CS 457 – Data modeling and Implementation Techniques

Instructor: Dr. Vidhyacharan Bhaskar

Homework 1A: (Due: Sep, 11, 2021 at 9 am)

1. What is a DBMS? What do we use it for?

Ans => A Computerized system that lets users to create and maintain database is called Database Management system (DBMS).

We use it for several reasons =>

1. To Define database with specific data types, structure and constraints of the data.
2. To Create the database is the process where the information is stored by the DBMS in form of database catalog or dictionary on some storage medium that is controlled by the DBMS.
3. To manipulate a database in the many ways like retrieve specific data, update database, generate reports from data.
4. To Share database with multiple users and programs.
5. To Protect system against hardware or software malfunction and to protect security against unauthorized access.
6. To maintain database system with allowing it to grow as per requirements change.
7. And to provide controlled access to data.
8. What is a Relational DBMS? How is it different from DBMS? Give examples for both.

Ans => Relational Database Management System (RDBMS) is a type of database that provides access to data points that are related to one another and way of representing data in tables as a record. In relational database, each row is record with unique ID key and each record usually has a value for each attribute (column), makes it easy to create the relationship among data points.

e.g. A small hospital might use relational database to process orders for its treatment.

a) In the first table you can add Patient’s contact info like Patient’s name, address, phone number, appointment time.

b) in Second table you can add a Patient’s ordered treatment info which includes ID(key) of the Patient, the medicine, from which doctor, copay, and so on.

RDBMS vs DBMS =>

RDBMS is an advanced version of a DBMS system. RDBMS system also allows the organization to access data more efficiently than DBMS. DBMS stores data as file while RDBMS used to store only data which is stored in tabular form.

E,g. for DBMS of small hospital =>

The database is organized as five files, Patients, contact information, treatment given, Doctor and copay.

1. Which user groups interact with a running DBMS?

Ans => Different user groups like the administrator, IT and database admins, application integrators and data consumers needs easily understandable query language and intuitive UI to use the DBMS system efficiently.

4) Describe the following terms:

a. Data model. What are its three parts?

A data model is a collection of concepts used to describe the structure of a database, also supports data abstraction so user can perceive data at their preferred level of detail.

\* High level or conceptual data model(based on entities and relationships)

\* Low level or physical data models

\* Representational or implementation data models(record based, object oriented)

b. Logical schema

Logical Schema represents the abstract structure of a domain of information and mostly used in business processes to capture things of importance to an organization and how they relate to one another. Physical schema is also called Logical schema

c. Conceptual schema

The area used to provide concepts that are close to the way many users perceive data is called Conceptual schema. (Also called entity-based or object-based data models.)

d. Physical schema

The area used to provide concepts that describe details of how data is stored in the computer is called Physical schema. These are usually specified in an ad-hoc manner through DBMS design and administration manuals.

1. How is SQL used in RDBMS?

Ans => **SQL** (Structured Query Language) is a programming language **used** to communicate with data stored in a **relational database**  management system.

SQL allows the joining of tables using a few lines of code, with a structure most nontechnical employees can learn quickly.

1. What is redundancy? Give one advantage and disadvantage of redundancy.

Ans => Data redundancy occurs when the same piece of data (duplicate) is stored in two or more separate places and**i**s a common occurrence in many businesses.

Advantages =>1. Alternative data backup method

2. Better data security

3. Faster data access and updates

4. Improved data reliability

Disadvantages =>1) Possible data inconsistency

2) Increase in data corruption

3) increase in database size

4)Increase in cost

Elmasri 7e book (page 28-29)

1. Problem 1.8

Identify some informal queries and update operations that you would expect to apply to the database shown in Figure 1.2

. Examples of queries are as follows

■ Retrieve the transcript—a list of all courses and grades—of ‘Smith’

■ List the names of students who took the section of the ‘Database’ course offered in fall 2008 and their grades in that section

■ List the prerequisites of the ‘Database’ course

Examples of updates include the following:

■ Change the class of ‘Smith’ to sophomore

■ Create a new section for the ‘Database’ course for this semester

■ Enter a grade of ‘A’ for ‘Smith’ in the ‘Database’ section of last semester

1. Problem 1.10

Specify all the relationships among the records of the database shown in Figure 1.2

* The record for ‘Brown’ in the STUDENT file is related to four records in the GRADE\_REPORT file. Similarly, each section record is related to one course record and to a number of GRADE\_REPORT records—one for each student who completed that section.

1. Problem 1.12

Cite some examples of integrity constraints that you think can apply to the database shown in Figure 1.2

=> we can specify that every section record must be related to a course record. This is known as a referential integrity constraint.

Another type of constraint specifies uniqueness on data item values, such as every course record must have a unique value for Course\_number. This is known as a key or uniqueness constraint.

1. Problem 1.14

Consider Figure 1.2

1. If the name of the CS (Computer Science) department changes to CSSE (Computer Science and Software Engineering) Department and the corresponding prefix for the course number also changes, identify the columns in the database that would need to be updated.

* COURSE\_NUMBER, DEPARTMENT and PrerequisiteNumber these columns in the database would need to be updated.

1. Can you restructure the columns in the COURSE, SECTION, and PREREQUISITE tables so that only one column will needs to be updated ?

* You can restructure the columns into only one column as follow

CourseDept and CourseNum

PreReqDept and PreReqNum