

UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

Department of Computing

B. Sc. Examination 2015

IS52027B

Databases, Networks and the Web

Duration: 3 hours

Date and time:

*This paper is in two parts, Part A and Part B. There are a total of three questions in each part. **You should answer TWO questions from Part A and TWO questions from Part B.** Your answers to Part A and Part B should be written in separate answer books. If you choose to answer Question 3, answer parts XI and XII on the additional sheet and include it in your answer book. Make sure to put your student number at the top.*

Full marks will be awarded for complete answers to a total of four questions, two from Part A and two from Part B. Each question carries 25 marks. The marks for each part of a question are indicated at the end of the part in [.] brackets.

There are 100 marks available on this paper.

Electronic calculators must not be programmed prior to the examination. Calculators which display graphics, text or algebraic equations are not allowed.

**THIS PAPER MUST NOT BE REMOVED
FROM THE EXAMINATION ROOM**

TURN OVER

PART A

Question 1

Consider the following university database (details of which can be found below), where the primary keys are underlined.

LOCATION (loc_id, bldg_code, room, capacity)
 FACULTY (f_id, f_last, f_first, loc_id)
 STUDENT (s_id, s_last, s_first, s_address)
 TERM (term_id, term_desc, status)
 COURSE (course_id, call_id, course_name, credits)
 COURSE_SECTION (c_sec_id, course_id, term_id, sec_num, f_id, loc_id, max_enrl)
 ENROLMENT (s_id, c_sec_id, grade)

The foreign keys are as follows:

FACULTY(loc_id) references LOCATION(loc_id) {note, this is the office location of the faculty}
 ENROLMENT(s_id) references STUDENT(s_id)
 ENROLMENT(c_sec_id) references COURSE_SECTION(c_sec_id)
 COURSE_SECTION(f_id) references FACULTY(f_id)
 COURSE_SECTION(term_id) references TERM(term_id)
 COURSE_SECTION(course_id) references COURSE(course_id)
 COURSE_SECTION(loc_id) references LOCATION(loc_id) {note, this is the classroom where the course is offered}

University database

LOCATION

<u>loc_id</u>	bldg_code	room	capacity
1	CR	101	150
2	CR	202	1
3	BUS	103	42
4	BUS	105	2
5	LIB	402	1

FACULTY

<u>f_id</u>	f_last	f_first	loc_id
1	Smith	John	2
2	Miller	Sarah	4
3	Black	Daniel	4

STUDENT

<u>s_id</u>	s_last	s_first	s_address
100098341	Brown	Philip	100 Main St.
100112981	Umato	Brian	4099 West 41 st Ave.
100099199	Connoly	Michael	200 Arbutus St.
100121111	White	Roy	122 Main St.

TERM

<u>term_id</u>	term_desc	status
1	Fall 2005	closed

2	Spring 2006	closed
3	Summer 2006	open

COURSE

course_id	call_id	course_name	credits
1	CPSC100	Essential Computer Appl.	3
2	CPSC1150	Program Design I	3
3	CSIS1300	System Analysis	3
4	BUSM1310	Employability Skills	3

COURSE_SECTION

c_sec_id	course_id	term_id	sec_num	f_id	loc_id	max_enrl
1	1	2	1	1	1	100
2	1	3	1	2	1	100
3	2	3	1	1	3	30
4	2	3	2	1	3	30

ENROLMENT

s_id	c_sec_id	grade
100098341	1	A
100098341	2	B
100099199	1	C
100099199	3	
100099199	4	

Construct SQL commands to solve the following questions:

- (a) Write an SQL query to retrieve the distinct list of room capacities that are available in the university. [2 marks]
- (b) Write the SQL command to create the FACULTY table with appropriate constraints. You may assume that the LOCATION table is already created. [4 marks]
- (c) Write an SQL query to list the total number of students enrolled in the following course section id numbers (c_sec_id = 1, 2, 3). [2 marks]
- (d) Write an SQL query to list the first name, last name and the office (bldg_code and room) of all the faculty members. [3 marks]

- (e) Write an SQL query to list the first and last names of all students who live on Main Street. [3 marks]
- (f) Write an SQL query to list the building code and room of every location that is either in use as a faculty office or in use as a classroom during the Summer 2006 term. [3 marks]
- (g) Write an SQL query to list the course name and credits of all the courses that John Smith has never taught (use a set operator). [4 marks]
- (h) Write an SQL query to find all the students (first name and last name) who enrolled in more than 2 course sections in the Summer 2006. [4 marks]

Question 2

(a) Consider the relation R (A, B, C, D, E, F). Capitalized letters are attribute names.

A	B	C	D	E	F
a1	b2	c2	d3	e2	f6
a1	b2	c2	d1	e4	f6
a2	b3	c2	d1	e5	f5
a2	b4	c5	d1	e5	f5

Which of the functional dependencies (from those listed below) hold in relation R? List them in your answer book.

- $A \rightarrow B$ $A \rightarrow F$ $B \rightarrow A$ $E \rightarrow A$
 $C \rightarrow D$ $D \rightarrow F$ $B \rightarrow C$ $E \rightarrow D$
 $B \rightarrow F$ $F \rightarrow B$ $F \rightarrow A$ $C \rightarrow A$

[8 marks]

(b) Consider the relation R(A, B, C, D, E) with functional dependencies:

$A \rightarrow B, B \rightarrow C, C \rightarrow A, E \rightarrow D, D \rightarrow E$

List all the candidate keys for R. Justify your answer.

[7 marks]

(c) Consider the following CAR_SALE relation:

CAR_SALE (Car_id, Option_type, Option_listprice, Sale_date, Option_discountprice)

This relation refers to options installed in cars (e.g. cruise control) that were sold at a dealership.

The CAR_SALE relation has the following functional dependencies:

Car_id \rightarrow Sale_date
Option_type \rightarrow Option_listprice
Car_id, Option_type \rightarrow Option_discountprice

- i. Explain why the CAR_SALE relation is not in 3NF
(Please note that the primary key is composed of Car_id and option_type)

[4 marks]

- ii. Normalize the CAR_SALE relation to 3NF relations.

[6 marks]

Question 3

General database questions

Select the correct answer to each of the following questions. For the multiple-choice questions, each question has one correct answer (do not choose multiple options). You will be penalized for incorrect answers.

- I. In which of the following situations would a file-based approach instead of a DBMS be more suitable for data storage and data access?
- a) Simple, well-defined applications that are not expected to change at all
 - b) Embedded system with limited storage capacity
 - c) Applications with no multiple-user access to data
 - d) All of the above

[1 mark]

- II. Which of the following are advantages of a database management system:
- a) Data consistency
 - b) Control of data redundancy
 - c) Improved security
 - d) All of the above

[1 mark]

- III. Which of the following statements is incorrect? There is only 1 correct answer.
- a) The external level of the 3-level ANSI-architecture describes what data is stored in the database.
 - b) The conceptual level of the 3-level ANSI-architecture contains the logical structure of the entire database.
 - c) The conceptual level of the 3-level ANSI-architecture is independent of any storage considerations.
 - d) The objective of the 3-level ANSI-architecture is to separate each user's view of the database from the way the database is physically represented.
 - e) The internal level of the 3-level ANSI-architecture is the physical representation of the database and describes how the data is stored.

[1 mark]

IV. Fill in the missing words in the sentences below from the following list of words: [attribute, degree, cardinality, superkey, candidate key, domain, data definition language, data manipulation language, tuple, DBMS]. Write the correct words for each section in your answer book.

- a) _____ is a named column of a relation.
- b) _____ is a row of a relation.
- c) _____ is the set of allowable values for one or more attributes.
- d) _____ is a language that provides a set of operations to support the basic data manipulation operations on the data held in the database.
- e) _____ is an attribute or set of attributes that uniquely identifies a tuple within a relation.

[5 marks]

V. Consider the following relations for a database that keeps track of student enrolment in courses and the books adopted for each course:

STUDENT(Student#, Name, Major, BirthDate)
COURSE(Course#, CourseName, Dept)
ENROL(Student#, Course#, Semester, Grade)
BOOK_ADOPTION(Course#, Semester, BookISBN)
TEXT(BookISBN, BookTitle, Publisher, Author)

Specify the foreign keys for this schema, stating any assumptions you make. Use the notation given in Question 1.

[3 marks]

VI. Which SQL statement is used to return only different values?

- a) SELECT IDENTITY
- b) SELECT DIFFERENT
- c) SELECT UNIQUE
- d) SELECT DISTINCT

[1 mark]

VII. Which SQL command is used to alter a value within a relation?

- a) ALTER
- b) UPDATE
- c) INSERT
- d) SET VALUE

[1 mark]

VIII. Given the tables T1 and T2 below, which of the following queries will return the following?

A B B C
 NULL NULL b6 c6

T1		T2	
A	B	B	C
a1	b1	b1	c1
a2	b2	b3	c3
a3	b3	b1	c1
a4	b4	b6	c6

- a) `SELECT * FROM `T1` LEFT JOIN `T2` ON T1.B = T2.B WHERE T1.B IS NULL`
- b) `SELECT * FROM `T1` RIGHT JOIN `T2` ON T1.B = T2.B WHERE T1.B IS NULL`
- c) `SELECT * FROM `T1` , `T2` WHERE T1.B = T2.B AND T1.B IS NULL`
- d) `SELECT * FROM `T1` INNER JOIN `T2` ON T1.B = T2.B AND T1.B IS NULL`

[2 marks]

IX. A table is in 2NF if the table is in 1NF and what other condition is met?

- a) There are no full functional dependencies.
- b) There are no null values in primary key fields.
- c) Every non-primary key attribute is fully functionally dependent on the relation's primary key.
- d) There are no non-primary key attributes that are transitively dependent on the primary key.

[1 mark]

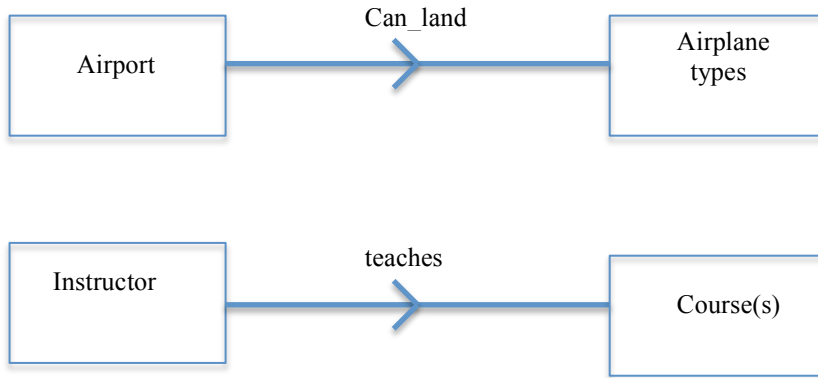
X. A table is in 3NF if the table is in 1NF, 2NF, and what other condition is met?

- a) There are no full functional dependencies.
- b) There are no null values in primary key fields.
- c) Every non-primary key attribute is fully functionally dependent on the relation's primary key.
- d) There are no non-primary key attributes that are transitively dependent on the primary key.

[1 mark]

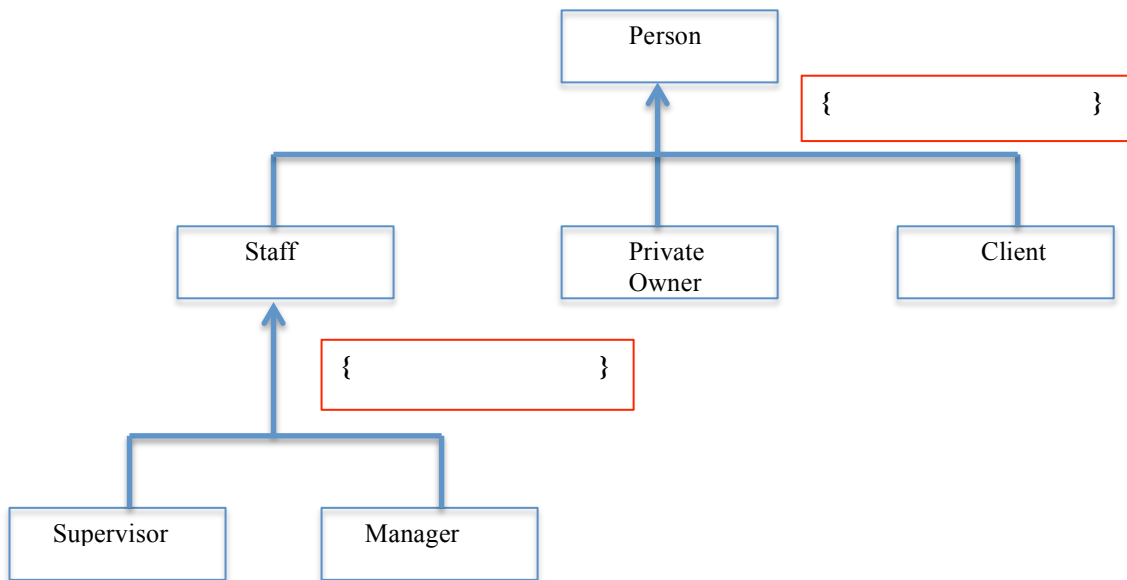
USE THE EXTRA ANSWER SHEET TO FILL IN THE ANSWERS OF THE NEXT 2 QUESTIONS (QUESTIONS XI and XII). SUBMIT WITH YOUR ANSWER BOOK. DO NOT FORGET TO FILL IN YOUR STUDENT NUMBER.

XI. Fill in the multiplicity ratio for the following E-R relations.



[4 marks]

XII. Fill in the Participation and Disjoint Constraint (in the 2 red boxes with curly brackets) for the EER diagram below.



[4 marks]

PART B

Question 4.

Question 4 concerns PHP, forms and arrays.

(a)

- i. What does it mean to say that 'PHP is a server-side scripting language'?
- ii. Name two aspects of dynamic websites that are made simpler by using PHP? [3 marks]
- iii. Explain the difference between 'For...' and 'While...' control structures in PHP and explain when each is appropriate to be used. [2 marks]

(b)

Consider this simple array:

```
$list = array ('Alan', 'Bob', 'Charlie');
```

- i. Write a statement to append the item 'Dave' to the end of the list
- ii. Write a statement to assign the element 'Charlie' to the variable \$user

[2 marks]

Consider the following multidimensional array;

```
$Books = array(  
    'PHP' => array(  
        'Title'=>'Programming PHP',  
        'Author'=>'Kevin Tatroe, Peter MacIntyre and Rasmus  
Lerdorf',  
        'Publication Date' => '2013'),  
    'Javascript' => array(  
        'Title'=>'JavaScript: The Good Parts',  
        'Author'=>' Douglas Crockford',  
        'Publication Date'=>'2008'),  
    'MySQL' => array(  
        'Title'=>'MySQL Cookbook',  
        'Author'=>'Paul DuBois',  
        'Publication Date'=>'2007'),  
    'Networking' => array(  
        'Title'=>'Computer Networking: A Top-Down Approach',  
        'Author'=>'James F. Kurose and Keith W. Ross',  
        'Publication Date'=>'2012')  
);
```

- iii. Write a statement that assigns the array element containing the author names for the networking book to the variable \$networking_authors.
- iv. Here is a nested loop that will print out the contents of the array. Find five errors in the code, and rewrite with the errors corrected.

```
foreach ($Books as $category =>book) {
```

```
        echo '$category:<br />';
        foreach ($book as $key=$value) {
            echo "$key: $value<br />"
            echo "<br />";
        }
    }
}
```

[5 marks]

- v. Which php function(s) print an array? When might this be useful?

[2 marks]

(c)

Describe the differences between the GET and POST methods for submitting form data, with particular reference to

- i. how data is sent to the server
- ii. when it is appropriate to use each method

[5 marks]

- iii. Here is the PHP code to create a simple online form:

```
print '
<form action="class.onepageform.php" method="post">
<p>first name: <input type="text" name="firstname" size="20" /></p>
<p>last name: <input type="text" name="lastname" size="20" /></p>
<p><input type="submit" name="submit" value="submit first & last
name" /></p>
</form>';
```

Write the code necessary for the same PHP script to handle the form submission and print out a greeting using the submitted values for firstname & lastname. Include comments to explain what you are doing.

[6 marks]

Question 5

Question 5 concerns PHP, databases, templates and functions.

(a)

- i. Draw a labelled diagram to represent the request / response communications between browser, web server, php processor and mysql database when the browser requests a web page containing dynamic content.

[5 marks]

A connection between PHP and a MySQL database can be established in the following way:

```
$dbc = mysqli_connect('A', 'B', 'C');  
mysqli_select_db($dbc, 'D');
```

- ii. What do the parameters A, B, C and D represent?

[4 marks]

- iii. You have a database query in your PHP script, but it is not working correctly. Describe the ways you would go about debugging it.

[3 marks]

(b)

You are using a form on a web page to modify values stored in database. The initial values are retrieved from the database and stored in \$row. The PHP print statement that creates the form is:

```
print '<form action="edit.php" method="post">  
  <p><label>title <input type="text" name="title" value="'.  
  $row['title'].'" /></label></p>  
  <p><label>content <textarea name="content" rows="5"  
  cols="30">'.$row['content'].'</textarea></label></p>  
  <input type="hidden" name="id" value="'. $row['id']. '" />  
  <p><input type="submit" name="submit" value="Edit the text"  
  /></p>  
</form>';
```

When the script handles the form submission and modifies the data, it constructs the following query:

```
$query = "LLL entries MMM title='$title', entry='$content' WHERE  
entry_id=NNN";
```

- i. What should replace LLL, MMM and NNN?

[4 marks]

- ii. What is the purpose of mysqli_real_escape_string(\$dbc, query) and what does it help to prevent?

[1 mark]

(c)

Consider the following page, which contains a form generated by PHP.

```
1 <!DOCTYPE html>  
2 <html>  
3 <head>  
4   <title>Erewhon Events</title>  
5 </head>  
6 <body>  
7   <div id="header">  
8     <ul id="nav">
```

```

9         <li><a href="home.php">Home</a></li>
10        <li><a href="cancellation.php">Cancellations</a></li>
11    </ul>
12 </div>
13
14 <?
15 print'Select days of the week: <br />';
16
17 print '<form action="event.php" method="post">';
18 $days =
array("Sunday","Monday","Tuesday","Wednesday","Thursday","Friday","Saturday");
19 foreach ($days as $day){
20     print "<input type='checkbox' name='days[]' value='$day' /> $day
<br />";
21 }
22 print '<input type="submit" name="submit" value="submit selection"
/></p>';
23 print '</form>';
24 ?>
25
26 </div>
27
28 <div id="footer">
29     <p>Design by Erewhon Ltd</p>
31     <p>&copy; 2015</p>
32 </div>
33 </body>
34 </html>

```

i. Referring to line numbers, label parts of the page which could be put in to separate template files.

[2 marks]

ii. Write a re-usable PHP function to generate the form.

[4 marks]

iii. Re-write the page using your templates and function.

[2 marks]

Question 6

Question 6 concerns internet protocols

(a)

The Internet Protocol Stack can be usefully described in terms of 4 layers that sit above the physical layer.

- i. Name these 4 layers and describe their essential function(s).

[8 marks]

- ii. Why do we need TCP in addition to IP?

[1 mark]

(b)

The network settings for a laptop on the Goldsmiths network includes the following:

address 158.223.59.103

netmask 255.255.0.0 gateway 158.223.2.89

- i. Which part of the IP address is important when traversing the internet, and why?

[2 marks]

- ii. What is special about IP addresses of the form 192.168.. ?

How are they used?

[2 marks]

- iii. Why does this traceroute listing *not* show the route of an individual IP packet? Describe what is happening.

traceroute google.com

traceroute to google.com (216.58.208.46), 30 hops max, 60 byte packets

```
1  router.gold.ac.uk (158.223.2.89)  1.082 ms  1.059 ms  1.031 ms
2  192.168.0.1 (192.168.0.1)  0.348 ms  0.335 ms  0.546 ms
3  gi0-6-0-9.londsh-rbr1.ja.net (146.97.136.237)  1.777 ms  1.762
ms  2.099 ms
4  be2.londic-rbr1.ja.net (146.97.66.34)  2.107 ms  2.786 ms  2.760
ms
5  ae24.londpg-sbr1.ja.net (146.97.37.197)  1.993 ms  1.899 ms
1.881 ms
6  ae29.londhx-sbr1.ja.net (146.97.33.1)  2.546 ms  2.489 ms  2.459
ms
7  po1.lond-ban3.ja.net (146.97.35.106)  2.381 ms  2.280 ms  2.235
ms
8  72.14.196.137 (72.14.196.137)  2.196 ms  2.118 ms  2.093 ms
9  209.85.246.246 (209.85.246.246)  2.021 ms  2.021 ms  1.982 ms
10 216.239.51.235 (216.239.51.235)  2.199 ms  2.123 ms  2.111 ms
11 lhr08s07-in-f14.1e100.net (216.58.208.46)  2.071 ms  2.049 ms
2.022 ms
```

[3 marks]

(c)

- i. What information is contained in TCP headers that, when combined with IP addresses, makes a 4-tuple (a combination of 4 numbers) that uniquely identifies the connection between a process on one computer connected to the internet and a process on another?

[2 marks]

- ii. TCP headers can be examined with a tool called tcpdump. Here is the output of a single packet captured with tcpdump. In this case, what are the 4 numbers referred to above?

```
IP 158.223.59.103.32889 > 159.182.165.92.80: Flags [.], ack
1853349429, win 2641, options [nop,nop,TS val 2377088 ecr
58663145], length 0
```

[2 marks]

- iii. What's the difference between TCP and UDP? Give an example where UDP is used.

[2 marks]

- iv. Draw a labelled diagram to show how nameservers are queried when looking up the domain www.hiig.de. Ignore cacheing.

[3 marks]

END OF EXAMINATION