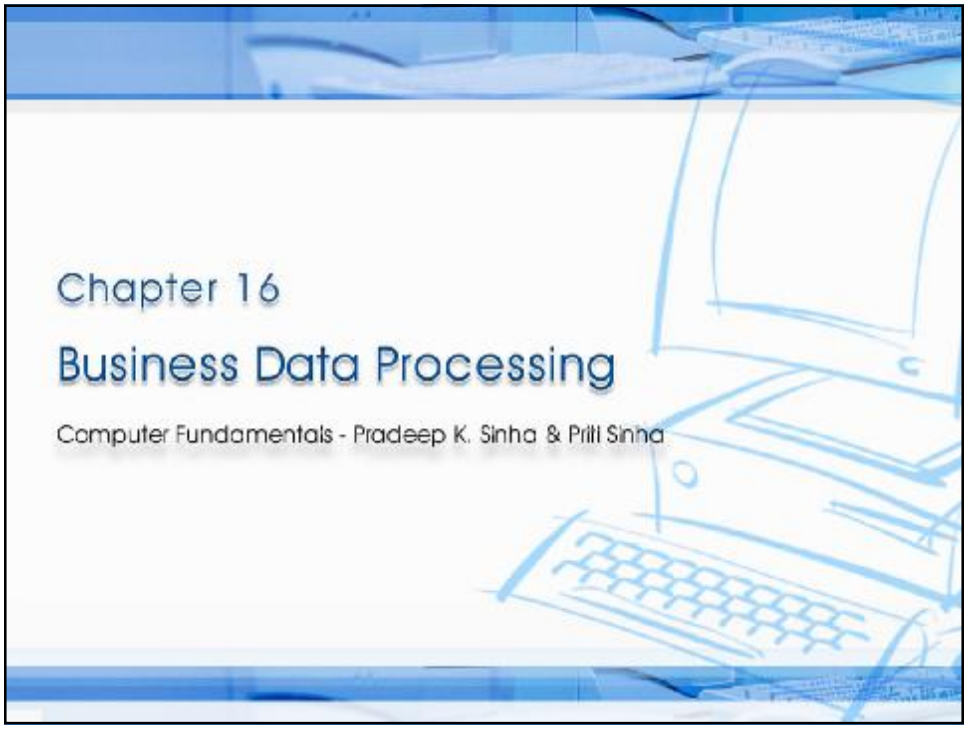


Learning Objectives

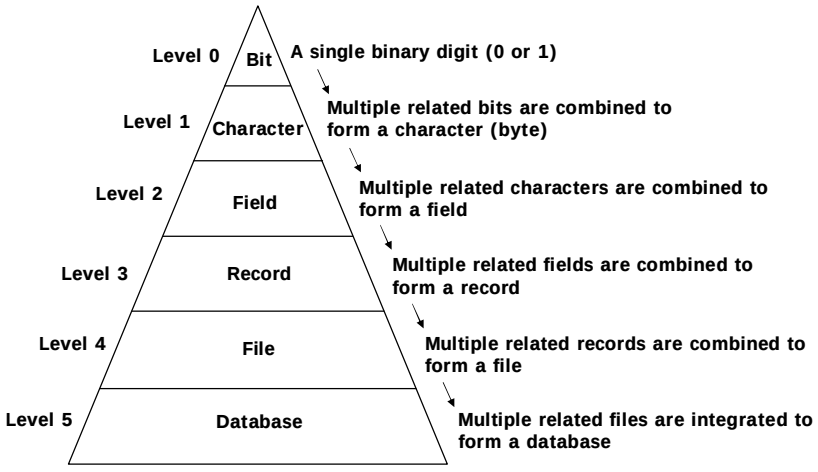
In this chapter you will learn about:

- Difference between data and information
- Data processing converts raw data into useful information
- Data storage hierarchy commonly used to facilitate data processing
- Standard methods of organizing data
- Basic concepts of database systems

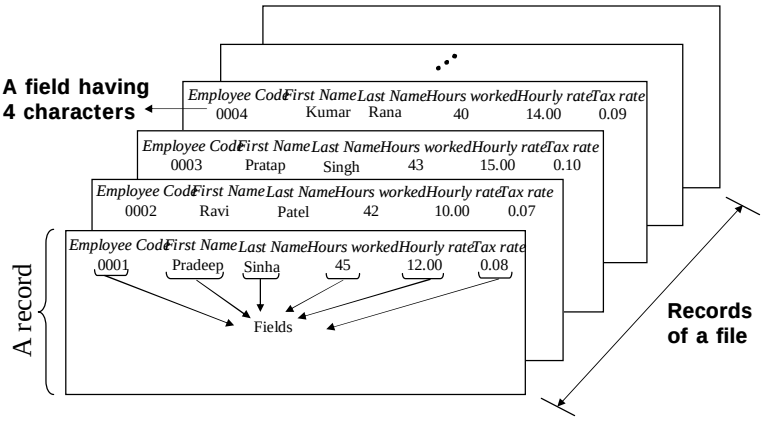
Data Processing

- Data is a collection of facts – unorganized but able to be organized into useful information
- Information is data arranged in an order and form that is useful to the people who receive it
- Data processing is a series of actions or operations that converts data into useful information
- A data processing system includes resources such as people, procedures, and devices used to process input data for producing desirable output

Data Storage Hierarchy



Relationship Among Character, Field, Record, and File



Standard Methods of Organizing Data

- **File-oriented approach: Application's data is organized into one or more files and application program processes them to generate the desired output**
- **Database-oriented approach: Data from multiple related files are integrated together to form a database:**
  - **Provides greater query flexibility**
  - **Reduces data redundancy**
  - **Solves data integrity (inconsistency) problem**
  - **Makes data independent of the application programs**
  - **Includes data security features at database level, record level, and field level**

(Continued on next slide)

File Management System

- In *file-oriented approach* of organizing data, an application's data is organized into one or more files
- Application program processes the data stored in these files to generate the desired output
- Set of programs is provided to facilitate the users in organizing, creating, deleting, updating, and manipulating their files
- All these programs together form a File Management System (FMS)

File Types

A file management system supports following file types:

- **Transaction file:** Stores input data until it can be processed
- **Master file:** Contains all current data relevant to an application
- **Output file:** Stores output produced by one program that is used as input to another program
- **Report file:** Holds a copy of a report generated by an application
- **Backup file:** Copy of a file, created as a safety precaution against loss of data

File Organizations

- File organization is the physical organization of the records of a file for convenience of storage and retrieval of data records
- Three commonly used file organizations are:
  - **Sequential:** Records are stored one after another in ascending or descending order determined by the value of the key field of the records
  - **Direct/random:** Desired record pertaining to current transaction can be directly located by its key field value without having to navigate through sequence of other records

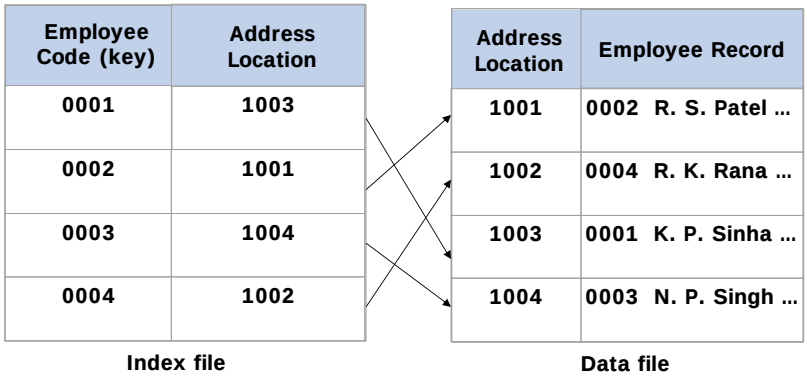
(Continued on next slide)

File Organizations

(Continued from previous slide)

- **Indexed sequential:** There are two files for every data file – the data file which contains the records stored in the file, and the smaller index file which contains the key and disk address of each record stored in the data file

Organization of An Indexed Sequential File



File Utilities

- Routines to perform a variety of generalized operations on data files
- Operations performed by some commonly used file utilities are Sorting, Searching, Merging, Copying, Printing, and Maintenance

Sorting On One Key

Employee Code	Department Code	Other fields (Name, Address, Qualification, Basic Salary, etc.)
101	2	---
123	3	---
124	1	---
176	2	---
178	1	---
202	3	---
213	1	---

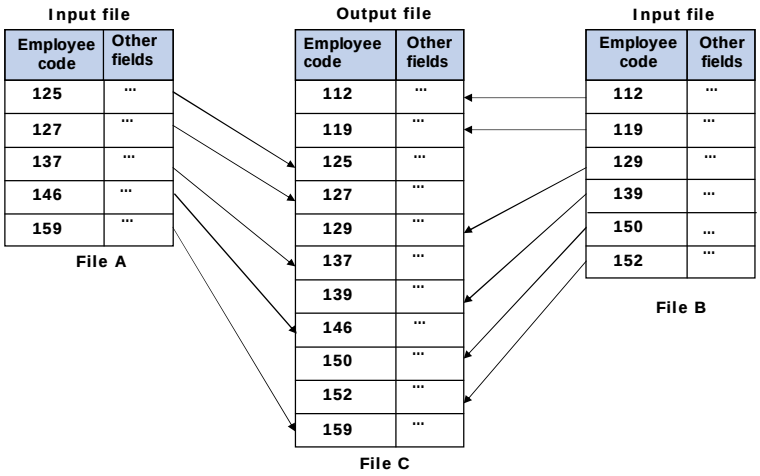
Sorting on ascending employee code sequence

Sorting On Two Key

Employee Code	Department Code	Other fields (Name, Address, Qualification, Basic Salary, etc.)
124	1	---
178	1	---
213	1	---
101	2	---
176	2	---
123	3	---
202	3	---

Sorting on a ascending employee code (secondary key) within ascending department code (primary key)

Merging of Two Files



Merging of files A and B to produce file C

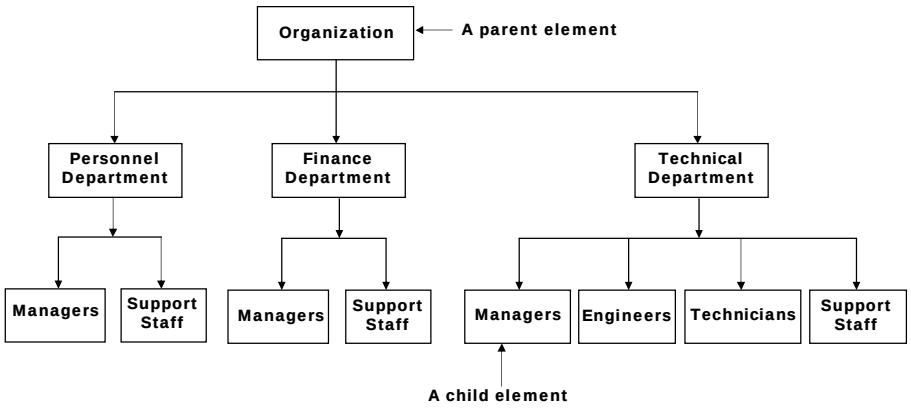
Database Management System

- In *database-oriented approach* of organizing data, a set of programs is provided to facilitate users in organizing, creating, deleting, updating, and manipulating data in a database
- All these programs together form a *Database Management System (DBMS)*

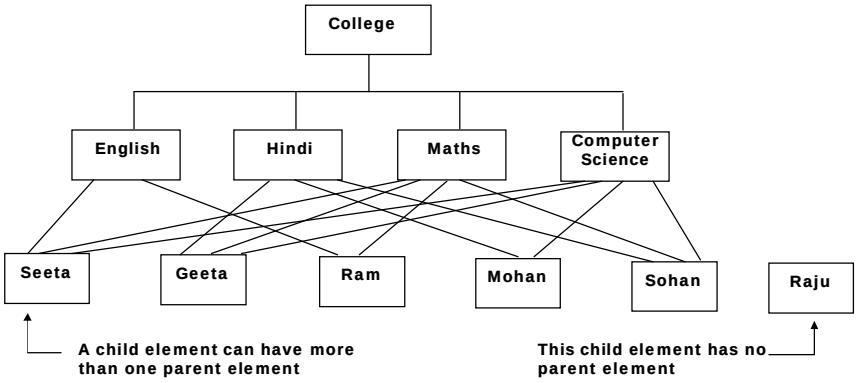
Database Models

- *Database model* defines the manner in which the various files of a database are linked together.
- Four commonly used database models are:
  - Hierarchical
  - Network
  - Relational
  - Object-oriented

Hierarchical Database



Network Database



Relational Database

Membership No.	Member's name	Member's Address
83569	K. N. Raina	C-15, Sarita Vihar, Pune-7
62853	D. P. Singh	A-22, Anand Park, Pune-5
12859	R. Pandey	D-18, Vrindavan, Pune-7
32228	R. S. Gupta	A-12, Nandanvan, Pune-2
23466	S. K. Ray	B-05, Royal Villa, Pune-3
11348	P. K. Sen	B-16, Anand Park, Pune-5
16185	T. N. Murli	A-11, Vrindavan, Pune-7

(a) Members data table.

Borrower (Membership No.)	Book No. (ISBN)	Due Date (DD-MM-YYYY)
12859	27-21675-2	10-12-2007
11348	89303-530-0	08-11-2007
32228	13-201702-5	10-11-2007
16185	22-68111-7	05-12-2007
12859	71606-214-0	06-11-2007
62853	13-48049-8	15-11-2007
11348	18-23614-1	12-11-2007

(b) Borrowed books data table

Book No. (ISBN)	Book Title	Author
13-201702-5	Concepts of Physics	H. C. Verma
13-48049-8	Concepts of Chemistry	S. S. Dubey
18-23614-1	Astrology for You	N. K. Sharma
22-68111-7	Fundamentals of Computers	K. Ramesh
27-21675-2	C++ Programming	R. P. Rajan
71606-214-0	Computer Networks	A. N. Rai
89303-530-0	Database Systems	P. N. Dixit

(c) Books data table

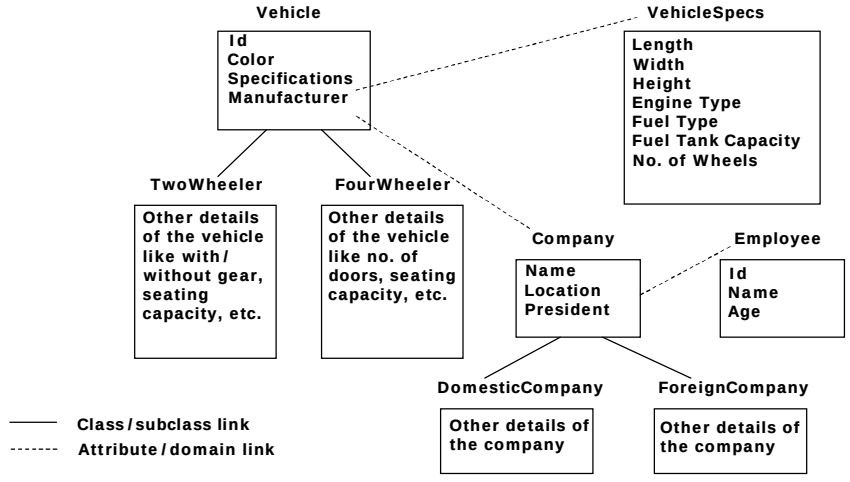
Sample Report

List of overdue books as on 10-11-2007

Membership No.	Member's Name	Member's Address	Due Date	Book No.	Book Title	Book Author
11348	P. K. Sen	B-16, Anand Park, Pune-5	08-11	89303-530-0	Database Systems	P. N. Dixit
32228	R. S. Gupta	A-12, Nandanvan, Pune-2	10-11	13-201702-5	Concepts of Physics	H. C. Verma
12859	R. Pandey	D-18, Vrindavan, Pune-7	06-11	71606-214-0	Computer Networks	A. N. Rai

A report of overdue books as of 10-11-2007 from the sample database of previous slide

Object-Oriented Database



Main Components of a DBMS

- DBMS allows users to organize, process and retrieve selected data from a database without knowing about the underlying database structure
- Four major components of a DBMS that enable this are:
  - *Data Definition Language (DDL):* Used to define the structure (schema) of a database
  - *Data Manipulation Language (DML):* Provides commands to enable the users to enter and manipulate the data

(Continued on next slide)

Main Components of a DBMS

(Continued from previous slide)

- *Query Language*: Enables users to define their requirements for extracting the desired information from the database in the form of queries
- *Report generator*: Enables the users of a database to design the layout of a report so that it can be presented in the desired format

Creating a Database

Creation of a database is a three step process:

- Defining its structure (schema)
- Designing forms (custom screens) for displaying and entering data
- Entering the data into it

Sample Database Form

EMPLOYEE DATABASE DATA ENTRY FORM

EMPLOYEE ID:	856392	SEX:	M	AGE:	42
EMPLOYEE NAME:		LAST NAME:	SINHA		
		FIRST NAME:	PRADEEP		
		MIDDLE NAME:	KUMAR		
CONTACT ADDRESS:		ADDRESS 1:	F/8, ANAND PARK		
		ADDRESS 2:	SOCIETY, AUNDH		
		CITY:	PUNE		
		STATE:	MH		
		POSTAL CODE:	411007		
TELEPHONE NO.:	(020) 5680-489				
ANY OTHER INFORMATION:	IS FLUENT IN JAPANESE LANGUAGE				

Viewing, Modifying, Deleting, and Adding Records

- All database systems provide commands to view, modify, delete, or add records of an already established database
- Many database systems also provide a facility to set up a filter allowing user to browse through and view only those records that meet some criterion

Searching a Database

Commonly supported features for enabling a user to search for desired information in a database are:

- *Find command*: Used for simple database queries
- *Query language*: Used for more complex database queries
- *Query By Example (QBE)*: Provides a simple user interface for specifying search criteria

Creating Reports

- Reports are generated by using report generator of a database system to assemble the output of a database query in desired format
- Report generator enables user to specify layout of the report, titles & subtitles for the report, column headings for various fields, and other elements to make the report appear more presentable

Sample Output of Report

LIST OF EMPLOYEES WHO BELONG TO PUNE

DATE: DECEMBER 15, 2007

LAST NAME	FIRST NAME	ADDRESS-1	ADDRESS-2	TELEPHONE NUMBER
Gupta	Rajiv	A-12, Nandanvan	M. G. Road	4623-4892
Murli	Tapan	A-11, Vrindavan	Pashan Road	5863-4905
Pandey	Rupa	D-18, Vrindana	Pashan Road	5865-3236
Raina	Pushpa	C-15, Sarita Vihar	Aundh Road	5755-8328
Ray	Suhas	B-05, Royal Villa	M. G. Road	4685-6356
Sen	Prakash	B-16, Anand Park	Aundh Road	5762-3333
Singh	Deepak	A-22, Anand Park	Aundh Road	5728-6287

The report is sorted to present the list in alphabetical order of their last name

Key Words/Phrases

- Activity ratio
- Backup file
- Collision
- Copying
- Data
- Data Definition Language (DDL)
- Data dependence
- Data dictionary
- Data file
- Data integrity
- Data Manipulation Language (DML)
- Data processing
- Data redundancy
- Data storage hierarchy
- Database
- Database administrator
- Database Management System (DBMS)
- Database model
- Direct file
- Field
- File
- File Management System (FMS)
- File utilities
- Filter
- Hashing
- Hashing algorithm
- Hierarchical database
- Index file
- Indexed sequential file
- Information
- Master file
- Merging
- Network database
- Output file
- Peripheral Interchange Program
- Primary key

(Continued on next slide)

Key Words/Phrases

(Continued from previous slide)

- Query By Example
- Query language
- Record
- Relational database
- Report file
- Report Generator
- Schema
- Searching
- Secondary key
- Secondary key
- Sequential file
- Sorting
- Transaction file
- Tuple