

# COMP2400/6240 - Relational Databases

## Assignment 2 (Database Theory)

Due date: 23:59, 12 October 2021

### Instructions:

- This assignment must be done individually (no group work).
- This assignment will count for 15% of the final grade. Marks are assigned for the process of finding a solution, not only for the result. Hence, include all essential ideas and steps that are necessary to derive a solution.
- You must submit a single PDF file named as “u1234567.pdf” (replace u1234567 with your UID). Make sure you only upload a PDF file, not a Word or text file.
- You should try your best to type the solutions. The scanned images of handwritten texts and equations can be unreadable for marking. As for the EER diagram, you are highly recommended to export a JPEG file from TerraER and include it in the PDF file.
- Late submission is not granted under any circumstance. You will be marked on whatever you have submitted at the time of the deadline. Please take careful note of deadlines and adhere to them. Of course, if you find yourself in a situation beyond your control that you believe significantly affects an assessment, you should follow the ANU’s special consideration process (<http://www.anu.edu.au/students/program-administration/assessments-exams/special-assessment-consideration>).
- Plagiarism will attract academic penalties in accordance with the ANU guidelines. A student in this course is expected to be able to explain and defend any submitted assessment item. The course convener can conduct or initiate an additional interview about any submitted assessment item for any student. If there is a significant discrepancy between the two forms of expression for assessment, it will be automatically treated as a case of suspected academic misconduct.

### Question 1

4 Marks

ACTScooter is a newly established company that offers scooter rental services in Canberra. The employees at ACTScooter are classified into the three non-overlapping categories, administrators, distributors and technicians. An administrator helps a customer to register their unique credit card number and home address at ACTScooter. The work type of administrators can be either remote, onsite or hybrid. Each distributor helps to move scooters around and can be reached through a mobile phone number. Each technician has a qualification level, conducts periodic inspections for the scooters and records the date and feedback of every inspection for a scooter. Each ACTScooter employee can be uniquely identified by their employee number. ACTScooter also stores the brand, the manufacture year and an internal distinct ID for each scooter. Every ride is booked by exactly one customer. The database should allow customers to retrieve the detailed information about each of their rides, including a unique reference number, the scooter ID, the fare, the start time and end time, the origin and the destination. Once a ride is completed, the customer’s credit card will be automatically charged by ACTScooter.

Your task is to design an Enhanced Entity Relationship (EER) diagram for the above database, which should include entities, relationships, attributes and constraints wherever appropriate (you can make more assumptions if necessary).

You also need to identify the requirements that cannot be captured in an EER-diagram.

**Question 2****5 Marks**

Consider the relation schema  $R = \{A, B, C, D, E\}$  and the following set  $\Sigma$  of FDs:

- $AB \rightarrow C$
- $BC \rightarrow A$
- $C \rightarrow DE$
- $DE \rightarrow B$

2.1 What are the candidate keys of  $R$ ? Justify your answer (i.e., include the main steps used for finding the candidate keys). (2 Mark)

2.2 Find a minimal cover of  $\Sigma$  and include the main steps used for finding a minimal cover. (2 Mark)

2.3 Consider another set  $\Sigma_1$  of FDs on  $R$ :

- $AB \rightarrow CDE$
- $C \rightarrow AB$
- $DE \rightarrow B$

Are  $\Sigma$  and  $\Sigma_1$  equivalent or not? Justify your answer. (1 Mark)

**Question 3****2 Marks**

Consider the relation schema  $APPOINTMENT = \{Patient, GP, Date, Time, Clinic, Room\}$  and the following set  $\Sigma$  of FDs:

- $Patient, Clinic, Date \rightarrow Time$
- $Patient \rightarrow GP$
- $GP \rightarrow Clinic$
- $Clinic, Date, Time, Room \rightarrow Patient$
- $Patient, Date, Time \rightarrow Clinic, Room$

Is the above relation schema  $APPOINTMENT$  in BCNF? If not, identify a BCNF decomposition for  $APPOINTMENT$ . You need to include the main steps used for identifying the BCNF decomposition. Check if this BCNF decomposition is dependency preserving. (2 Mark)

**Question 4****4 Marks**

The following table contains the relational algebra operators covered in our course. You should only use these operators to answer the following questions.

$\sigma_{\varphi} R$	Selection by condition $\varphi$
$\pi_{A_1, \dots, A_n} R$	Projection onto the set of attributes $\{A_1 \dots, A_n\}$
$\rho_{R'(A_1, \dots, A_n)} R$	Renaming the relation name to $R'$ and attribute names to $A_1, \dots, A_n$
$\rho_{R'} R$	Renaming the relation name to $R'$
$\rho_{(A_1, \dots, A_n)} R$	Renaming the attribute names to $A_1, \dots, A_n$
$R_1 \cup R_2$	Union of two relations $R_1$ and $R_2$
$R_1 \cap R_2$	Intersection of two relations $R_1$ and $R_2$
$R_1 - R_2$	Difference of two relations $R_1$ and $R_2$
$R_1 \times R_2$	Cartesian product of two relations $R_1$ and $R_2$
$R_1 \bowtie_{\varphi} R_2$	Join of two relations $R_1$ and $R_2$ with the join condition $\varphi$
$R_1 \bowtie R_2$	Natural join of two relations $R_1$ and $R_2$
$\varphi_1 \wedge \varphi_2$	condition $\varphi_1$ AND condition $\varphi_2$
$\varphi_1 \vee \varphi_2$	condition $\varphi_1$ OR condition $\varphi_2$

Consider the following relation schemas:

STUDENT={SID, Name, Degree, College, Address, Phone} with the primary key {SID},

COURSE={CourseNo, College, Semester} with the primary key {CourseNo, Semester},

TUTOR={TID, Email, CourseNo, Semester} with the primary key {TID, CourseNo, Semester} and the foreign keys: [CourseNo, Semester]  $\subseteq$  COURSE[CourseNo, Semester] and [TID]  $\subseteq$  STUDENT[SID],

ENROL={SID, CourseNo, Semester, Unit, Status} with the primary key {SID, CourseNo, Semester} and the foreign keys: [CourseNo, Semester]  $\subseteq$  COURSE[CourseNo, Semester] and [SID]  $\subseteq$  STUDENT[SID].

- 4.1 Answer the following questions using relational algebra queries. You should only use the relational algebra operators in the above table. You are encouraged to use relational algebra expressions to represent intermediate results if needed. (2 Mark)

[a] List the phone numbers of students who studied COMP2400 in 'S2 2020' and became a tutor for COMP2400 in 'S2 2021'. (1 Mark)

[b] List the TIDs of tutors who had tutored exactly one course in 'S2 2021'. (1 Mark)

- 4.2 Optimise the following relational algebra query (Your marks will depend on how well you optimise the query in your solution). Additionally, draw the query trees of the query before and after your optimisation. (2 Mark)

$\pi_{SID, Address, Phone}(\sigma_{CourseNo='COMP2400'}((COURSE \bowtie ENROL) \bowtie_{ENROL.SID=STUDENT.SID} STUDENT))$   
(2 Mark)

+++++