

Characteristics of the Database Approach

In this lesson we will discuss the characteristics of a database.

WE'LL COVER THE FOLLOWING

- Characteristics of a database
 - 1. Self-describing nature of a database system
 - 2. Insulation between program and data
 - 3. Support for multiple views of data
 - 4. Sharing data and multiuser system

Characteristics of a database

Several characteristics distinguish the database approach from the file-based system. These include:

1. Self-describing nature of a database system

A database system is self-describing because it not only contains the database itself, but also the **meta-data** which defines and describes the data and relationships between tables in the database. This information is stored in a catalog by the DBMS software. The separation of data and information about the data makes the database system different from the traditional file-based system in which the data definition is part of the application programs.

Let's consider the university database example in the previous lesson. The DBMS catalog will store the definitions of all the tables in that database. The figure below shows part of the database catalog:

Tables

Table_name	Number_columns
------------	----------------

Student	3
Course	3
Department	2
Instructor	3

Columns

Attributes	Datatype	Belongs_to
ID	Integer (4)	Student
First_Name	Character (50)	Student
Last_Name	Character (50)	Student
Course_name	Character (50)	Course
Course_Section	Integer (4)	Course
Department_Name	Character (50)	Department
...

Now consider the scenario in which a user wants to access the information of a student, i.e., the student personal information(ID, first name, and last name), the department to which the student belongs, etc. Although the information is stored in separate tables, the DBMS software will refer to the catalog in order to determine the structure of these tables as well as the relationships between them. In this way, the DBMS will retrieve the information relevant to the user.

2. Insulation between program and data

In the file-based system, the structure of the data files is defined in the application programs so if a user wants to change the structure of a file, all the programs that access that file might need to be changed as well.

On the other hand, in the database approach, the data structure is stored in the system catalog and not in the programs. Therefore, one change is all that is needed to change the structure of a file. This insulation between the programs and data is also called **program-data independence**.

For example, we have an application program that is designed to work with the current structure of the STUDENT table. If we want to add another piece of data to the STUDENT table, say a column called `Home_Address`, such a program will no longer work and would have to be changed.

On the other hand, in a DBMS environment, we only need to change the description of the STUDENT table in the catalog to reflect the inclusion of the new data item `Home_Address` in each student record. So the next time a DBMS program refers to the catalog, the new structure of the STUDENT table will be used.

3. Support for multiple views of data

A database supports multiple views of data. A **view** is a subset of the database, which is defined and dedicated for particular users of the system. Multiple users in the system might have different views of the system. Each view might contain only the data of interest to a user or group of users. For example, one user of the university database may be interested only in accessing the names of the professors in each department. The view for this user is shown below:

Department_code	Department_name	Instructor_name
1	Computer Science	Sam Winchester
		Doug Waterman
		Glen Orsburn
		Dean Wright

2	Electrical Engineering	Dean Wright Terri Keane James Avery
---	------------------------	---

The user will only be presented with the relevant information rather than the whole database.

4. Sharing data and multiuser system

Current database systems are designed for multiple users. That is, they allow many users to access the same database at the same time. This access is achieved through features called **concurrency control strategies**. These strategies ensure that the data accessed are always correct and that data integrity is maintained. Consider the case when multiple reservation agents try to assign a seat on an airline flight, the DBMS should ensure that each seat can be accessed by only one agent at a time so that the same seat is not assigned to multiple passengers.

In the next lesson, we will go over the benefits of using the database approach.