines, with the same hardware configuration as the machine in the centralised database. Parallel database has been created by the horizontal partitioning using the use hash partitioning technique. It is assumed that that each node has about $1/6^{th}$ of the tuples after partitioning. He plans to run queries over a set of tea buyers. His database consists of 3 relations, a **buyers** table, listing the information about each buyer that he has contacts with, a **quotes** table, listing information about the quotations he sent, and an **orders** table listing the order information from buyers.

The tables are as follows:

Buyers (buyer ID, Name, address)
Quotes (buyer ID, time, price)
Orders (buyer ID, time, quantity, price)

He runs three different types of queries:

A. A query to compute the most recent quote from a given buyer ID
B. A query to compute the purchase orders where the price is higher than the price of the recent quotation
C. A query to compute the average price of tea over the last week, ordered by price

i) When the student runs query A on the centralised database, he finds that it takes x seconds to complete. When he runs it on all 6 machines, approximately how long do you expect it will take? Justify your answer. [5 Marks]

ii) When he runs query B on all 6 machines, he finds that the performance is 8 times faster than on the centralised database. Briefly describe one reason why he might be seeing a speedup greater than the number of machines in the parallel system (superlinear speed up). [5 Marks]

iii) When he runs the query C on all 6 machines, he finds that the performance is only 3 times faster than on the centralised database. Briefly explain one reason why he might be seeing a speedup less than the number of machines in the parallel system (sublinear speed up). [5 Marks]