# CPT140 Database Concepts

# Study Period 4, 2011

# Assignment 2

Due: 11:59pm, Sunday 12th February 2012.
(Please Note : This date is a week later than the date published in the Weekly Schedule. This is to offset any timing isssued cause by the problems with Assignment 1)

Worth: 100 marks in total, which is worth 20% of the overall assessment for the course.

Submission: Submission is via the WebLearn assignment submission system.

- To access WebLearn, log in to myRMIT and go to the Studies tab. Look under Courses and find Database Concepts. Click the Weblearn link.
- In WebLearn, click on the Assignments button on the right-hand-side and follow the instructions
- Select "Ian Baker" as tutor.
- Submit only one file in pdf format or a zip file containing pdfs.
- Each time you submit, the old submission is over-written with the new submission.
- Make as many partial submissions as you like, up to the due date.
- Keep a backup of your submission.
- It is important to keep your confirmation email from WebLearn. It is your proof of submission.

Assessment: This is an individual assignment. The minimum penalty for plagiarism is loss of marks for this assignment. If this means that a hurdle requirement is not met, the student fails the course. Please read http://www.cs.rmit.edu.au/students/integrity/ for further information.

(1) This is an individual assignment. The minimum penalty for plagiarism is loss of marks for this assignment. If this means that a hurdle requirement is not met, the student fails the course. Please read http://www.cs.rmit.edu.au/students/integrity/ for further information.

(2) For SQL; In the course of marking, most queries will be tested. Queries that do not execute without produceing errors will receive very low marks (ie <50%).

(3) For SQL; It is frequently possible to write two queries that will produce the same result. Just getting the correct output will not get full marks. It is important that the logic of your query is also correct. Incorrect logic will result in a poor mark.

(4) For SQL; It is also possible to write two queries that produce the correct result and are also logically correct but one is significantly more complex and convoluted. Excessive/un-necessary complexity results in reduced marks.

Late Submission Penalty: WebLearn will remain open for 5 days after the Due Date for late submissions. A late submission incurrs a penatly of 10% per day or part day late. As an Example, an assignment submitted at 2AM on Tuesday morning will incurr a 20% penalty. Submission more than 5 days late will not be accepted.

Extensions and Special Considerations:

To request an extension or Special Consideration, please email the details to

Email : ouacsit@rmit.edu.au

Please note the following points about extensions and special consideration;

- I am unable to (as in, not allowed to) grant extensions. The Admin people are the only ones who can.
- You must apply at least three days before the due date of the assignment.
- You must apply on the appropriate form.
- You must have supporting documentary evidence.
- The grounds for extensions are very limited and do not include excessive workload or tiredness.

Other Information : The aim of the assignment is to practice the concepts learnt from the material and to allow you to test your understanding and learn more about each topic.

The questions as presented are not meant to represent real-world situations. They are contrived in a manner that tests your knowledge. You are expected to review the text and Blackboard materials to work out the answers. There are no 'trick' questions.

Not every situation in the assignment (or in real life) is presented in the references. You may have to ask questions (of me or others) in the Blackboard discussion forums while doing the assignment.

Start each answer with the number of the question you are answering.

Please set out SQL answers as you see them in the text and lectures, that is using multiple lines, without line numbers. This makes testing your query and analysing it for part marks easier.

I do not do 'reviews' of submissions before the due date. I will not answer "Is this right" type questions. I will answer requests for clarifications for any of the questions. I will also answer specific questions like, "I can get this far, but I do not know what the next step is". I will typically give a clue or hint rather than solve it for you.

Students often require clarification for some questions. This is done in the Discussion Forums in the course Blackboard. Please check the Discussion Forums regularly for information.

You ***MUST*** put your name and student number on your submission.

The ***ONLY*** acceptable format is PDF.

Good Luck with your assignment!

**Question 1. Short Answer** (10 Marks)

Briefly explain each of the following database terms. Use no more than 2-3 lines for each term.

1. DBMS
2. Update Anomaly
3. Alias (referring to tables in SQL queries)
4. DDL
5. Domian

**Question 2. ER** (25 Marks)

Read the description below and draw an Entity-Relationship diagram for it. You must use the symbols described in the course materials.

This database will help the Parents Assocciation of the local school to manage their upcoming Summertime Fete. The main aim of the database is to manage aspects of the operation of the stalls.

- PA members are represented in the database. Their name is used to identify them and their email address is also recorded.
- All stall are given a unique title ("Bring and Buy", "Cakes", etc). The stall also has a location and the number of people required to operate it.
- Stalls are sponsored by companies and local businesses. Each sponsor has a company name, contact number, contact person and their advertising slogan is included in the database (to be announced over the Public Address system on the day of the fete)
- Companies usually sponsor many different stalls. The amount donated to each stall is recorded. Most stalls need sponsorsip from more than one company. Every company in the database must sponsor at least one stall.
- Work times are 1 hour in duration. Each slot is known by its commencement time (eg "the 11AM slot"). No other information is recorded about time slots.
- Members work on stalls in selected time slots. There are not restrictions on members working on stalls, etc except for one condition; for a given time slot, a member can only work on one stall.
- At the fete, there are range of administrative jobs that need attending to. They are identified by the task. There is also a record of the equipment required for the task.
- Some members will be assigned a job but no more than one job. Some jobs are complex and will need many people to complete them.

**Question 3. Relational Design** ( 15 + 15 Marks)

Parts marks for this question are based on intermediate steps shown and stated reasoning.

**3(A)** For the relation;

CarSales (RegNo, BuyerName, Address, SalesPerson, Make, Model, Commission)

The following functional dependences hold,

RegNo _ID → Make, Model

RegNo→ Commission

RegNo, BuyerName → SalesPerson

BuyerName → Address

Make, Model → Commission

RegNo → SalesPerson

a. Use Inference rules to find the minimal basis and determine the Key of the relation
b. Based on this key, determine if the relation is in BCNF.
c. If it is not in BCNF then decompose the relation into BCNF.

**3(B)** The following relation holds details about cars ;

.

CAR_DETAILS(Registration_No , Make , Model , Tow_Load , Engine_Size, Colour )

And has the following functional dependencies;

Make , Model -> Engine_Size

Registration_No -> Make, Colour, Model

Engine_Size -> Tow_Load

The relation is decomposed into ;

CAR_DETAILS1(Registration_No , Colour, Make , Engine_Size* ,Model)

CAR_DETAILS2(Engine_Size, Tow_Load)

Prove that the decomposition is incomplete **and** complete the decomposition.

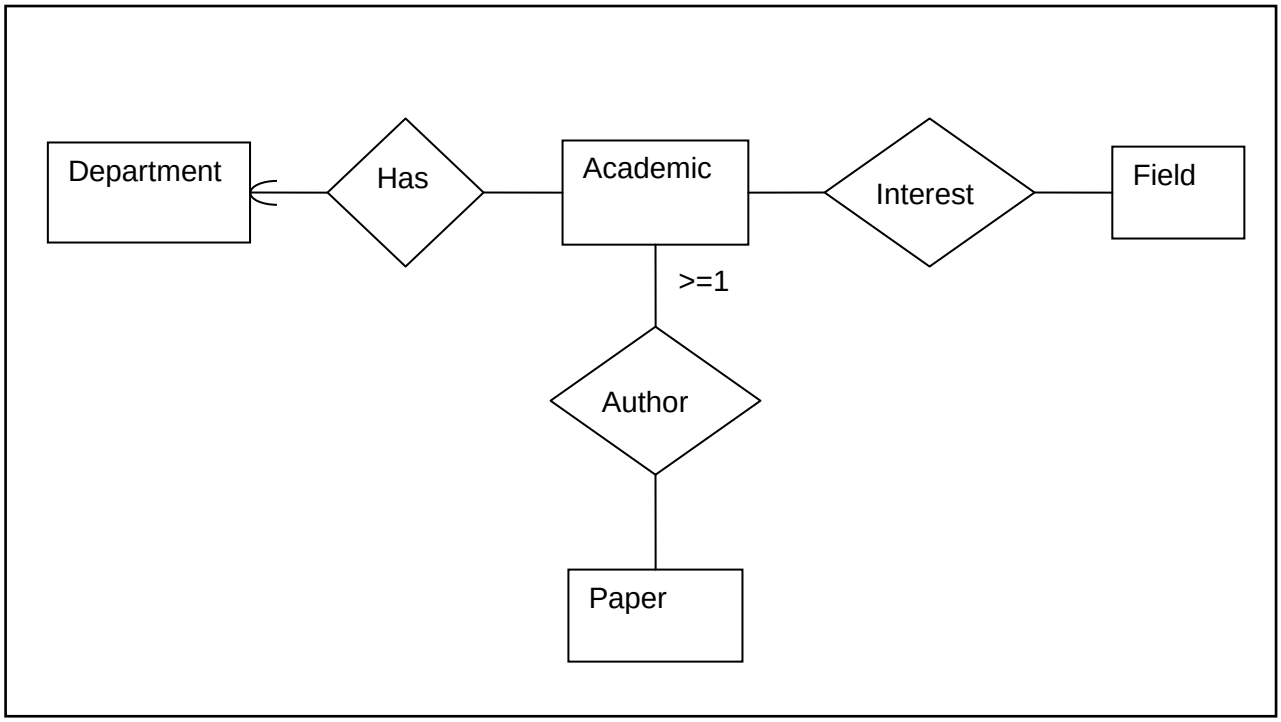
**Question 4. Advanced SQL** (25 + 10 Marks)

You are required to write SQL queries based on the database described below.

The database is available for you on the RMIT Oracle system to develop your solutions. You can access the database from the yallara command line prompt. SQLDeveloper is also a satisfactory way to test your queries. The details for accessing the research database are;

## ResearchDatabase:

- Username: roresearch
- Password: research
- Yallara Command : sqlresearch




This assignment concerns the Research database. The database ER model is shown above, where attributes are omitted for clarity. The relational model for the database is:

Department(DeptNum, Descrip, Instname, DeptName, State, Postcode)

Academic(AcNum, DeptNum*, FamName, GiveName, Initials, Title)

Paper(PaNum, Title)

Author(PaNum*, AcNum*)

Field(FieldNum, ID, Title)

Interest(FieldNum*, AcNum*, Descrip)

**DeptNum**, **AcNum**, **PaNum**, and **FieldNum** are the identifiers of departments, academics, papers, and fields of interest respectively. Each Academic belongs to one Department and is the author of some (or no) Papers. Each Paper has at least one author. The table Field

describes areas of research work; each Academic works in some (or no) Fields, and may provide a description Descrip of their work.

The research database is used by the Department of Education, Science and Training of the Australian government on a regular basis to report on academic departments of education institutions. To this end, you are asked to write SQL queries, and present the report based on the output from your queries. Note that data inconsistency has been introduced during the data entry process, which you have to take care of in producing your report. It is known that the same state information may be represented in different ways. For example, Victoria may be kept as VIC, vic, or Vic. It is also known there are missing values for some attributes. For example, not all academic departments have their State information recorded in the database.

The output of your queries does not need to be submitted. The queries you write shouls be included in as part of your pdf document.

There are 6 questions in total. The first 5 questions are worth 5 marks each. Write ONE SQL query for each question. The last questions are worth 10 marks each. Output of your queries should not contain duplicates but use DISTINCT only if necessary.

Five questions worth 5 marks each;

1. How many institutions contain the word “Technology” in their name. Your query must produce a single number as its output.
2. Make a list of academics that have more than five interests and have written or co-written less than ten papers. Use one IN construct and one EXISTS construct in your query. Output all details of the academics in the list.
3. Use a Set operator to create a list of academics who have written or co-written more than 5 papers and also have greater then 3 interests. List their academic number in the output.

Continued next Page.

4. Explain the following SQL query in English;

```
select givename, famname, instname
from academic natural join department
where acnum in
        (select acnum
        from author
        where acnum not in
                (select acnum
                from interest
                group by acnum))
and   deptNum in  (select deptNum
                        from academic
                        where deptname = 'Computer Science');
```

5. Find all the academics who have co-written a paper with academic number 151. Use only the Author table in a join query. You can use multiple instances of the table. The output should list the academic number of each co-author just once in the output. The output should not include academic number 151.

This questions is worth10 marks;

6. Write a query to find papers with more than five authors where the authors all come from different departments..

==========