

1. The following schema is used to record information about student option choices:

Course (courseId, courseName)

Option (optionId, optionName, level)

Availability (optionId, courseID, maximumNumbers)

Teaching (optionId, roomNo, capacity)

Course contains information about named degree paths, Option contains individual option modules and Availability indicates which option runs on which course, along with the maximum number of students that the option can have. The Teaching table indicates which option has been scheduled in which room along with the maximum number of students possible.

(a) Staff members may teach more than one option, a single option may be taught by many staff members and students may take up to four option choices. Explain how the schema can be altered to store this information whilst maintaining 3NF. Write down any assumptions you have made. **(8 marks)**

(b) Two queries are required from the student option schema. The first must list the option names, level and room numbers for all options that will be taught on the BSc Business Information Technology course. The second must show all option choices that have not been selected by any students. Express these reports in Relational Algebra and then implement them in SQL. **(8 marks)**

(c) The results of the first query will be used as a base for a new query that will generate options that have more students than the capacity of the room. Construct this as a (non-optimised) query and then optimise it using a query tree. Write down any assumptions you have made. **(9 marks)**

~~(d) Temporal extensions to the database will have to be manually implemented as they are not naturally supported by the database. Redraft the structure and then use example queries to show how valid time and transaction time support could be implemented. **(8 marks)**~~

~~2. (a) A small cinema wishes to use a ticket booking system with a web interface. Use diagrams to explain what happens when a customer requests a ticket from an HTML form. You should describe the relationship between the browser and the web server, along with interactions between the server and the database backend. Use sample code extracts or pseudo code where appropriate. **(16 marks)**~~

~~(b) Previous versions of mySQL did not support transactions (nor the rigours of the ACID test). Using examples from the booking schema, explain the ACID test, the difficulties that this may introduce and possible solutions. **(17 marks)**~~

Continued ...

3. (a) Using the following deductive rules and facts about railway stations, write and evaluate a query to answer “Is Kingston station connected to Raynes Park station?”. Write down any assumptions you have made. **(8 marks)**

```
nextto(Kingston, Norbiton) ←  
nextto(Norbiton, NewMalden) ←  
nextto(NewMalden, RaynesPark) ←  
nextto(Kingston, HamptonWick) ←  
connected(x,y) ← nextto(x,z),connected(z,y)  
connected(x,y) ← nextto(x,y)
```

- (b) The following rule is added:

```
nextto(y,x) ← nextto(x,y)
```

Explain the differences between NAF and CWA and show how the evaluation of the query “is Hampton Wick connected to Norbiton?” can result in different conclusions.

(8 marks)

- (c) Using the appropriate notation, model the following situation: Homer likes beer and TV. Marge likes bowling and cooking. Lisa likes school and Bart likes TV. Bart does not like school, nor anything that Homer likes except TV. Lisa likes everything that Marge likes but not that Grandpa likes. Grandpa likes everything that Homer likes.

Use the situation to show the evaluation of a safe and an unsafe query. Add rules or facts for your example as necessary. **(8 marks)**

- ~~(d) Using examples, explain how the deductive notation can be applied to integrity constraints, as seen in Active Databases. **(9 marks)**~~

Continued ...

4. (a) Write down and explain the formal definitions of 3NF and BCNF and then use them to show whether the following schema is 3NF or BCNF. Why would a full proof be difficult to show? **(12 marks)**

$$R = (ABCDE)$$

$$F = \{A \rightarrow BD, A \rightarrow C, D \rightarrow A, C \rightarrow E, E \rightarrow A\}$$

- (b) With the tuple instances $\{\langle a, b, c, d, e \rangle, \langle a, b1, c1, d, e \rangle\}$ explain whether a decomposition into ABC and ADE is a lossless join decomposition or a dependency preserving decomposition. **(9 marks)**

- (c) ~~Given the relation scheme~~

~~$$R = (ABCDEF)$$~~

~~$$X = \{A \rightarrow BCD, B \rightarrow E, BC \rightarrow AF, *[BC, BDEF, A, CD]\}$$~~

~~Using a formal definition show why this scheme fails the 5th Normal Form requirement. Explain a change to the scheme that could place it in 5th Normal Form.~~

~~**(12 marks)**~~

5. The following relationships exist in a large family:

Peggy is the mother of Phil, Grant and Sam. Pauline and Arthur are the parents of Mark, Martin and Michelle. Michelle and Den are the parents of Vicky. Michelle and Grant are the parents of Mark Jr. Mark married Lisa who had a child, Louise with Phil. Phil was married to Cathy (before Lisa) and had a son Ben.

Model this information using relational and deductive approaches. What are the advantages and disadvantages of each approach? Show the problems (or otherwise) of evaluating a query requiring a list of a given parent's children and grandchildren. **(33 marks)**

END OF EXAMINATION

1. A relation scheme for a video rental shop has the following design:

FILM Film No, Film Name, Film Director, Date of Release, Principle Star, Category.

CUSTOMERS Customer No, Customer Name, Address, Phone.

RENTALS Film No, Customer No, Date Rented

Field descriptions as follows

Film No Key field for FILMS.

Film Name Full Title of film

Film Director Full name of Film Director

Date of Release Date of release on video

Principle Star Name of Actor listed first in credits

Category Type of film i.e. Horror, Science Fiction, Romance, Thriller

Customer No Key field for CUSTOMERS

Customer Name Name of rental customers

Address Address of rental customers

Phone Phone number of rental customers

Film No Key field from FILMS

Customer No Key field from CUSTOMERS

Date Rented Date film rented

- (a) Write the following queries using relational algebra. If a query cannot be expressed, explain why. Write down any assumptions you have made. **(8 marks)**
- (i) List all customer names and addresses
 - (ii) List all film names currently rented
 - (iii) List all film directors in alphabetical order
 - (iv) List the principle stars of films rented by 'Mr Smith'
- (b) Optimise the following queries, showing each query in a parse tree. Clearly explain any heuristic you use for this purpose. **(8 marks)**
- (i) List all Films Names rented by people living in Norbiton
 - (ii) List all Romance films rented by people whose name is "Mr Smith"
- ~~(c) The scheme only stores information on films that are currently being rented. Explain changes that could be implemented to allow storage of information on previous rental activities. Use your discussion to explain the difference in the implementations of the queries "List all Films currently rented by 'Mr Smith'" and "List all films ever rented by 'Mr Smith'" **(9 marks)**~~

Continued ...

2. (a) Using suitable predicates, represent the following information as a definite database:
Lou is the mother of Pauline and Pete. Pauline and Arthur are the parents of Michelle, Mark and Martin. Arthur, Mark, Martin and Den are male. Pauline, Lou, Victoria and Michelle are female. Michelle and Den are the parents of Victoria. If you are a male parent you are a father. If you are a female parent you are a mother. **(6 marks)**

- (b) The following rule is used to define brother-
 $\text{brother}(x,y) \leftarrow \text{mother}(z,x), \text{mother}(z,y)$

Use this rule to evaluate 'who are the brothers of Mark'. Show the answer being evaluated using an SLD tree (leftmost selection) and then comment on the validity of the rule definition. **(8 marks)**

- (c) Define a rule that describes a grandparent in terms of parent and then show the Naïve evaluation of the query 'who are the grandparents of Victoria'. **(6 marks)**

- ~~(d) Describe an active database, using your answer to explain why deductive databases have not been as popular as previously predicted. **(5 marks)**~~

3. (a) Use a formal definition of 3rd Normal Form to analyse the following scheme and then use this to provide an explanation of the common definition of 3rd Normal Form which involves the use of the term 'transitive dependency'. **(6 marks)**

$$R = (ABD)$$
$$F = \{A \rightarrow B, B \rightarrow D\}$$

- (b) Show that the following scheme is in 3NF but not in BCNF. Explain why F1 can be tested for 3NF without having to generate the closure of the functional dependency set. **(6 marks)**

$$R1 = (ABC)$$
$$F1 = \{AB \rightarrow C, C \rightarrow A\}$$

- ~~(c) Explain the difference between a synthesis algorithm and a decomposition algorithm. Why are synthesis algorithms regarded as being inefficient? **(6 marks)**~~

- ~~(d) Given the relation scheme~~

$$R = (ABCDEF)$$
$$X = \{A \rightarrow BCD, B \rightarrow E, BC \rightarrow AF, *[BC, BDEF, A, CD]\}$$

~~Use a definition to show why this scheme fails the 5NF requirement. Explain a change to the scheme that could place it in 5th Normal Form. **(7 marks)**~~

Continued ...

4. Consider the following database D

married_to(bob, joy) \leftarrow
person(bob) \leftarrow
person(joy) \leftarrow
person(tom) \leftarrow
married_to(x,y) \leftarrow married_to(y,x)

- (a) Draw SLDNF trees to show the safe leftmost and unsafe leftmost evaluations of the following query on D **(7 marks)**
 $\leftarrow \sim \text{married_to}(x,y)$
- (b) Briefly explain why the result produced by the unsafe evaluation of the previous query is a counter-intuitive one. **(5 marks)**
- ~~(c) Use an example to describe Clarks Soundness rule for safe computation? **(5 marks)**~~
- (d) Suppose the definite query, married_to(tom, tom) were evaluated on D using the following methods
- SLD-resolution with Negation as Failure
 - SLD-resolution with a Closed World Assumption
- Explain why the answers produced by these queries are different. **(8 marks)**

5. ~~(a) Explain the four integrity rules that should hold in a relational database system **(8 marks)**~~
- ~~(b) Describe the Bell-LaPadula security model. Include an explanation of the following Axioms **(8 marks)**~~
- ~~Simple Security property~~
 - ~~Star property~~
 - ~~Tranquillity principle~~
 - ~~Discretionary security property~~
 - ~~Non-accessibility of inactive objects~~
 - ~~Rewriting of inactive objects~~
- ~~(c) Explain Polyinstantiation using an example relation with sample data. **(9 marks)**~~

Continued ...

1. (a) Write a formal definition of 3rd Normal Form and explain the difference between this and Boyce-Codd Normal Form. Why are most organisations satisfied with normalisation to 3NF or BCNF? **(5 marks)**
- (b) Explain why the following relation scheme could be in 3NF. Is it in BCNF? **(7 marks)**
 $R = (ABCD)$
 $F = \{A \rightarrow BD, A \rightarrow C, D \rightarrow A\}$
- (c) What difference does it make if a relation scheme is in right reduced form? **(7 marks)**
- ~~(d) Explain the properties of a join dependency using an example. **(6 marks)**~~

2. (a) Write the following statements as deductive database facts and rules. **(8 marks)**
 An actor gets a credit in a movie if they starred in it. A successful actor gets top billing. Bruce Willis is a successful actor who starred in the movies Unbreakable and The 6th Sense. Samuel L Jackson starred in the movies Unbreakable and Shaft.
- (b) Given the following rules and facts, formulate the query “is it true that barbara does not fly?”. Use a query tree to evaluate the expression, using a safe leftmost computation rule. **(8 marks)**
 $\text{bird}(\text{brenda}) \leftarrow$
 $\text{bird}(\text{barbara}) \leftarrow$
 $\text{penguin}(\text{brenda}) \leftarrow$
 $\text{insect}(\text{bruce}) \leftarrow$
 $\text{flies}(x) \leftarrow \text{bird}(x), \sim \text{abnormal}(x)$
 $\text{abnormal}(x) \leftarrow \text{penguin}(x)$
- (c) Explain what happens in an evaluation if the computation rule is not safe. **(5 marks)**
- (d) Use examples to explain the difference between Negation as Failure and the Closed World Assumption. **(4 marks)**

- ~~3. (a) Every transaction in a database should pass the ACID test. Explain what this term means. **(5 marks)**~~
- ~~(b) Use examples to explain the difference between the following 'optimistic' concurrent access control mechanisms. Explain any assumptions you have made. **(10 marks)**~~
~~Transaction scheduling~~
~~Exclusive locking with 2PL and WAIT-DIE deadlock assumption policy~~
~~Shared locks with immediate lock release and WOUND-WAIT policy~~
- ~~(c) Define the term Deadlock and explain the criteria that could be used to determine which transactions should be rolled back. **(10 marks)**~~

Continued ...

4. A bank uses the following relational scheme for its customer details

ACCOUNT (account-type, interest-rate, minimum-deposit, withdrawal-notice)

CUST (cust-id, address1, address2, city, phone, occupation)

BALANCE (cust-id, account-type, balance)

CUST stores information about each customer, BALANCE stores the current balance, ACCOUNT stores information about each type of account.

- (a) Define a relational algebra expression that will return the cust-id, name, address1, address2, city, balance, account-type, withdrawal-notice attributes, for accounts that have greater than £100 in them. **(4 marks)**
- (b) A special financial offer will be made to people having accounts with a withdrawal period of 30 days or less. Write a relational algebra expression that will return the appropriate names and addresses. **(4 marks)**
- (c) Use query trees to explain how the expression from (b) could be optimised. Clearly show each stage and explain your working along with any assumptions you have made. **(8 marks)**
- (d) The bank wishes to store audit information so that previous balance information could be retained. Describe ways in which this could be implemented. **(9 marks)**
5. ~~(a) What would be the principle requirements for any system which integrates database applications with the Web? **(5 marks)**~~
- ~~(b) MySQL does not have support for transactions. Use examples to explain the consequences of this. **(5 marks)**~~
- ~~(c) Explain the advantages of a three tier web based system over a traditional two tier style system. Use Apache, PHP and MySQL sample code to explain how pages would be generated in a dynamic environment. **(15 marks)**~~
6. ~~(a) Describe the features proposed in the Object Oriented Database System Manifesto. **(10 marks)**~~
- ~~(b) Using examples, describe how SQL has been extended to enable processing on an Object Oriented Database. What other features are present in SQL3? **(10 marks)**~~
- ~~(c) Describe how several Object Oriented Database products have been repositioned as XML repositories. **(5 marks)**~~

Continued ...

4. A garage is to use a relational database to store information on jobs and mechanics for costing purposes as follows

Each mechanic has a unique identifier and competence level, along with the typical information stored in a personnel table such as address, phone number etc. Each car has a make, model, registration no and owner. A booking consists of work on a single car and may consist of several jobs. A job is performed by a specific mechanic with an associated labour cost and time.

- (a) Define a relational scheme for the garage justifying the level of normalisation used. Clearly explain any assumptions you have made. **(10 marks)**
- (b) Create relational algebra expressions for the following queries, and use your answers to explain how common relational algebra operators are performed in SQL. If a query cannot be expressed in relational algebra, explain why.
- (i) Find the owner of a car worked on by a specified mechanic at a specific date.
 - (ii) Find the total cost of the repair for one car.
 - (iii) Discover if any mechanic works on every car.
- (12 marks)**
- (c) Use query trees to explain how the expression from (i) could be optimised. Clearly show each stage and explain your working along with any assumptions you have made. **(6 marks)**
- ~~(d) The garage wishes to store previous repair information for service histories. Explain ways in which this could be implemented. **(5 marks)**~~

5. ~~(a) Every transaction in a database should pass the ACID test. Explain what this means and why this is especially important for a data repository that has a web interface. What other factors would influence the choice of a database in such a system? **(15 marks)**~~
- ~~(b) An artist is looking to preview and sell items over the internet. Design an architecture (i.e. protocols, scripting language and database) that could be used. You should justify your design by outlining the advantages of your structure and use an example transaction to show how the technologies interact. **(18 marks)**~~

Continued ...

6. (a) Write the following statements as deductive facts and rules. **(10 marks)**

Mary, Mungo and Midge live together in a flat. Mary is a girl and Midge is a mouse. If it is morning, it is time for a walk. Dogs are only allowed out for a walk if their owner is with them. If Mungo doesn't watch Midge, Midge can get lost on a walk. Mungo is a dog owned by Mary.

- (b) Given the following rules and facts, formulate the queries “are there any undergraduates?” and “is Bob an undergraduate?”. Use query trees to evaluate your expressions, once with an unsafe leftmost computation rule and once where the rule is safe. Explain what happens. **(10 marks)**

```
postgrad(mary) ←  
student(bob) ←  
student(mary) ←  
student(bruce) ←  
undergrad(x) ← ~postgrad(x), student(x)
```

- (c) Use examples to compare the use of the Closed World Assumption in a relational database to Negation As Failure as used in a deductive database. **(13 marks)**

Continued ...

1. A garage is to use a relational database to store information on jobs and mechanics for costing purposes as follows

Each mechanic has a unique identifier and competence level, along with the typical information stored in a personnel table such as address, phone number etc. Each car has a make, model, registration no and owner. A booking consists of work on a single car and may consist of several jobs. A job is performed by a specific mechanic with an associated labour cost and time.

- (a) Define a relational scheme for the garage justifying the level of normalisation used. Clearly explain any assumptions you have made. **(8 marks)**

- (b) Create relational algebra expressions for the following queries, and use your answers to explain how common relational algebra operators are performed in SQL. If a query cannot be expressed in relational algebra, explain why.

- (i) Find the owner of a car worked on by a specified mechanic at a specific date.
- (ii) Find the total cost of the repair for one car.
- (iii) Discover if any mechanic works on every car.

(6 marks)

- (c) Use query trees to explain how the expression from (i) could be optimised. Clearly show each stage and explain your working along with any assumptions you have made. **(6 marks)**

- ~~(d) The garage wishes to store previous repair information for service histories. Explain ways in which this could be implemented. **(5 marks)**~~

- ~~2. (a) An artist is looking to preview and sell items over the internet. Design an architecture (i.e. protocols, scripting language and database) that could be used. You should justify your design by outlining the advantages of your structure and use an example transaction to show how the technologies interact. **(15 marks)**~~

- ~~(b) The current (stable) version of mySQL does not fully support transactions, subqueries or standard SQL-92 implementations of set operations (such as UNION, INTERSECTION etc). Explain each of these, showing the problems that these deficiencies cause, along with possible solutions. **(10 marks)**~~

Continued ...

3. (a) Using the following deductive rules and facts, write and evaluate a query to answer “who are vicki’s uncles?”. Ensure that all assumptions are clearly stated. **(10 marks)**

uncle(x,y)← parent(z,y),brother(x,z)
auntie(x,y)← parent(z,y),sister(x,z)
married(x,y)← married(y,x)
brother(mark,martin)←
brother(mark,michelle)←
brother(martin,michelle)←
brother(robbie,sonia)←
parent(pauline,mark)←
parent(pauline,martin)←
parent(pauline,michelle)←
parent(michelle,vicki)←
parent(martin,chleo)←
parent(sonia,chleo)←
parent(mark,louise)←
married(mark,lisa)←

- (b) Use a query tree to explain why the evaluation of “who are louise’s Uncles?” fails to produce the correct result. What changes would be required for this to correctly evaluate? **(10 marks)**

~~(c) Using the deductive notation, write an integrity constraint to state that a brother must never marry his sister. Explain your answer. **(5 marks)**~~

4. (a) Write a formal definition of 3rd Normal Form and explain the difference between this and Boyce-Codd Normal Form. Why are most organisations satisfied with normalisation to 3NF or BCNF? **(5 marks)**

- (b) Explain why the following relation scheme could be in 3NF. Is it in BCNF? **(7 marks)**

~~$R = (ABCD)$
 $F = \{A \rightarrow BD, A \rightarrow C, D \rightarrow A\}$~~

- (c) What difference does it make if a relation scheme is in right reduced form? **(7 marks)**

~~(d) Explain the properties of a join dependency using an example. **(6 marks)**~~

Continued ...

~~5. (a) Evaluate the features proposed in the Object Oriented Database System Manifesto. (12 marks)~~

~~(b) Using examples, explain the proposed extensions in SQL3 and SQL4 that will allow manipulation of alternative data models. (13 marks)~~

6. Consider the following database D

married_to(bob, joy) \leftarrow

person(bob) \leftarrow

person(joy) \leftarrow

person(tom) \leftarrow

married_to(x,y) \leftarrow married_to(y,x)

(a) Draw and explain an evaluation tree showing the unsafe leftmost evaluation of the following query on D (7 marks)

$\leftarrow \sim \text{married_to}(x, y)$

(b) Briefly explain why the result produced by the unsafe evaluation of the previous query is a counter-intuitive one. (5 marks)

(c) Suppose the definite query, married_to(tom, tom) were evaluated on D using the following methods

SLD-resolution with Negation as Failure

SLD-resolution with a Closed World Assumption

Explain why the answers produced by these queries are different. (8 marks)

~~(d) Explain the reasons why there has been a slow uptake of the deductive model in the marketplace. (5 marks)~~

Continued ...

7. Taught modules in Kingston Business School have prerequisites that indicate which modules a student must have been passed before the next module may be begun. The following relationships exist:

Information Systems for Business 1 (Year 1) and Software Development (Year 1) are prerequisites for Information Systems for Business 2 (Year 2)

Software Development (Year 1) is a prerequisite for Business PC (Year 2)

Business Mathematics (Year 1) is a prerequisite for Statistics (Year 2) and Management Science (Year 2)

Information Systems for Business 2 and Business Mathematics are prerequisites for Database Systems (Year 4)

Software Development (Year 1) and Business PC (Year 2) are prerequisites for Computing Environments (Year 4)

Model this information using relational and deductive approaches. What are the advantages and disadvantages of each approach? Show the problems (or otherwise) of evaluating a query requiring a list of a given modules prerequisites. **(25 marks)**

END OF EXAMINATION