Birkbeck

(University of London)

BSc/FD EXAMINATION

Department of Computer Science and Information Systems

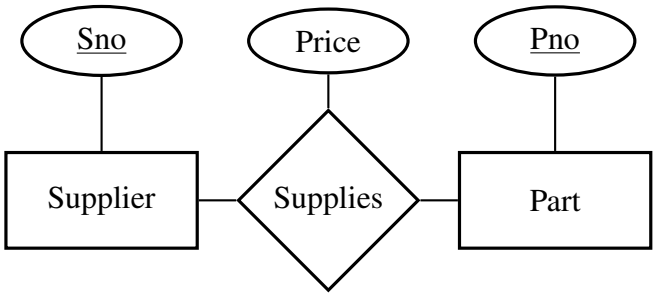
Database Management (COIY028H6)

CREDIT VALUE: 15 credits

Date of examination: 9 June 2016
Duration of paper: (14:30–16:30)

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 the following fragment of an *Entity-Relationship Diagram*:



i. What is the *cardinality* (or *multiplicity*) of the Supplies relationship type?

(**1** mark)

ii. Describe the participation constraints represented above.

(**2** marks)

iii. How would the Supplies relationship type be represented in the relational model (i.e., give the relation schema and any necessary constraints)? You can assume that the Supplier and Part entity types are represented by relation schemas of the same names.

(**6** marks)

(b) Consider a relation schema Course, with schema(Course) = {Code, Name, Level, Textbook}, where Code is the course code, Name its name, Level its FHEQ level (4, 5, 6, 7), and Textbook the name of a textbook recommended for the course. Assume we have established that the following functional dependencies (FDs) hold:

Code $\rightarrow$ Name
Code $\rightarrow$ Level
Name $\rightarrow$ Code
Name $\rightarrow$ Level

i. Identify a *redundant* FD among the four given above, explaining why it is redundant.

(**3** marks)

ii. Identify two *keys* for the Course schema, demonstrating why they are keys.

(**4** marks)

iii. Decompose the Course schema into *Boyce-Codd Normal Form* (BCNF), explaining each step of the decomposition.

(**6** marks)

(c) When comparing the value of an attribute, say Spouse, to a NULL value in SQL, one is required to use the syntax `Spouse IS NULL` or `Spouse IS NOT NULL` (rather than using = or !=). Explain why this special syntax is required.

(**3** marks)

2. (a) Assume that we have relation schemas `Visits(person,store)`, `Sells(store,item)` and `Likes(person,item)`, representing the stores people visit, the items sold by stores, and the items liked by people, respectively. Now consider the following SQL query:

```
select person, store
from Visits V
where store not in
      (select store
       from Sells S, Likes L
       where S.item=L.item and L.person=V.person);
```

i. What is the name given to constructs such as those represented by `V`, `S` and `L` in the above SQL query?

(**1** mark)

ii. The subquery in the above SQL query is known as a *correlated* subquery. Explain what this term means, with reference to the above example.

(**2** marks)

iii. Describe what will be returned by the query, i.e., what the relationship is between each person and store returned in the answer.

(**5** marks)

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

$$AB \rightarrow C$$
$$AD \rightarrow E$$
$$B \rightarrow D$$
$$AG \rightarrow B$$

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

(**3** marks)

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

(**3** marks)

(c) Describe the four steps needed to arrive at a *Third Normal Form* decomposition of a set of attributes U, given a set of functional dependencies F.

(**4** marks)

(d) Name the three different types of *update anomaly* which can occur in a relational database. Describe two different situations which can give rise to two of these anomalies.

(**5** marks)

(e) How are null values handled by aggregation operators in SQL? How are null values handled by the GROUP BY construct in SQL?

(**2** marks)

3. (a) Assume that we have the following three relation schemas

```
Module(m_code, m_name)
Student(s_id, s_name)
Result(s_id, m_code, grade)
```

with primary keys `m_code`, `s_id` and (`s_id`, `m_code`), respectively. Relation `Module` associates each module code with its name, relation `Student` associates each student id with a student name, and relation `Result` stores the grade achieved by each student in each module taken.

i. Describe (in full) any *foreign key* constraints you would expect to be specified. What purpose would these constraints serve?

(**6** marks)

ii. Now consider the following view definition:

```
create view Transcript(student, module, mark) as
select s_name, m_name, grade
from Student S, Module M, Result R
where S.s_id=R.s_id and M.m_code=R.m_code;
```

Describe (in full) the output returned by the following SQL query:

```
select * from Transcript where student like '%Jones';
```

(**5** marks)

iii. The view defined in 3(a)ii above is not *updatable* according to the SQL specification. State why, from a syntactic point of view, `Transcript` is not updatable. Now explain the problems a database system would have if it *did* try to execute the following statement:

```
insert into Transcript
       values ('A.N. Other','Database Management',85);
```

You should consider all possible scenarios in your answer.

(**6** marks)

(b) Consider the following fragment of PHP code:

```
while ($row = $result->fetch()) {
  print ("<tr>");
  for ($i = 0; $i < $result->columnCount(); $i++) {
    print ("<td> $row[$i] </td>");
  }
  print ("</tr>");
}
```

where the variable `$result` contains the result returned by some query to a database. Explain *in detail* what the above fragment of code does, also explaining the meaning and use of `fetch()` and `columnCount()`.

(**8** marks)

4. (a) One of the properties of 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 transactions do not execute atomically.

(**8** marks)

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

(**7** marks)

(c) Consider a relation over the schema $contributes(artist, song, role)$, where a tuple such as (a, b, c) denotes the fact that artist a contributes to song b in the role c (e.g., "singer" or "drummer"). Explain precisely what each of the following statements means (do not say whether or not you believe they are true).

i. The functional dependency $artist \rightarrow role$ is ***satified*** by the relation.

(**2** marks)

ii. The functional dependency $artist \rightarrow song$ is ***violated*** by the relation.

(**2** marks)

Now for each of the following functional dependencies

iv. $artist, song \rightarrow role$

v. $artist, role \rightarrow song$

vi. $song, role \rightarrow artist$

explain why you would ***not*** expect them to ***hold*** on the $contributes$ relation schema.

(**6** marks)

5. (a) Database systems provide three levels of abstraction which support data independence. Describe the three levels and the extent to which they can guarantee data independence.

(**5** marks)

(b) One desirable property of relational database design is that the relation schemas should *preserve* dependencies.

i. What does it mean to say that a design is dependency preserving?

(**2** marks)

ii. Which normal form can guarantee that dependencies are preserved?

(**1** mark)

iii. Assume that the dependency $AB \rightarrow C$ is not preserved. What would the database system have to do to ensure that the constraint specified by the dependency is satisfied at all times?

(**2** marks)

(c) When defining a foreign key constraint in SQL, we can specify the actions the database system should take when a corresponding primary key value is deleted or updated. Explain what actions can be specified as well as what happens in each case when a primary key value is deleted or updated.

(**5** marks)

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

i. first normal form

ii. weak entity type

iii. lossless join

iv. transaction commit

v. dynamic SQL

(**10** marks)

