

Chapter 1

Introduction to Database Management System (DBMS)

-Niraj K.C.

Table of Content

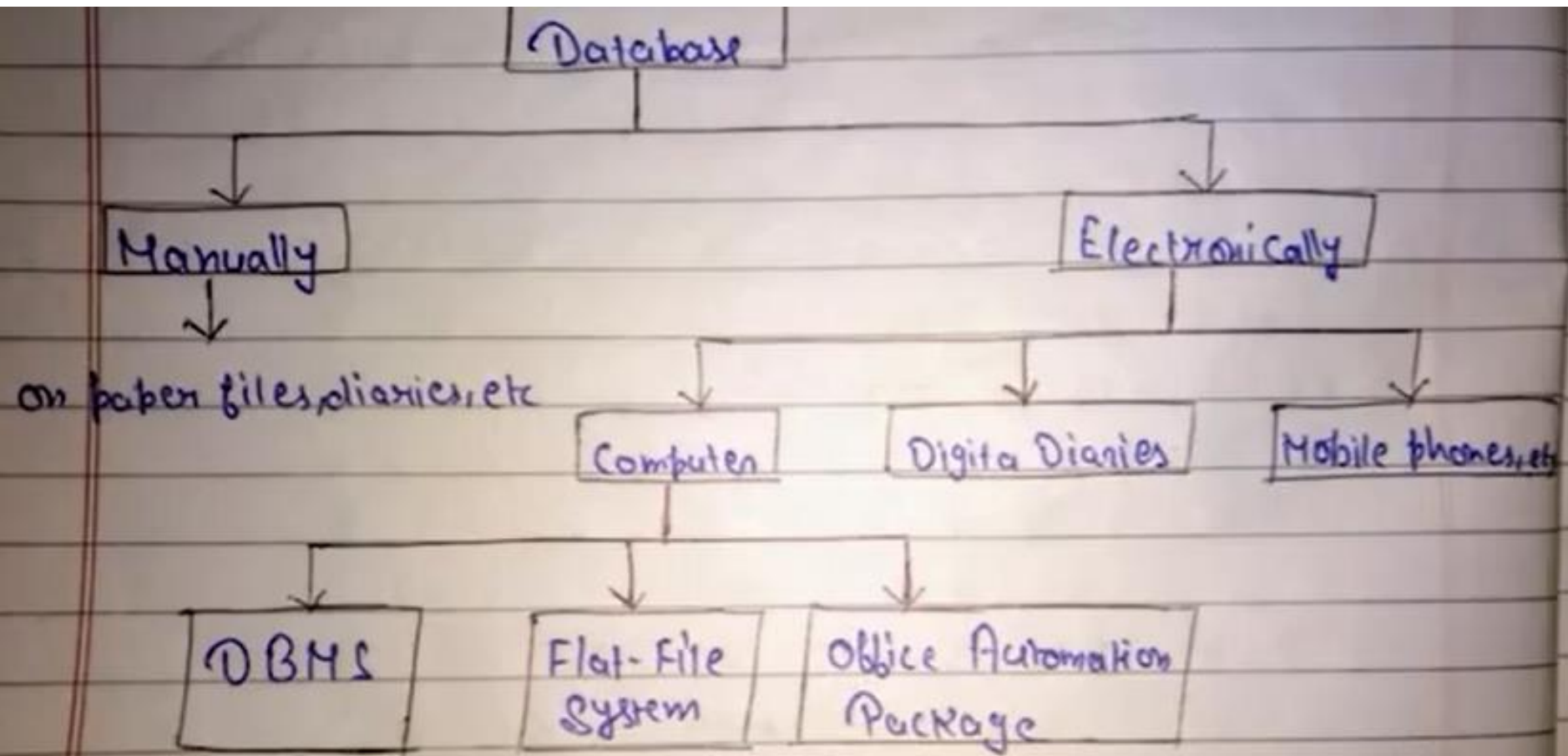
- General overview of DBMS, Terms and Terminologies.
- Application of DBMS: Where and how it is used, Comparison with other data technologies.
- Database technology for geospatial data, Basic characteristics of a DBMS.
- Components of DBMS, function of DBMS, interaction with DBMS

Database Introduction:

DBMS: Database Management System

Database + Management + System

- Database is organized shared collection of logically related data in a systematic manner , that is stored to meet the requirement of different users of an organization, institution , government bodies , that can be easily accessed, managed and updated.
- Database is actually a place where related piece of information is stored and various operation can be performed on it.
- Database can be maintained manually or through electronic devices such as : digital diaries, Mobile phones, Computers etc.



DBMS

- A DBMS is a Software (or a collection of programs) that enables users to create, store, modify, and extract information from database as per the requirements.
- A DBMS is a tool, that is used to perform any kind of operation on data in database.
- Some DBMS examples: MY SQL, Oracle, System 2000, MS Excess, MY SQL Server, etc.

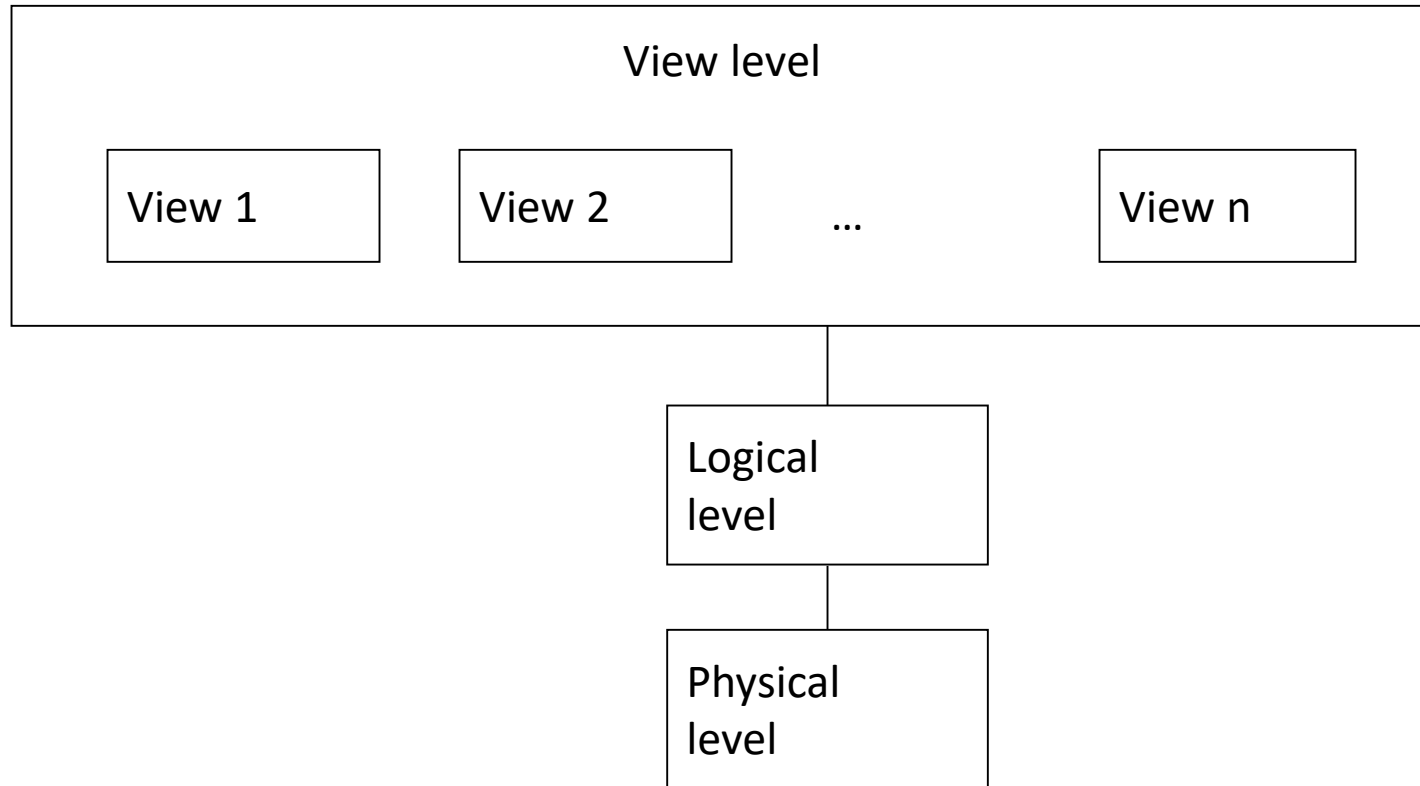
Levels of Abstraction: View of Data

- **Physical level:** describes how a record (e.g. customer) is stored.
- **Logical level:** describes data stored in database, and the relationships among the data.

```
type customer = record
    name: string;
    street: string;
    city: integer;
end;
```

- **View level:** application programs hide details of data types. Views can also hide information (e.g. salary) for security purposes.

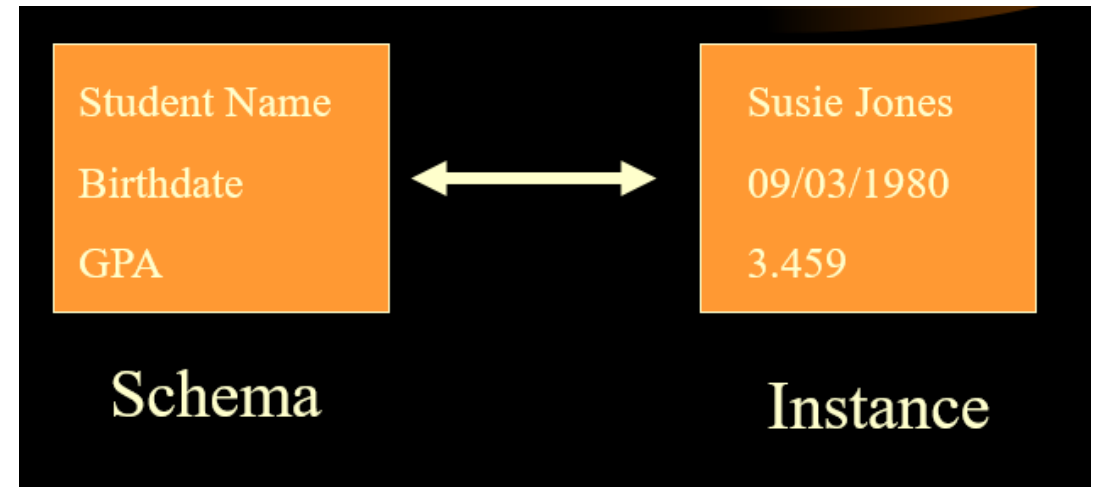
An architecture for a database system



Terms and Terminologies

Instance and Schemas:

- Similar to types and variables in programming languages
- The collection of information stored in the database at a particular moment is called an **instance** of the database. The overall design of the database is called the **Schema** of database
- Schema – the logical structure of the database (e.g., set of customers and accounts and the relationship between them)
- Instance – the actual content of the database at a particular point in time



Columns: Columns are similar to fields, that is, individual items of data that we wish to store. A Student's Roll Number, Name, Address etc. are all examples of columns. They are also similar to the columns found in spreadsheets (the A, B, C etc. along the top).

Rows: Rows are similar to records as they contain data of multiple columns (like the 1, 2, 3 etc. in a spreadsheet). A row can be made up of as many or as few columns as you want. This makes reading data much more efficient - you fetch what you want.

Tables: A table is a logical group of columns. For example, you may have a table that stores details of customers' names and addresses. Another table would be used to store details of parts and yet another would be used for supplier's names and addresses.

Example:

Student Table

Roll Number	Name	City
1	Ajay	Amritsar
2	Rajesh	Patiala
3	Rahat	Delhi

Column

Row

Characteristics of database

The data in a database should have the following features:

- **Organized/Related:** It should be well organized and related.
- **Shared:** Data in a database are shared among different users and applications.
- **Permanent or Persistence:** Data in a database exist permanently in the sense the data can live beyond the scope of the process that created it.
- **Validity/integrity/Correctness:** Data should be correct with respect to the **real world entity** that they represent.
- **Security:** Data should be protected from unauthorized access.
- **Consistency:** Whenever more than one data element in a database represents related real world values, the values should be consistent with respect to the relationship.
- **Non-redundancy:** No two data items in a database should represent the same real world entity.
- **Independence:** Data at different levels should be independent of each other so that the changes in one level should not affect the other levels.
- **Easily Accessible:** It should be available when and where it is needed i.e. it should be easily accessible.
- **Recoverable:** It should be recoverable in case of damage.
- **Flexible :** It should be flexible to change

Components of DBMS

Software

The main component of a DBMS is the software. It is the set of programs used to handle the database and to control and manage the overall computerized database

Hardware

Hardware consists of a set of physical electronic devices such as computers (together with associated I/O devices like disk drives), storage devices, I/O channels, electromechanical devices that make interface between computers and the real world systems etc, and so on. It is impossible to implement the DBMS without the hardware devices.

Data

Data is the most important component of the DBMS. The main purpose of DBMS is to process the data. In DBMS, databases are defined, constructed and then data is stored, updated and retrieved to and from the databases. The database contains both the actual (or operational) data and the metadata (data about data or description about data).

Procedures

Procedures refer to the instructions and rules that help to design the database and to use the DBMS. The users that operate and manage the DBMS require documented procedures on how to use or run the database management system. These may include. 1. Procedure to install the new DBMS. 2. To log on to the DBMS. 3. To use the DBMS or application program. 4. To make backup copies of database. 5. To change the structure of database. 6. To generate the reports of data retrieved from database.

Database Access Language

Database Access Language

The database access language is used to access the data to and from the database. The users use the database access language to enter new data, change the existing data in database and to retrieve required data from databases. The user write a set of appropriate commands in a database access language and submits these to the DBMS. The DBMS translates the user commands and sends it to a specific part of the DBMS called the Database Jet Engine. The database engine generates a set of results according to the commands submitted by user, converts these into a user readable form called an Inquiry Report and then displays them on the screen. The administrators may also use the database access language to create and maintain the databases. The most popular database access language is SQL (Structured Query Language). Relational databases are required to have a database query language.

Users:

Users The users are the people who manage the databases and perform different operations on the databases in the database system. There are three kinds of people who play different roles in database system

1. Application Programmers

The people who write application programs in programming languages (such as Visual Basic, Java, or C++) to interact with databases are called Application Programmer.

2. Database Administrators

A person who is responsible for managing the overall database management system is called database administrator or simply DBA.

3. End-Users

The end-users are the people who interact with database management system to perform different operations on database such as retrieving, updating, inserting, deleting data etc.

Database Language:

A DBMS must provide appropriate languages and interfaces for each category of users to express database queries and updates. Database Languages are used to create and maintain database on computer. There are large numbers of database languages like Oracle, MySQL, MS Access, dBase, FoxPro etc. SQL statements commonly used in Oracle and MS Access can be categorized as data definition language (DDL), data control language (DCL) and data manipulation language (DML).

Data Definition Language (DDL)

It is a language that allows the users to define data and their relationship to other types of data. It is mainly used to create files, data dictionary and tables within databases. It is also used to specify the structure of each table, set of associated values with each attribute, integrity constraints, security and authorization information for each table and physical storage structure of each table on disk. The following table gives an overview about usage of DDL statements in SQL

Specify a database schema by a set of definition expressed by a special language called a Data-Definition Language.

For instance, the following statement in the SQL language defines the account table:

e.g. create table account(account-number
char (10),balance integer)

Execution of the above DDL statement create the account table.

S.No	Need and Usage	The SQL DDL statement
1	Create schema objects	CREATE
2	Alter schema objects	ALTER
3	Delete schema objects	DROP
4	Reneme schema objects	RENAME

Data Manipulation Language (DML)

It is a language that provides a set of operations to support the basic data manipulation operations on the data held in the databases. It allows users to insert, update, delete and retrieve data from the database. The part of DML that involves data retrieval is called a query language.

The following table gives an overview about the usage of DML statements in SQL:

S. No	Need and Usage	The SQL DML statement
1	Remove rows from tables or views	DELETE
2	Add new rows of data into table or view	INSERT
3	Retrieve data from one or more tables	SELECT
4	change column values in existing rows of a table or view	UPDATE

- The retrieval of information stored in the database.
- The insertion of new information into the database.
- The deletion of information from the database.
- The modification of information stored in the database.
- A **data-manipulation language(DML)** is a language that enables users to access or manipulate data as organized by the appropriate data model. There are basically **two types** :

1. Procedural DMLs

2. Declerative DMLs

- The query in the SQL language finds the name of the customer whose customer-id is 1999:

```
Select Customer.customer-name  
From customer  
Where customer.customer-id = 1999
```

Data Control Language (DCL)

DCL statements control access to data and the database using statements such as GRANT and REVOKE. A privilege can either be granted to a User with the help of GRANT statement. The privileges assigned can be SELECT, ALTER, DELETE, EXECUTE, INSERT, INDEX etc. In addition to granting of privileges, you can also revoke (taken back) it by using REVOKE command. The following table gives an overview about the usage of DCL statements in SQL:

S. No.	Need And Usage	Age
1	Grant and take away privileges and roles	Grant Revoke
2	Add a comment to the data dictionary	Comment

Note:

- In practice, the data definition and data manipulation languages are not two separate languages. Instead they simply form parts of a single database language such as Structured Query Language (SQL). SQL represents combination of DDL and DML, as well as statements for constraints specification and schema evaluation.
- A database object in a relational database is a data structure used to either store or reference data. The most common object that most people interact with is the table. Other objects are indexes, stored procedures, sequences, views and many more.

Advantage Of DBMS

- Controlled redundancy
- Reduced inconsistency in the data
- Shared access to data
- Standards enforced
- Security restrictions maintained
- Integrity maintained easier
- Providing Backup and recovery

Disadvantage of DBMS

- Increased complexity in concurrency control
- Increased complexity in centralized access control
- Security needed to allow the sharing of data
- Needed redundancies cause updating complexities

Assignment 1:

1. Compare DBMS with other Data technologies like Flat File system, Office Automation Package etc.
2. Where and How DBMS can be used? (Application Of DBMS)
3. What is Geospatial Data? Describe Database Technology for Geospatial Data in detail with Suitable example.
4. List out the function of DBMS.

...Thank You...