



**UNIVERSITY  
OF LONDON**

**CO2209 ZB**

**BSc EXAMINATION**

**COMPUTING AND INFORMATION SYSTEMS, CREATIVE COMPUTING and  
COMBINED DEGREE SCHEME**

**Database Systems**

Thursday 9 May 2019: 10.00 – 13.00

Time allowed: 3 hours

**DO NOT TURN OVER UNTIL TOLD TO BEGIN**

There are **FIVE** questions in this paper. Candidates should answer **FOUR** questions. All questions carry equal marks and full marks can be obtained for complete answers to **FOUR** questions. The marks for each part of a question are indicated at the end of the part in [ ] brackets.

Only your first **FOUR** answers, in the order that they appear in your answer book, will be marked.

There are 100 marks available on this paper.

A handheld calculator may be used when answering questions on this paper but it must not be pre-programmed or able to display graphics, text or algebraic equations. The make and type of machine must be stated clearly on the front cover of the answer book.

### Question 1

The Lorentz-Fitzgerald Company manufactures Temporal Dilators. The following table records the PerCentPass rate for newly-manufactured batches of Temporal-Dilators, identified by BatchNo, which are undergoing a preliminary screening by an Inspector. If a batch is rated at less than a 99% pass rate, for a particular type of test, it is re-inspected by a different Inspector, on the same or later date. (Thus, no particular batch is ever re-inspected by the same Inspector for the same TestType.)

Part of a relation that holds this information might look like this:

BatchNo	Inspector	TestType	InspectionDate	PerCentPass
B203-3A	SMITH897	Time dilation	2019-01-23	98.7
B203-3A	SINGH766	Time dilation	2019-01-24	98.2
B203-3A	SMITH897	Space contraction	2019-02-01	99.7
C216-5C	CHANG034	Space contraction	2019-02-01	99.3
C216-5C	SMITH897	Time dilation	2019-01-24	99.2
P189-3B	CHANG034	Space contraction	2019-02-10	97.2
P189-3B	SMITH897	Space contraction	2019-02-10	99.1

The Primary Key of this relation is BatchNo + Inspector + TestType.

- A. What bad consequences could follow if, in creating the table – before we added the data – we defined the primary key of the table as

- (a) BatchNo alone?
- (b) Inspector alone?
- (c) InspectionDate Alone?
- (d) BatchNo + Inspector + InspectionDate?
- (e) BatchNo + Inspector?

[5 marks]

- B. Suppose it is decided to allow Batches to re-inspected by the same Inspector when they were initially failed, at a later date. Would we need to change the definition of the Primary Key? If we didn't change it, what problem might arise? If we changed it, what would the new key be? Explain your answer.

[5 marks]

- C. Write brief definitions of each of the following terms:

- (1) Attribute Integrity
- (2) Referential Integrity
- (3) Entity Integrity
- (4) Foreign Key
- (5) Candidate Key

[5 marks]

- D. (1) The elements can be divided into three mutually-exclusive groups: Metals, Non-Metals, and Metalloids. Consider a chemical database that consists of three relations which hold information on these three groups, called Metals, Non-Metals and Metalloids. Each relation contains identical attributes (for example, Name, Symbol, AtomicNumber, AtomicMass, Electronegativity). If we wished to merge these three relations into one, but to retain the same information that they represent at present, what changes, if any, would we have to make to the new relation?

**[3 marks]**

- (2) What is the most common reason for having two or more relations with identical attributes?

**[2 marks]**

- E. A hospital stores information about patients (such as names, addresses, phone/email numbers, doctor, *etc.*) in a table with PatientNum as primary key. Occasionally, it receives full information about individuals, but without a PatientNum yet being allocated. It was proposed simply to enter the information into the database when received, and to add the PatientNum some time later when known. Comment fully on this approach. Suggest an alternative approach.

**[5 marks]**

**TOTAL = 25 marks**



## Question 2

- A. All major DBMS packages record a variety of statistics about the data in their databases as they are used over time. What is the purpose of these statistics? Give a plausible example of what one of them might be.

**[5 marks]**

- B. The T. Soprano Construction Supplies Company supplies several dozen different materials to builders who are working on construction projects all over a particular city. At the end of each day, the builders ring in and report on how many of each material they want to re-order. The company employee who answers the telephone updates a product database held on a company computer. This allows the company to know when a particular material needs to be re-ordered. When deliveries of these items arrive from the company's suppliers, the database is also updated.

A "snapshot" of part of the main table in the database might look like this:

Code	Description	Re-Order Level	Re-Order Quantity	Price	Quantity on Hand
K3998	Cement, 25kg bag	100	250	\$80	210

A typical update of the database: when a builder orders fifteen 25kg bags of cement, the employee would alter the "Quantity on Hand" field in the first row from 210 to 195.

As the company has grown, it has expanded the number of materials it sells, and also the number of builders it sells to. As a result, it has decided to put in a second telephone line and has hired a second employee to take calls from builders and to update the database. Both employees could be updating the database at the same time.

Are there any possible problems with this? If so, describe what they are, and propose a solution to them.

**[10 marks]**

- C. When processes need to access *more* than one record at a time, and are doing so while other processes are *also* trying to access more than one record at a time, the solution to the problem of concurrent access (as proposed by you in answer to part B of this question) can itself give rise to a new problem. Describe in detail this problem and discuss a way in which it might be resolved.

**[5 marks]**

- D. In the database context a "transaction" is defined as a "logical unit of work". What does this mean? Explain the consequences of a transaction being only partially completed.

**[5 marks]**

**TOTAL = 25 Marks**

### Question 3

- A. In an Entity-Relationship diagram, what do we mean by the “participation” of an entity in a relationship between two entity-types? Illustrate your answer with a diagram showing that a Student must have a (single) Lecturer as an Advisor, but that a Lecturer is not required to advise any Students.

**[3 marks]**

- B. A data modeller wants to draw an Entity-Relationship Diagram to show the relationships amongst the three entities School Districts, Supervisors employed by a District, and the Schools in the District. Every School is in just one District. A School is normally supervised by a single Supervisor (who supervises only one School, and belongs to only one District). A School may not have a Supervisor assigned to it at the moment.

(1) Draw an E/R diagram based on this situation which contains a “Chasm Trap”.

(2) Draw a second diagram which resolves the Chasm Trap.

**[4 marks]**

- C. Briefly define the following terms.

- (1) Composite Key
- (2) Query Optimization
- (3) Functional Dependency
- (4) Determinant
- (5) Null Value

**[10 marks]**

- D. One way of understanding database management systems is to conceive of them as having a “three-level architecture” or three “schemas”. Describe the three levels of the ANSI / SPARC architecture of a DBMS.

**[6 marks]**

- E. What is the distinction between database security and database integrity?

**[2 marks]**

**TOTAL = 25 marks**

#### Question 4

- A. Consider a (very simplified) company database, which holds information on the company's Employees, the Projects to which they are assigned, and the after-work Clubs sponsored by the company to which the Employees may belong.

**Relation: EMPLOYEE** Employees, their birthdates, and their monthly salaries.

**Attributes:**

PK: ENum    An Employee Number  
Surname  
BDate    The Employee's Birthdate  
Sal       Employee's monthly salary

<u>ENum</u>	<u>Surname</u>	<u>BDate</u>	<u>Sal</u>
E21748	Johansen	1981-02-14	3000
E25432	Olimbi	1984-11-05	3000
E38769	Namamura	1990-06-21	2000
E343667	Deng	1977-05-24	5000

**Relation: BELONGSTO** Clubs, Employees, and their level of membership. Employees may belong to more than one club.

**Attributes:**

PK: ClubName  
PK: ENum       Employees associated with the club at a particular level;  
Level       Member, Probationary, Applicant

FK: ENum references ENum in EMPLOYEE

<u>ClubName</u>	<u>ENum</u>	<u>Level</u>
JOGGING	E21748	Member
JOGGING	E25432	Probationary
JOGGING	E38769	Member
CHESS	E25432	Member
BRIDGE	E21748	Probationary
BRIDGE	E25432	Applicant
BRIDGE	E38769	Applicant
HIKING	E38769	Member

**Relation: PROJECT** – Employees assigned to particular Projects, and the hours per week they are assigned to it for. Employees may be assigned to more than one Project.

**Attributes:**

PK: ENum    Employee  
PK: PNum    Project  
HrsPerWeek    How many hours per week the Employee spends on the Project.



FK: ENum references ENum in **EMPLOYEE**

<u>ENum</u>	<u>PNum</u>	HrsPerWeek
E21748	P034	20
E21748	P922	20
E25432	P922	40
E38769	P034	10

- (1) What is the query that will list the Surnames of the Employees, without duplication, who have a monthly Salary greater than \$1500?  
[1 mark]
- (2) What is the query that will list the total monthly Salaries for all Employees?  
[1 mark]
- (3) What is the query that will list Surnames of the Employees who (according to the database) are full Members (not Applicants or Probationary members) of the Jogging Club?  
[2 marks]
- (4) What is the query that will list the Surnames of the Employees who are members of either the Jogging Club or the Chess Club, or both?  
[2 marks]
- (5) What is the query that will list the Surnames of the Employees who are members of both the Jogging Club and the Chess Club?  
[2 marks]
- (6) What is the query that will list the Surnames of Employees who are members of the Jogging Club but not of the Chess Club?  
[3 marks]
- (7) What is the query that will list the Surnames of the Employees with salaries equal to the highest monthly salary?  
[2 marks]
- (8) What is the query that will list all the Project Numbers, and for each one, the total number of Employees assigned to it?  
[2 marks]
- (9) What is the query that will list the Projects which have at least 10 Employees assigned to them, and for each Project, the total number of Employees assigned?  
[2 marks]

**B.** Database Management Systems allow designers to “index” attributes of the relations they design. **(1)** What is an “index”? **(2)** What is its purpose? **(3)** What are the disadvantages to indexing?

**[3 marks]**

**C. (1)** What is a database “view”?  
What kind of alterations to data can take place as the result of creating a view?

**(2)** What is the SQL for creating a view?

**(3)** What are two reasons for creating views?

**[5 marks]**

**TOTAL = 25 marks**



### Question 5

The following table holds information about the annually-recorded weight of particular prize sheep, and about the veterinarians (Vets) who weigh them. Each sheep belongs to a particular owner. No sheep is owned by more than one owner. Each sheep's birthdate and current owner is recorded. Every year, a veterinarian weighs each prize sheep and records its weight. The date of the weighing, and the ID number of the vet, plus the mobile phone number of the vet, is recorded. Vets have only one mobile phone number. The Primary Key of the table recording this information is **SheepID+WeighDate**. The table has not been normalized beyond First Normal Form. (That is, there are no "repeating groups", but there may be Partial and Transitive Dependencies.) Information about Vets is not recorded in any other table.

SheepID	Owner	Birthdate	WeighDate	Vet	Weight	VetPhoneNum
K3922	McNab013	2013-05-12	2013-08-14	M330	22	7633088852
K3922	McNab013	2013-05-12	2014-06-02	S929	34	7609865463
K3922	McNab013	2013-05-12	2015-08-02	M330	43	7633088852
K3922	McNab013	2013-05-12	2016-07-30	P301	53	7682907965
K3922	McNab013	2013-05-12	2017-08-12	P301	52	7682907965
K3922	McNab013	2013-05-12	2018-07-22	S929	51	7609865463
T8832	McNab013	2012-03-26	2012-08-14	K339	19	7602907550
T8832	McNab013	2012-03-26	2013-09-01	S929	26	7609865463
T8832	McNab013	2012-03-26	2014-08-15	K339	32	7602907550
T8832	McNab013	2012-03-26	2015-07-28	K339	40	7602907550
T8832	McNab013	2012-03-26	2016-08-11	T975	42	7646746741
P9742	Smith002	2014-05-10	2014-09-11	K339	14	7602907550
P9742	Smith002	2014-05-10	2015-08-10	S300	25	7629920821
P9742	Smith002	2014-05-10	2016-08-29	K339	35	7602907550
P9742	Smith002	2014-05-10	2017-09-01	S929	44	7609865463
P9742	Smith002	2014-05-10	2018-08-22	S929	51	7609865463
M3110	Smith002	2015-05-12	2015-09-12	S300	15	7629920821
M3110	Smith002	2015-05-12	2016-09-10	S300	26	7629920821
M3110	Smith002	2015-05-12	2017-08-17	K339	34	7602907550

A. Identify the Functional Dependencies in this table. If Attribute A functionally determines Attribute B, show it this way:  $A \rightarrow B$ .

[10 marks]

B. This table is susceptible to insertion, deletion, and update anomalies. Give an example of each kind, using the table above.

[6 marks]

C. Split the original table into tables in BCNF, specifying the Primary Key of each table.

[7 marks]

D. Suppose a Vet changes their phone number. What are the implications for your new schema?

[2 marks]

**TOTAL = 25 marks**

**END OF PAPER**