

2	2	0	2	4	4	1	3
---	---	---	---	---	---	---	---

Academic year: 2002/23

Exam Session: Semester B

School/Dept: Computer Science

Module code: 4WCM2002

Module title: Data Modelling for Databases

Duration of Test: Starts 5 May at 9:00 and Ends 5 May at 13:00 (UK time)

INSTRUCTIONS TO CANDIDATES:

- Please make sure to include your student ID number on any sheets that you plan to submit.
- **All answers should be added to this document and uploaded to Canvas by 13:00 on May 5th.**
- Please note that email submissions to the module team will not be accepted.
- This paper consists of 4 questions.
- Answer ALL 4 questions.

QUESTION 1 – Entity-Relationship Modelling 1 (20 Marks)

The Hertfordshire Restaurant Guide website has an underlying database which provides up to date reviews of restaurants in Hertfordshire. Details about the restaurant are recorded, such as, its location, capacity, opening hours and the price range of its menu using a scale 1-4 (4 being expensive). The types of cuisines that restaurants can serve is recorded. The website has several reviewers, with various levels of experience, working for it. They visit restaurants regularly and may visit the same restaurant more than once, but never on the same day. These reviewers rank the restaurant against a few criteria on each visit. These criteria are:

- Quality of food
- Timeliness of service
- Cleanliness of restaurant
- Value for Money

Using the specification above, answer the following questions:

a) List 4 entities identified in the specification describing the objects they contain.

[8 marks]

b) Describe any two relationships defining for both, the:

- i. Name (in one direction)
- ii. Cardinality
- iii. Participation

ensure you detail the assumptions you make in each case

[8 marks]

c) Identify one of the following within any of the entities identified:

- i. A possible simple key with explanation
- ii. A possible composite key with explanation

[4 marks]

Your Answers: (pleased use as much space/pages as you need)

QUESTION 2 – Entity-Relationship Modelling 2 (20 Marks)

The Hatfield Apple Tree Suppliers (HATS) grow apple trees for supply to the public via an on-line ordering system. HATS produce many different varieties of trees for which there are details such as the variety of tree by name, when it flowers and when the fruit is available, plus, the size and colour of the fruit. Individual trees are grown of differing varieties and HATS maintains a stock of trees of differing ages and sizes of all varieties.

Customers details are recorded including contact details ie, address and phone number and personal details, i.e. DoB. Customers can place orders for individual trees from the list of trees in stock, each order potentially contains a set of order items, each order item details a single tree ordered.

Using the specification above, answer the following questions:

a) List 4 entities identified in the specification describing the objects they contain.

[8 marks]

b) Describe any two relationships defining for both, the:

- i. Name (in one direction)
- ii. Cardinality
- iii. Participation

ensure you detail the assumptions you make in each case

[8 marks]

c) Identify one of the following within any of the entities identified:


- i. A possible simple key with explanation
- ii. A possible composite key with explanation

[4 marks]

Your Answers: (pleased use as much space/pages as you need)

QUESTION 3 – Normalisation (30 Marks)

- a) Badgeston Stadium supports a number of cricket events throughout the summer, customers can purchase tickets for these events via an on-line ordering system. An invoice is emailed to the customer reflecting each order which can contain ordered for a number of tickets to a number of events. Here is an invoice example.

Badgeston Stadium 

Customer ID: B02309234

Customer Name: Michael Spooner
17, Maple Drive,
Streetley,
Yorkshire,
LE5 8DR

Test Match vs New Zealand – Day 1

July 1st 2023

Stand	Row	Seat	Unit Price
Midwicket	R	56	£50.00
Midwicket	R	57	£50.00
Midwicket	R	58	£50.00

Test Match vs New Zealand – Day 2

July 2nd 2023

Stand	Row	Seat	Unit Price
Pavilion	G	9	£74.00
Pavilion	G	10	£74.00
Pavilion	G	11	£74.00

Total £366.00

Using the details and invoice example above, the Un-normalised form (UNF) has been extracted as follows:

(Customer_ID, Customer_Name, Customer Address, (Event, Event_Date, (Stand, Row, Seat, Unit_Price)))

Identify and write out the First Normal Form (1NF) in relational schema form

- b) A car showroom employs a number of sales staff, who are able to sell any car to any customer, in fact a sale may be undertaken by more than one sales staff, in which case the sales commission is shared. A car is sold to a single keeper (customer) and the sales people earn commission for each sale as a proportion of the agreed price of sale, which may be less than the recommended retail price. The commission proportion varies for each car sold.

The following First Normal Form (1NF) has been derived

(Salesperson_ID, Salesperson_Name, Date_Sold, Car_ID, Car_Make, Car_Model, Car_Colour, Car_Retail_Price, Car_Sale_Price, Customer_ID, Customer_Name, Customer_Phone_Number, Commission)

Identify and write out the Second Normal Form (2NF) in relational schema form

- c) A logical model for a Lending Library has been developed. The database records which books have been loaned, creating a record for each loan and uniquely identifying it. The loan record also identifies, when, which book and by whom the loan was made for, with dates of the loan starting and how long the loan is for.

The First Normal Form (1NF) is as follows:

Loan (Book_ID, ISBN, Title_of_Book, Member_ID, Member_Name, Member Address, Member_Phone_No, Member_Join_Date, Loan_ID, Loan_Date, Length_of Loan)

If we know the Loan_ID, we can determine the Book and the Member that loaned it, so as this is a Simple Primary Key, it is also in 2NF.

Identify and write out the Third Normal Form (3NF) in relational schema form

- d) A scientist is working to determine different sugar content of various fruits. They record each experiment which involves taking a sample of a particular fruit and how much the sample weighs, then recording the findings of the experiment, which determines (by weight) how much of various types of sugars (fructose, sucrose, maltose, etc) the sample contains. Each record indicates the content of a specific sugar in a specific sample experiment.

The UNF is derived as follows:

(Experiment_No, Date of Experiment, Fruit_Type, Sample_Weight, (Sugar_Type, Sugar_Description, Weight_in Sample))

Identify and write out the First Normal Form (1NF) in relational schema form.

- e) Richard Osman's House of Games TV quiz show follows a particular format. There is a new competition each week with 4 competitors. Over 5 days (Monday to Friday), they play a number of games for points. On each day there are 5 games played in sequence. The order these games are played in is random, except for the last game each day (the fifth game each day) which is always the same (AnswerSmash). For the purposes of the database, we can record the name of these separate games as AnswerSmash1, AnswerSmash2, ... AnswerSmash5. The score of each individual game by each individual competitor needs to be recorded. Occasionally, a competitor may return for a second time to compete in a special competition, "Champion of Champions". The First Normal Form (1NF) has been identified as below:

Game_Score (Competition_No, Start_Date, Competitor_No, Competitor_Name, Competitor_Occupation, Game_Name, Day_Game_Played, Sequence_No_of_Game_on_Day, Score)

Identify, and write out in Relational Schema form, the Second Normal Form (2NF).

- f) Doggy Dips is a local company that offers swimming sessions for dogs. There are organised sessions in which up to 4 dogs can attend at any one time. They all last for 45 minutes. Owners can bring more than one dog at a time and can choose to swim with the dogs or not during a particular session. Each session is supervised by a company employee. The price of the session varies on the size of the dog and the popularity of the session.

Given the comments above and constructed First Normal Form (1NF) below:

Booking (Session Number, Date, Time, Supervisor_No, Supervisor Training_Level, Dog Number, Dog_Name, Breed_of_Dog, Age_of_Dog, Owner_Number, Owner_Name, Owner_Swimming, Price)

The Second Normal Form is evaluated as:

Session (Session Number, Date, Time, Supervisor_No, Supervisor_Name, Supervisor Training_Level)

Dog (Dog Number, Dog_Name, Breed_of_Dog, Age_of_Dog, Owner_Number, Owner_Name)

Booking (Session Number*, Dog Number*, Owner_Swimming, Price)

Identify and write out the Third Normal Form (3NF) in relational schema form.

Your Answers: (pleased use as much space/pages as you need)

QUESTION 4 – Relational Algebra (30 Marks)

Given the following set of table definitions:

Book (Book_Id, ISBN, Title, Publisher, Year)

Student (Student_Id, Student Name, Course Name, Age)

Author (Author_Name, Date of Birth, Email)

Borrow (Book_Id*, Student_Id*, Borrowed on)

Wrote (Book_Id*, Author_Name*)

Classification (Book_Id*, Genre_Name*)

Genre (Genre_Name, Description)

Produce the relational algebra definitions (NOT SQL) for the following queries:

a) List all students, by name and age, who are on course "Further Mathematics"

[4 marks]

b) List the name and course name of all students who borrowed books in the "Sci-Fi" genre.

[7 marks]

c) List author names and the book title written, for those authors who had a book published between 2018 and 2021 inclusively. Note: Any sensible date format can be used [5 marks]

d) List the students (by name) on the course "Database Concepts", who have borrowed a book written by any author born in 1988. We only require books that were published after 2015 and that are categorised as "Computing".

[8 marks]

e) List all titles, genre and descriptions of books written by authors born between 1st September 1975 and 24th March 1982 inclusively

[6 marks]

Your Answers: (pleased use as much space/pages as you need)

Question 4

Relational Algebra

- a. $\pi_{\text{Student Name, Age}} (\sigma_{\text{course Name} = \text{"Further Mathematics"}} (\text{Student}))$
- b. $\pi_{\text{Student Name, course Name}}$
 $\sigma_{\text{Genre_Name} = \text{"Sci-Fi"}}$
 $(\text{Student} \bowtie \text{student ID}$
 $\text{Borrow} \bowtie \text{Book ID Classification})$
- c. $\pi_{\text{Author_Name, Title}}$
 $\sigma_{\langle \text{year} \rangle = 2018 \wedge \text{year} \leq 2021}$
 $(\text{Author} \bowtie \text{Author Name}$
 $\text{Wrote} \bowtie \text{Book ID Book})$
- d. $\pi_{\text{student name}}$
 $\sigma_{\langle \text{course Name} = \text{"Database Concepts"} \wedge$
 $\text{Date of Birth} = \text{"01/01/1988"} \wedge$
 $\text{Date of Birth} = \text{"30/12/1988"} \wedge$
 $\text{year} > \text{"2015"} \wedge$
 $\text{Genre_Name} = \text{"Computing"} \rangle$
 $(\text{Student} \bowtie \text{student ID}$
 $\text{Borrow} \bowtie \text{Book ID}$
 $\text{Book} \bowtie \text{Book ID}$
 $\text{Classification} \bowtie \text{Book ID}$
 $\text{Wrote} \bowtie \text{Author Name Author})$

e. Π Title, Genre Name, Description

\bigcirc \langle Date of Birth $\rangle =$ "1st September 1975" \wedge

Date of Birth \leq "24th March 1982" \rangle

(Author \bowtie Author Name

Wrote \bowtie Book Id

Book \bowtie Book Id

Classification \bowtie Genre Name Genre)

Q3 Normalization

-a-

(Customer-ID, Customer-Name, Customer-Address,
Event, Event-Date, Stand, Row, Seat,
Unit-Price)

-b-

Salesperson (Salesperson-ID, Salesperson-Name, Date-Sold)
Car (Car-ID, Car-Make, Car-Model, Car-colour,
Car-Retail-Price, Car-Sale-Price,
Customer-ID*)

Customer (Customer-ID, Customer-Name, Customer-
Phone-Number, ~~Commission~~)

Commission (Customer-ID*, Salesperson-ID*,
Commission)

-c-

Loan (Loan-ID, Loan-Date, length-of-Loan,
Member-ID*, Book-ID*)

Member (Member-ID, Member-Name, Member-Address,
Member-Phone-Number, ~~Member-Phone-Number~~ Member-Join-Date)

Book (Book-ID, ISBN, Title-of-Book)

- d -

(Experiment - No, Date of Experiment, Fruit Type,
Sample Weight, Sugar Type, Sugar Description,
Weight in Sample)

- e -

Competitor (Competitor No, Competitor - Name,
Competitor Occupation).

Game Score (Game Name, Day Game Played,
Sequence No of Game On Day, Score,
competitor - No*)

Competition (Competition - No, Start Date,
Competitor - No*)

- f -

Session (Session - Number, Date, Time, Supervisor - No*)

Supervisor (Supervisor - No, Supervisor - Name,
Supervisor - Training - Level)

Dog (Dog - Number, Dog - Name, Breed of - Dog,
Age of Dog, owner - Number*)

Owner (Owner - Number, owner - name, owner -
swimming)

Booking (Session - Number*, Dog - Number*,
Owner - Number*, Price)

Q1.

a. 1. Restaurant

(Location, capacity, opening hours, Price Range, ^{Name} Restaurant)

2. Menu

(Price Range, Types of cuisines, Restaurant)

3. Reviewer

(Name, Level of Experience)

4. Ranking Criteria

(Timeliness of Service, Quality of food, cleanliness of restaurant, value for money)

b. Relationship 1:

Name: The relationship between reviewer and restaurant is "visits".

Reviewer $\xrightarrow{\text{"visits"}} \text{Restaurant}$.

This relationship's cardinality is

1 - to - many as one reviewer can visit a restaurant multiple time just not on the same day.

The participation is that the reviewer "may" visit one restaurant multiple times.

Relationship 2:

Name: The relationship between Restaurant and Ranking Criteria is that all restaurants are "ranked according to" a singular ranking Criteria. -X- continued on next page.

c. Simple / Primary Key :- The Primary key is "Name" in the restaurant entity as restaurant names are unique and hence, can uniquely identify the rows. This kind of primary key is called a "Natural key".
Composite key :-

•X continued Part (b) of Q1
Hence, Cardinality :- many to 1

Restaurant $\xrightarrow[\text{according to}]{\text{ranked}}$ Ranking Criteria.

Participations- All restaurants "must" be ranked according to a specified ranking criteria.

Q2

a. 1. Customer

(Customer-ID, Contact, Address, Phone-No, DOB)

2. Order

(Order-ID, Order-Date, status, Total, Customer-ID*)

3. Fruit

(Fruit-ID, Fruit-Name, Fruit-Size, Fruit-Colour)

4. Tree

(Tree-ID, Variety, Flowering, Fruit-ID*)

b. Relationship 1:

Name: Fruit $\xrightarrow[\text{"grows on"}]{} \text{Tree}$

Cardinality:- Multiple fruit ~~can~~ grows on one tree hence cardinality is many-to-1.

Participation:- Naturally, Many fruit "must" grow on a single ~~tree~~ ^{tree} for feasibility of business.

Relationship 2:

Name: Customer $\xrightarrow[\text{"Places"}]{} \text{Order}$

Cardinality:- One customer can place many orders hence cardinality is 1-to-many.

Participation: A customer "may" place multiple orders.

- C. In customer entity Customer ID is a Primary / Simple key as Customer IDs are unique and can identify each row uniquely. This kind of primary key is called a surrogate key.

Composite Key:-