CS-A1153 Databases, Homework 2

Deadline: May 4, 2022 at 18:00 (late submission until May 11, 2022 at 18:00 with 50% of the points)

The relational algebra statements and SQL queries will be written and submitted in A+ environment. The problem descriptions and schemas can also be found in A+. Please submit the solutions for the problems 5–10 to the designated folder in A+ as a PDF file.

These exercises utilize the same example database schema from round 1. Please, refer to Homework 1 assignment if needed. Also, the A+ assignments contain the needed schema per problem.

SQL queries

1. (1p.) Consider the hospital database. Write a SQL query to answer the following query: Find the appointments (with date, patient name, and doctor name) from year 2021, ordered by date (latest first) and then by patient name (from A to Z).

2. (1p.) Consider the hospital database. Write a SQL query to answer the following query: Find the name and BMI value of top 20 patients with the lowest BMI value. BMI value can be calculated as following (note that height of the database in centimeters, but the formula requires meters):

$$\frac{weight}{height \times height}$$

3. (1p.) Consider the hospital database. Write a SQL query to answer the following query: Find all the names of patients who have visited doctor *Dr. Nova Shanahan* during year 2021. List each patient only once.

4. (2p.) Consider the library database. You should use at least one subquery and one of the operators EXISTS and IN in your query. Write a SQL query to answer the following query: Find the IDs of the items that have never been loaned.

5. (5p.) Consider the following database schema.

   ```
   Student (name, studentID, email, studyProgram)
   Course (code, name, credits, teacher, MOOC)
   Grade (studentID, courseCode, grade)
   ```

   Write `CREATE TABLE` commands for corresponding tables with following constraints and default values:

   - The name and email of a student are string values and cannot be `NULL`.
   - The student ID is a numerical value and cannot be `NULL`.
   - The study program of a student is a string value, and it has to be one of the following: "data science", "computer science", "mathematics", or "industrial engineering".
   - The code, name, and teacher of a course are string values. The code and name cannot be `NULL`.
   - The amount of course credits is an positive numerical value and cannot be `NULL`. Due to historical reasons in university management, old courses might have also one decimal in the credit amount, e.g. 3.5.
   - The attribute `MOOC` defines whether certain course is an open online course. By default this should be false.
   - Grades are integer values between 0–5 and cannot be `NULL`.
   - Remember to define primary keys as well

UML modeling

6. (3p.) In this exercise you need to create small examples of different modelling structures. Use the UML notation taught on the course and remember to mark the primary keys. Write brief explanations about the examples you came up with.

   (a) Create an example of a many-many relationship. Give at least 3 attributes for both of the classes.

   (b) Create an example of a subclass structure with at least two subclasses. Give at least 3 attributes for the superclass and at least 1 attribute for each of the subclasses.

   (c) Create an example of an association class between two classes. Give at least 3 attributes for both of the classes.

7. (7 p.) Give an UML diagram for a database recording information about companies, employees, and customers, including:

   - For each company: its name, its business id, its employees, its CEO (one of its employees) and the cities where its offices are located.
   - For each employee: his/her name, social security number, position, date of birth, salary, and office where he/she is mainly working from.
   - For each customer: his/her name, his/her phone number, companies he/she is a customer for, and a contact person in each company (one of the employees).

   Make the following assumptions:

   - Each employee belongs to at most one company.
   - Each company has exactly one CEO.
   - All companies, employees, and customers have unique names.
   - Each customer may be a customer of several companies and has different contact person in each company. One employee can serve as contact person for several customers.
   - The companies may have offices in several cities. Each company has at least one office.

Functional dependencies

8. (3p.) Consider the library database.

```
LibraryCustomer (name, phone, email, libraryCardNumber)
LibraryItem (ID, loanPeriod)
Book (ID, title, author, language, year)
CD (ID, title, artist, year, length)
DVD (ID, title, year, length)
Loan (itemID, libraryCardNumber, startDate, endDate, returned)
```

   List all the non-trivial functional dependencies. You may assume that the emails and phone numbers are unique, and that one person can only own one library card.

9. (3p.) Consider the relation $R(A, B, C, D, E, F)$, and the functional dependencies $A \to F$, $D \to C$, $CF \to BE$, $B \to DE$. Which of the following functional dependencies can be derived from the listed ones? Justify your answer by computing the closures of the left-hand-sides.

   (a) $AD \to E$.

   (b) $BC \to ABCDEF$.

   (c) $AC \to BDEF$.

10. (4p.) Consider the following table with schema $R(A, B, C, D, E)$ and tuples

A	B	C	D	E
"123"	14	1	"A"	0.12
"123"	42	0	"A"	0.12
"417"	32	0	"G"	1.52
"510"	14	1	"G"	0.75
"810"	12	1	"C"	0.75
"811"	12	1	"E"	0.12

Which of the following functional dependencies are possible considering the contents of the table? Justify your answers.

(a) $A \rightarrow E$.

(b) $E \rightarrow A$.

(c) $D \rightarrow ABE$.

(d) $DC \rightarrow ABCDE$.

(e) $B \rightarrow C$.

(f) $BD \rightarrow E$.