Ollscoil na hÉireann, Gaillimh GX 1470
National University of Ireland, Galway

Semester I Examinations, 2003/2004

Exam Code(s)	2IF 2IF External Repeat Erasmus
Exam(s)	2nd Year Examination in Information Technology 2nd Year Examination in Information Technology (Repeat) Erasmus
Module Code(s)	CT230
Module(s)	Database Systems I
Repeat Paper	
External Examiner(s)	Professor P. Nixon
Internal Examiner(s)	Professor G. Lyons Ms. J. Griffith

Instructions: Answer **THREE** questions.
All questions carry equal marks.

Duration **2hrs**

No. of Pages 4
Department(s) Information Technology

Q. 1. (i) Distinguish between the traditional file processing approach and the database approach, describing the advantages and disadvantages of each approach. *(10)*

(ii) With the aid of examples, distinguish between:

a) Database schema and database instance. *(5)*

b) External schema and physical schema. *(5)*

(iii) Define the relational model, describing the concepts of *relation*, *attribute*, *tuple* and *domain* in your answer. *(5)*
Discuss the integrity constraints that are considered part of the relational model and give examples of cases where the DBMS must check to ensure that such integrity constraints are not violated by update operations. *(5)*

Q. 2. (i) Outline the motivations for normalisation. *(2)*
With the aid of examples, explain the following terms:

a) Functional dependency. *(4)*

b) Partial dependency. *(4)*

(ii) Give definitions for:

a) First normal form. *(3)*

b) Second normal form. *(3)*

c) Third normal form. *(4)*

(iii) An un-normalised relation for orders in a company has the following schema:

```
order(o_no, o_date, customer_no, customer_name,
customer_address, customer_balance, product_no,
p_description, quantity_required, unit_price)
```

where associated with each order is an order number, order date, name, number and address of customer who placed order as well as customer balance, and details on the product ordered (product number, product description, quantity of product ordered and unit price of product).

a) Show where duplication may result given this relation design. *(4)*

b) Normalise the relation to third normal form, showing the stages involved. *(6)*

Q. 3. Given the following relational schema (with keys underlined) and interpretation:

```
TOURS(Code, Type, Location, DepartDate, Duration)
LOCATION(LocationName, SummerTemp, WinterTemp, TimeZone)
CUSTOMER(ID, Surname, Firstname, Address, PhoneNum)
BOOKINGS(CustID, TourCode, No_of_People)
```

A travel company offers a number of TOURS, where each tour has an associated, unique code (Code), is of a certain type (e.g. skiing, beach, etc.), is located in a certain location and has a fixed departure date and duration (number of days of the tour). The LOCATION relation holds information on tour locations: a unique location name (LocationName), an associated average summer and winter temperature and a time zone. The details held on customers in the CUSTOMER relation are: ID, first name, surname, address and phone number. The relation BOOKINGS holds details of bookings made by a customer: the primary key is the customer ID and the code of the tour they have booked. Details are also kept on the number of people booked by the customer for that tour (e.g., 1, 2, etc.).

(i) With the aid of examples based on the relational schema given, describe each of the following, identifying, if applicable, an example of the attributes in each of the above relations:

a) Composite attribute. *(2)*
b) Single-valued attribute. *(2)*
c) Primary key. *(3)*
d) Foreign key. *(3)*

(ii) Give a relational algebra solution and an SQL solution for the following information need: *(10)*

> List the name and address of all customers who are booked for a tour of type 'beach' where the tour duration is longer than 7 days.

(iii) Develop SQL queries to satisfy the following information needs:

a) List the location and departure date of all tours of type 'beach' whose duration is greater than 5 days and less than 10 days. *(3)*

b) List the total number of people who are booked for tours of type 'hiking'. *(3)*

c) List the name and address of all customers who have booked the same tour as customer 'Jasper Welsh'. *(4)*

Q.4. (i) With the aid of examples, distinguish between the following:

a) Primary and clustering indexes. *(3)*

b) Dense and non-dense (sparse) indexes. *(3)*

c) Sequential and hashed file organisation. *(4)*

(ii) What is meant by a view? *(3)*
How can a view be implemented? *(3)*
Discuss the advantages and disadvantages of using views. *(4)*

(iii) With the aid of examples, explain why the SQL command DISTINCT is required. *(5)*
Outline an approach that can be used to implement the SQL DISTINCT command. *(5)*

Q. 5. (i) What is meant by query processing? *(4)*
With the aid of a diagram, outline the main steps involved in processing a query. *(6)*

(ii) With respect to query processing and optimisation, describe the following terms:

a) Query tree. *(3)*

b) Evaluation plan. *(3)*

c) Cost estimates. *(4)*

(iii) Using the relational schema and interpretation defined in **Q. 3.**, develop an SQL query to satisfy the following information need: *(2)*

> List the type, location, average summer and winter temperatures of all tours booked by customer 'Bill Smith'.

Discuss query optimisation heuristics, and using these heuristics develop a query tree which represents an efficient evaluation strategy for the developed query. *(8)*