SCHOOL OF MATHEMATICAL AND COMPUTER SCIENCES

Department of Computer Science

F21DF

Database and Information Systems

Semester 1 2018/19

Duration: Two Hours

ANSWER THREE QUESTIONS

Q1   The following tables keep track of research projects, the staff working on
            these projects and the roles they play in the project (pi = principal
            investigator, ra = research associate, admin = administrator), and the
            number of hours each person has worked on each project each day.

```
Staff:
+-----+-------------+-------+
| sid | name        | role  |
+-----+-------------+-------+
|  12 | Prof Smith  | pi    |
|  86 | Ken Bond    | ra    |
|  87 | Karen Pitt  | ra    |
|  99 | Mike Elder  | admin |
| 195 | Prof Conner | pi    |
+-----+-------------+-------+
ResearchProject:
+-----+---------+--------+-----+
| pid | title   | funder | pi  |
+-----+---------+--------+-----+
|  48 | Takar   | BBSRC  | 195 |
| 149 | Tetra   | BBSRC  |  12 |
| 199 | HelpTak | EPSRC  |  12 |
+-----+---------+--------+-----+
TimeRecords:
+-----+-----+------------+-----------+
| pid | sid | workDate   | workHours |
+-----+-----+------------+-----------+
|  48 |  87 | 2013-02-18 |         8 |
|  48 |  87 | 2013-02-19 |         3 |
| 149 |  12 | 2013-02-18 |         2 |
| 149 |  86 | 2013-02-19 |         8 |
| 199 |  99 | 2013-02-18 |         2 |
| 199 |  99 | 2013-02-19 |         2 |
+-----+-----+------------+-----------+
```

(a)   Develop SQL queries to answer the following questions:

(i)   Return the title and funder of projects on which Prof Conner is the pi.

(4 marks)

(ii)   Create a view called `projectHours` which returns the names of projects
        and the total number of hours that have been worked on each project.

(6 marks)

(b)

(i)   Draw the query plan tree for the following query:

```
SELECT S.name, P.title
FROM   Staff as S, ResearchProjects as P,
       TimeRecords as T
WHERE  S.sid = T.sid
AND    P.pid = T.pid
AND    P.funder = 'BBSRC';
```

(8 marks)

(ii)   There are 5 funders of roughly similar size. Explain why an index on the
        funder column would not help with query answering.

(2 marks)

Q2

(a)

(i) Explain and provide an example for each of these terms: entity, attribute, and relationship

(3 marks)

(ii) Draw an ER diagram to capture the following requirements:
- The University stores the names (first and last) of a lecturer, these are not unique
- Every Course has a unique course code and unique title
- A Course can be taught by more than one Lecturer
- All Courses must be taught by at least one Lecturer
- The number of hours a Lecturer teaches on a specific course needs to be captured
- Not all Lecturers teach on Courses

(7 marks)

(b) Consider the following two transactions taking place in parallel; one to move £10 from one account to another and a second to sum up the total at each branch.

Time	T_1	T_2
1	Start T_1	
2	$R_1(A)$	
3	$W_1(A)$	
4		Start T_2
5		$R_2(A)$
6		$R_2(B)$
7	$R_1(B)$	
8	$W_1(B)$	

(i) State what it means for two transactions to be serializable and explain whether or not T_1 and T_2 are serializable.

(3 marks)

(ii) What guarantees does a Transaction Manager provide?

(4 marks)

(iii) State what lock T_1 acquires in step 3. Explain the effect on T_2 and the remainder of the steps in T_1. *(Assume that a transaction will pause execution when waiting for locks.)*

(3 marks)

Q3

(a)   Neo4J has a flexible schema. Define *flexible schema* and compare
        it to the schema offered by MySQL.

                                                                         (2 marks)

(b)   State 2 advantages and 2 disadvantages with a *flexible schema*?

                                                                         (4 marks)

(c)   Explain what *impedence mismatch* is and why it causes problems
        for programmers.

                                                                         (2 marks)

(d)   Discuss a typical persistence workflow using Hibernate ORM.
        What does the programmer do, and what is done for the
        programmer by Hibernate? You do not need to write code, but
        please make sure to mention the Persistence Context.

                                                                         (8 marks)

(e)   Define a *Full Text Search Engine* (sometimes known as a *Search
        Engine*). List common features.

                                                                         (2 marks)

(f)   Define *faceted search*. You should include an example to illustrate
        your explanation.

                                                                         (2 marks)

Q4

Consider the following information:
Nissan Leaf
Range: 270.4 km
RRP: 27235
Battery: 40 kWh
Cargo volume: 420L

(a)   Write the above information (on the Nissan Leaf) in YAML.

(2 marks)

(b)   Write the above information (on the Nissan Leaf) in JSON.

(2 marks)

(c)   Write the above information (on the Nissan Leaf) in XML.

(4 marks)

(d)   Write an XML DTD for the above information (on the Nissan Leaf).

(4 marks)

(e)   Write an XML Schema for the above information (on the Nissan Leaf).

(4 marks)

(f)   Extend the XML Schema (from part e) such that the *Range* must be provided as a decimal number with only 1 place after the decimal point.

(4 marks)

END OF PAPER