SOF202

Database

Dr Subashini Ganapathy
subashini.ganapathy@xmu.edu.my

Chapter 1
Database

A general introduction

(Part 1)

Learning outcomes

1. Define a data, database, database management system and type of database.
2. Define the components of a database management system
3. A simplified database system environment.
4. Example of Database (University)

WHAT IS DATA?

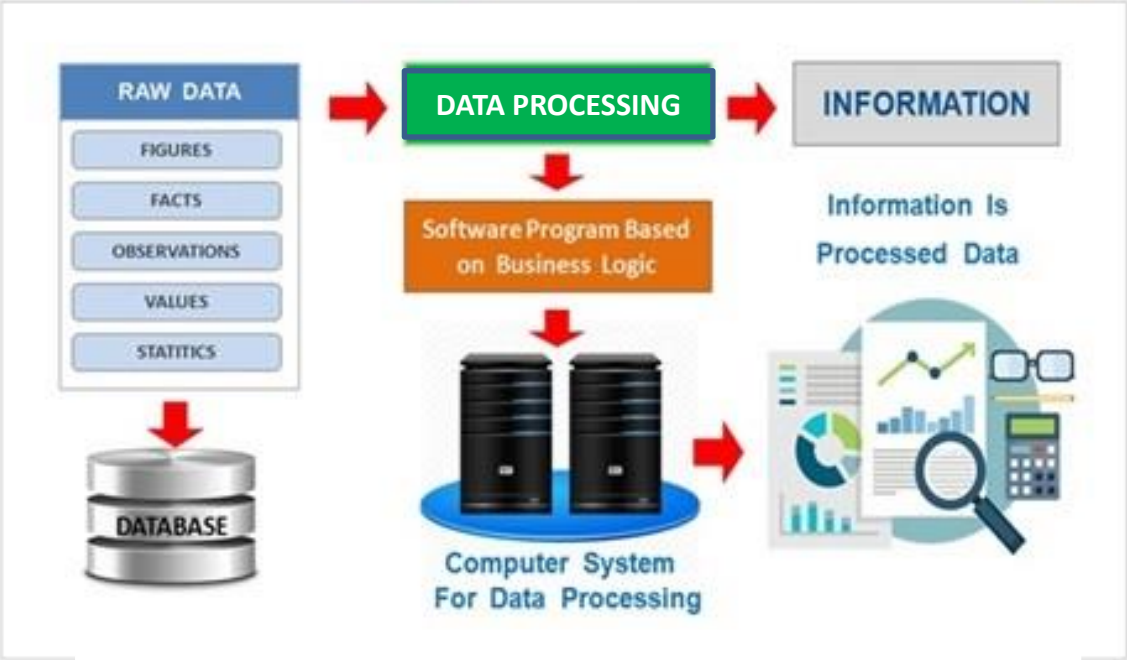
- Data is just the data and information that can be stored
- it is typically **unprocessed** and raw.
- Once we have put our data into context, **data is transformed into information**, which is then used to make decisions.
- Lets say we have a data of number 23, 50, and 70. These are the number of students registered for other classes.
  - Data are not all that valuable on their own. To be useful, it needs to be given **context**.
  - **Adding the context** that the numbers represent the number of students who have registered for a particular class makes a difference → **Converted data into information**.

Databases, Data, and Information

Database	Data	Information

• Collection of data organized in a manner that allows access, retrieval, and use of that data	• Collection of unprocessed items
	  • Text
	  • Numbers
	  • Images
	  • Audio
	  • Video
		• Processed data
		  • Organized
		  • Meaningful
		  • Useful

Data Vs. Information Vs. Knowledge

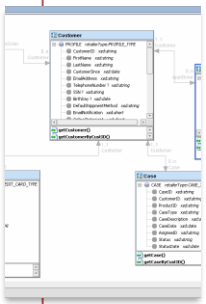


WHAT'S THAT DATABASE?

Database = A **large collection** of related data

A **set of data** as a whole, in most cases, some way to update it.

The database is a collection of **interconnected data** that is used to efficiently **retrieve**, **insert**, and **remove** data.



Classically, a DB models a real-world organization (e.g., enterprise, university)

- Entities (e.g., students, courses)
- Relationships (e.g., “Joanne is taking SOF202 in Semester 2023/04”)

Changes in the organization = changes in the database

- Examples: – employee’s detail, banking, airline reservations

DATABASE MANAGEMENT SYSTEM (DBMS)

- **Compilation** of related data
- A **collection of programs for accessing data**
- A database management system (DBMS) **stores information about a specific organization**.
- **Reliable** and **easy** to use

Financial transactions

Reservations and schedules for airlines

Registration and grades at universities

Customers, goods, and orders are all forms of sales

Production, inventory, orders, and the supply chain are all aspects of manufacturing

Employee reports, wages, and tax deductions are all part of human resources

DATABASE MANAGEMENT SYSTEM (DBMS)

DBMS gives us an **interface** or tool to carry out a variety of tasks, including **building databases**, **storing data** in them, **updating** data, **creating tables** in the databases, and much more.

Defining a particular database in terms of

- **its data types ,structures , and constraints .**

Manipulating the database

- **Retrieval: Querying, generating reports**
- **Modification: Insertions, deletions and updates to its content**
- **Accessing the database through Web applications**

Processing and sharing by a set of concurrent users and application programs yet, keeping all data valid and consistent.

DATABASE MANAGEMENT SYSTEM (DBMS)

- Additionally, DBMS offers **security** and **protection** to the databases.
- **Maintaining data consistency** when **multiple users** are present.
- Example of DBMS software:
  - MySql
  - Oracle
  - Ms SQL Server

TYPES OF DATABASE

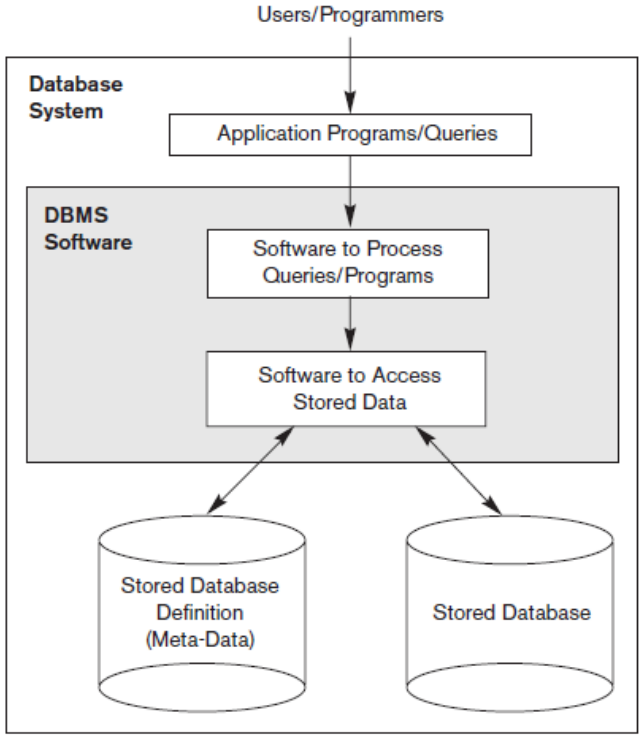
- Traditional Applications:
  - Numeric and Textual Databases
- More Recent Applications:
  - Multimedia Database
    - (images, audio, video…)
  - Geographic Information Systems (GIS)
    - Store and analyze maps, weather data, and satellite images

TYPES OF DATABASE

- Data Warehouses and online analytical processing (OLAP) systems
  - **Extract and analyze** useful business information **from very large databases**.
  - Support decision making.
- Real-time and Active Databases
  - A **real-time database** is a database system that **processes data in real time** to **handle workloads** that are **constantly changing**.

A simplified database system environment

is a collective system of elements that comprises and regulates the group of data, management, and use of data. This system includes **hardware**, **software**, **people**, database handling **techniques**, and the **data** itself.



Con’t

Hardware

- The computers and other devices used to handle databases.

Software

- refers to the entire system, including the operating system (OS) and application programmes like SQL Server, MySQL which are database administration programmes.

People

- Who use and administer the database system.

Technique / procedures

- guidelines, ideas, and directives provided to people, software, and a collection of facts and information positioned within a database system.

Data

- Data to be applied

EXAMPLE

Part of a UNIVERSITY environment

- Entities:-
  - STUDENT
  - COURSE
  - SECTION
  - GRADE_REPORT
  - PERQUISITES

STUDENT

Name	Student_number	Class	Major
Smith	17	1	CS
Brown	8	2	CS

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	07	King
92	CS1310	Fall	07	Anderson
102	CS3320	Spring	08	Knuth
112	MATH2410	Fall	08	Chang
119	CS1310	Fall	08	Anderson
135	CS3380	Fall	08	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Figure 1.2
A database that stores student and course information.

Define the database

Specify the structure of the records of each file by specifying the different **types of data elements** to be stored in each record.

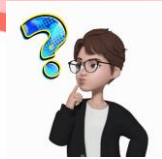
- Integer (Student_numberof STUDENT)
- String of Alphabetic characters (Name of STUDENT)
- Single character (Grade of GRADE_REPORT)
- Coding scheme to represent the values of a data item. (the Class of a STUDENT as 1 for Newbie, 2 for Senior, …)

Construct the database

- ***Store data to represent each student, course, section, grade report, and prerequisite as a record in the appropriate file.***
- ***Define the relationships among the records.***

- ## *Manipulate the database*
  - ***Querying***
  - ***Updating***

QUESTIONS



1. List the **prerequisites** of the '**Database**' course.
2. List the names of students who took the section of the '**Database**' course offered in **fall 2008** and **their grades** in that section.
3. Display the transcript of student name '**Smith**' with his **grade** and **course name**.

STUDENT

Name	Student_number	Class	Major
Smith	17	1	CS
Brown	8	2	CS

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	07	King
92	CS1310	Fall	07	Anderson
102	CS3320	Spring	08	Knuth
112	MATH2410	Fall	08	Chang
119	CS1310	Fall	08	Anderson
135	CS3380	Fall	08	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Figure 1.2
A database that stores student and course information.

THE DATABASE APPROACH'S MAIN CHARACTERISTICS

Database Catalog

A database system's self-descriptive nature is as follows:

- The **description** of a database is stored in a DBMS **catalog** (e.g. data structures, types, and constraints).
- Meta data* is the term for the **description**.

RELATIONS

Relation_name	No_of_columns
STUDENT	4
COURSE	4
SECTION	5
GRADE_REPORT	3
PREREQUISITE	2

COLUMNS

Column_name	Data_type	Belongs_to_relation
Name	Character (30)	STUDENT
Student_number	Character (4)	STUDENT
Class	Integer (1)	STUDENT
Major	Major_type	STUDENT
Course_name	Character (10)	COURSE
Course_number	XXXXNNNN	COURSE
….	….	…..
….	….	…..
….	….	…..
Prerequisite_number	XXXXNNNN	PREREQUISITE

Note: Major_type is defined as an enumerared type with all known majors. XXXXNNNN is used to define a type with four alpha characters followed by four digits

Figure 1.3
An example of a database catalog for the database in Figure 1.2.

THE DATABASE APPROACH'S MAIN CHARACTERISTICS

Data Abstraction

- A data model **hides storage details** while providing users with a **conceptual view** of the database.
- Programs refer to the data model constructs rather than data storage details.

Multiple data views are supported

- **Each user may see a different view** of the database that only shows the data that is relevant to them.

Data sharing and transaction processing for multiple users

- **Allowing multiple users to access and update the database** at the same time.
- The **recovery subsystem** ensures that the effect of each completed transaction is permanently recorded in the database.
- Database applications rely heavily on OLTP (Online Transaction Processing). Hundreds of **concurrent transactions** can be executed per second as a result of this.

Tutorial 1

Actors in Database environment

1. Who is actors behind the scene?
2. What about actors on the scene?