Assignment 1

DBMS Overview and Relational Model

Assigned: Aug 30, 2024
Due: Sep 11, 2024. 11:59:59PM
Total Points: 50

Objectives

The purpose of this assignment is to help you reinforce the concepts we have been talking in class for databases, database management system, relational model. There will be a few questions asking for a concrete SQL statement. You should be fine with just the syntax we talked in the class. However, if you want a very concrete specification from a real-world DBMS, please refer to SQLite documentation as it will serve as the preferred SQL dialect in this course.

Please submit your solution to GradeScope before due date above. If you have further questions, please contact the instructor or any TA.

Part I. DBMS Overview [20 pts]

1. [5 pts] When designing an application, one important decision is whether to include a DBMS in the application stack, or to manually implement the necessary data access features as part of the application (i.e., store the data in some file format and write code for manipulating the file). Give two advantages and one disadvantage to each approach.

2. [5 pts] The word "consistency" is famously, and annoyingly, overloaded. *In the context of RDBMS*, what does it mean for the database to be consistent?

3. [5 pts] In an RDBMS, the data in a table can be stored in several different ways. For example, it might be sorted and stored as an array, or laid out as a search tree. From the end-user perspective, this layout need not be considered to construct correct queries. Is this an example of physical or logical data independence? Explain why this property is useful.

4. [5 pts] When creating a schema, one very important decision is what data type to use to best represent a certain piece of data. Say you are modeling some data that includes US ZIP codes. Do a little bit of research on ZIP codes and determine what data-type would be the best to properly model this information, out of INT, CHAR, FLOAT, VARCHAR. Explain the reason for your decision.

Part II. Relational Model [30 pts]

Throughout this part, all the questions are by default based on the relation instance provided below, which stores information pertaining to the faculty of a university.

ID	FNAME	LNAME	MNAME	SSN	BDAY	DEPT	COLLEGE
2	Michael	Stonebraker	Ralph	371926461	10/11/1943	CS	ENG
1	Srinivasa	Ramanujan	NULL	859128361	12/22/1887	MATH	SCI
4	Charles	de Gaulle	Andre	193475831	11/22/1890	EGO	LAW
2	Thomas	Aquinas	NULL	837495038	NULL	REL	ART
2	Alan	Turing	Mathison	176484938	6/23/1912	CS	SCI
1	Giacomo	Puccini	Antonio	137401834	12/22/1858	MUS	ART
7	Muhammad	al-Khwarizmi	Musa	173710274	7/3/780	MATH	SCI
3	Martin	Luther	NULL	172947152	11/10/1483	REL	ART
5	Fyodor	Dostoevsky	Mikhailovich	184618284	11/11/1821	LIT	ART
3	Michael	Faraday	NULL	123821827	9/22/1791	PHYS	SCI

1. [5 pts] What is the schema for this relation? Please include value domain for each attribute as well. You may choose your own relation name, and the value domain in your answer must be compatible with the instance given above. If multiple value domains are possible for an attribute, use the most strict one that fits the given database instance.

Note: You should use the relation name you choose here in other questions of this part.

2. [5 pts] Name five different super keys for the given relation.

3. [6 pts] Name three candidate keys with two attributes in the given relation.

4. [4 pts] Consider a general dataset with the same schema as you answered in the first question **(NOT the instance provided above)**. Choose a reasonable three-attribute primary key and briefly explain why it makes sense.

5. [2 pts] Provide a SQL statement for inserting a new record to the relation for a faculty member named "John Doe" who has an appointment in the "ENG" college and "ECE" department with id number "15" and SSN "987364830".

6. [2 pts] Provide a SQL statement that remove all records for faculty members from the "SCI" college.

7. [2 pts] Provide a SQL statement that updates the birthday of "John Doe" to December 11, 2001.

8. [2 pts] Provide a SQL statement that drops the attribute "BDAY" of the relation.

9. [2 pts] Suppose we specify "(COLLEGE, DEPT)" is a foreign key referring to another relation called "Departments" and set its referential action as "ON DELETE SET NULL". What will happen when we try to remove the record with primary key "(ENG, CS)" in the "Departments" relation.