CS-A1155 Databases, Homework 1

Deadline: April 23, 2024 at 16:00 (late submission until April 24, 2024 at 10:00 with 75% of the points)

Please submit the solutions for the theoretical problems to the designated folder in A+ as a pdf file.

Please attend an exercise session (schedule can be found here).

EXERCISES MUST BE DONE INDIVIDUALLY!

At the start of the exercise session you can indicate wether you are willing to present/explain your solutions or not and a student will be picked at random to present, so **please make sure you arrive to your session on time**. If you won't be willing to do so your points for the theoretical problems will not count.

We are not trying to punish you by presenting, just making sure you actually did the homework yourself and creating a space for discussions. Thus, you need not have a perfect solution, nor even a correct one.

Note: If you have social anxiety or are unable to present for any reason, you can instead explain your working to the TA and in that case the TA will present the model solution!

Good luck with your homework!

Theoretical problems

1. Consider the following schema of restaurants and suppliers.

   ```
   Restaurants(restaurantID, name, address, owner)
   Suppliers(ID, name, phone, email)
   Contracts(restaurantID, supplierID, payment)
   ```

   Each restaurant may form a contract only once with each supplier, but a restaurant can have many suppliers, and a supplier can sign a contract with multiple restaurants. Suppose that there are 100 rows in Restaurants, 50 rows in Suppliers and 100 rows in Contracts. You may suppose that there are no NULL values in the tables.

   (a) (2p.) Explain briefly, what is the purpose of the following relational algebra expression, and what does it return.

   $$\pi_{restaurantID,owner}(Restaurants) - \pi_{restaurantID,owner}(Restaurants \bowtie Contracts).$$

   (b) (2p.) Explain briefly, what is the purpose of the following SQL query, and what does it return.

   ```
   SELECT DISTINCT Suppliers.ID, Suppliers.name
   FROM Suppliers, Contracts AS C1, Contracts AS C2
   WHERE C1.supplierID = C2.supplierID AND C1.restaurantID <> C2.restaurantID AND
   ID = C1.supplierID
   ```

   (c) (2p.) What is the maximal possible number of rows in the queries (a) and (b)? What is the minimal number of rows? Please justify your answers briefly.

2. Consider the following relation and explain what is **wrong** in the proposed solutions.

   ```
   Teacher (name, employeeNumber, email, phone)
   Course (name, credits, teacher)
   Student (name, studentID, email)
   Grade (studentID, courseName, grade)
   ```

   (a) (2p.) Task: List all the courses and their teachers (relational algebra).

   Proposed solution: Course $\bowtie$ Teacher.

(b) (2p.) Task: Find the IDs of all students who have completed both courses Programming 1 and Programming 2.

Proposed solution:

```
SELECT studentID
FROM Grade, Course
WHERE Grade.courseName = Course.name AND (Course.name = 'Programming 1' OR
Course.name = 'Programming 2')
```