COMP5112
Data Structures and Database Systems
Department of Computing
The Hong Kong Polytechnic University
Quiz 2 --- Solution Guide
Subject Lecturer: Dr. Kevin Yuen
Tuesday, 26 October 2024

Question 1 [45 marks]

Given part of the project selection database schema that contains two relations:

- Project (ProjectID, InstructorID, Title, Abstract)
- Instructor (InstructorID, FirstName, LastName, Phone, Email)

The *InstructorID* attribute in the *project* relation is a foreign key to the *Instructor* relation. All data fields are in string format. Answer the questions below.

1(a) (20 marks)

At the current timestamp, all project relation data is given in Table 1a. **Write** the relational algebra expressions and the corresponding SQL statements, respectively, that will result in Table 1b. The display order of tuples listed in the table can be omitted.

ProjectID	InstructorID	Title	Abstract
P1	N1	T1	A1
P2	N1	T2	A2
P3	N2	T3	A3
P4	N3	T4	A4
P5	N2	T5	A5
P6	N3	T6	A6

Table 1a: *Project* relation data (given)

ProjectID	InstructorID	Title	Abstract
P1	N1	T1	A1
P2	N1	T2	A2
P3	N2	T3	A3
P5	N2	T5	A5
P7	N4	T7	A7

Table 1b: *Project* relation data (final)

Solution:

relational algebra expressions	SQL statements
Project ← *Project* – $\sigma_{\text{InstructorID='N3'}}$ (*Project*) *Project* ← *Project* ∪ { ('P7','N4','T7','A7') }	DELETE FROM Project WHERE InstructorID = 'N3'; INSERT INTO Project VALUES ('P7','N4','T7','A7');

- The order for insert and delete is not important.
- Accept the other solutions, e.g.. :

  DELETE FROM Project WHERE ProjectID = 'P4';
  DELETE FROM Project WHERE ProjectID = 'P6';
  INSERT INTO Project VALUES ('P7','N4','T7','A7');

  Project ← *Project* – $\sigma_{\text{projectID='P4'}}$ (*Project*)
  Project ← *Project* – $\sigma_{\text{projectID='P6'}}$ (*Project*)
  Project ← *Project* ∪ { ('P7','N4','T7','A7') }

1(b) (10 marks)

Write a relational algebra expression to find the number of projects offered by each instructor.

Solution

$_{\textbf{InstructorID}}\mathbf{G}_{\textbf{count(ProjectID)}}(\textbf{\textit{Project}})$

1(c) **(15 marks)**

Write an SQL statement to find the last name, phone and email address of each instructor, who does not offer any projects.

Solution:

There are no more than one instructors who have the same phone and email address.

```
select lastName, phone, email
from instructor
where instructorID in
(select instructorID from instructor
minus
select Distinct instructorID from Project);
```

- **Accept the other solutions. For example,**

```
(select lastName, phone, email from instructor)
minus
(select lastName, phone, email
from instructor (inner/right)* join Project
on instructor.instructorID = Project.instructorID);
```

*optional, but left join is wrong.

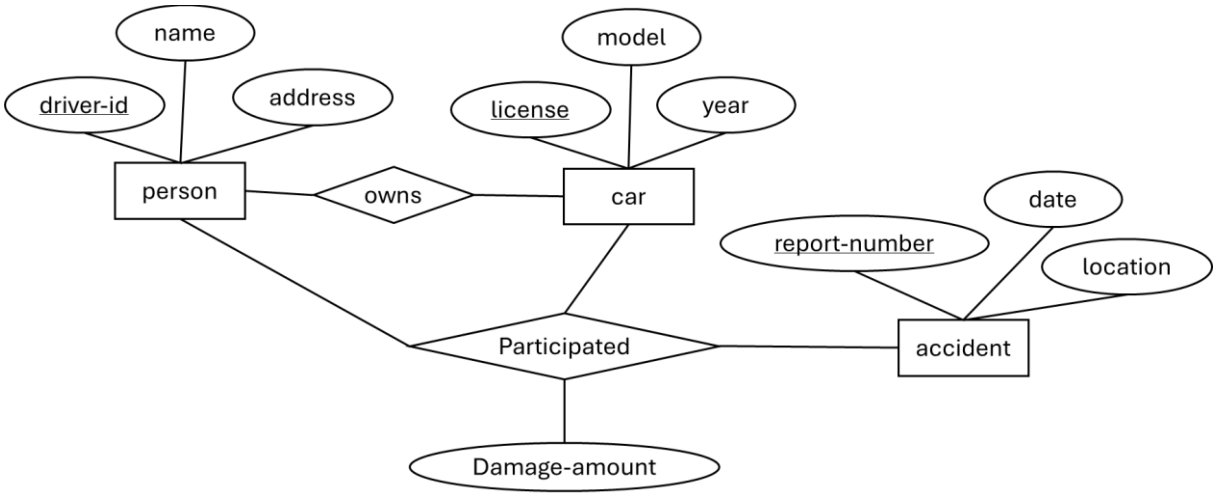
Question 2 [25 marks]

A database schema of a car insurance company contains the following relation schemas.

- person (driver-id, name, address)
- car (license, year, model)
- accident (report-number, date, location)
- owns (driver-id, license)
- participated (driver-id, license, report-number, damage-amount)

Construct an E-R diagram for a car insurance company whose customers own one or more cars each. Each car is associated with from zero to any number of recorded accidents.

Solution:



Question 3 **[30 marks]**

You are given a relational schema *R*= (A, B, C, D, E, F, G, H) and the following set $\mathcal{F}$ of functional dependencies.

- A B → C D
- B → E F
- B → G H
- A → B
- A → C
- A B → E F G H
- A → C D
- H → B
- G → B

3(a) **(10 marks)**

Find a candidate key for the relational schema *R*. **Show** that it is a candidate key.

Solution

If the closure of the **single** attribute A, denoted by A^+, is ABCDEFGH, A is a candidate key.
By using A → B, A^+ contains AB;
By using AB → CD, A^+ contains ABCD;
By using AB → EFGH, A^+ contains ABCDEFGH;
Thus, A is a candidate key.

- If the answer is AB/AH/AG, up to 5 marks will be awarded. AB/AH/AG is a super key but not a candidate key.
- Accept the other solutions.

3(b) **(20 marks)**

Find the canonical cover for $\mathcal{F}$. **Show** your steps.

Solution.

- A→BCD can replace A→B, A→C and A→CD.
- B→EFGH can replace B→EF and B→GH.
- A B → C D E F G H can replace A B → C D , A B → E F G H and B→EFGH
- B is extraneous in A B → C D E F G H
  Compute (A B – B) = A^+ = ABCDEFGH, which contains B. (Proof of A^+ is shown in 3-(a))
  Thus, A → C D E F G H, as B is removed from A B → C D E F G H
- A → B C D E F G H can replace A→BCD and A → C D E F G H.

Thus, the canonical cover for $\mathcal{F}$ is
{A → B C D E F G H , H → B , G → B }

- Accept the other solutions.