



# M358/F

## Mathematics and Computing: A Third Level Course M358 Relational Databases

Wednesday 11 October 2000

2.30 pm – 5.30 pm

---

Time allowed: 3 hours

---

There are **TWO** parts to this paper and you should attempt **BOTH PARTS**. You should attempt **ALL** questions in Part I and **TWO** questions from Part II. Your answers to Part I questions should be written *in the spaces provided* inside this examination paper. Your answers to Part II questions should be written in the answer book(s) provided. Please begin each answer to a Part II question on a new page.

Part I carries 60% of the total examination marks; each question is worth 5 marks.

Part II carries 40% of the total examination marks; each question is worth 20 marks.

The marks for each part of each question are shown thus: [4].

### At the end of the examination

Check that you have written your personal identifier and examination number on this examination paper and on each answer book used. **Failure to do so will mean that your work cannot be identified.**

Attach this examination paper to the front of the answer book(s) in which you have answered questions from Part II with the paper fastener provided.

Examination Number								
Personal Identifier								

## PART I

Answer ALL questions from this part.

### Question 1

The three schema architecture described in the course involves **logical**, **storage** and **external** schemas. Complete each of the following sentences with the appropriate schema name in each underlined space required to make the sentence consistent with the architecture.

- (a) The data required by a user process is defined by a(n) \_\_\_\_\_ schema. [1]
- (b) A mapping is required from a(n) \_\_\_\_\_ schema to each of the associated \_\_\_\_\_ schemas. [2]
- (c) The organization of data on a disk is specified by a(n) \_\_\_\_\_ schema. [1]
- (d) Physical data independence means that a(n) \_\_\_\_\_ schema can be changed without affecting user processes. [1]

### Question 2

Figure 1 shows part of the E-R diagram for the Hospital data model used in the course.

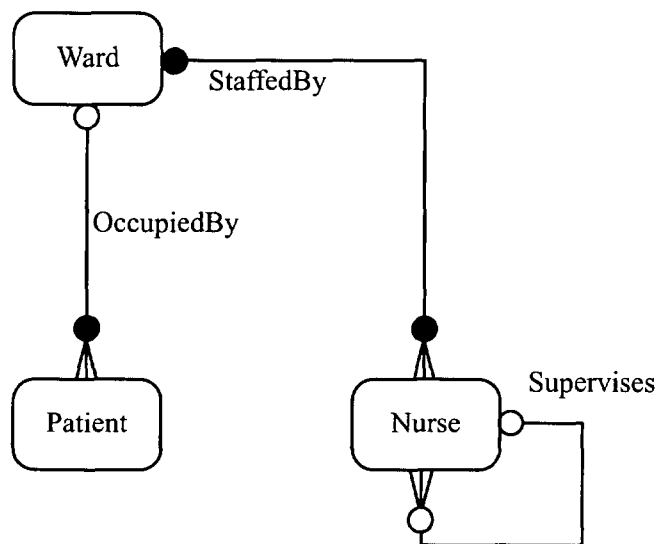


Figure 1 Part of the Hospital E-R diagram

Show how each of the sentences (a) to (e) express the meaning of a relationship in Figure 1 by writing one of the letters W, X, Y or Z in the underlined space in each sentence to indicate which of the following phrases is required:

- W zero or one  
X exactly one  
Y one or many  
Z zero, one or many

- (a) A ward is occupied by \_\_\_\_\_ patient(s). [1]
- (b) A ward is staffed by \_\_\_\_\_ nurse(s). [1]
- (c) A nurse staffs \_\_\_\_\_ ward(s). [1]
- (d) A nurse supervises \_\_\_\_\_ nurse(s). [1]
- (e) A nurse is supervised by \_\_\_\_\_ nurse(s). [1]

### Question 3

- (a) Consider the following proposed extension of a relation,  $R$ .

$R$

$A$	$B$	$C$	$D$	$E$
a	5	12	1	red
b	3	10	2	red
a	4	11	2	blue
b	2	10	1	unknown
b	null	14	3	blue
b	3	10	2	red
a	1	13	3	red

Explain why this extension is not valid.

[2]

- (b) Four relations,  $R$ ,  $S$ ,  $T$  and  $U$ , represent the three entity types,  $R$ ,  $S$  and  $U$ , and the relationships between these entity types. Relations  $R$ ,  $S$ ,  $T$  and  $U$ , have the following headings.

$R$  ( $A$ ,  $B$ ,  $C$ )

$S$  ( $D$ ,  $E$ ,  $F$ )

$T$  ( $G$ ,  $H$ )

$U$  ( $K$ ,  $L$ ,  $M$ )

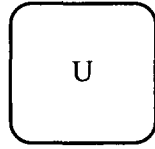
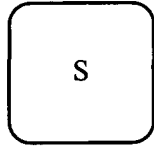
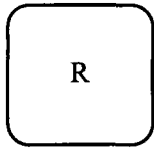
$A$ ,  $C$ ,  $D$ , ( $G$ ,  $H$ ) and  $K$  are each a candidate key in their respective relations.

Attributes  $A$  and  $F$  are defined on the same domain and represent a relationship between  $R$  and  $S$ .

Attributes  $C$ ,  $D$  and  $G$  are defined on the same domain and represent relationships between  $R$  and  $S$ , and  $S$  and  $T$ .

Attributes  $H$  and  $K$  are defined on the same domain and represent a relationship between  $T$  and  $U$ .

Complete the following E-R diagram by drawing the relationships between the entity types. Do *NOT* include participation conditions.



[3]

#### Question 4

Consider the following three relations.

<i>Student</i>	<i>Course</i>	<i>Studies</i>	
<i>StudentId</i>	<i>CourseCode</i>	<i>StudentId</i>	<i>CourseCode</i>
s05	c1	s05	c4
s07	c4	s05	c7
s09	c2	s07	c4
s10	c9	s09	c4
	c7	s09	c2
		s09	c7
		s10	c7
		s10	c4

For each of the following relational algebra expressions, give the relation that results from its execution

(a)

***select*** Studies ***where*** CourseCode = 'C2'

[1]

(b)

***project*** (***join*** Course ***and*** Studies ***where*** CourseCode = CourseCode) ***over*** CourseCode

[1]

(c)

(***project*** Studies ***over*** CourseCode) ***difference*** Course

[1]

(d)

*(project* Studies *over* CourseCode) *intersection* Course

[1]

(e)

*divide* Studies *by* Student *over* StudentId

[1]

### Question 5

The following table represents details of nurses, based on the Hospital example in the course text. For this question the sample data is given only to help your understanding of the questions.

#### Nurse

<i>staff_no</i>	<i>nurse_name</i>	<i>ward_no</i>	<i>supervisor_no</i>
126	Sesonske	w4	null
458	Galvin	w5	null
153	Cooke	w3	700
256	Locke	w4	523
823	Sadiq	w2	null
192	Wilkie	w2	823
206	Reay	w3	700
523	Descartes	w4	126
700	Birch	w3	null

For this table, write an SQL query to answer each of the following requests.

(a) Give the name and staff number of the nurses in wards w4 and w5.

[1]

(b) Give the name and staff number of the nurses in ward w2 who are supervised by a nurse with staff number 823.

[1]

(c) Give the name and staff number of the nurses who do not have a supervisor.

[1]

(d) Give the name and staff number of the nurses who supervise nurses on ward w3.

[2]



### Question 6

This question refers to the **nurse** table with the data given in Question 5.

- (a) Give an integrity constraint that ensures each value of **supervisor\_no** in the **nurse** table is a valid staff number.

[1]

- (b) Supposing you have defined the constraint specified in part (a), explain, with a reason, what happens when you execute the following statement:

```
DELETE FROM nurse WHERE staff_no = '126'
```

[1]

- (c) For this part of the question the **patient** table has the following columns and data.

*patient*

<i>patient_id</i>	<i>patient_name</i>	<i>consultant_no</i>	<i>ward_no</i>
p04	Anarkali	232	w3
p12	Hook	231	w2
p22	Parsi	232	w2
p31	Rubinstein	231	w4
p37	Seaward	232	w3
p38	Ming	110	w2
p67	Jarvis	232	w4
p73	Purdy	200	w3
p91	Fellows	232	w2

Give the request that the following query answers and the table that results from executing it.

```
SELECT n.ward_no, COUNT(*)  
FROM nurse n, patient p  
WHERE n.ward_no = p.ward_no  
GROUP BY n.ward_no
```

[3]

**Question 7**

- (a) Give an SQL statement to create a view called **nurse\_patient** based on the SQL query given in question 6(c).

[1]

- (b) Give two reasons why it is not possible to update the view **nurse\_patient**.

[2]

- (c) Give two differences between a view and a snapshot table

[2]

**Question 8**

The E-R Diagram in Figure 2 shows the relationships and participation conditions of the entity types for part of a conceptual data model.

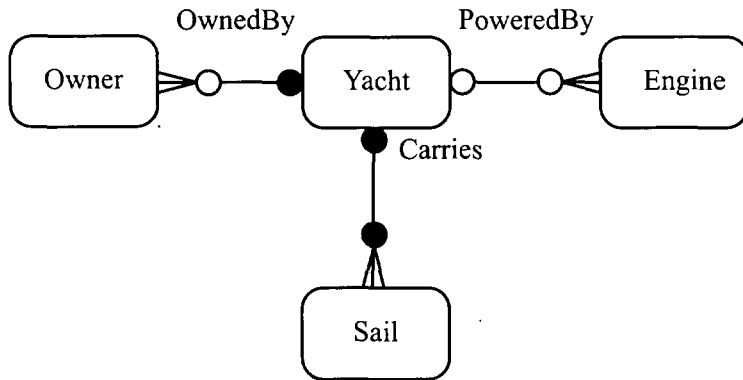


Figure 2 Example E-R diagram

- (a) For each of the following statements about occurrences of the entity types in this diagram, write 'Yes' if the statement can be represented according to the diagram or 'No' if it cannot be represented according to the diagram.
- (i) A group of friends share the ownership of a yacht. \_\_\_\_\_ [1]
- (ii) Mrs Onedin owns a luxury yacht which has twin engines and four sails as well as a yacht with a single sail and no engine. \_\_\_\_\_ [1]
- (b) Suppose the data requirements relating to the E-R diagram in Figure 2 are revised to include the statement: 'Some owners have a shared ownership in several yachts and the fraction of each yacht they own needs to be recorded.'

Give the additions or changes to Figure 2 that are needed to support this revision of the data requirements. Include the degree and participation conditions for relationships. (You do not need to redraw any part of the diagram that does not change.)

[3]

### Question 9

The following relation, which is not in BCNF, records data about parts. The parts are supplied in packs which are stored in locations in warehouses.

*Parts*(PartNumber, PackSize, PartDescription, Price, FloorNumber, StorageLocation)

The following functional dependencies hold for the attributes of this relation:

*PartNumber*  $\twoheadrightarrow$  *PartDescription*

*PackSize*, *PartNumber*  $\twoheadrightarrow$  *Price*

*PackSize*, *PartNumber*  $\twoheadrightarrow$  *FloorNumber*

*PackSize*, *PartNumber*  $\twoheadrightarrow$  *StorageLocation*

*StorageLocation*  $\twoheadrightarrow$  *FloorNumber*

Decompose this relation into an equivalent set of BCNF relations.

[5]

### Question 10

- (a) A company wants to record employee data in an SQL database, including the telephone extension numbers used to contact them. The telephone extension numbers allocated to employees are in the range 0100 to 1999 and the switchboard operator has the number 0001. The contact telephone extension number for each employee is in a column of a table where it is not the primary key or a foreign key and the recorded value is either the telephone number allocated to that employee or the number for the switchboard operator.
- (i) Give the most suitable data type for this column, and one reason to justify your choice.

[1]

- (ii) Give two properties, other than the data type, of the employee telephone extension number that might be taken into account when defining this column in the SQL CREATE TABLE statement for the table. (You do not need to write SQL statements.)

[2]

- (b) Write True or False for each of the following sentences to show whether it is a valid statement about triggers and the use of triggers.

Triggers are developed as part of an application process using a database. \_\_\_\_\_

A trigger must be defined when a column contains derived data. \_\_\_\_\_

The action for a trigger associated with the update of a column in a table is always fired before the update takes place. \_\_\_\_\_

A trigger may be associated with updates to specific rows of a table. \_\_\_\_\_

[2]

### **Question 11**

- (a) Identify the three properties of an index that are required for its creation (one of which is optional in a CREATE INDEX statement).

[3]

- (b) Give one advantage and one disadvantage of using an index.

[2]

**Question 12**

- (a) Suppose you have a requirement to include dates as a column in a table. Give three reasons why you should define this column using the DATE data type rather than a CHARACTER data type.

[3]

- (b) Explain why the DATE data type is important for data warehousing.

[2]

**PART II**

*Answer TWO questions from this part.*

### Question 13

The administrator of an IT training centre needs to maintain data about each member of staff working for the centre, and about each student attending a course at the centre. Staff working at the centre carry out a number of tasks, including organizing and teaching courses and managing other staff. Some staff may be attending as students, since the company encourages staff to enrol on its courses as part of its staff development programme. The relational model for representing the data requirements for the administration of courses is given in Figure 3.

**model** TrainingCentre

#### **domains**

CourseIdentifiers = c1..c9  
NamesOfCourses = string  
NamesOfPeople = string  
StaffIdentifiers = 1000..9999  
StudentIdentifiers = 1000..9999

#### **relation** Staff

StaffNo: StaffIdentifiers  
Name: NamesOfPeople  
ManagerStaffNo: StaffIdentifiers  
CourseCode: CourseIdentifiers  
StudentId: StudentIdentifiers

**primary key** StaffNo

{relationship Manages}

**foreign key** ManagerStaffNo **references** Staff **allowed null**

{relationship Teaches}

**foreign key** CourseCode **references** Course **allowed null**

{relationship Is}

**alternate key** StudentId

**foreign key** StudentId **references** Student **allowed null**

#### **relation** Student

StudentId: StudentIdentifiers  
Name: NamesOfPeople  
CourseCode: CourseIdentifiers

**primary key** StudentId

{relationship Attends}

**foreign key** CourseCode **references** Course **not allowed null**

#### **relation** Course

CourseCode: CourseIdentifiers  
Name: NamesOfCourses  
OrganiserStaffNo: StaffIdentifiers

**primary key** CourseCode

{relationship Organises}

**foreign key** OrganiserStaffNo **references** Staff **not allowed null**

**constraint (project** Course **over** CourseCode) **difference**  
**(project** Staff **over** CourseCode) **is empty**

Figure 3 The relational model of the administration of residential courses



- (a) Give an E–R diagram and entity types of the E–R data model that corresponds to the relational model given in Figure 3. [15]
- (b) For the relational model to be consistent with the requirements, the following constraints must be represented.
- (i) Staff who organize courses must manage staff.
  - (ii) Course managers must manage the staff teaching the courses that they organize.
- Give the modifications to the relational model in Figure 3 that represent each of these constraints. [5]

### Question 14

The Study database used in the course contains the table **assignment**, which has the columns **student\_id**, **course\_code**, **assignment\_no** and **grade**. An SQL query based on this table is as follows:

```
SELECT course_code, assignment_no, COUNT (*),
       CAST (AVG (grade) AS DECIMAL (3,1))
FROM assignment
GROUP BY course_code, assignment_no
```

- (a) Describe the logical processing for this query according to the model described in the course. [4]
- (b) Give the request that this query answers. [2]
- (c) Suppose there is a requirement for the result given by this query for a particular course code and assignment number. Give the definition of a procedure called **count\_grade** that has input parameters for a value of course code and assignment number and two output parameters that give the same result as the last two items in the SELECT clause of the query. [6]
- (d) Give a statement to show the use of the procedure you defined in your answer to part (c) for course code c2 and assignment number 2 and explain how the results obtained by executing this statement can be displayed. [2]
- (e) Explain why the requirement described in part (c) cannot be fulfilled by a function. [2]
- (f) Describe two benefits of using stored procedures. [4]

### Question 15

- (a) The data requirements for a financial services company are as follows.

The company is structured as a number of departments. Each of the services offered by the company is the responsibility of one department, and a department may offer many services. Currently there are 12 departments, with names like Fund Management, Personal Finance, etc., and each department is classified as being either Corporate or Personal. The services offered include Fund Management, Portfolio Advice and Personal Stock Management. Each service is described by the company as long, medium or short term investment and as high or low risk.

Each employee within the company is either a manager, a consultant or support staff, and no employee has more than one position in the company. The same information of name, address, contact telephone number and unique employee number is recorded for all employees. In addition, the management grade of each manager and the specialism and hourly charge rate for each consultant is recorded. For support staff their salary grade is recorded. Managers and consultants are given a finance code which they need to include on all financial documents they sign.

Each manager and each consultant work for a single department. A department must be managed by a single manager but not all managers manage a department. Each department has at least one consultant working in it. Support staff do not work in a specific department but are appointed to work in a department for periods of a minimum of 1 week. For internal budget purposes, records of the start and end date for each support appointment are kept.

Give an E-R model to represent these data requirements. The model should include an E-R diagram (showing the degree and participation conditions for all relationships), entity types definitions, constraints and any assumptions required to justify your model (you do not need to include assumptions about uniqueness of codes.) [15]

- (b) The company described in part (a) has further data requirements relating to sales.

Staff in a telephone call-centre receive calls from potential clients enquiring about the company's services. For legal reasons detailed records of each telephone call and the result it achieved (follow-up call requested, financial advisor interview appointment made, request for literature, no follow-up required, etc.) are recorded. In the call-centre all staff can see a large display that shows, for each department, the total number of client calls so far that day and the number of clients who have requested some follow-up.

- (i) Explain why the data being displayed is not included in the conceptual data model that represents the sales data requirements. [2]
- (ii) Give three benefits of creating a table that stores the data required for this display. [3]

**Question 16**

- (a) Describe the purpose of the SQL CONNECT statement. Explain why a user process does not include this statement when using a distributed database management system (DDBMS). [4]
- (b) Describe three different ways in which a DDBMS may be used to distribute the data in a table amongst the different locations of a distributed database. [6]
- (c) A DDBMS creates the optimum execution plan for a query for a distributed database in two stages. Explain what is done in each stage and describe the main criterion used to determine the optimum plan. [6]
- (d) Explain why single-phase commit is not acceptable for a distributed database and briefly describe the steps of two-phase commit. [4]

**[END OF QUESTION PAPER]**