UNIT I

INTRODUCTION

Main Characteristics of the Database Approach

- A number of characteristics distinguish the database approach from the much older approach of programming with files.
- The main characteristics of the database approach versus the file-processing approach are the following:
  - **Self-describing nature of a database system**
  - **Insulation between programs and data**
  - **Support of multiple views of the data**
  - **Sharing of Data and Multiuser Transaction Processing**

Data Models

- **Data Model**:
  - A set of concepts to describe the *structure* of a database.
  - By structure of a database we mean the data types, relationships, and constraints that apply to the data.

- **Data Model Operations**:
  - Operations for specifying database retrievals and updates on the database
  - Operations on the data model may include *basic operations* and *user-defined operations*.

Categories of Data Models

Many data models have been proposed, which we can categorize according to the types of concepts they use to describe the database structure.

1) **Conceptual** (**high-level**, **semantic**) data models: Provide concepts that are close to the way many users *perceive* data. (Also called **entity-based** or **object-based** data models.)

2) **Physical** (**low-level**, **internal**) data models: Provide concepts that describe details of how data is stored in the computer. *Concepts provided by low-level data models are generally meant for computer specialists, not for end users.*

3) **Implementation** (**representational**) data models: Provide concepts that fall between the above two, balancing user views with some computer storage details. Representational data models hide many details of data storage on disk but can be implemented on a computer system directly

Schemas versus Instances

- **Database Schema**:

  The *description* of a database. Includes descriptions of the database structure and the constraints that should hold on the database.

- **Schema Diagram**:

  A diagrammatic display of (some aspects of) a database schema.

- **Schema Construct**:

  A component of the schema or an object within the schema, e.g., STUDENT, COURSE.

- **Database Instance**:

  The actual data stored in a database at a *particular moment in time*. Also called **database state** (or **occurrence**).

Database Schema Vs. Database State

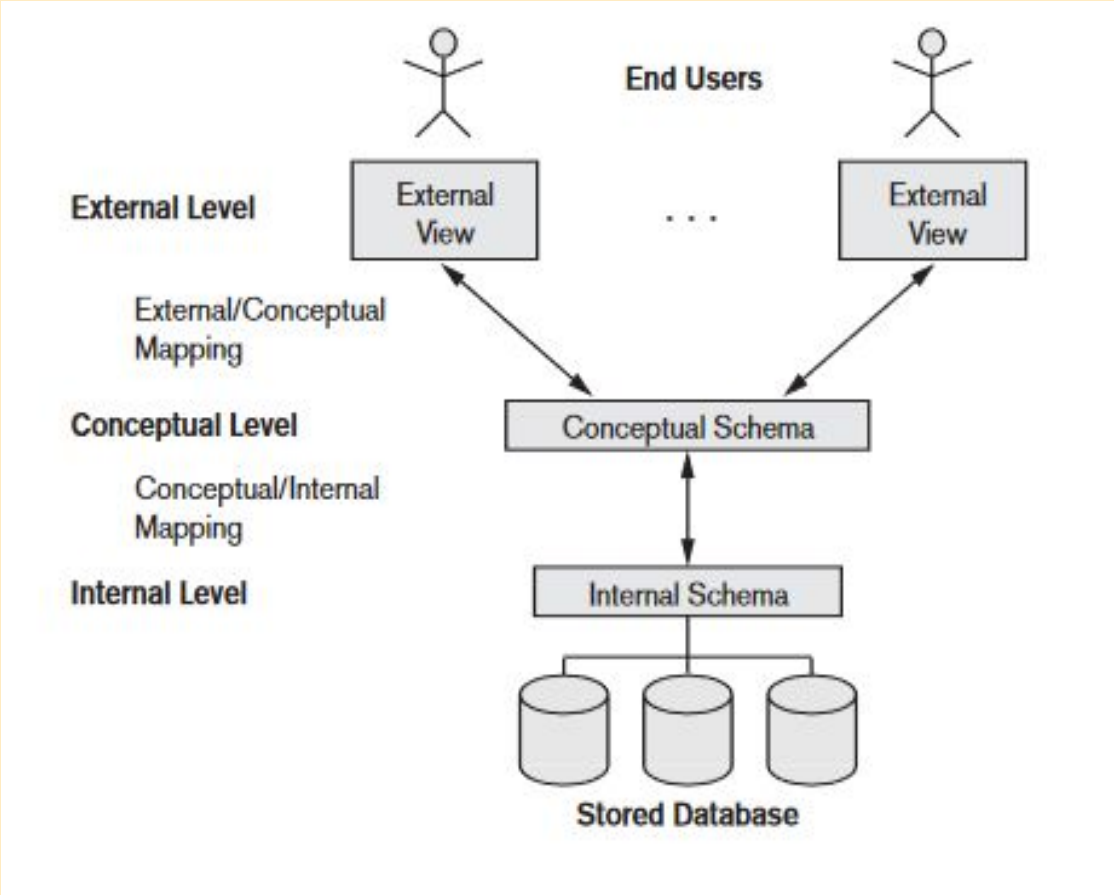
- **Database State:** Refers to the content of a database at a moment in time.
- **Initial Database State:** Refers to the database when it is loaded
- **Valid State:** A state that satisfies the structure and constraints of the database.
- **Distinction**
  - The **database schema** changes *very infrequently*. The **database state** changes *every time the database is updated.*
  - **Schema** is also called **intension**, whereas **state** is called **extension**.

The Three Schema Architecture

- The three-schema architecture, was proposed to help achieve and visualize three of the four important characteristics of the database approach viz.,
  (1) use of a catalog to store the database description (schema) so as to make it self-describing,
  (2) insulation of programs and data (program-data and program-operation independence), and
  (3) support of multiple user views.

- **The goal of the three-schema architecture, is to separate the user applications from the physical database.**

The Three Schema Architecture



- In this architecture, schemas can be defined at the following three levels:
  - Internal level.
  - Conceptual level and
  - External level

The Three Schema Architecture

1) The Internal Level :

- Has an **internal schema** which **describes the physical storage structure of the database**.
- The internal schema uses a **physical data model** and describes the complete details of data storage and access paths for the database.

2) The Conceptual Level:

- Has a **conceptual schema**, which **describes the structure of the whole database for a community of users**.
- Hides the details of physical storage structures and concentrates on describing entities, data types, relationships, user operations, and constraints.
- A **representational data model** is used to describe the conceptual schema when a database system is implemented

The Three Schema Architecture

3) The External or View level:

- Includes a number of **external schemas or user views**.
- Each external schema describes the part of the database that a particular user group is interested in and hides the rest of the database from that user group.
- Each external schema is typically implemented using a **representational data model**.

Note:

The three schemas are only descriptions of data; the stored data that actually exists is at the physical level only.

Advantages of Using the Database Approach

- Controlling redundancy in data storage and in development and maintenance efforts.
- Sharing of data among multiple users.
- Restricting unauthorized access to data.
- Providing persistent storage for program Objects
- Providing Storage Structures for efficient Query Processing
- Providing backup and recovery services.
- Providing multiple interfaces to different classes of users.
- Representing complex relationships among data.
- Enforcing integrity constraints on the database.
- Drawing Inferences and Actions using rules