Birkbeck

(University of London)

BSc/FD EXAMINATION

Department of Computer Science and Information Systems

Database Management (COIY028H6)

CREDIT VALUE: 15 credits

Date of examination: 30 May 2018
Duration of paper: (10:00–12:00)

There are **five** questions on this paper.
Answer only **four** of the five questions.
If you answer more than four questions, only the best four answers will count.
Each question carries **25** marks in total.
The paper is not prior-disclosed.
The use of electronic calculators is not permitted.

1. (a) Consider a relationship type R between entity types E_1 and E_2 in the entity-relationship (ER) model. Assume that the primary keys of E_1 and E_2 are K_1 and K_2, respectively.

   i. Name the different multiplicity constraints that can be placed on (binary) relationship types in the entity-relationship model.

   (3 marks)

   ii. For each different multiplicity constraint on R, describe how R might be translated in to an equivalent relational design (remember to include information about primary keys).

   (6 marks)

   (b) Assume that a restaurant offers "meal deals" to its customers. There are starters, main dishes and drinks. For a meal, you choose one of each type. Assume that the ordered meals are stored in a relation *Meals* as follows:

(N)umber	(D)ish	(T)ype
1	soup of the day	starter
1	roast duck	main
1	red wine	drink
2	garlic mushrooms	starter
2	butternut casserole	main
2	red wine	drink
...	...	...

   We have the following functional dependencies (where attributes have been abbreviated to their initial letters): $NT \rightarrow D$ and $D \rightarrow T$. That is, each order can have only one dish of any type, and the dish determines the type.

   i. What are the (candidate) keys of the relation *Meals*?

   (2 marks)

   ii. Is the *Meals* relation schema in Third Normal Form? Justify your answer.

   (3 marks)

   iii. Is the *Meals* relation schema in Boyce-Codd Normal Form? Justify your answer.

   (3 marks)

   iv. What redundancy problem (if any) exists in the *Meals* relation schema?

   (2 marks)

   (c) Give three limitations of the relational model which have led to alternative data models being proposed. Name three of these alternative models.

   (6 marks)

2. (a) With the help of a diagram, describe the three levels of abstraction provided by database management systems, highlighting the advantages to users and/or application programs of having these levels of abstraction.

(8 marks)

(b) Consider the following relation representing information about lecturers, the courses they teach, and the textbooks recommended for the courses:

Lecturer	Course	Textbook
AC	ISM	BMF
PW	DM	UW
PW	DM	SKS
RK	ISC	BMF

i. Give three functional dependencies *violated* by the relation.

(3 marks)

ii. Give two functional dependencies *satisfied* by the relation.

(2 marks)

iii. For each of the functional dependencies satisfied by the relation, state whether or not you would expect them to *hold* on the relation schema. Explain your answers.

(4 marks)

(c) One of the properties of database transactions is that they should execute *atomically*. What is meant by the term "atomically"? Use an example, consisting of a specific sequence of instructions, to illustrate what can go wrong if a transaction does not execute atomically.

(8 marks)

3. (a) Consider a database which is to contain personal information about people. There is to be an attribute ID which will uniquely identify each person, as well as attributes such as Name, Phone and Email. Suppose that, for each person who is married, you would like to record who their spouse is as well as their spouse's personal information (e.g., Name, Phone and Email). Suggest two different ways of modelling this situation in a relational database (i.e., different relational database designs), stating which you think would be preferable and why. Remember to include information about constraints in your designs.

(7 marks)

(b) Explain what the HAVING clause is used for in SQL.

(4 marks)

(c) SQL specifies what restrictions must be placed on view definitions in order for them to be *updatable*.

i. Explain in general why there is a potential problem with updating a database through a view.

(2 marks)

ii. Describe three of the syntactic restrictions imposed by SQL.

(6 marks)

(d) Database systems keep various statistics about the data stored in them in order for the query processing system to determine an efficient query execution plan for each query. Two such statistics are the number of rows in a relation, and the number of distinct values for a column in a relation. Explain how these statistics are used in estimating the size of the result of the following queries:

i. `select * from R where R.A = 'value';`
where `A` is a column of relation `R` and `value` is a value.

(2 marks)

ii. `select * from R, S where R.A=S.B;`
where `S.B` is a foreign key referencing `R.A`

(2 marks)

Explain what assumption is made when using the number of distinct values to estimate the size of results, as well as what can happen when this assumption is incorrect.

(2 marks)

4. (a) Explain why *aliases* are useful in SQL query expressions, as well as what syntax is used to introduce them.

(3 marks)

(b) Let U be a relation schema with $schema(U) = ABCDEI$. Let F be the following set of functional dependencies (FDs) over U:

$$AB \rightarrow C$$
$$BE \rightarrow I$$
$$E \rightarrow C$$
$$CI \rightarrow D$$

i. Compute the *closure* of the set of attributes BE, explaining each step of the computation.

(3 marks)

ii. Find a *key* for U, explaining *in detail* why it is a key.

(3 marks)

iii. Now consider a decomposition of U into schemas (A, B, C, I) and (C, D, E, I). Give four reasons why this decomposition is a bad decomposition, justifying your claims in each case.

(8 marks)

(c) Consider the following fragment of PHP code:

```
$query = "select max(quantity) from supplies where product=?";
$stmt = $db->prepare($query);
$product = $_GET['product'];
$stmt->execute(array($product));
$row = $stmt->fetch();
```

Explain *in detail* what each line of code means and any action resulting from its execution.

(8 marks)

5. (a) What are the two fundamental constraints of the relational model? Explain precisely what each constraint demands.

(6 marks)

(b) MySQL does not provide the standard SQL INTERSECT operator. Give an equivalent query to the following

```
select A from R intersect select A from S;
```

which will work in MySQL.

(3 marks)

(c) The use of NULL values in relational databases has resulted in SQL using 3-valued logic in determining the results of Boolean expressions.

i. Explain what is meant by the term "3-valued logic".

(3 marks)

ii. Consider the expression

```
Part.city = 'London' and Supplier.status > 10
```

which might appear in the `where` clause of an SQL query. Assuming that the condition `Part.city = 'London'` evaluates to TRUE, explain all the possible values to which the above expression could evaluate, depending on the value of the condition `Supplier.status > 10`.

(3 marks)

(d) Give a short explanation of what is meant by each of the following terms (2 marks each):

i. weak entity type
ii. composite key
iii. Cartesian product
iv. correlated subquery
v. dirty read

(10 marks)

