Assignment 1

Please note that we have different deadline for each question. Please make sure you have submitted each question by its own deadline:

Question 1 : ~~Fri 10~~ Sun 12 Mar, 5:00 pm

Questions 2, 3 : Fri 31 Mar, 5:00 pm

Question 1 (4 marks)

A government founded project aims to collect statistics about university professors and their research projects. As a database system student, the first idea comes to your mind is to build a database to assist this purpose. Now, here is your task:

- Each research project has to be supported by a funding. A funding is uniquely identified by its ID, and we also need to know its amount.
- For a research project, we also need to know the name, area, expected outcome, starting time and expected duration. A project may touch several different areas.
- A research project is carried out by several researchers, and a researcher can work on several research projects. The time a researcher works on a project is in need.
- A researcher can be uniquely identified by his/her email. The total number of papers he/she has published is of interest, so as his/her address, which is composed of city and street.
- A researcher can be either a Professor or a PhD student. A PhD student has to be supervised by exactly one professor. A professor can supervise multiple PhD students.
- We need to know the year a PhD student started his/her PhD program.
- We need to know the number of students a professor is supervising.
- A research project can have a leading professor, and a professor can lead at most one project.
- A paper must be the result of cooperation between several researchers.
- A paper can be uniquely identified by its DOI. Also of interest are the title, publisher and date published.

Draw an ER diagram to represent this scenario, and clearly state the assumptions you make if any.

Question 2 (2 marks)

Convert your ER-diagram from Question 1 into a relational model.

Question 3 (4 marks)

Jeffrey Ullman is a well know researcher in the field of database. Assume that we have the following schema: Paper(doi, title, year), Researcher(name, area), Authoring(doi, name).

1) (1 mark) Write relational algebra expressions to find the titles of the papers for which Jeffrey Ullman is an author.

2) (1 mark) Write relational algebra expressions to find all the research areas the researchers who have co-authored all the papers Jeffrey Ullman had published in year 2011 have touched.

3) (2 marks) Write relational algebra expressions to find all the research areas the researchers who have co-authored all the papers Jeffrey Ullman had published in year 2011 have touched. Note that you are not allowed to use division.

Assignment Submission

We accept electronic submissions only. Please submit your assignments as follows:

- Ensure that you are in the directory containing the file to be submitted. (note: we only accept files with .pdf extension)
- For question 1, type "give cs9311 ass1q1 ass1q1.pdf" to submit.
- For questions 2 and 3, type "give cs9311 ass1q23 ass1q23.pdf" to submit.
- Please keep a screen capture (including **timestamp** and the **size** of submitted file) for your submissions as proof in case that the system is not working properly. If you are not sure how, please have a look here.

Note:

1. We do not accept e-mail submissions, and the submission system will be immediately closed after the deadline.
2. If the size of your pdf file is larger than 2MB, the system will not accept the submission. If you face this problem, try converting to compressed pdf.
3. If you have any problems in submissions, please email to swan398@cse.unsw.edu.au or xwang@cse.unsw.edu.au.

Late Submission Penalty

Zero mark