# JARAMOGI OGINGA ODINGA UNIVERSITY

# OF SCIENCE & TECHNOLOGY

## UNIVERSITY EXAMINATIONS 2012/2013

## 2ND YEAR 1ST SEMESTER EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE (COMPUTER SECURITY AND AUDIT)

## (KISUMU L.CENTRE)

**COURSE CODE: IIT 3213**

**COURSE TITLE: DATABASE ADMINISTRATION & DESIGN**

**DATE: 13/8/2013** **TIME: 9.00-11.00 AM**

**DURATION: 2 HOURS**

**INSTRUCTIONS**

1. **This paper consists of 5 Questions.**
2. **Answer Question 1 (Compulsory) and any other 2 questions.**
3. **Write your answers on the answer booklet provided.**

**QUESTION ONE [30 MARKS]**

(a) Below is an excerpt about Database Administrator (DBA) extracted from Acme Corporation, US:

*The CIO of Acme Corporation hires a management consulting company to streamline their information technology (IT) operations. The consultant, determined to understand the way Acme works, begins by interviewing the CIO. One of his first questions is: "So, I see that you have a DBA on staff. What does he do?" The CIO replies, "Well, I'm told that we need a DBA to make sure our Oracle databases stay online. I know that some of our critical business processes like order entry and inventory use Oracle, but I really don't know what the DBA does. But please don't tell me I need another one, because we can barely afford to pay the one we have!"*

(i) In your own words, describe a DBA and the role he plays in an organization? [4 Marks]

(ii) Why is it important to study Database Administration? [2 Marks]

(b) Denormalization is the process of deliberately introducing redundancy to date. Give TWO considerations that can necessitate denormalization of table databases. [2 Marks]

(c) Why should database systems maintain integrity constraints? [2 Marks]

(d) Do all relations have at least one superkey? Support your answer. [2 Marks]

(e) Explain any TWO techniques that can be used to optimize the performance of database structures. [2 Marks]

(f) Consider the 1NF relation R(A,B,C,D,E,F,G,H,J,K).

1) Suppose the dependencies are A, B →C; A → D,E; B → F; F → G,H; D → J,K. What is the key for R? Decompose R into 2NF relations. Then decompose R into 3NF relations. Can R be decomposed into BCNF? [2 Marks]

2) Suppose the dependencies are A, B → C; B,D →E,F; A,D → G,H; A → J; H →K. What is the key for R? Decompose R into 2NF relations. Then decompose R into 3NF relations. Can R be decomposed into BCNF? [2 Marks]

(g) Cite TWO important reasons for learning Relational Algebra and Relational Calculus in database design. [2 Marks]

(h) Consider the schema below;

Sailors (<u>sid: integer</u>, sname: string, rating: integer, age: real)

Boats (<u>bid: integer</u>, bname: string, color: string)

Reserves (sid: integer, bid: integer, day: date)

Using the schema above, provide relation algebra, relational calculus and SQL equivalent of the question below;

*Find the names of sailors who have reserved all boats* [6 Marks]

(i) Consider the SQL query provided below; [4 Marks]

**SELECT** E.title, COUNT (*) **FROM** Executives E **WHERE** E.dname > 'W%'
**GROUP BY** E.title

Suppose that a clustered B+ tree index on *title* is available, determine the cost of best plan.

**QUESTION TWO [20 MARKS]**

(a) Use suitable example to distinguish between the following as applies to database design; [4 Marks]

(i) Database Schema and Database Instance

(ii) Conceptual Model and Logical Model

(b) What is *Integrity Constraints*? Explain the FOUR types of integrity constraints. [5 Marks]

(c) "Constant database monitoring and tuning are valuable recipes for effective database performance management". What are the components of database tuning? [3 Marks]

(d) "One of the **SQL Rules of Thumb** is *Judicious Use of OR*. However, it is argued that when OR is swapped with IN in an SQL statement, the performance is likely to improve" Use SQL statements to support this statement. List any other THREE SQL Rules of Thumb you know. [5 Marks]

(e) "A database administrator never requires business metadata to his/her job and the businessman never needs technology metadata". Do you agree? Explain. [3 Marks]

**QUESTION THREE [20 MARKS]**

(a) Give TWO reason support existence of relationship between ER Model and the relational model. [2 Marks]

(b) Identify TWO factors to be considered when designing an ER diagram. [2 Marks]

(c) Draw an ER diagram for a physical therapy center database from the description below.

In your ER diagram, you must properly denote all applicable concepts, including weak or strong entities, and keys, composite or multi-valued attributes; relationships and their cardinality and participation constraints. Then map developed ER diagram into set of relations, provide list of functional dependencies and underline primary keys and asterisk foreign keys. Give its normal form.

A database is needed to keep track of the operation of a physical therapy center.

Every patient must be referred by a physician and have a prescription for physical therapy in order to receive treatments. A patient may have different physicians at different times. The database keeps all information about prescriptions and treatments, both past and current. When appointments are made, the information about scheduled date and time is recorded. No patient is ever scheduled for two visits on one day. The center has several physical therapists, and a patient may be treated by different physical therapist at different visits. When a patient makes a visit at an appointed time, the name of the therapist, the treatment, the date, time, and the equipment used are all recorded for that visit. Each of these has only one value for the visit. This information will be used later for insurance billing, which is not part of the database.

[16 Marks]

**QUESTION FOUR [20 MARKS]**

(a) "A database is both data portion and log partition and this call for constant monitoring of database space usage". Explain the meaning of underlined terms ti qualify the statement. [2 Marks]

(b) When a page in the buffer pool is modified, how does the database system ensure that this change is propagated to the disk? [3 Marks]

(c) Consider a relation stored as a randomly ordered file for which the only index is an unclustered index on a field called *marks*. If you want to retrieve all records with *marks* > 60, is using the index always the best alternative? Support your answer. [4 Marks]

(d) On hard disks, each file is stored as a sequence of fixed-size pages. A typical page size is 4 kilobytes. The cost of accessing a file is often measured as the number of pages that must be accessed. Randomly fetching a single page from disk typically takes about 10 milliseconds (roughly the time required to execute 5,000,000 instructions on a current desktop computer). Sequentially fetching a page is much cheaper; times vary widely, but 1 millisecond is a reasonable rough approximation. The number of bytes read from the page has little impact on access costs.

1. Consider a relation of 1,000,000 records, each of 200 bytes. Suppose that records are fetched in random order. If 10,000 randomly-selected records are to be fetched, how long might it take to fetch them? What are the implications of these results for use of indexes? [5 Marks]
2. In nested loop join, for each record in the first, "outer" relation (table *R*, say), every record and hence every page is sequentially fetched from the second, "inner" relation (table *S*). Suppose both *R* and *S* have 1,000,000 records of 200 bytes each. How long will it take to compute the join of these two relations? Suppose one relation has 1,000,000 records of 200 bytes each and the other has100,000 records of 2,000 bytes each. Which should be the inner relation and which should be the outer? [6 Marks]

**QUESTION FIVE [20 MARKS]**

(a) While using practical examples, explain your understanding of the following; [8 marks]

(i) Temporal Databases
(ii) Statistical Databases
(iii) Text Databases
(iv) Spatial Databases

(b) Explain how a company offering services over the internet could use public-key encryption to make its order process more secure. [3 Marks]

(c) Compare the relative merits of centralized and hierarchical deadlock detection in a distributed database management system (DBMS). [4 Marks]

(d) Give an example that shows the correct application of deductive queries. [2 Marks]

(e) Use a well labeled diagram to illustrate a typical data warehousing architecture [3 Marks]