System user specifies the business process in terms of process requirements for a new system. Process requirements mean users expectations of the processing requirements for business process and its information systems. Business process requirements are frequently defined in terms of policies and procedures. Policies are explicit rules that govern business process. In other words, we can adopt a business process that match with business policies. Procedures are the precise steps to be followed in completing business processes.

Process requirements are also frequently specified in terms of work flow that means flow of transactions through business process to ensure that the appropriate checks and approval are implemented. In the process of system development, this is a challenge to identify, express and analyze business process requirements exclusively in business term which is well understood by system users.

System designers view: Business processes in terms of application development environment and the software technology is used to develop the system. Many business purchase readymade commercial software solutions instead of building or developing their own in-house softwares. In case of building their own software, business processes are usually designed first, so the business process specifications are supplemented with software specifications. The designer must develop software specifications which fulfill the business process requirement of system users as well as provide sufficient details and consistency for communicating the software design to system builders.

System builders focus on custom built application programme that automate business processes.

At this stage, again the question arises whether to build or buy a commercial software package as a system solution. Sometime, they buy a customize software and system builder makes an application programme to enhance the package's functional capabilities. System builders also focus on program utilities that helps with the conversion and integration of the commercial program and existing systems.

2.6.3 Communication

A common goal of most business organizations is to improve their communication and cooperation among employers and various components of the information system. The need of communication improvement can be at two ends:

1. Information system must provide an effective and efficient communication interface to the system's users. These interfaces lead to teamwork.
2. Information systems must act as an effective & efficient interface with other information systems either in the business itself or outside the business organization.

System owners define the communication scope of an information system development project. Actually, this scope gives a simple list of business locations or systems with which the information system must interface. The relevant problems, opportunities and constraints may be identified and analyzed to obtain system owner's view: a proper solution.

System user view: Communication in terms of information system is done through inputs and outputs. The inputs and outputs represent how these systems would interact with users, employees, customers, divisions of business itself and with the other business houses. So, the details of both input and output are essential as well as important.

System designers are concerned with the technical design of both user and system to system communication interfaces. They are known as interface specification. Interface specification are technical designs that document how system users are interact with a system and how a system

interacts with other systems. Both user and designer involved in interface design. Earlier is interested in requirement and format while the latter is interested in consistency, completeness and compatibility with other systems.

System builder construct, install, test and implement both user and system-to-system interface solutions using interface technology. For user interface, the interface technology is frequently embedded into the application development environment used to construct software for the system. System-to-system interface are considerably more complex than user interface. One system-to-system interface technology is middle ware. Middleware is a layer of utility software that sits in between application software and system's software to transparently integrate different technologies that they can inter-operate.

Finally, today's information system are built on networks. Network technology allows properly designed information systems to separate the knowledge, process and communication and force them to communicate across the network.

2.7 WORKING OF INFORMATION SYSTEM

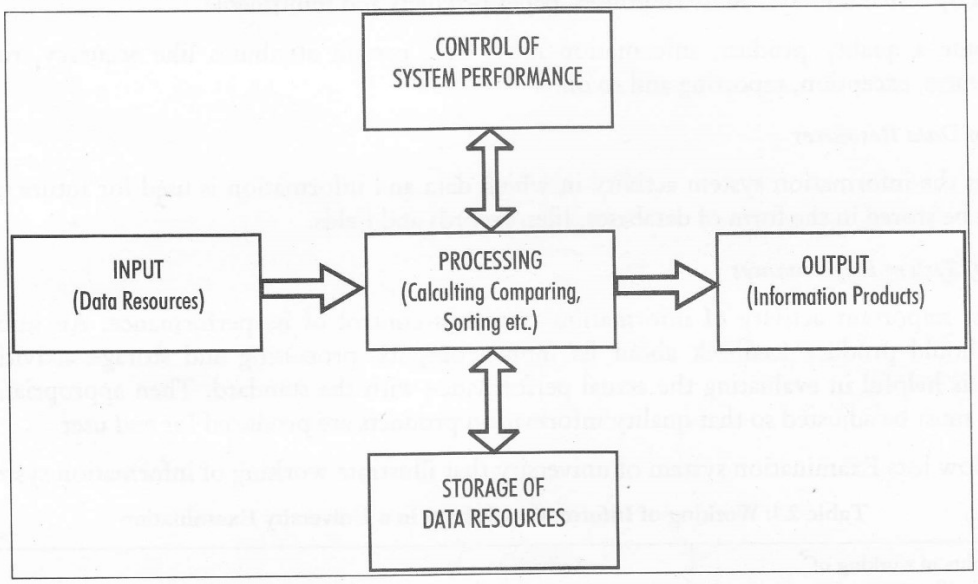

Figure 2.6: Working of Information System

In this section, we are going to study how a information system works. As we all know, the working of any system consists of Input, processing, output, storage of that input and output and controlling each and every activity of the system. Alike any system, the information system also accepts input and after data processing give output and these output and program are stored for future use and the whole working is controlled by some controlling measure. The working of information system is shown in Figure 2.6

Let us discuss each activity of information system in details.

Input

Data related to each transactions and events must be captured and prepared for processing by the input activity. This activity comprises recording and editing of data. Recording of data can be done manually

by using a pen and paper or directly enter into the computer by the data entry operator. These are number of editing activities which ensure that data recording is correct and complete. Computerized data may be stored on some storage media like magnetic tape etc. Method such as optical scanning and displays of menus, prompts and fill-in-the-blanks formats make it convenient for end user to enter data correctly into information system.

Processing

After the input of data, it is subject to processing activities such as calculating, comparing, sorting, classifying and summarizing. These activities convert data into useful information for system stakeholder like system owners and system users. The quality of data must be maintained by a continuous process of updating and correcting activities.

Output

Through output activity, the end user may able to get information in desired formats. The ultimate goal of information system is the production of accurate, timely and appropriate information for end users. Example of information product are messages, reports, forms and graphic images which may be provided by visual displays, audio responses, paper products and multimedia.

To become a quality product, information must have certain attributes like accuracy, timeliness, completeness, exception, reporting and so on.

Storage of Data Resources

Storage is the information system activity in which data and information is used for future purposes. Data can be stored in the form of databases, files, records and fields.

Control of System Performances

The most important activity of information system is control of its performance. An information system should produce feedback about its inputs, outputs, processing and storage activities. This feedback is helpful in evaluating the actual performance with the standard. Then appropriate system activities must be adjusted so that quality information products are produced for end user.

Table below lists Examination system of university that illustrate working of information system.

Table 2.1: Working of Information System in a University Examination

Components of working of Information System	Activities
Input	Optical scanning of bar-coded answer-sheets.
Processing	Adding marks in each subject, calculating total marks obtained, depicting division and soon.
Output	Producing detailed marksheet, degree.
Storage	Maintaining records of students personal details, marks in each subject, knowledge base for result preparation.
Control	Generating audible signals to indicate proper entry of marks, check on any variations, checks on omitting any rules.

Check Your Progress

1. Choose the appropriate answer:
   An information system always:
   i. Requires hardware even if only a pencil
   ii. Transforms information to the unit
   iii. Is computer-based
   iv. None of these
2. What is system?

2.8 LET US SUM UP

Information is a valuable resource akin to any material, money or human resource. Information systems boost profitability and improve efficiency in the business processes of an organisation. A system is the set of elements in form of ideas, things and people which are interrelated and are part of a cohesive set-up, that synergises to achieve a specific goal or goals. Management information system enables managers to tie planning and control procedures to operational systems of implementation.The building blocks of information systems evolve from needs of business processes. Business model comprises sum total of business activities. The aim of the Information systems is to provide support for management at all levels viz. strategic planning, management control and operational control. System analysis is customised approach to the use of the computer for problem solving on the basis of a study about operations within an organization as a system and the requirements of an organization. MIS has become a generic term of information systems for management. In 1970s MIS supplanted the data processing systems. The stage of data processing systems was set in 1960s with the advent of computers. The revolutionary upsurge in computing and communication technologies in 1980s led to up gradation of MIS and the evolution of decision support systems. Decision support systems ought to be necessarily and entirely computer based whereas the MIS need not necessarily be so.

2.9 KEYWORDS

System: A system is the set of elements in the form of ideas, things and people which are inter-related and part of a cohesive set-up, that synergize to achieve a specific goal or goals.

Knowledge: knowledge is derived from data information.

2.10 QUESTIONS FOR DISCUSSION

1. What is Information System? Discuss the need for information system.
2. How the business organization can be described as a system? What is the concept of subsystem?
3. What is systems approach? Discuss with examples.
4. Discuss the evolution of information system.

Check Your Progress: Model Answers

1. Requires hardware even if only a pencil.
2. The term 'system' is derived from Greek word system which means an organized relationship among functioning units or components

2.11 SUGGESTED READINGS

Ashok Arora & Akshaya Bhatia, *Management Information Systems*, Excel Books.

G. V. Satya Sekhar, *Management Information Systems*, Excel Books.

UNIT II

LESSON

3

INFORMATION AND DATABASES

CONTENTS

3.0 AIMS AND OBJECTIVES

- After studying this lesson, you will be able to:
- Understand the concept of information
- Understand DBMS and data modeling
- Discuss DBMS models

3.1 INTRODUCTION

Data are the facts and figures which can take the form of the historical records pertaining to the operations of the enterprise. These can be filed appropriately and act as a source of the fixed documents, accounting ledgers, stock registers and so on. These can be processed to derive the information of paramount importance to an organization such as the profit and loss statements, budgeting reports, marketing survey analysis, payroll and so on.

It has been very difficult for organizations to manage their data effectively. In trying to do so, we have to meet two very big challenges which are standing out. Implementing a database requires a widespread organizational change in the role of information and information managers, the allocation of power al senior levels, the ownership and sharing of information, and patterns of organizational agreement.

3.2 INFORMATION

Data are the facts and figures which can take the form of the historical records pertaining to the operations of the enterprise. These can be filed appropriately and act as a source of the fixed documents, accounting ledgers, stock registers and so on. These can be processed to derive the information of paramount importance to an organization such as the profit and loss statements, budgeting reports, marketing survey analysis, payroll and so on. Data have to be collated carefully and presented ethically and truthfully according to the accounting norms set by the regulatory bodies. The files and books consisting data are subject to the validation by the external auditors of the organization who are responsible for scrutinizing the books of accounts of the company and authorized legally to endorse the balance sheet of the company after the completion of the financial year of its operation. Data is retrieved and processed, merged with the ingredients from the variety of data files (databases) to form an information which can be utilized for inference, statistical argumentation processes and typically become a basis for forecasting, making decisions and formulating strategies.

To understand this concept, let us take an example. While conducting a market survey about the demand of company products in the market by a consumable goods producing company, the surveyor reports the following facts:

1. There is increasing demand of industrial products.
2. The consumer durable goods also have a good demand.
3. Among durable goods the demand of company's product is very high.
4. Some consumers want changes in the existing products.

On the basis of this report, the manager as a information process uses the information for immediate action and start producing more goods. He discards the information regarding the industrial good because his company is not dealing with it. He uses the above information regarding the change for future use and incorporates the features demanded by the customer to retain his clients. This survey report can be used by some other companies as reference material.

3.2.1 Need for Information

Need for information in a management system can be extrapolated reviewing:

- The fundamental tasks or the organizational goals of the management;
- The fundamental functions of the management seeking the objectives to fulfill and implement these fundamental tasks;
- Deciphering the role of the flow of information (communication) within a management process.

The fundamental tasks or the organizational goals of the management:

1. Looking forward to mark the opportunities from the business environment and deciding upon the specific mission of the organization vis-à-vis the capabilities of the organization and the needs of the social environment.
2. Revamping the organizational structure for attainment of the fundamental goals of the organization.
3. Allocation of the resources for smooth functioning of the organization towards the goal of the optimum productivity.
4. Maintaining and managing responsibility towards customers, shareholders, employees, vendors, distributors, franchisees, retailers, industry and other social segments.
5. Coping with and augmenting technological sophistication pragmatically and growing public criticism and political hindrances ethically and tactically.
6. Coping with the ripened aspirations of the customers as well as the employees.
7. Maintaining and managing the cordial relationships with the social segments viz. labour markets, recruitment agencies, suppliers of machines, technologies and services, banks and financial institutions, franchisees, statutory auditors, Government, employees, customers, the press and other advertising and news propagating media.

3.2.2 Dimensions of Information

Information has many dimensions as Business Dimension, Technical Dimension, Economic Dimension, Behavioral Dimension etc. which are discussed bellow.

Business Dimension

Information can be analyzed from its business dimension. In a business organization, there are different levels of management. Each level of management requires different type of information.

The information need of a top manager is altogether different from an operation manager. It is due to the fact that manager at different level has different responsibility and functions to perform.

Economic Dimension

Economic Dimension of information refers to the cost and related benefit of information. Transformation or economic dimension leads to optimum use of resources like time and money. At the same time benefit analysis help in determining kind of information required. However, it is difficult to measure both because of intangible characteristics of information.

The cost of information is associated with collection, processing and storage of date. It may be also related to the response time required to generate and communicate the information as well as the speed of generation and accuracy of the information. For example, a system with low response time is more reliable and accurate having cost higher value as compared to the system with relatively low accuracy.

Technical Dimension

Technical relates to the technical aspect of the information. Various aspect of data base are considered like the type of data, capacity of database, security, validity etc., as well as whether the information is computerized or manual, format of its representation etc.

Behavioral Dimension

Behavioral dimension relates to behavioral aspect of the information. Behavioral aspect means the impact of information on individual, group and organization. The support of top management in acquiring information and the support of operational management

External Information

The information generated from outside business organization is known as external information. For e.g. the data regarding industry sales and trends may be collected through chamber of commerce. External information is further divided into two parts. One is information regarding environment and other is of competitors.

Environmental Information

Environment information includes the following:

1. ***Government policies:*** Information about government policies with regard to tax benefits, concessions quota's etc.
2. ***Factor of Production:*** Information regarding various factors of production like labour, capital, material, about their sources, cost and availability.
3. ***Economic trends:*** It includes information relating to economic issues like consumer disposable income, capital investment, monetary policy, Exim policy etc.
4. ***Technological Information:*** Information regarding the present and future technology available, its impact upon firm and cost of acquiring it.
5. ***Sociological Information:*** Information regarding taste, preferences and consumer moments, customer value and norms of the customers.
6. ***Political Information:*** Information regarding the political scene of the country, what is the political manifesto of the ruling party etc.

Competitive Information

This category include information regarding the competitor's product and their demand, their strengths and weakness, different policies regarding pricing, promotion etc.

Internal Information

The internal information is generated with in the company e.g., sales information is retrieved through invoice data. Internal There are mainly 3 level of management in every business organization viz top level, middle level and lower level management.

3.2.3 Determinants of Organization's Information Needs

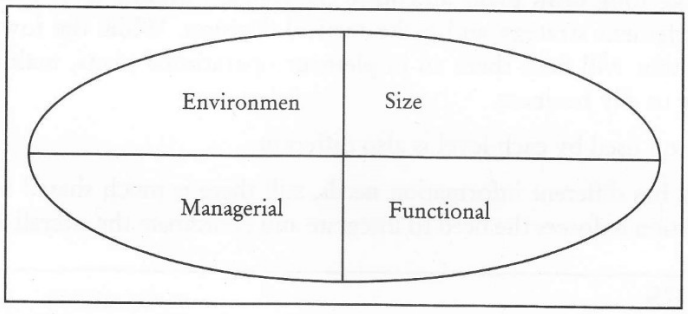

Figure 3.1: Determinants of Information Need

To determine the need of an effective information system many factors have to be considered. Some of them are internal and other are external. There is no hard and fast rule about these factors but as shown in Figure above, four common factors that should be considered while determining the information needs are:

Environment

Every business organization has two types of environment, i.e., micro and macro environment. Micro environment is inside the organization which comprises of value, mission, goals etc. of an organization and which is unique in nature as every business organization has its own value system. Macro environment relates to the external world. It comprises competitor, political, economic and technological environment.

The environment affects the operation of any business organization. There is a greater need to formally organize and manage the information system as the organization must operate in an uncertain and complex environment.

Size

The next general determinant of information need is the relative size of the organization. Generally, the larger the organization, the greater is the need to systematically manage the information system. Also, it can generally be assumed that the greater the diversity within the organization, the more is the need to manage and systematically control the information system.

Functional Area

A business organization has many functional areas viz. marketing, finance, production, R&D, personnel and so on. Each functional area has its own unique set of information needs. For example;

finance needs information regarding source of finance, debt, payment, cash in flows etc. While marketing department needs information regarding sales, pricing, market share and details of customers. Contrasted to these unique information needs, there are certain information needs which are shared by each functional area, e.g. projected sales data can have impact on management plan for all function areas. Therefore, the information system is such that it takes a co-ordinated and integrated approach to serve and satisfy the information need of each functional area to achieve the common goal of the business organization.

Managerial Level

The information need at different levels of management is different. Top management needs information to make long term plans and strategic policies. Middle level management wants tactical information to implement strategy and make tactical decision. While the lower level of management needs information that will help them to implement operational plans, making short term decisions and conducting day to day business.

Source of information used by each level is also different.

Although each level has different information needs, still there is much shared information which is also required. This condition enforces the need to integrate and coordinate the overall information system.

3.3 DATABASES

3.3.1 Introduction

Effective use of information depends on how data are stored, organized and accessed in an organization. Proper delivery of information not only depends upon the capabilities of computer hardware and software but also on the organization's ability to manage data as an important resource. After seeing the various methods and approaches of System Development, now, let us have some understanding of how important are the data resources in an organization and how the information system we developed is going to handle them.

It has been very difficult for organizations to manage their data effectively. In trying to do so, we have to meet two very big challenges which are standing out. Implementing a database requires a widespread organizational change in the role of information and information managers, the allocation of power al senior levels, the ownership and sharing of information, and patterns of organizational agreement. A database management system (DBMS) challenges the existing power arrangements in an organization and for that reason often generate political resistance. In a traditional file environment, each department constructed files and programs to fulfill its specific needs. Now, with a database, files and programs must be built that take into account the full organization's interest in data. Although the organization has spent the money on hardware and software for a database environment, it may not reap the benefits it should if it is unwilling to make the requisite organizational changes.

3.3.2 Functions of Database

A database is a collection of data, which is organized in a way that allows for easy data retrieval and manipulation. While a folder with several files in it may be viewed as a simple database, database professionals usually require that a database have a Database Management System (DBMS). A DBMS is a software tool, which stores data in a specified form and provides access to this data for a user or an application. Specifically, a DBMS provides some or all of the following functionality:

1. ***Data Definition:*** A DBMS must define a structure for stored data, and provide a means for a user to define and organize their data within that structure.
2. ***Data Retrieval:*** A DBMS must provide a toolset that allows a user to retrieve data stored in the database (for instance, query tools).
3. ***Access Control:*** The database administrator should be able to define data access for an individual or a group.
4. ***Data Sharing:*** More than one user should be able to use the database at the same time without a danger of overwriting each other's data changes.
5. ***Data Integrity:*** A DBMS should provide mechanisms for maintaining data integrity through system failures and inconsistent, or incomplete, updates.

3.4 TRADITIONAL VS DATABASE APPROACH

3.4.1 Files vs Traditional Approach

Traditionally, data files were developed and maintained separately for individual applications. Thus, the file processing system relied on the piecemeal approach of data across the organization where every functional unit like marketing, finance, production, etc. used to maintain their own set of application programs and data files.

No doubt such an organization was simple to operate and had better local control but the data of the organization is dispersed throughout the functional sub-systems.

This approach was rendered inadequate, especially when organizations started developing organization-wide integrated application. The major drawbacks of file processing system may be outlined due to the following reasons:

1. Data duplication
2. Data inconsistency
3. Lack of data integration
4. Data dependence
5. Program dependence

1. ***Data Duplication:*** Since each application has its own data file, the same data may have to be recorded and stored in several files.

   Example: Payroll application and personnel application, both will have data on employee name, designation, etc. This results in unnecessary duplication/redundancy of common data items.

2. ***Data Inconsistency:*** Data inconsistency leads to the data inconsistency especially when data is to be updated. Data inconsistency occurs because the same data items which appear in more than one file do not get updated simultaneously in all the data files.

   Example: Employee's designation, which is immediately updated in the payroll system may not necessarily be updated in the provident fund application. This results in two different designations of an employee at the same time.

3. ***Lack of Data***: Integration Because of independent data files, users face difficulty in getting information on any ad hoc query that requires accessing data stored in more than one file. Thus, either complicated programs have to be developed to retrieve data from each independent data file or users have to manually collect the required information from various outputs of separate applications.

4. ***Data Dependence:*** The applications in file processing systems are data dependent, i.e., the file organization, its physical location and retrieval from the storage media are dedicated by the needs of the particular application.

   Example: In order processing application, the file may be organized on customers records stored on their last name, which implies that retrieval of any customer's records has to be through his/her last name only.

5. ***Program Dependence:*** The report produced by the file processing system are program dependent, which implies that if any change in the format or structure of data and records in the file is to be made, a corresponding change in the programs have to be made. Similarly, if any new report si to be produced, a new program will have to be developed.

   It is because of these drawbacks in the traditional files approach of organizing data that led to the development of databases.

3.4.2 Database – The Modern Approach

An alternative approach to the file processing system in the modern approach is known as the database approach. A database is an organized collection of records and files which are related to each other. In a database system, a common pool of data can be shared by a number of applications as it is data and program independent. Thus, unlike a file processing system, data redundancy and data inconsistency in the database system approach are minimized. The user is free from the detailed and complicated task of keeping up with the physical structure of the data.

3.5 DATA MODELLING

Data modelling is the analysis of data objects that are used in a business or other context and the identification of the relationships among these data objects. A data model is a collection of logical constructs used to represent the data structure & the data relationships found within the database. The two different categories of data models are:

1. ***Conceptual models:*** Focus on the logical nature of the data representation. They are concerned with what is represented rather than how it is represented. The Entity-relationship model is a popular conceptual model.

2. ***Implementation models:*** Place the emphasis on how the data are represented in the database or on how the data structures are implemented. The examples are:
   (a) Hierarchical
   (b) Network
   (c) Relational
   (d) Object-oriented
   (e) Semi-structured.

Before we move on to the study of the common data models, let us understand the concepts of data relationships. Two pieces of data may have any of the following three relationships:

1. ***One-to-one (1:1):*** For example, each store-is managed by a single employee & each store manager (employee) only manages a single store, i.e., employee (1) manages store (1).
2. ***One-to-many (1:M):*** For example, a painter paints many different paintings but each one of them is painted by only that painter, i.e., painter (1) paints painting (M).
3. ***Many-to-many (M:N):*** For example, an employee may learn many job skills & each job skill might me learned by many employees, i.e., employee (M) learns skill (N).

3.5.1 Hierarchical Model

In 1966, IBM released the first commercially available DBMS - IMS (Information Management System) based on the hierarchical data model. The basic structure of this model is:

1. Collection of records is logically organized to conform to the upside-down tree (hierarchical) structure.
2. The top layer is perceived as the parent of the segment directly beneath it.
3. The segments below other segments are the children of the segment above them.
4. A tree structure is represented as a hierarchical path on the computer's storage media.

The following diagram shows the organization of a hierarchical DBMS.

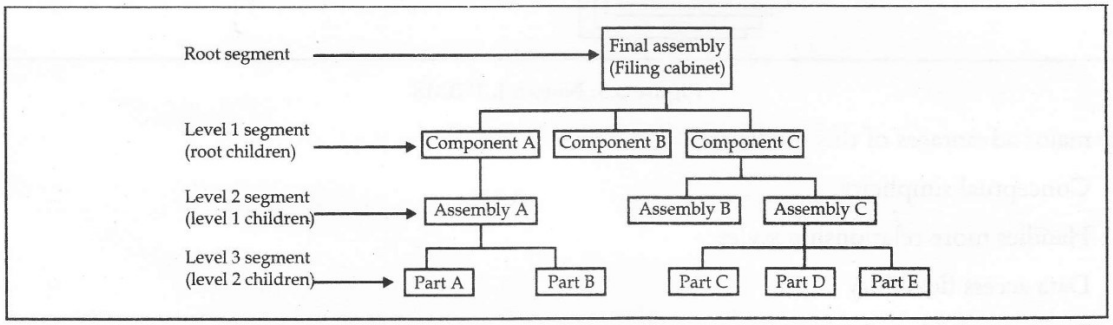

Figure 3.2: Organization of a Hierarchical DBMS

The major advantages of this model are:

1. Conceptual simplicity
2. Database security
3. Data independence
4. Database integrity
5. Efficiency dealing with a large database

The model went out of favour due to the following limitations:

1. Complex implementation
2. Difficult to manage
3. Lacks structural independence

4. Complexity of application programming & use
5. Implementation limitations
6. Lack of standards

3.5.2 Network Model

General Electric (GE) developed a DBMS called IDS (Integrated Data System) based on the network data model in 1967. In this model, a relationship is called a set. Each set is composed of at least two record types: an owner (parent) record & a member (child) record. A set represents a 1:M relationship between the owner & the member. The following is a visual representation of a network DBMS.

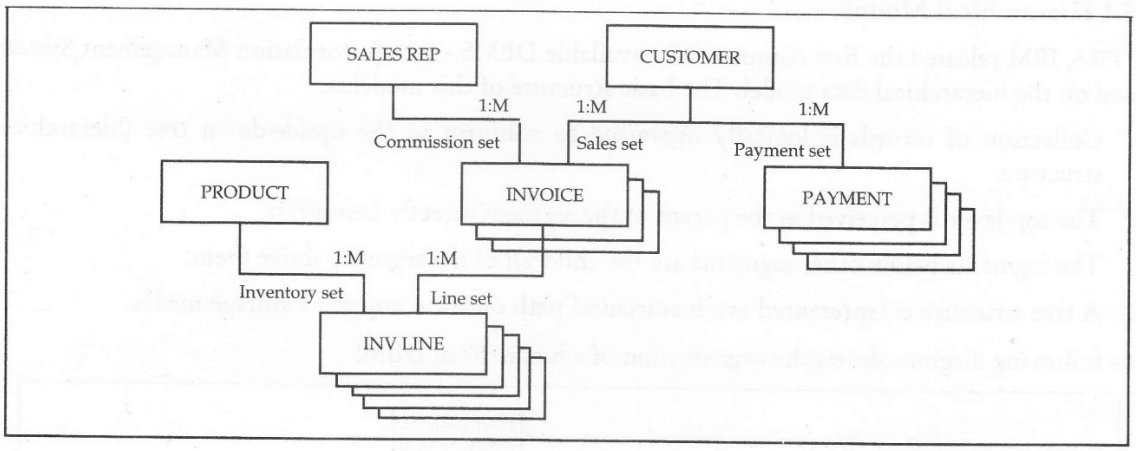

Figure 3.3: Network DBMS

The major advantages of this model are:

1. Conceptual simplicity
2. Handles more relationship styles
3. Data access flexibility
4. Promotes database integrity
5. Data independence
6. Conformance to standards

The model was discarded owing to the following disadvantages:

1. System complexity
2. Lack of structural independence

3.5.3 Entity-relationship Model

It is one of the most widely accepted graphical data modelling tools. It represents data as entities (e.g., customers, accounts, and bank branch) & their relationships (e.g., account 23456 is held by Reetesh) in a database. A relationship set depositor associates customers with accounts. It complements the relational model concepts. The database design in the E-R model is usually converted into the

relational model that is used for storage & processing. E-R models are normally represented in an Entity-relationship Diagram (ERD). In these diagrams, an entity is represented by a rectangle; a diamond connected to the related entities defines a relationship. Each entity is also described by a set of attributes. An attribute describes a particular characteristic of the entity.

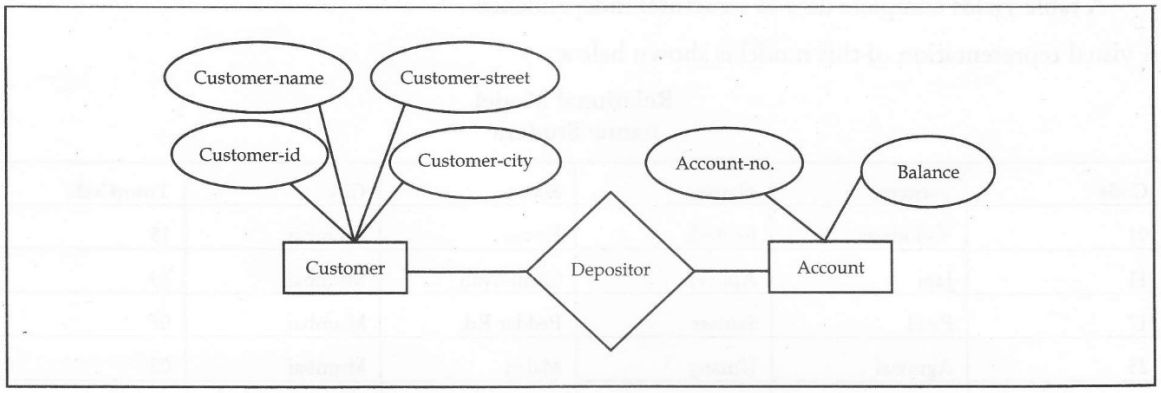

Figure 3.4: ER Model

The advantages of the ER model are:

1. Exceptional conceptual simplicity
2. Visual representation
3. Effective communication tool
4. Integrated with the relational database model

However, this model suffers from serious limitations:

1. Limited constraint representation
2. Limited relationship representation
3. No data manipulation language
4. Loss of information content

3.5.4 Relational Database Model

In 1970, Dr. E.F. Codd, IBM researcher, proposed the relational data model in a theoretical paper. In this model the data are organized as a set of formally described tables from which the data can be accessed or assembled in many different ways without having to reorganize the database Tables.

The standard user and application program interface to a relational database is the SQL (Structured Query Language). SQL statements are used both for interactive queries for information from a relational database and for gathering data for reports.

The basic structure of the relational model is as follows:

1. RDBMS allows operations in a human logical environment
2. The relational database is perceived as a collection of tables
3. Each table consists of a series of row/column intersections

4. Tables (or relations) are related to each other by sharing a common entity characteristic (called a key)
5. The relationship type is often shown in a relational schema
6. A table yields complete data & structural independence

A visual representation of this model is shown below.

Relational Model
name: Student

Code	Surname	Name	Street	City	TutorCode
01	Vadodaria	Reetesh	Powai	Mumbai	15
11	Jani	Apurva	Chandivali	Mumbai	19
17	Patel	Sameer	Peddar Rd.	Mumbai	02
25	Agrawal	Umang	Malad	Mumbai	05
30	Bansal	Shilpi	Borivli	Mumbai	22

Link through Tutor Code
name: Faculty

TutorCode	Surname	Name	Street	City
02	Kumta	Gita	Goregaon	Mumbai
05	Acharya	Neeta	Santacruz	Mumbai
15	Agnihotri	Prafulla	Dindoshi	Mumbai
19	Sirur	Pratap	Malabar Hill	Mumbai
22	Sunder Ram	K.	Mira Rd.	Mumbai

The advantages of the relational model are:

1. Structural independence
2. Improved conceptual simplicity
3. Easier database design, implementation, management & use
4. Ad hoc querying (a query that is not pre-defined) capability through SQL
5. Powerful data management system

The disadvantages, on the other hand, are:

1. Substantial hardware & system software overhead
2. Possibility of poor design & implementation
3. Potential "islands of information" problem

3.6 DATABASE ON THE WEB

Traditionally databases were available to users with login access to the CDS server itself or by means of Client/Server software. More recently we have been able to provide an increasing number of systems directly via the CDS web site.

3.6.1 Web Database

A web database is a database for the internet. You can get access to your data via the internet from anywhere in the world or you can create a database driven website by using such databases. A web database stores large amount of information in an organized format that is easily accessible from scripting languages (like PHP).

In web databases, modifications can be performed without any hassles. That makes the dense editing of the html code obsolete. Repeated types of data, such as contact information, can be generated automatically in your website through a web database.

3.6.2 Daffodil DB: A Web Database

Daffodil DB is a SQL-99 and JDBC standards compliant Java RDBMS. It is the first java database that is compatible with PHP. With the release of PHP extension module, Daffodil DB has joined the select list of Java enabled web databases.

Daffodil DB is an ideal for anyone who is looking for a web database. It can be embedded within any application and delivers high performance with minimum system resource usage - thereby striking the right balance between size, features, and performance. At less than 3 MB, Daffodil DB is a small/compact database.

Daffodil DB can run virtually on any platform as it supports all platforms for which a Java Virtual Machine (JVM) is available. These include Windows, Linux, Solaris, and UNIX. With the help of Daffodil DB, developers can build applications for any platform and these can be compiled and delivered on all other major platforms.

3.6.3 PHP

PHP (Hypertext Preprocessor) is an open source server-side scripting language that is particularly suited for web development and can be used to generate dynamic web content.

PHP scripts are embedded within web pages along with HTML, similar to other web scripting languages, such as Microsoft's ASP or Sun Microsystems's JSP. Like ASP and JSP, PHP runs on a web server (rather than on the Web browser or other client) when a page is requested via HTTP.

PHP enables you to insert instructions into your Web pages, which your Web server software (be it Apache, Personal Web Server, or whatever) executes before sending it to a browser that requests them.

PHP is easy to use for web development because it has been designed from the onset for the web environment. PHP has many built-in functions that make web programming simpler, so that programmers can focus on the logic of programming without wasting precious development time.

PHP is one of the fastest growing server side scripting languages and you need to add it to your current arsenal of Perl, ASP, JSP, JavaScript, VBScript, and Java.

3.6.4 PHP and Daffodil DB: Together?

Daffodil has come up with a PHP extension module that makes Daffodil DB/One$DB compatible with PHP. It is a database extension module driver that is required to work with Daffodil DB in PHP. This extension module uses the capabilities of a Java database and extends it to your web infrastructure.

By using this extensive module, PHP developers can create sophisticated data-driven web applications having Daffodil DB as a web database in a fraction of the time and with reduced cost.

PHP extension for Daffodil DB provides a simple, yet effective means for connecting to Daffodil DB within PHP. JVM is created using JNI invocation interface and everything runs in-process.

This PHP Extension module boosts open source activities around Daffodil DB. Moreover, open Source stack -Linux, Apache, One$DB, and PHP; all being open source products give an ultimate combination to web development community.

PHP as a language has its own way of doing things, and it has borrowed features from other languages. But for many people, the main reason for learning a scripting language like PHP is because of the interaction with databases it can offer. Moreover, PHP is cross platform means it allows you to develop/use your solution on multiple OSs.

PHP is platform independent and Daffodil DB is also platform independent. PHP is free and Daffodil DB is also free (having open source version). So What about using the powers of two most promising technologies, PHP and Daffodil DB together?

The capabilities of Daffodil DB and PHP can make up what must be the best blend for data-driven Web sites on the globe.

Users can use PHP and the Daffodil DB database to access/store information on the web and include it into their website. By using PHP as a front-end and Daffodil DB as a back-end web database, users can benefit from huge savings on the licensing costs of commercial alternatives.

3.7 DATABASE MODELS

A database model or database schema is the structure or format of a database, described in a formal language supported by the database management system, In other words, a "database model" is the application of a data model when used in conjunction with a database management system.

Different types of database models are:

1. Relational model
2. Object-oriented model
3. Hierarchical model
4. Network model
5. Flat file model

A data model is not just a way of structuring data: it also defines a set of operations that can be performed on the data. The relational model, for example, defines operations such as select, project, and join. Although these operations may not be explicit in a particular query language, they provide the foundation on which a query language is built.

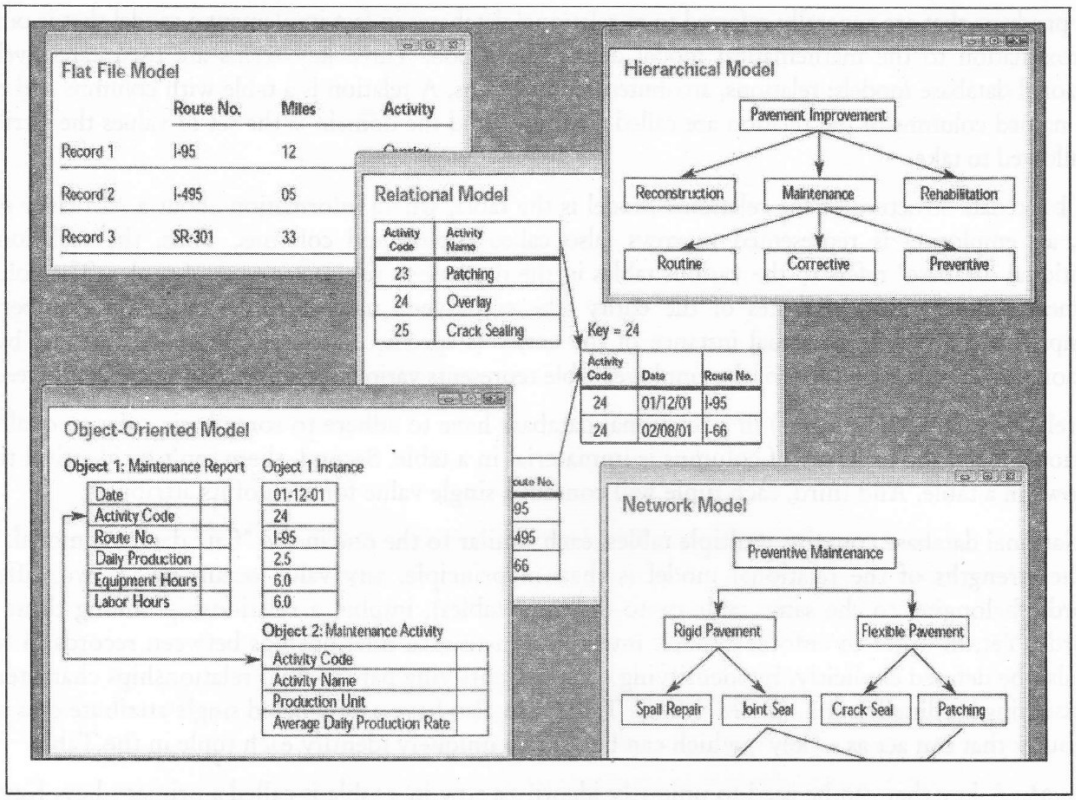

Figure 3.5: Database Model

3.7.1 Relational Model

The relational model was introduced by E.F. Codd in 1970 as a way to make database management systems more independent of any particular application. It is a mathematical model defined in terms of predicate logic and set theory.

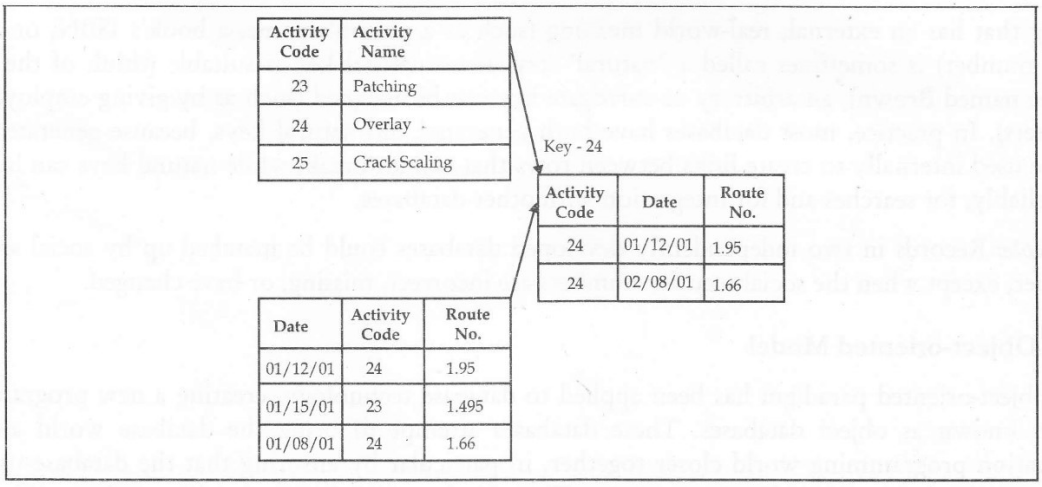

Figure 3.6: Database Model

The products that are generally referred to as relational databases in fact implement a model that is only an approximation to the mathematical model defined by Codd. Three key terms are used extensively in relational database models: relations, attributes, and domains. A relation is a table with columns and rows. The named columns of the relation are called attributes, and the domain is the set of values the attributes are allowed to take.

The basic data structure of the relational model is the table, where information about a particular entity (say, an employee) is represented in rows (also called tuples) and columns. Thus, the "relation" in "relational database" refers to the various tables in the database; a relation is a set of tuples. The columns enumerate the various attributes of the entity (the employee's name, address or phone number, for example), and a row is an actual instance of the entity (a specific employee) that is represented by the relation. As a result, each tuple of the employee table represents various attributes of a single employee.

All relations (and, thus, tables) in a relational database have to adhere to some basic rules to qualify as relations. First, the ordering of columns is immaterial in a table. Second, there can't be identical tuples or rows in a table. And third, each tuple will contain a single value for each of its attributes.

A relational database contains multiple tables, each similar to the one in the "flat" database model. One of the strengths of the relational model is that, in principle, any value occurring in two different records (belonging to the same table or to different tables), implies a relationship among those two records. Yet, in order to enforce explicit integrity constraints, relationships between records in tables can also be defined explicitly, by identifying or non-identifying parent-child relationships characterized by assigning cardinality (1:1, (0)1:M, M:M). Tables can also have a designated single attribute or a set of attributes that can act as a "key", which can be used to uniquely identify each tuple in the Table.

Example: A key that can be used to uniquely identify a row in a table is called a primary key. Keys are commonly used to join or combine data from two or more Tables.

Example: An Employee Table may contain a column named Location which contains a value that matches the key of a Location table.Keys are also critical in the creation of indexes, which facilitate fast retrieval of data from large tables. Any column can be a key, or multiple columns can be grouped together into a compound key. It is not necessary to define all the keys in advance; a column can be used as a key even if it was not originally intended to be one.

A key that has an external, real-world meaning (such as a person's name, a book's ISBN, or a car's serial number) is sometimes called a "natural" key. If no natural key is suitable (think of the many people named Brown), an arbitrary or surrogate key can be assigned (such as by giving employees ID numbers). In practice, most databases have both generated and natural keys, because generated keys can be used internally to create links between rows that cannot break, while natural keys can be used, less reliably, for searches and for integration with other databases.

Example: Records in two independently developed databases could be matched up by social security number, except when the social security numbers are incorrect, missing, or have changed.

3.7.2 Object-oriented Model

The object-oriented paradigm has been applied to database technology, creating a new programming model known as object databases. These databases attempt to bring the database world and the application programming world closer together, in particular by ensuring that the database uses the same type system as the application program. This aims to avoid the overhead (sometimes referred to

as the impedance mismatch) of converting information between its representation in the database (for example as rows in tables) and its representation in the application program (typically as objects). At the same time, object databases attempt to introduce the key ideas of object programming, such as encapsulation and polymorphism, into the world of databases.

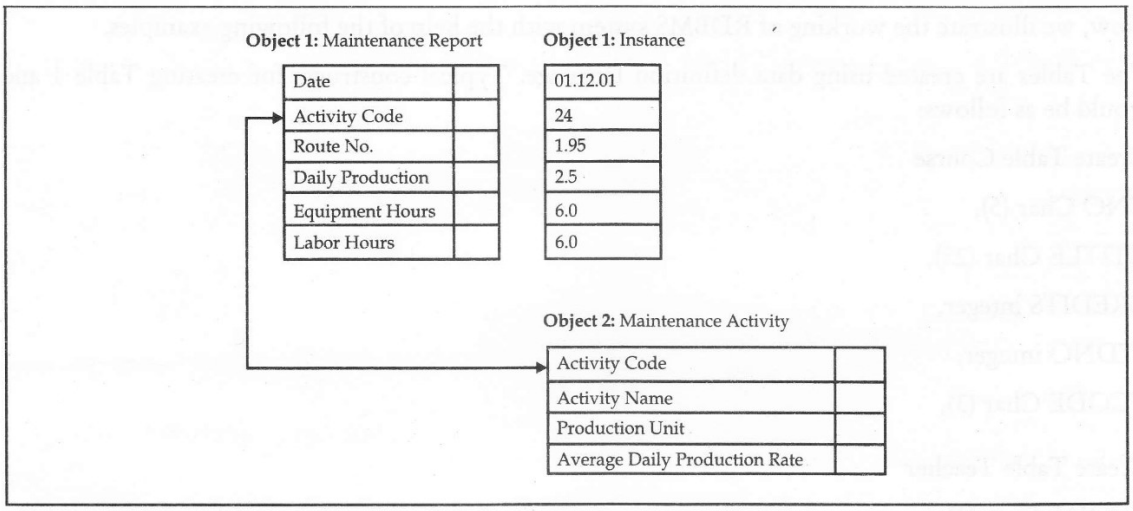

Figure 3.7: Object-oriented Model

A variety of these ways have been tried for storing objects in a database. Some products have approached the problem from the application programming end, by making the objects manipulated by the program persistent. This also typically requires the addition of some kind of query language, since conventional programming languages do not have the ability to find objects based on their information content. Others have attacked the problem from the database end, by defining an object-oriented data model for the database, and defining a database programming language that allows full programming capabilities as well as traditional query facilities.

Object databases suffered because of a lack of standardization: Although standards were defined by ODMG, they were never implemented well enough to ensure interoperability between products. Nevertheless, object databases have been used successfully in many applications: usually specialized applications such as engineering databases or molecular biology databases rather than mainstream commercial data processing. However, object database ideas were picked up by the relational vendors and influenced extensions made to these products and indeed to the SQL language.

3.7.3 Relational Operations SQL

Structured Query Language, popularly known as SQL is the language that is used in most relational database systems. It is called structured query language because it follows a rigorous set of rules and procedures in answering queries. SQL is also termed as 4GL to distinguish it from other 3GL programming languages like, PASCAL, COBOL or C.

SQL is a simple and powerful query language that is capable of answering simple to most complex queries. Any query on a single table can be performed by using only two basic operators, namely SELECT and PROJECT.

The select operator selects a set of records from the table, whereas PROJECT takes out selected fields from the table. The two operators may be understood, in the user's view, as a horizontal cut and vertical cut, respectively of the table. Another operator JOIN is also used in SQL when the query requires more than one table. JOIN links or combines two tables together over a common field.

Now, we illustrate the working of RDBMS system with the help of the following examples.

The Tables are created using data definition language. Typical constructs for creating Table 1 and 2 would be as follows:

Create Table Course

CNO Char (5),

CTITLE Char (25),

CREDITS integer,

STDNO integer,

TCODE Char (3),

Create Table Teacher

TCODE Char (3),

NAME Char (20),

DEPTT Char (5),

DESIG Char (12),

PHONE Char (8),

Table 1: Data Table Course

CNO	CTITLE	CREDITS	STDNO	TCODE
CS101	MIS	6	25	07
CS201	SAD	4	25	15
CS304	Software Engineering	4	25	30
CS406	Information Technology	3	20	06
CS303	Strategic Systems	4	20	11

Table 2: Data Table Teacher

TCODE	NAME	DEPTT	DESIG	PHONE
07	R.Balekar	MGT	Reader	235467
06	S.C. Sharma	IT	Reader	219834
30	P. Gupta	IT	Professor	230956
15	Rakesh Kumar	ENGG	Reader	327654
11	Anuj Saxena	MGT	Professor	351087