Database Management Systems (DBMS)

CSC1213/COM122β

Anusha Indika Walisadeera
Department of Computer Science
University of Ruhuna

Evaluation Criteria:

- Written examination (70%) + Practical examination (30%)
- For BCS Students (CSC1213) ➔ 2 hrs paper – answer all (4) the questions
- For BSC Students (COM122β) ➔ 2 hrs paper - answer 4 questions including Q1 (FO) (5questions available)
- Requirement ➔ 80% attendance

DBMS - Objectives

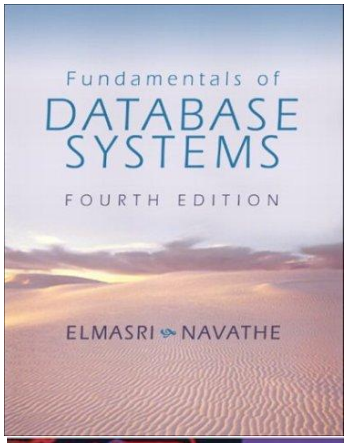
- To be able to understand
  - Identify database requirements
  - Analyze database requirements
  - Design databases
  - Create and use databases
  - Apply security

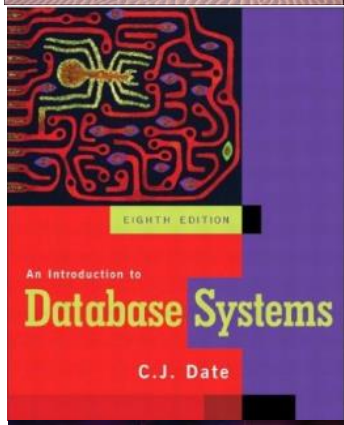
Brief Syllabus - DBMS

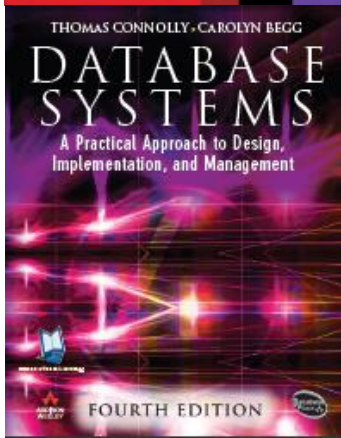
- Introduction to DBMS
- Relational Model
- Database Design
- Data Normalization process and the normal forms
- Structured Query Language (SQL)
- DB Security

References

- *Fundamentals of Database Systems* by Ramez Elmasri, Shamkant B. Navathe
- *An Introduction to Database Systems* by C.J. Date
- *Database Systems: A Practical Approach to Design, Implementation, and Management*, Third Edition by Thomas Connolly, Carolyn Begg







Why use a Database?

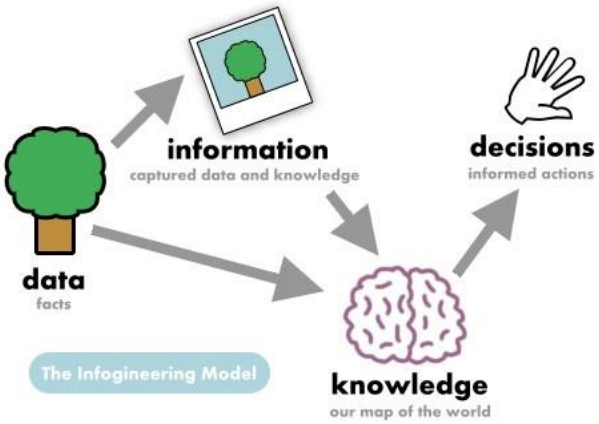
- Many people collect things
  - How about you?
- If you collect anything, you probably are familiar with some of the problems of managing a collection
  - Ex: stamps, photos, paper cuttings
- One way to keep track of a collection is to create a <u>database</u>

Some application of database systems…

- Purchases from the supermarket
- Purchases using your credit card
- Booking a holiday at the travel agents (Air Line Reservation)
- Using the local library
- Studying at university
- Hospital admissions
- Using the Internet

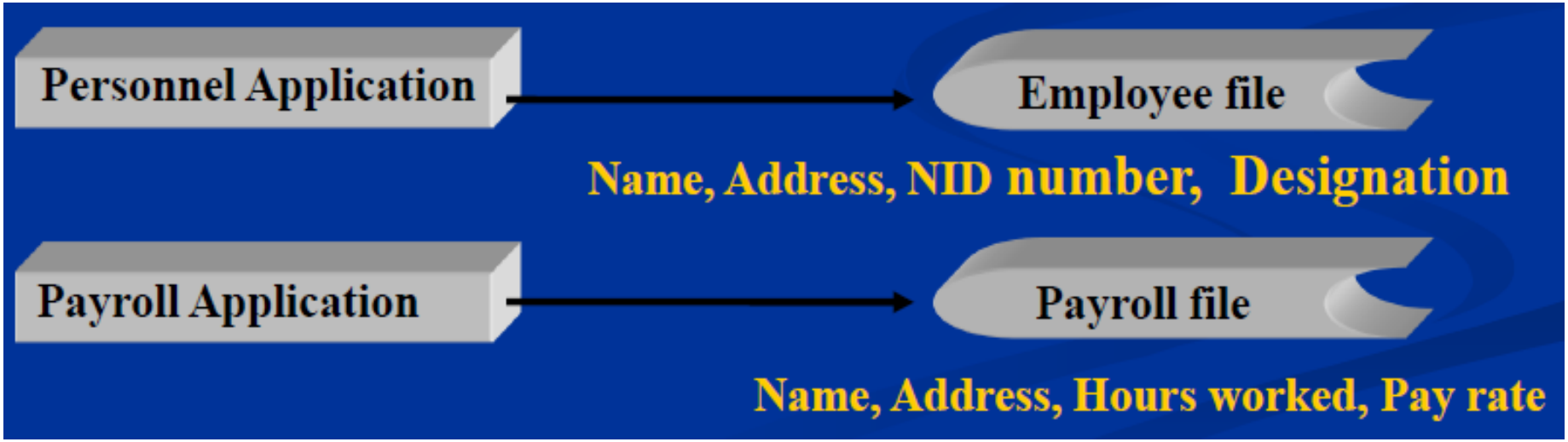
Data, Information and Knowledge

- **Data:**
  - Raw values relating to facts, events or transactions
  - Raw facts and figures that on their own have no meaning, and used to generate information
  - *Ex: 81000*
- **Information:**
  - Data that has been processed within a context to give it meaning
    - Data which has meaning and context
  - *Ex: 81000 Postal code of Matara*
- **Knowledge:**
  - Use of information in context, to make decisions or come to a conclusion
  - Knowledge is the understanding of rules needed to interpret information
  - Facts, feelings or experience known by a person or group
  - *Ex: Manager could use the information to decide whether or not to raise or lower price*



File-Based Systems

- A collection of application programs that perform services for the end users (e.g. report generation). Each program defines and manages its own data.



Limitations of File-Based Approach

- Separation and isolation of data
  - Each program maintains its own set of data.
  - Users of one program may be unaware of potentially useful data held by other programs.
- Duplication of data
  - Same data is held by different programs.
  - Wasted space and potentially different values and/or different formats for the same item.

Limitations of File-Based Approach Cont…

- Incompatible file formats
  - Programs are written in different languages, and so cannot easily access each other's files.
- Fixed queries of application programs
  - Programs are written to satisfy particular functions.
  - Any new requirement needs a new program.

Limitations of File-Based Approach Cont…

- Data Dependence
  - File structure is defined in the program code.

```
program Customer-Entry
....
type customer = record
        customer-id: string;
        customer-name: string;
        customer-street: string;
        customer-city: string;
    end;
....
```

```
program Customer-Orders
....
type customer = record
        customer-id: string;
        customer-name: string;
        customer-street: string;
        customer-city: string;
    end;
....
```

```
10 | Perera | Galle Road | Colombo 04
12 | Silva   | Reid Drive  |  Colombo 07
```

Database Approach

- Why the database and Database Management System (DBMS)?
  - Overcome the limitations in file-based approach

What is a Database?

- A database is an organized collection of data that are related in a meaningful way and stored in a common pool which can be accessed by one or many persons for many purposes.
- A database is a group of related files.
- Ex: UNIVERSITY database for maintaining information concerning students, courses, and grades in a university environment.

University Database

STUDENT	Name	StudentNumber	Class	Major
	Smith	17	1	CS
	Brown	8	2	CS

COURSE	CourseName	CourseNumber	CreditHours	Department
	Intro to Computer Science	CS1310	4	CS
	Data Structures	CS3320	4	CS
	Discrete Mathematics	MATH2410	3	MATH
	Database	CS3380	3	CS

SECTION	SectionIdentifier	CourseNumber	Semester	Year	Instructor
	85	MATH2410	Fall	98	King
	92	CS1310	Fall	98	Anderson
	102	CS3320	Spring	99	Knuth
	112	MATH2410	Fall	99	Chang
	119	CS1310	Fall	99	Anderson
	135	CS3380	Fall	99	Stone

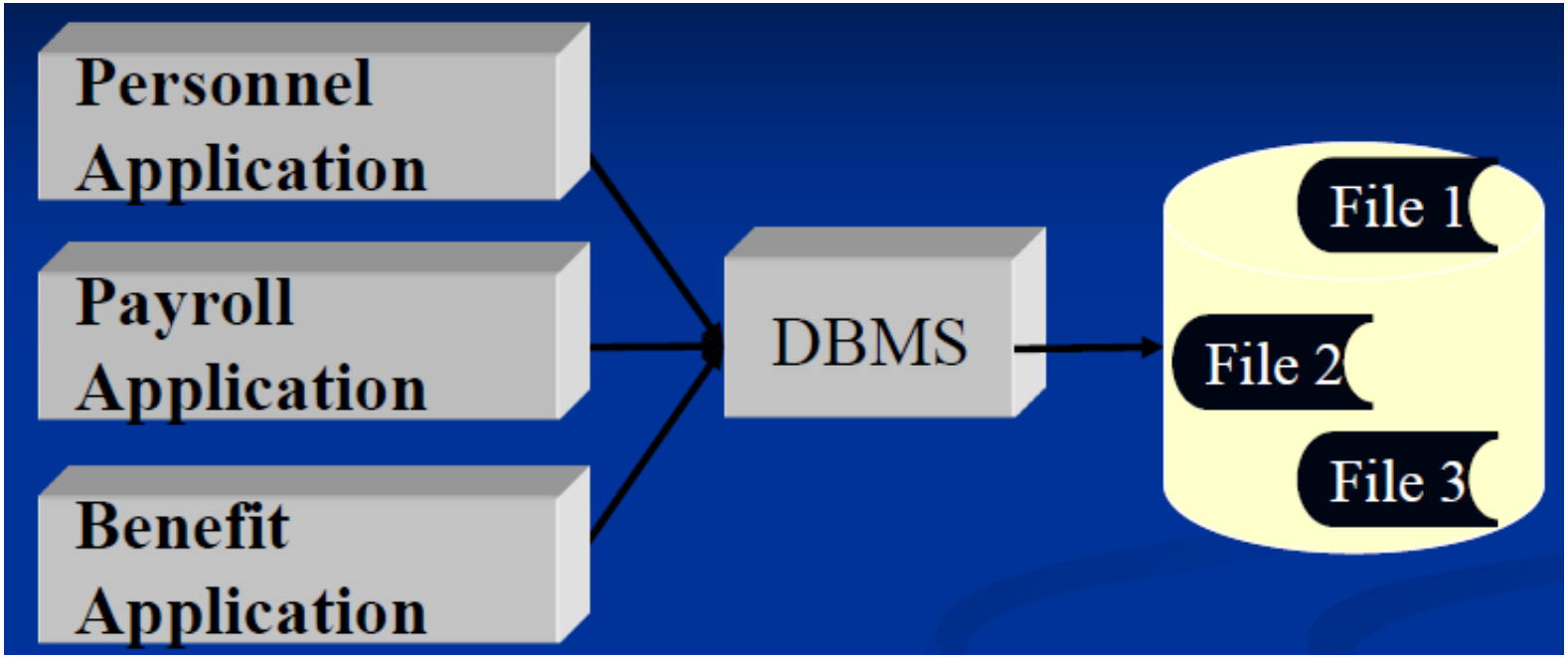
GRADE_REPORT	StudentNumber	SectionIdentifier	Grade
	17	112	B
	17	119	C
	8	85	A
	8	92	A
	8	102	B
	8	135	A

PREREQUISITE	CourseNumber	PrerequisiteNumber
	CS3380	CS3320
	CS3380	MATH2410
	CS3320	CS1310

Database Management System (DBMS)

- DBMS is a software system that enables users to define, create, and maintain the database and that provides *controlled access* to this database.
- DBMS is a general-purpose software system that facilitates the processes of defining, constructing, manipulating, and sharing databases among various users and applications.
- E.g. Oracle, Access, SQL Server, DB2, MySQL

Database Approach



Ex: Integrated human resources database

Employees: *Name, Address, NID number, Designation*

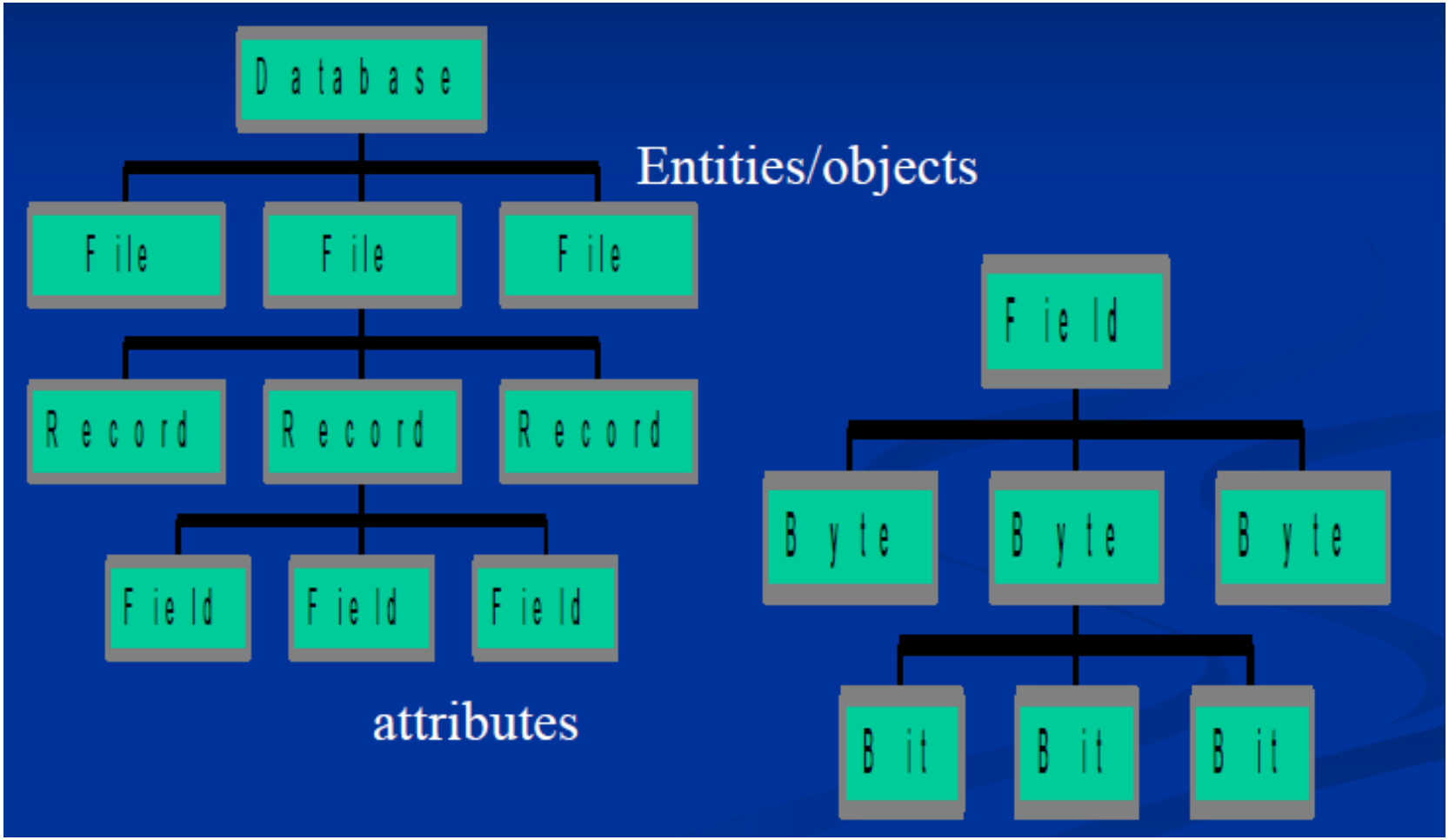
Payroll: *Hours worked, Pay rate*

Benefit: *Insurance, Pension plan*

Database Building Blocks

- File/Table
  - A collection of related records
- Record
  - A record is a group of related fields
- Field
  - Field represents an attribute, or a characteristic, or a piece of information

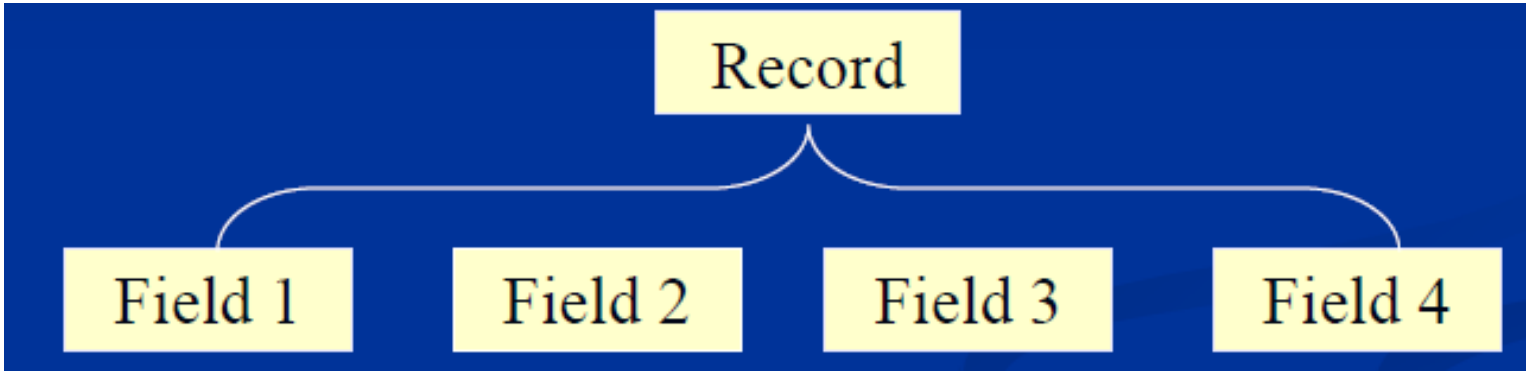
Database Building Blocks Cont…



Database Building Blocks Cont…

(Empno, name, designation, salary, department)

2 Perera Manager 35,000 Personnel



Major Components of a Database System

- Data
  - The data in the database; integrated and shared
- Hardware
  - Physically stores data, it can be secondary storage on which the database physically resides together with associated I/O devices
- Software
  - Between the physical database and the users is a layer of software which is known as the DBMS
- Users
  - Database Administrators
  - Database Designers
  - End Users
  - System Analysts and Application Programmers (Software Engineers)

Database Users

- Database Administrators (DBA)
  - Administering resources (Database, DBMS and related software)
  - Responsible for authorizing access to the database
  - Coordinating and monitoring database's (its) use
  - Acquiring software and hardware resources as needed
  - Accountable for problems such as breach of security or poor system response time

- Database Designers
  - Identify the data to be stored in the database
  - Choose appropriate structures to represent and store this data

Database Users Cont…

- End Users
  - database primarily exists for their use
  - people whose jobs require access to the database for querying, updating, and generating reports;
  - End-users can be classified according to the way they use the system:
    - Casual end users
    - Naive end users
    - Sophisticated end users
    - Stand-alone users

Database Users Cont…

- Casual end users
  - occasionally access the database
  - may need different information each time
  - use a query language to specify their requests

- Naive end users
  - constantly querying and updating the database
  - use standard types of queries and updates
  - need to know very little about the DBMS

Database Users Cont…

- Sophisticated end users
  - thoroughly familiarize themselves with the facilities of the DBMS
  - implement their applications to meet their complex requirements
- Stand-alone users
  - maintain personal databases
  - use ready-made program packages that provide easy-to-use menu-based or graphics-based interfaces

Data Dictionary/System Catalog (or Mata data)

- A subsystem that keeps track of the definitions of data items in the database which includes
  - elementary-level data items (fields/attributes)
  - relationships that exists between various data structures
  - files or relational tables
  - indexes that are used to access data quickly

Meta Data

- Data that describe the properties or characteristics of other data.
- Some of these properties include the name of the data item, data type, length, minimum and maximum allowable values (where appropriate),rules or constraints and a brief description of each data item.
- Meta data allow database designers and users to understand what data exist, what the data mean.
- Data without clear meaning can be confusing, misinterpreted or erroneous.

Example for Meta Data

- E.g. Employee

Name	Type	Length	Min	Max	Description
EmpNo	Number	9			Employee No.
Name	Character	30			Employee Name
Dept	Character	10			Dept. No.
Salary	Number	8	5000	60000	Employee Salary

Employee No. (ID) unique

Table Definition

Employees : Table

Field Name	Data Type	
EmployeeID	AutoNumber	Number automatically assigned to new employee.
LastName	Text	
FirstName	Text	
Title	Text	Employee's title.
TitleOfCourtesy	Text	Title used in salutations.
BirthDate	Date/Time	
HireDate	Date/Time	
Address	Text	Street or post-office box.
City	Text	
Region	Text	State or province.
PostalCode	Text	
Country	Text	
HomePhone	Text	Phone number includes country code or area code.
Extension	Text	Internal telephone extension number.
Photo	OLE Object	Picture of employee.
Notes	Memo	General information about employee's background.
ReportsTo	Number	Employee's supervisor.

General | Lookup

Field Size	20
Format	
Input Mask	
Caption	Last Name
Default Value	
Validation Rule	
Validation Text	
Required	Yes
Allow Zero Length	No
Indexed	Yes (Duplicates OK)
Unicode Compression	Yes
IME Mode	No Control
IME Sentence Mode	None

Advantages of the database approach

Improved maintenance through Program-Data/Data Independence

- The separation of data descriptions (meta data) from the application programs that use the data.
  - This simplifies database application maintenance.
- In the database approach, data descriptions are stored in a central location called the data dictionary. This property allows an organization's data to change and evolve (within limits) without changing the application program that process the data.

Minimal Data Redundancy

- Data files are integrated into a single, logical structure. Each primary fact is recorded in only one place in the database.

  E.g. Employee data are not with the payroll and benefit files.

  Note: Data redundancy is not eliminated entirely. Some data items will appear in more than one place (e.g. employee no.) to represent the relationship with others.

Improved Data Consistency

- By eliminating (or controlling) data redundancy, we greatly reduce the opportunities for inconsistency.

  E.g. employee address is stored only once and hence we cannot have disagreement on the stored values.

  Also, updating data values is greatly simplified and have avoided the wasted storage space.

Improved Data Sharing

- A database is designed as a shared corporate resource and can be shared by all authorized users. In this way more users share more of the data.

  E.g. Insurance scheme for overworked employees.

  New applications can be built on the existing data in the database.

Enforcement of Standards

- Integration allows the DBA to define and enforce the necessary standards.
- Standards include naming conventions, data quality standards and uniform procedures for accessing, updating and protecting data.

Improved Data Integrity

- Integrity can be expressed in terms of constraints, which are consistency rules that the database is not permitted to violate.

- Eg: A member of staff's salary cannot be greater than 60,000.

Improved Data Accessibility and Responsiveness

- With relational database, end users without programming experience can often retrieve and display data, even when it crosses traditional departmental boundaries.

- English-like query language SQL and query tools such as Query-By-Example provide such facilities.

Improved Security

- DBMS can be used to enforce database security. This may take the form of user names and passwords to identify people authorized to use the database.
- The access that the authorized user is allowed on the data can also be restricted by the operation type (retrieval, delete, update, insert).

Increased concurrency

- Many DBMSs allow users to undertake simultaneous operations on the database. The DBMS implements a concurrency control mechanism that prevents database accesses from interfering with one another.

Disadvantages of DBMS

- Complexity
- Size
- Cost of DBMS
- Additional hardware costs
- Cost of conversion
- Performance
- Higher impact of a failure

SQL (Structured Query Language)

- A query language for relational systems
- A powerful data manipulation language
- **Data Definition Language** (DDL) enables to define the schema .
  - Permits specification of data types, structures and any data constraints.
  - All specifications are stored in the database.
- **Data Manipulation Language** (DML) to manipulate the data.
  - General enquiry facility (query language) of the data.

Lesson 2

Introduction to SQL

Questions

1. What is a Database?
2. What is a Database System?
3. Define Database Management System.
4. Discuss the advantages and disadvantages of a Database Management System over the traditional file processing approach.