VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wananga o te Upoko o te Ika a Maui

Introduction to Database Systems
Part I

Coordinator: Dr Hui Ma
Lecturer: Dr Pavle Mogin

SWEN 304
Database System Engineering

# *Plan for Intro to DB Systems I*

- ## What is:
  - A database
  - Data
  - A database management system (DBMS)

  - ## *Reading:*
    - *Chapter 1*

Definition of a Database
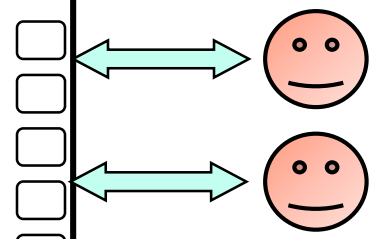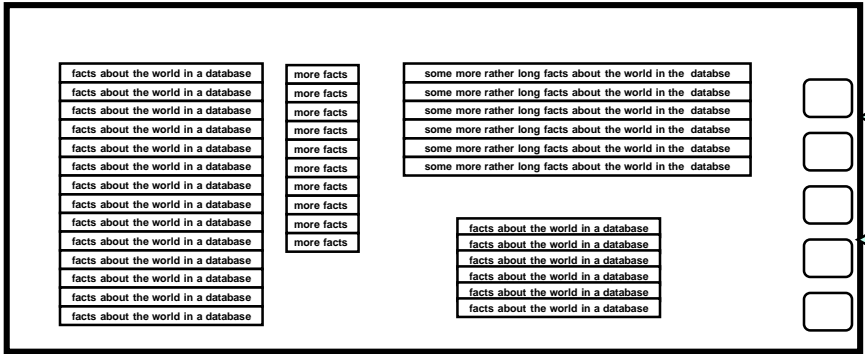
- A **database** is a collection of related data
- Essential database characteristics are:
  - Represents an aspect of the real world (miniworld, UoD),
  - We shall suppose it is well structured (even has a strict regular structure),
  - Reflects (or should reflect) current state of the UoD,
  - Has users and applications,
  - Stored in a permanent (persistent) computer memory, and
  - Managed by a Database Management System (DBMS)
- All these characteristics have to be met

What Is a Database?

Real World
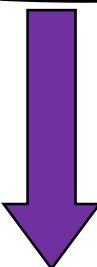
Facts about the world and its behaviour represented in the content and operations of a database

Database

facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database

more facts
more facts
more facts
more facts
more facts
more facts
more facts
more facts
more facts
more facts

some more rather long facts about the world in the databse
some more rather long facts about the world in the databse
some more rather long facts about the world in the databse
some more rather long facts about the world in the databse
some more rather long facts about the world in the databse
some more rather long facts about the world in the databse

facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database
facts about the world in a database

Users

A Sample Database

Student

LName	FName	StudId	Major
Smith	Susan	131313	Comp
Bond	James	007007	Math
Smith	Susan	555555	Comp

Course

CName	CourId	Hours	Dept
DB Sys	C302	2+1	Comp
SofEng	C301	2+0	Comp
DisMat	M214	4+1	Math

Grade

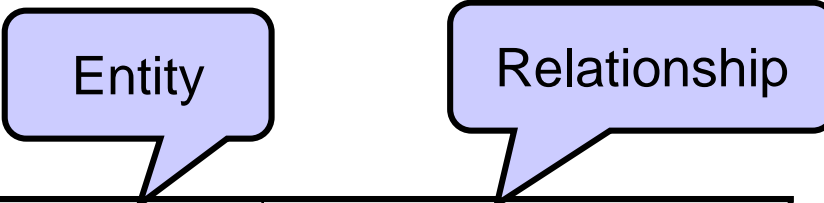| StudId	CourId	Grade
007007	C302	A+
007007	C301	A
007007	M214	A+
131313	C301	B-
555555	C301	C
131313	C302	D
555555	C302	E

Questions for You

1. Is a book (like "Fundamentals of Database Systems") a database?

2. Is an old style library card catalog a database?

3. Is a bank statement a database

Definition of Data (Datum)

- **Data** is a value of a property of an individual UoD entity or a relationship (between two UoD objects) at a particular period of time

- Example

UoD object(s)	James	James & CompSci
Property	Age	Number of Points
Time	July 2010	July 2010
Value	21	240

Entity → James

Relationship → James & CompSci

- If time is not recorded, it is assumed the value relates to the current time

Assumptions

- Fundamental Assumptions of Data Management:
  - Databases provide data for multiple application programs
  - Data in databases is accessed and manipulated concurrently
  - Data in databases is dynamic, that is, may change over time
  - Data in databases is persistent
  - The amount of data in databases can be huge

Amazon.com Database

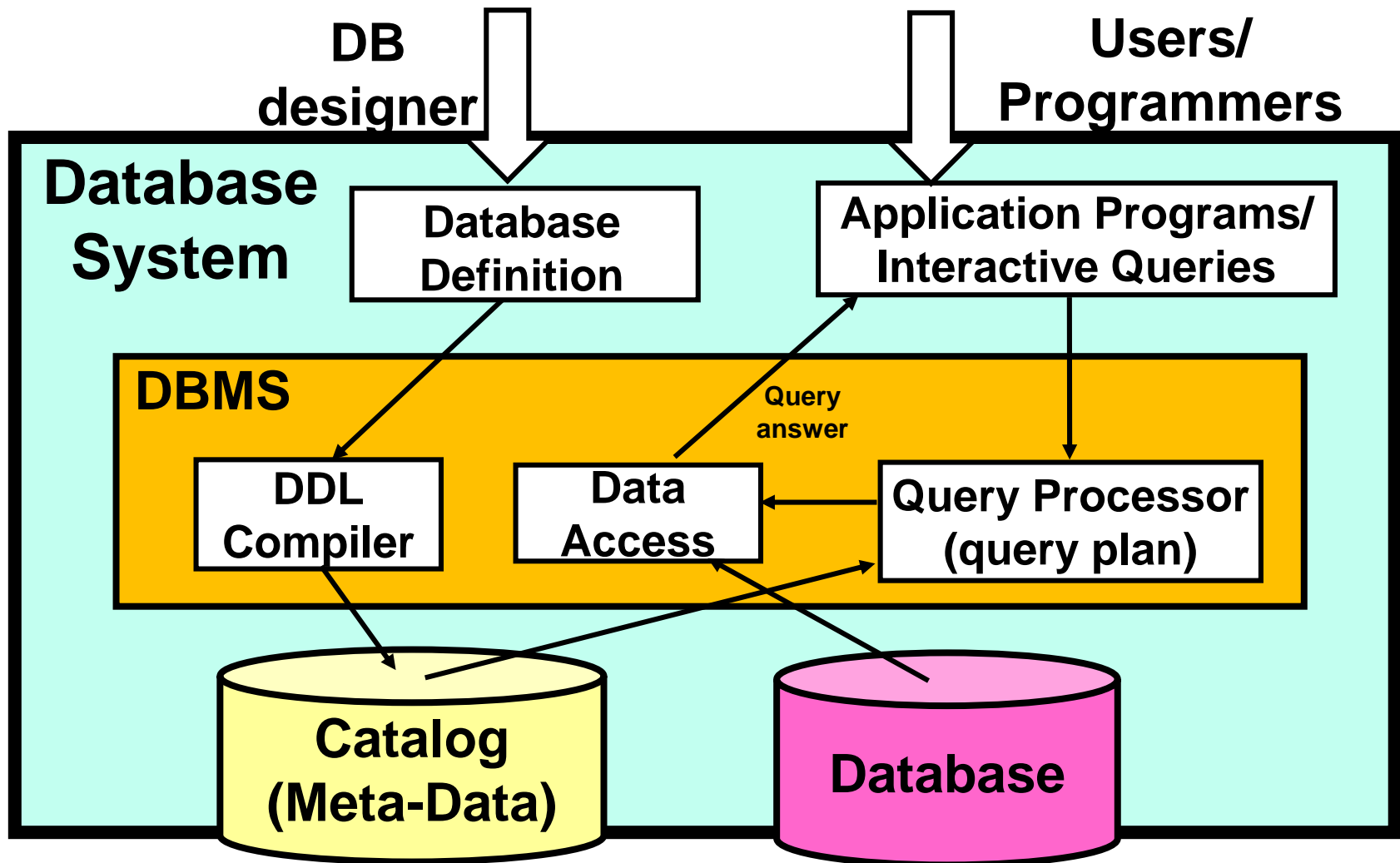- 20 million books, CDs, videos, DVDs, electronices, apparel and other items

- Occupies over 42 terabytes (1 terabytes = 1024GB)

- Stored on 200 different computers

- 15 million visitors access Amazon.com each day

- The database is continually updated as new books/items are added to the inventory and purchases are transacted

- 100 people are responsible for keeping the database up-to-date

Database Management System (DBMS)

- A Database Management System (DBMS) is a general purpose software system that facilitates:
  - Defining (describing the structure),
  - Populating by data (Constructing),
  - Manipulating (querying, updating),
  - Preserving consistency,
  - Protecting from misuse,
  - Recovering from failure, and
  - Concurrent using

  of a database

- The product of database defining is called meta-data
  - It is stored by DBMS as a small database called catalog or dictionary

A Simplified Database System Layout

DB designer

Users/ Programmers

Database System

Database Definition

Application Programs/ Interactive Queries

DBMS

Query answer

DDL Compiler

Data Access

Query Processor (query plan)

Catalog (Meta-Data)

Database

Typical DBMS Functionality

- Supports:

  - **Defining** of a particular database in terms of its data types, structures, and constraints (contains DDL)

  - **Loading** the initial database contents on a secondary storage medium

  - **Manipulating** the database (contains DML):

    - Retrieval: querying, generating reports

    - Modification: insertions, deletions and updates

  - **Processing** and **Sharing** by a set of concurrent users and application programs

    - Keeps all data valid and consistent

  - **Protection** or **Security** measures to prevent unauthorized access

Data Definition Example

Defining a table in SQL:

CREATE TABLE *Course* (

    CourId   char(4)   CONSTRAINT cspk PRIMARY KEY,

    CName   varchar   NOT NULL,

    Points   int   NOT NULL CHECK (Points >= 0),

    Dept   varchar

  );

Query and Insert Examples

SELECT   *LName* AS *SURNAME*, *CName*, *Grade*
FROM   *Student s*, *Grades g*, *Course c*
WHERE   *FName* = 'James' AND
s.StudId = *g.StudId* AND *c.CourId* = *g.CourId* ;

INSERT INTO *Student* (*FName*, *LName*, *StudId)*
VALUES   ('Ann', 'Bolen', 111111),
             ('Sharon', 'King' 121212);

Concurrency Control Example

prg #1	*A* in prg #1	*A* in db	prg #2	*A* in prg #2
Read *A*	2	2		
	2	2	Read *A*	2
A = A + 1	3	2		2
	3	2	*A = A + 1*	3
Write *A*	3	3		3
		3	Write *A*	3

Time

The net result of executing these two programs is: 2 + 1 + 1 = 3

A Question for You

- What caused the error in the database processing example:

  a) Pavle made a mistake
  b) DBMS does not know to do arithmetic
  c) Wrong program design

Advantages of Using Databases (1)

- Helps in:

  - Controlling redundancy in data storage

  - Investing less efforts in development and maintenance

  - Sharing of data among multiple users

  - Restricting unauthorized access to data

- Also, allows using sophisticated design techniques:

  - Data normalization

  - Denomalization: sometimes it is necessary to use **controlled** redundancy to improve the performance of queries

Advantages of Using Databases   (2)

- Helps in:
  - Providing backup and recovery services
  - Providing multiple interfaces to different classes of users
  - Representing complex relationships among data
  - Enforcing integrity constraints on the database
    - Referential integrity constraint
    - Key or uniqueness constraint
    - Attribute constraints
  - Drawing inferences and actions using rules (active databases)
    - E.g. triggers and stored procedures

Summary

- A database is a collection of related data that is well structured and stored permanently

- A data (datum) is a value of a real object's (or a relationship's) property in a perceived moment of time

- A DBMS is a set of programs that supports:
  - Defining,
  - Populating by data,
  - Querying,
  - Preserving consistency,
  - Protecting from misuse,
  - Recovering from failure, and
  - Concurrent using

  of a database

Plan for the next lecture

- Data models

- Schemas and instances

- The three schema architecture

- Data independence

- Database users and languages

  - *Reading:*
    - *chapter 2 of the textbook*