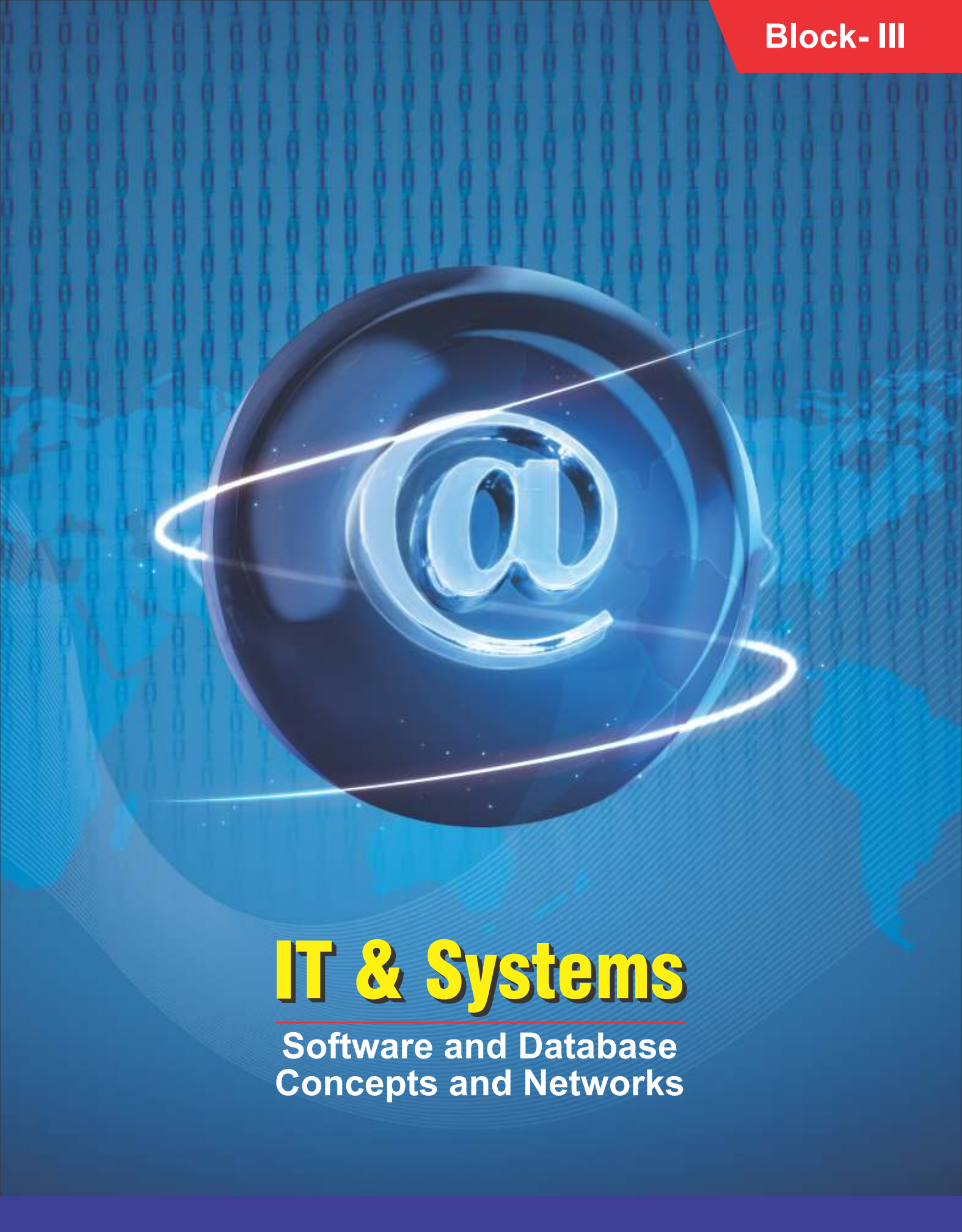
Block- III



IT & Systems

Software and Database Concepts and Networks

Block

III

SOFTWARE AND DATABASE CONCEPTS, AND NETWORKS

Ref. No. ITM SLM 102021B3

For any clarification regarding this book, the students may please write to The ICFAI Foundation for Higher Education (IFHE), Hyderabad specifying the unit and page number.

While every possible care has been taken in type-setting and printing this book, The ICFAI Foundation for Higher Education (IFHE), Hyderabad welcomes suggestions from students for improvement in future editions.

Our E-mail id: cwfeedback@icfaiuniversity.in

Centre for Distance and Online Education (CDOE)
The ICFAI Foundation for Higher Education
(Deemed-to-be-University Under Section 3 of UGC Act, 1956)
Donthanapally, Shankarapalli Road, Hyderabad- 501203.

BLOCK III: SOFTWARE AND DATABASE CONCEPTS, AND NETWORKS

The third block to the course on IT & Systems deals with the software and database concepts, and networks. The block contains four units. The first two units discuss certain software concepts and database concepts while the next two units discuss computer networks and telecommunication networks.

The first unit, *Program Design and Programming Languages*, introduces you to the process of developing programs and to the various tools used for designing programs like flow charts, decision tables, and pseudocode. This unit also discusses the programming languages that are used for instructing a computer to carry out specific tasks.

The second unit, *Database Management*, deals with the database management system that has replaced the file-processing system used earlier by organizations to store data. The database management system ensures easy storage and retrieval of information. It also ensures system security, and prevents unauthorized access and loss of data due to system failure. The relationship between MIS and RDBMS is highlighted in the unit.

The third unit, *Computer Networks*, examines the key role played by computer networks in facilitating information sharing and distribution. Computer networks have become indispensable today, and have affected the various ways in which people/organizations operate and communicate. Components of unified communication are discussed in the unit.

The fourth unit, *Telecommunication Networks*, examines the role played by telecommunications for the purpose of communication. Lots of changes have taken place in the field of telecommunications. As a result of these changes, distance is no longer a barrier to communication. They have also immensely benefited the conduct of business operations.

Unit 7

Program Design and Programming Languages

Structure

7.1 Introduction

In the previous unit, we studied Information Systems for Business Operations, Transaction Processing Systems, Financial Information Systems, Management Information Systems, Decision Support Systems, Executive Information Systems and Knowledge Based Expert Systems.

In this unit, we introduce you to program, and process through which it is developed and the various tools used.

Computer software consists of a set of programming instructions, called as programming languages. These languages are used by software professionals for writing programs. Programming languages consist of a set of rules called syntax that the programmer should follow while writing programs. Proper care should be taken while planning the logic and writing the program in order to avoid errors. The various program design tools are flow charts, pseudocode, decision tables, etc. Computer programs are generally designed and developed in sets. The Program Development Life Cycle (PDLC) is an organized plan that consists of modules. These modules are units into which the various program developed tasks are divided.

In this unit, we would discuss about the various tools used for designing programs, the program development life cycle, and the different programming languages.

7.2 Objectives

By the end of this unit, you should be able to:

- State the various stages in the program development lifecycle.
- Compare the various program design tools.
- Define the generations of programming languages.
- Explain the concept of language translators and some popular programming languages.
- Recall the basic concepts of object-oriented programming.

7.3 Program Development Lifecycle

The development of computer programs enables automation of the existing processes and creation of new functionalities. Computer programs are usually designed and developed in sets that together constitute a whole system. Proper care should be taken while writing computer programs as leaving out of any program instructions or writing the instructions in the wrong sequence would result in incorrect answer. Therefore, proper sequence should be used for writing each and every instruction to generate an effective computer program without any omissions. The basic objective of a program is to generate output. Program preparation begins with deciding upon what comprises the output, how this information will be formatted, and how it will appear. The PDLC, developed in early 1970s is similar to the System Development Life Cycle (SDLC). It provides an organized plan for breaking down the task of program development into manageable units called modules. Each of these modules should be successfully completed before the programmer moves on to the next phase.

7.3.1 Defining the Problem

The first step in developing a program is to define the problem. The system analysts give program specifications like the input data, the processing that needs to be done, the output features, and the user interface design to the programmers. Depending on the size of the job, program development can be handled by a single individual or a team of programmers. In program development, documentation is an ongoing process and these documents are maintained for reference in the subsequent stages, if required.

7.3.2 Designing the Program

Program designs are developed from program specifications. These designs elaborately specify the path of program execution. In this step, programmers identify the output and then break the program into manageable modules. This approach is called structured programming or

top-down program design. Here, the first step involves identifying the main routine, followed by breaking down the various components of the main routine into smaller subroutines or modules until each of them is highly focused and accomplishes only one single task. Algorithms are developed to solve problems. An algorithm is a step-by-step description of a method for carrying out a task or for arriving at a solution. It can be defined as a sequence of instructions which when executed will give the desired results. Algorithms are combined with structured English, flow charts, or with pseudocode.

7.3.3 Coding the Program

Coding refers to the process of translating the algorithm into specific program language instructions. A suitable programming language should be selected and the program should be created by typing the code. Programmers should carefully follow the language syntax. Syntax refers to rules that specify how certain operations should be expressed to arrive at appropriate results. A particular language suits a particular program. For instance, assembly language is used for writing scientific programs and COBOL is mostly used for writing business programs.

7.3.4 Testing and Debugging the Program

Testing and debugging of the program is done so as to ensure that there are no obvious errors in the program. These processes are done for detecting and eliminating any syntactical and logical errors in the program code. The program testing tools check for syntax errors while the program is being written. The program will be executable after eliminating the syntactical errors. The output may still be error free as the language-testing tool cannot detect logic errors. A logic error is a mistake made by the programmer while designing the program. The program output must be carefully examined to find and correct the logic errors. Both syntactical and logical errors are called as bugs, and the process of eliminating these errors is called debugging.

7.3.5 Formalizing the Solution

The program is almost complete at this stage. The structure chart and the pseudocode (or flow charts) which are developed during the design phase are used as manuals for modifying the program when required. While the program is being coded, other documents like a list of variable names and definitions, descriptions of files that the program needs to work with, and layouts of output that the program produces, should be created. All these

documents should be saved and should be available for future reference to the people who will maintain the information system.

The program developed should be tested by the users before implementing it. Manuals should be developed for explaining thoroughly the working of the program. These manuals give a general idea about the program's functionality, tutorials for new users, in-depth explanations of major program features, reference documentation of all program commands, and a detailed description of the error messages generated by the program.

7.3.6 Maintaining the Program

After implementation, the program should be maintained and evaluated on a regular basis. This process involves correcting the errors identified by the users. Periodic evaluation may result in program modification.

Activity: Rahul Chaudhury (Rahul) works as a Systems Analyst at RiskSoft Ltd, which specializes in risk management solutions for life insurance companies. Recently, an insurance company asked RiskSoft Ltd to design a program which would help them find out the insurees who were not regularly paying their premium to the company within the due date. As part of the program development team, Rahul is responsible for designing the program. Suggest measures to Rahul on how he should go about developing the program design. What are the various stages in the program development lifecycle?

Answer:

Check Your Progress-1

1. What are the various steps in the program designing process?
2. In which of the following stages in the Program Development Life Cycle (PDLC) are the output features defined?
   a. Defining the problem
   b. Designing the program
   c. Coding the program
   d. Formalizing the solution

3. What are manuals? Why are they developed?

4. What is coding?

5. The process of error elimination in a software program is called:
   a. Translating
   b. Debugging
   c. Coding
   d. Looping

6. _____ refers to the approach in which programmers identify the main goal of the program (output) and then break the program into manageable modules, each of which contributes to goal achievement.
   a. Structured programming
   b. Object-oriented programming
   c. Cross-platform programming
   d. None of the above

7. Once the problem has been defined, the following activities are undertaken as part of the program development life cycle.
   i. Testing and debugging the program
   ii. Coding the program
   iii. Maintaining the program
   iv. Formalizing the solution
   v. Designing the program

What is the sequence in which the above activities should take place?

7.4 Program Design Tools

Before developing the program structure, programmers identify the user requirements. Then, they develop the program logic in a step by step manner. Some of the program designing tools that help programmers arrive at a logical flow are flow charts, decision tables and pseudocode.

7.4.1 Flow Charts

A flow chart is a diagrammatic representation of the logic to be incorporated in the computer program. It is a planning document or 'blueprint' of a program's logic that specifies a processing plan. It helps

the programmer in organizing the sequence of steps required by a computer for solving a problem.

Program flow charts use symbols for identifying each step in the process. These flow charting symbols makes the understanding of the program or the process easier and also helps in finding out any loopholes in them. Since all software developers are familiar with these symbols, program flow charts can be easily understood even by the new employees. The basic processes common to all programs occur in this program flow chart:

- **Beginning/ending the program logic –** The program logic starts with the oval symbol at the top marked by 'start' and continues until a 'stop' symbol.
- **Data – input or output** – The input or output actions in a flow chart are depicted by parallelogram-shaped input/output symbols.
- **Decision making** – Questions are represented by diamond-shaped decision symbols. They indicate decision points in a flow chart. They present two decision alternatives, Yes or No. This is known as branching.
- **Processing** – Processing refers to any action performed on the input data. For instance, arithmetic computation is shown by the rectangular processing symbol in the flow chart.
- **Other symbols –** Four other symbols which are commonly used for preparing flow charts are the oval terminal symbol, on-page connector symbol, rectangular predefined processing symbol, and off-page connector symbol. The oval terminal symbol is used for representing the beginning and end points of a program. The on-page connector symbol is a small, round symbol used for indicating the place where the computer logic leads. This symbol eliminates drawing of long connector lines between the distant parts of a flow chart. The rectangular predefined processing symbol represents the programming logic which is defined elsewhere. An off-page connector symbol is an arrow-shaped symbol used in place of the on-page connector symbol for showing the entry to or exit from a page.

7.4.1.1 Flow Charting Rules

A number of rules and guidelines have been developed by the American National Standards Institute (ANSI) for standardizing the flow charting process. Similar types of flow charting standards have been established by various computer manufacturers and data processing departments.

Some of the guidelines or standards for drawing flow charts are:

- While drawing a flow chart, all the requirements should be listed in a logical order.

- The flow chart should be easy to comprehend without any ambiguity.
- Programmers should use conventional symbols to increase user understanding of program flow charts.
- The logic of a program flow chart should flow from top to bottom and from left to right. This follows standard English convention and leads to consistency in drawing.
- Each symbol used in a flow chart should have an entry point and an exit point.
- The instructions within the symbols of a program flow chart should be independent of any particular programming language. Sometimes, it is not known which computer language will be used to write a particular program. At other times, a program will be written in one language and later rewritten in another language.
- All decision branches should be labeled. Labeling of decisions will make the understanding of alternatives easier.
- If a flow chart becomes complex, it should be simplified by using connector symbols. The connector symbols bring down the number of flow lines connecting the various flow charting symbols.
- The flow chart's validity should be tested by putting it through a simple set of data.

7.4.1.2 Advantages and Limitations of Flow Charts

The benefits of flow charts for program planning include:

- Since a flow chart is a pictorial representation of a program, it is easier for a programmer to explain the logic of a program to another programmer or to his/her superior.
- A flow chart can be used for performing detailed analysis by breaking down each component of the flow chart. Flow charts can be used for analyzing the problems in a more effective way.
- When a team of programmers work together, each person prepares the design for a different module. By integrating all the flow charts, the total system design can be finalized.
- Flow charts help programmers understand the user requirements. It guides them from the starting point of the program to the final point, ensuring that no steps are omitted.
- Flow charts make it easy to detect bugs in programs. Even though a program has been designed with care, there may be some errors in the program. Sometimes, designers overlook issues that may come up in

exceptional cases. These errors are detected only when the program is executed on a computer. Once a bug has been detected, it is easy to find out the reason for the error by reviewing the logic of the program given in the flow chart.

The limitations of flow charts include:

- It is very time-consuming and laborious to draw flow charts with proper symbols and spacing, especially for large, complex programs.
- Owing to the symbol-string nature of flow charting, any changes or modifications in the program logic will usually require a completely new flow chart. Redrawing the flow chart is again a tedious and time consuming job.
- There are no standards regarding the amount of detail that should be included in a flow chart.
- In certain situations when the program logic is more problematic, the flow chart also becomes more complex.

Because of such limitations, many organizations are now reducing the use of flow charts. They are using alternative tools for program analysis.

7.4.2 Decision Tables

Decision tables are used for outlining a set of conditions that a computer program might encounter and the actions corresponding to those. These tables show the true or false status of a particular condition, exhibit the dependency of the condition on the values of certain attributes, and indicate the action to be performed on each condition. Decision tables are presented in a tabular form and they clearly and concisely state the problem. These tables are used when numerous decisions are to be taken, when lot of branching is involved, or for modeling complicated program logics. These are used in areas like programming, testing, business analysis, hardware designing, etc.

Activity: You have been hired as a software programmer by Bansal High School. A new progress report template has been designed by the school management; you have to develop a reporting program to print the year-end progress report on pre-printed stationery.

The database already has data about the students, including their names, roll numbers, classes/sections, and the marks obtained in various subjects in the annual examination. Draw a flow chart for calculating the percentage of marks obtained by each student in the annual examination and then printing it along with the name of the school, the name of the

Contd. .....

student, roll number, and the marks in each subject. The report should also indicate that the student has been promoted to the next class if he/she has scored at least 40% marks in each subject.

Answer:

The various steps in constructing a decision table are:

1. Clearly state the problem,
2. List all the conditions to be tested in the problem,
3. List the corresponding actions that should be taken with the combination of different conditions, and
4. Form a decision table using the two lists.

A typical decision table consists of rows and columns that are divided into four parts or quadrants. The first part of the table contains the condition stub and the condition entries. The second part of the table contains the action stub and the action entries. The 'condition stub' describes the conditions that exist in the program, i.e., the conditions in the decision symbols of a flow chart. The 'action stub' contains action statements that correspond to the respective conditions. Condition statements or causes reveal the probable state of input data while action statements or effects describe the possible actions by the computer system. Action entries contain the actual actions taken by the computer while the condition entries correspondingly branch out from the decision symbols.

Compared to flow charts, decision tables are easier to draw and edit. They provide more compact documentation. A small table can replace several pages of charts. It is easier to follow a particular path down one column than through several flow chart pages. However, flow charts are better to explain the total sequence of events that are required to solve a problem than decision tables. Also, programmers, especially those who are beginners, are familiar with flow charts.

7.4.3 Control Structures

A program does not always stick to a linear string of instructions as it may diverge, reiterate the code, or take decisions during the process. Control structures are therefore used for specifying what the program needs to do, when, and under what conditions. There are two types of control structures: branching structures and looping structures.

- Branching structures: If, Nested If, If Else, and Switch are used for transferring controls from one step to another in a program code.
- Looping structures: On the other hand, looping structures (like While, Do While, and For) are used in programs for performing repetitive tasks.

7.4.4 Pseudocode

Pseudocode is a type of structured English that is used to describe algorithms. Also known as imitation code, it is a summarized form of a program that is written in a form that can be easily converted into real programming statements. Pseudo instructions are written in natural languages like English, French, German, etc., and do not need any language syntax or formatting rules. The pseudocode uses a structure that resembles computer instructions. It lays emphasis on the program design of the program; therefore, it is also called Program Design Language (PDL). 'Sequence selection' and 'iteration' are the common processing procedures that are specially used for designing program logic.

7.4.4.1 Sequence Logic

Sequence logic, also called decision logic, is used to perform instructions one after another, in a sequence. They are written in the order (or sequence) in which they should be performed and are used for making decisions. The logic flows from top to bottom. It is used for selecting the appropriate path out of two or more alternative paths in the program logic. Selection logic has either an IF.THEN or IF.THEN.ELSE structure. According to the IF.THEN.ELSE structure, if the condition is true, then process 1 is executed, else process 2. In both the structures, process 1 and process 2 can consist of one or more processes. ENDIF is used to indicate the end of a decision structure.

7.4.4.2 Iteration Logic

Iteration logic is used for looping when one or more instructions are required to be executed several times on the basis of same condition. It uses two structures called DO WHILE and REPEAT UNTIL. Though both DO WHILE and REPEAT UNTIL are used for looping, in the DO WHILE

structure, the looping continues as long as the condition is true, whereas in the REPEAT UNTIL structure, the looping continues until the condition becomes true. In the DO WHILE structure, the condition is tested at the top of the loop, whereas in the REPEAT UNTIL structure, it is tested at the bottom of the loop. Moreover, END DO marks the end of a DO WHILE structure, whereas UNTIL followed by a condition, marks the end of the REPEAT UNTIL structure.

7.4.4.3 Advantages and limitations of Pseudocode

A pseudocode has the following advantages:

- Converting a pseudocode to a programming language is much simpler and easier than converting a flow chart or a decision table.
- Modifying the pseudocode to match changes in the program is easy.
- Less time and effort are required to write a pseudocode than draw a flow chart.
- A pseudocode can be written in any language. It is language independent and flexible.
- It is very easy to understand a pseudocode as it does not require the knowledge of any particular programming language to understand it.

However, a pseudocode suffers from the following limitations:

- A graphical representation of program logic is not available.
- There are no standard rules to follow while using a pseudocode. This could result in communication problems.
- For a beginner, it is more difficult to follow the logic of or to write a pseudocode, as compared to developing a flow chart.

As soon as the logic of the program is developed and designed with the help of a design tool, the programmer has to choose a language which the computer can understand to accomplish the given task.

Check Your Progress-2

8. What are the various program design tools? Explain them briefly.
9. A step-by-step, non-technical description (in plain English) of solving a problem is called a/an:
   a. Flow chart
   b. Algorithm
   c. Pseudocode
   d. Decision table

10. Which of the following statements is **false** regarding flow chart rules?
    a. Programmers should use conventional symbols to increase user understanding of program flow charts.
    b. The logic of a program flow chart should flow from bottom to top and right to left.
    c. Because a computer 'thinks' in a linear manner, a flow chart should depict a process sequentially.
    d. The instructions within the symbols of a program flow chart should be independent of any particular programming language.
11. The general rules and guidelines to help standardize the flow chart processes are given by _____.
    a. Advanced Research Project Agency
    b. American National Standards Institute
    c. National Science Foundation
    d. US Department of Defense
12. What are decision tables? When are they used?
13. Condition stub, condition entries, action stub, and action entries are part of
    a. Pseudocode
    b. Flow charts
    c. Decision tables
    d. Control structures
14. _____ is used when one or more instructions are required to be executed several times on the basis of the same condition.
15. Which of the following statements is **false** regarding decision tables and flow charts?
    a. Decision tables are easier to draw and edit compared to flow charts.
    b. Decision tables provide more compact documentation than flow charts.
    c. Programmers, especially beginners, are more familiar with flow charts than with decision tables.
    d. Decision tables are better able to express total sequence of events needed to solve a problem than flow charts.

7.5 Generations of Programming Languages

A programming language is a vocabulary and a set of grammatical rules for instructing a computer to perform specific tasks.

All computer languages have a vocabulary of their own. Every symbol of the vocabulary has a definite unambiguous meaning and instructs the computer to do a particular job. The symbols of a particular language must be used as per the set rules, called the syntax of the language. Else, the computer will not be able to understand the instructions.

Over the years, there have been considerable changes in the programming languages. They have progressed from machine-oriented languages that use strings of binary 0s and 1s to problem-oriented languages that use common mathematical and/or English terms. Computer languages can be classified into: low-level languages and high-level languages.

Low-level languages are hardware friendly. The machine and assembly languages are low-level languages. First generation languages and second generation languages are considered as low level languages. Languages of higher generations (i.e. third, fourth and fifth generations) are regarded as high-level languages.

7.5.1 First Generation Languages

A machine language (or first generation language) consists of binary numbers (0s and 1s) that directly correspond to the on and off electrical pulses generated by the computer. It is the only language that the computer can understand without any translation. Hence it should be written in accordance with the unique characteristics of a given processor. These languages are written in numbers and are therefore difficult to be comprehended by the humans. Machine languages are said to be machine-dependent since the programs need to be rewritten or recompiled for running them on other computers.

An instruction prepared in any machine language has a two-part format, opcode (operation code), the first part, and operand, the second part. Opcode instructs the computer what function to perform. Operand instructs the computer where to find or store the data that has to be manipulated. Thus, each instruction tells the control unit of the CPU what to do, and also the length and location of the data fields that are involved in the operation.

Despite the fact that newer programming languages give programmers easier ways to write programs, these programs have to be translated into machine language to get executed on a given processor.

7.5.2 Second Generation Languages

Also known as the second generation programming language (2GL), an assembly language is similar to a machine language except the fact that it

uses names instead of numbers. It is used especially for writing short programs. It allows programmers to use brief abbreviations called mnemonics, for program instructions and base-10 (decimal) numbers. For example, the mnemonic COMPARE x, y, tells the processor to compare the data stored in locations x and y.

Assembly language is also known as symbolic language. A program written in symbolic language is called an assembly code. An assembler translates an assembly code into a machine code. A symbolic program written by a programmer in assembly language is called a source code which is translated into a machine code called object code.

7.5.3 Third Generation Languages

A high-level language, called procedural language, instructs the computer what to do and how to do it. It is also known as the third generation programming language (3GL). In 3GL, the programmer can write instructions using familiar words like DISPLAY, ADD, PRINT, etc. Here, each line of instruction sums up many lines of assembly or machine codes.

While low-level languages are more hardware friendly, high-level languages are closer to human languages. These languages provide an instruction set that is similar to English language, and are then translated by a program called the compiler into the machine mode. Some of the high-level programming languages are Ada, Algol, BASIC, COBOL, C, C++, FORTRAN, LISP, and Pascal.

Example: FORTRAN – The First High Level Programming Language

Invented in 1954 and released commercially in 1957, FORTRAN or Formula Translation was the first high-level programming language. It was invented by John Backus (Backus) for IBM. FORTRAN started as a digital code interpreter for the IBM 701 (IBM's first commercially successful general-purpose computer). Backus wanted a programming language that would be closer to the language of human beings.

FORTRAN has applications in several areas. It is used in programming scientific and mathematical applications. It is also used to program video games, air traffic control systems, payroll calculations, numerous scientific and military applications, and parallel computer research.

Adapted from Mary Bellis, "Inventors of the Modern Computer: FORTRAN – The First Successful High Level Programming Language – Invented by John Backus and IBM," http://inventors.about.com/ library/weekly/aa072198.htm.

7.5.4 Fourth Generation Languages

Fourth generation languages (4GLs) are also known as non-procedural languages since they do not require the programmer to specify procedures in detail. Non-procedural languages are application specific languages and have commands similar to English. Most of these languages are used for accessing databases. 4GLs may be used for executing a series of commands entered by the programmer, like sort records, summarize them on a field, and print a report.

Two well known fourth generation languages which have the capability of creating and manipulating databases are FOCUS and NATURAL. Programming languages which are regarded as 4GLs include report generators (used for printing the database reports), and query languages (used for getting information from the databases).

However, according to some firms, 4GLs are less efficient compared to the procedural languages and cannot be used in a heavy transaction processing environment.

7.5.4.1 Application Generators

An application generator is a type of language designed to facilitate the development of a complete application. It includes defining input transactions, editing such transactions, creating a database, updating files, generating reports, and processing queries. These are higher level languages which are easier to use than other high-level programming languages like COBOL, C, or C++. These languages are meant for use in business information systems. Since the definition and creation of a database is a very critical aspect of the system design process, most application generators are built around database management systems. They significantly reduce the time required for programming and testing and are hence faster than procedural programs.

Application generators are domain-specific languages used for performing specialized tasks relating to implementation of communication protocols, solving partial differential equations, etc. Since application generators are domain-specific, they can detect errors at an advanced level and also have a greater potential for optimization.

However, these languages usually have lines of code which makes them very slow or disables them from carrying out certain functions. Hence, they are rarely used for writing commercial programs.

7.5.5 Fifth Generation Languages

The fifth generation programming languages or 5GLs are programming languages designed to enable the computer to solve a problem. With 5GLs,

the programmer can focus on the type of problems that can be solved and the conditions that are to be fulfilled so as to solve them. These languages incorporate the concepts of artificial intelligence, natural languages, expert systems, etc. Some of the 5GLs include Visual Basic, OPS5, Mercury, etc. The limitation of 5GL is that certain programs require the help of a human programmer to operate.

Programming languages have evolved over the years starting with the much less versatile first-generation languages (of machine language) to the presently popular fifth generation languages to deal with the most complex of problems being tackled. Exhibit 7.1 presents the evolution of programming languages

Exhibit 7.1: Evolution of Programming Languages

Programming languages have been evolved ever since Fortran (acronym for FORmula TRANslation) first appeared as a commercially available language. Business needs have driven this evolution with the present focus being popularity among programmers as well as suitability for web and mobile app development. A decision on which is a popular programming language has inherent element of subjectivity built into it but an objective criteria can be put together based on the language being open-source or not, prevailing trend, popularity of technology, career prospects etc. IEEE spectrum comes up with the list on ten most popular programming languages. For 2021, the list, spectrum ranking and important features of some are as follows:

1. Python (100)
2. C (99.7):
3. Java (99.5)
4. C++ (97.1)
5. C# (87.7)
6. R (87.7)
7. Javascript (85.6)
8. PHP (81.2)
9. Go (75.1)
10. Swift (73.7)

Python owes its immense rise in popularity to its very wide applications, being simple, extensible feature, having GUI programming support and being an object oriented portable language. Java too finds very wide application, is architecture neutral, and has excellent libraries and automatic memory management.

Contd. .....

C++ has gained immense popularity on account of its use for mobile app development besides being a general purpose, object oriented language having a rich library and memory management. C# on the other hand is mainly used on Windows to create apps (desktop apps, mobile apps, web apps, games etc). R, like Python finds great use among data scientists with adequate data handling and storage facility, graphical features and support for calculations on lists, arrays, vectors & matrices)

Source: https://becominghuman.ai/top-20-most-popular-programming-languages-for-2021-and-beyond-735ee8370c61

Activity: Rahul Malhotra (Rahul) was recently appointed as the professor to teach IT subjects at the Hyderabad-based ABC Institute of Computer Engineering. In the process of introducing the five generations of programming languages to the students, he said, "Among the five generations of programming languages, this hardware-friendly programming language uses names instead of numbers and is similar to machine language." Which programming language was Rahul talking about and explain some of its key characteristics? Explain briefly about the other programming languages that should be discussed by Rahul in the class.

Answer:

_____
_____
_____
_____
_____

Check Your Progress-3

16. _____ is the only language that the computer can understand without translation?

17. A symbolic program written by a programmer in assembly language is called a _____, which is translated into a machine code called _____.

    a. Object code, Source code
    b. Source code, Object code
    c. Pseudocode, Object code
    d. Object code, Pseudocode

18. Fourth generation programming languages (4GLs) include report generators and query languages. What are query languages?

19. _____ is a programming language which is also called as symbolic language while fourth generation languages are also called _____.

20. Which of the following statements is **false** regarding application generators?
    a. Application generators are higher level languages designed to facilitate development of a complete application.
    b. These languages are usually built around database management systems.
    c. Application generators are easier to use than other high-level programming languages such as C++.
    d. Application generators usually increase the time required for programming and testing.

21. In assembly language programming, abbreviations are used for program instructions. These abbreviations are referred to as _____.
    a. OPERAND
    b. OPCODE
    c. MNEMONIC
    d. COMPILER

7.6 Language Translators and Programming Languages

This section deals with the concept of language translators and then discusses some of the popular programming languages.

7.6.1 Language Translators

Language translators convert a source code into an object code. The two types of language translators are compilers and interpreters.

7.6.1.1 Compiler

A compiler is a translation program that translates the entire source code of a high-level programming language into an object code and creates an object file. This file contains instructions in a specific machine language. A source code is the code in which the programmers write programs and it has to undergo a series of steps to become an executable program. The first step in the process involves passing the source code through a compiler which translates the instructions given in the high-level language into the

object code. A liner or assembler program is then used for transforming the object code into an executable program that runs on a certain type of computer. The compiler checks the code for errors when it translates the code. In case errors are detected, the compiler identifies the error location and prints an output that highlights the location and likely cause of the error.

7.6.1.2 Interpreter

An interpreter is a program that executes the instructions that are written in a high-level language. It translates one line of the source code at a time and executes the translated instruction. Despite the fact that interpreters run programs more slowly than compilers, they are helpful tools for learning and easier for debugging. Moreover, interpreters do not undergo the compilation phase. Machine instructions are produced during this phase and if the program is lengthy, this process consumes a lot of time. An interpreter therefore, can be used for executing the high-level language programs immediately.

7.6.2 Programming Languages

Different types of programming languages cater to different types of needs. Some of the popularly used programming languages are:

7.6.2.1 Common Business Oriented Language (COBOL)

Developed in 1959, COBOL is one of the earliest high-level programming languages. However, due to incompatibility between the various versions that came out henceforth because of modifications and enhancements, ANSI developed the ANS COBOL, a standardized version of the language in 1968. Advanced versions of ANS COBOL were developed in 1974 and again in 1985. COBOL is useful for handling accounting information, including inventory control, billing, and payroll. Moreover, it was applied in the areas of finance, business, and administration by organizations and governments. This language requires programmers to write in a style which would explain what the program is doing at each step. Recently, COBOL has been upgraded to structured and object-oriented versions.

7.6.2.2 BASIC

BASIC (Beginner's All-Purpose Symbolic Instruction Code), developed in 1963 is an easy-to-use high-level programming language. Available in many versions, the language has wide applications. It is an interpreted language which allows the beginners to create a program in an interactive mode, run the program, test it, and debug it. It is taught in many high schools and colleges in elementary programming courses. However, professional programmers do not use BASIC since it is slower than

compiled programs. More recent versions of BASIC have incorporated the principles of structured, modular, and object-oriented programming, with added compilers.

7.6.2.3 Pascal

Developed by Professor Niklaus Wirth and named after French mathematician Blaise Pascal, Pascal is a very popular high-level programming language. Pascal was developed with the basic objective of creating a language that would be efficient and would have all the features that were not available in the other languages. Pascal is simple and expressive when compared to other languages and is hence used as a basis for teaching the concepts of computer programming. The language also encourages programmers to write well-structured programs. Pascal language is available in both interpreter and compiler versions.

7.6.2.4 C

Developed by AT&T's Bell Laboratories in 1972, C is a high-level programming language to be used along with the UNIX operating system. Later, the application of C has spread to other operating systems. The C programming language is an easy to learn, all-purpose, procedural language which blends the advantages of a high-level programming language with the efficiency of an assembly language. Since programmers are able to directly manipulate bits of data inside the processing unit, well-written C programs run very fast in comparison to the programs written in other high-level programming languages.

7.6.2.5 Ada

Ada, introduced in 1983, was named after August Ada Byron Lovelace (1815-1852), who was regarded as the first computer programmer and was also known to have helped Charles Babbage conceptualize the analytical engine. Ada is a high-level programming language which incorporates the principles of modular programming. Ada is suitable for controlling mission-critical systems. Till the year 1996, Ada was used for most US Defense Department projects.

7.6.2.6 C++

Developed in 1983 as an enhancement over the C programming language by Bell Laboratories, C++ is also an all-purpose programming language especially popular for its usage in object-oriented programming. It supports features like object-oriented programming, procedural programming, data abstraction, and generic programming. C++ is very popular among professional programmers.

7.6.2.7 Visual Basic

Introduced by Microsoft 1987, Visual Basic (VB) is one of the most widely used program development software across the world. It is the most successful Rapid Application Development (RAD) tool for Microsoft Windows applications. Visual Basic enables a programmer to develop an application quickly by designing the graphical user interface on-screen as the first step in program development. It also makes use of an event-driven programming model. In this model, the flow of program is determined by the actions initiated by the user or the messages received from other programs. Each on-screen control device (like a radio button or a text box) can be linked to a brief BASIC program that performs an action. The VB compiler automatically creates an executable program which can run on its own.

7.6.2.8 Java

Developed by Sun Microsystems, Java is the world's first truly cross-platform programming language, i.e., a language that makes possible the "write once, run anywhere" phenomenon. Java was designed based on the first version of C++. It excludes the inconsistent and confounding features of C and C++ languages and is an all-purpose language with numerous features that makes it suitable to be used on the World Wide Web. Java programs are designed to run on any type of computer that is commonly used. This capability is achieved through the Java Virtual Machine (JVM), which can process the compiled Java programs into machine language. JVM plays a central role in making Java portable. It creates a layer of abstraction between the compiled Java program and the underlying hardware platform and operating system.

Activity: Gurukul High School is about to introduce computer education for students from Class VII onwards. Which programming languages should the students be taught first, keeping in mind that they are totally new to this subject? What are the other programming languages which could probably be taught as the students move on to higher classes?

Answer:

7.7 Object-Oriented Programming – An Overview

In object-oriented programming, each program is developed as a component that can be used in more than one project. This enables component reusability. Component reusability is the capability to develop well-designed program modules that carry out a specific task. This feature enables the programmer to quickly develop a program without needing to develop the modules from scratch. In traditional programming, program and data are kept separate. Whenever the data is changed, the programs that access the data should also be modified. In object-oriented programming, the problem of updating data and programs is solved as there is no distinction between programs and data. Data is stored along with other programs and procedures required to access and use this data. If another program accesses the data, it immediately learns the procedures available, including any new method. Also, whenever a new object is added, the modules created need not be modified. The programmer has to create a new object that inherits the features of the existing objects. This process would eliminate the need to change the programs whenever there is a change in the data. Object-oriented programming is a more flexible system as the system objects change occasionally unlike the processes and procedures that vary too often. Java, .NET, etc are built on the object-oriented programming principles and programming languages like Java, C++, etc support object-oriented programming.

7.7.1 Basic Concepts in Object-Oriented Programming

Object-oriented programming has brought several benefits to software development, and integrates a lot of new concepts, development approaches, and vocabulary. It handles complex problems in a better way than the procedural approach. Following are some of the basic concepts of object-oriented programming.

7.7.1.1 Class

Most of the real world objects are similar in one way or the other. For example, there are many vehicles that exist in the entire world and most of them also turn out to be of the same model. These vehicles might have been built by using similar blueprints and therefore contain identical components. A vehicle is considered to be a class in object-oriented programming from which the individual objects are created. A class contains a set of objects that have a common structure and behavior.

7.7.1.2 Object

An object makes a request for an operation. This request invokes the operation that defines some action to be performed. Objects like book, car, television, etc have their own state, behavior, and these in turn define their relationships. All the objects that exist in the real world share the two

common characteristics of state and behavior. For instance, people have states like angry, happy, and hungry and these states are shown through behaviors of shouting, smiling, and eating respectively. It is therefore, important to identify the states and behaviors of the objects.

7.7.1.3 Abstraction

Abstraction is the process of hiding the complexity and internal functionality of objects.

7.7.1.4 Encapsulation

Encapsulation combines the code and the data it manipulates. It acts as a warehouse of the data and functions that are all organized in a single unit called class. It helps prevent arbitrary data access and prevents access of data to the outer world. Access is allowed only to the functions stored in the class. Due to encapsulation, the code can be accessed without bothering about the implementation details.

7.7.1.5 Polymorphism

It is the property of a function by which it can exhibit different behaviors based on different input values. Different sets of code in the function will be executed based on the inputs.

7.7.1.6 Inheritance

Inheritance refers to the process by which one object acquires the general properties such as the conditions and actions of other objects. There can be one or more levels of hierarchy. Inheritance helps in reuse of the code.

Check Your Progress-4

22. _____ is a translation program that translates the entire source code into an object code and creates an object file.

23. Which of the following statements is **true**?

    a. The interpreter translates one line of source code at a time.
    b. Interpreters are usually faster than compilers in running programs.
    c. Both interpreters and compilers involve the compilation phase.
    d. Compilers cannot be used to identify the error location.

24. What is Java?

25. In object-oriented programming, _____ refers to the process of hiding the complexity and internal functionality of the objects.
    a. Encapsulation
    b. Polymorphism
    c. Abstraction
    d. Inheritance

26. From the class 'vehicle', different categories of vehicles can be derived like scooter, car, bus, van, etc. All these share some common features of the generic class vehicle. This is an example of _____.

27. _____ helps bind together the code and the data that it manipulates.
    a. Encapsulation
    b. Abstraction
    c. Polymorphism
    d. Inheritance

7.8 Summary

- Software developers use the Program Development Lifecycle (PDLC) to define the program structure.
- PDLC consists of six phases. These include defining the problem, designing the program, coding the program, testing and debugging the program, formalizing the solution, and maintaining the program.
- Flow charts, decision tables and pseudocode are some of the program designing tools that help programmers to arrive at a logical flow.
- Computer programming languages can be classified into low-level languages and high-level languages. In a broader sense, the first generation and second generation programming languages are referred to as low-level languages and the third, fourth and fifth generation languages are referred to as high-level languages.
- Language translators convert a source code into an object code. The two types of language translators are compiler and interpreter.
- Different types of programming languages serve different purposes. Some of the programming languages which are popularly used include COBOL, BASIC, Pascal, C, Ada, C++, Visual Basic, and Java.

- Object-oriented programming is a programming technique designed to provide component reusability. Class, object, abstraction, encapsulation, polymorphism and inheritance are the basic concepts of object-oriented programming.

7.9 Glossary

- **Abstraction:** The process of hiding the complexity and internal functionality of objects.
- **Application generator:** A type of high-level language designed to facilitate the development of a complete application.
- **Assembly language:** A machine language that uses abbreviations, called mnemonics, to represent program instructions.
- **Beginner's All-Purpose Symbolic Instruction Code (BASIC):** A high-level programming language developed in 1963.
- **Branching structures:** These structures are used to transfer controls from one step to another in a program code.
- **COBOL:** Stands for Common Business Oriented Language. A COBOL is a third-generation language developed in 1959.
- **Compiler:** A translation program that translates the entire source code of a high-level programming language into an object code and creates an object file. This file contains instructions in a specific machine language.
- **Control structures:** A program does not always stick to a linear string of instructions. During the process, it may diverge, reiterate the code, or take decisions. Control structures are therefore used for specifying what the program needs to do, when, and under what conditions.
- **Decision tables:** These are used to outline a set of conditions that a computer program might encounter and the actions corresponding to them. Decision tables specify the true or false status of a particular condition and exhibit the dependency of the condition on the values of certain attributes. They also indicate the action to be performed for each condition. Presented in a tabular form, they state the problem clearly and concisely.
- **Encapsulation:** It binds together the code and the data it manipulates. It acts as a warehouse of the data and functions which are all organized into a single unit called class. It helps prevent arbitrary data access.
- **Fifth generation programming languages (5GLs):** These programming languages are designed in such a way so as to enable the computer to solve a problem. With 5GLs, the programmer can

concentrate on the type of problems that can be solved and the conditions to be met to solve those problems.

- **Flow chart:** A block diagram that outlines, step by step, the computational procedures necessary for accomplishing the processing requirements of a computer program. It is a planning document or 'blueprint' of a program's logic that specifies a processing plan.
- **Fourth generation languages (4GLs):** Non-procedural programming languages which do not require the programmer to specify procedures in detail, as required in a third-generation language. These languages are application specific languages and have commands that are more akin to English.
- **High-level language:** A procedural language that eliminates the need for programmers to understand the details of how a computer processes data. It enables programmers to create programs at a higher level of abstraction than machine or assembly languages.
- **Inheritance:** Inheritance refers to the process by which one object acquires the general properties such as the conditions and actions of other objects. There can be one or more levels of hierarchy. Inheritance helps in reuse of the code.
- **Interpreter:** A program that executes the instructions that are written in a high-level language. It translates one line of the source code at a time and executes the translated instruction. As the program is executed line-by-line, the programmer can see exactly what each line does.
- **Iteration logic:** It is used for looping when one or more instructions are required to be executed several times on the basis of the same condition. Iteration logic makes use of two structures called DO WHILE and REPEAT UNTIL.
- **Java Virtual Machine (JVM):** A program which can process the compiled Java programs into machine language. JVM plays a central role in making Java portable. It provides a layer of abstraction between the compiled Java program and the underlying hardware platform and operating system.
- **Java:** An object-oriented high-level programming language developed by Sun Microsystems. Java is the world's first truly cross-platform programming language, a language that makes possible the "write once, run anywhere" phenomenon.
- **Looping structures:** These structures are used in programs to perform repetitive tasks.

- **Looping:** In programming, a loop is a series of instructions that is repeated until a certain condition is met. Each pass through the loop is called an iteration.
- **Low-level languages:** These languages are hardware friendly. That is, they are closer to the hardware than to the human languages. The machine and assembly languages are low-level languages.
- **Machine language:** A programming language that consists of binary numbers (0s and 1s) that directly correspond to the on and off electrical pulses generated by the computer. Also called the first generation programming language (1GL), the machine language is the only language that the computer can understand without translation.
- **Non-procedural language:** A type of language in which the user requests the desired information but does not give a procedure for obtaining it.
- **Polymorphism:** The property of a function by which it can exhibit different behaviors based on different input values. Different sets of code in the function will be executed based on the inputs.
- **Procedural language:** A type of language for which the user provides instructions to the system for performing a sequence of operations on the database.
- **Program Design Language (PDL):** Instead of using symbols to describe the logical steps of a program (as in flow charting), the pseudocode uses a structure that resembles computer instructions. As it lays emphasis on the design of the program, the pseudocode is also called Program Design Language (PDL).
- **Program Development Life Cycle (PDLC):** An organized plan for breaking down the task of program development into manageable units called modules. Each of these modules must be successfully completed before moving on to the next phase.
- **Programming language:** The vocabulary and set of grammatical rules used for instructing a computer to perform specific tasks. With the help of a computer language, a program tells a computer what needs to be done.
- **Pseudocode:** A tool used for planning program logic. Pseudocode means “imitation code.” It is a type of structured English that is used for describing algorithms.
- **Sequence logic/decision logic:** A pseudocode used for performing instructions one after another, in a sequence. They are written in the order (or sequence) in which they are to be performed.

- **Source code:** The code in which the programmers write programs. This code has to go through a series of steps to become an executable program.
- **Syntax:** A set of language rules that the programmer should follow while writing programs. All computer languages have a vocabulary of their own. The symbols of a particular language must be used according to the set rules, which are known as the syntax of the language.

7.10 Self-Assessment Test

1. The Program Development Life Cycle (PDLC), which was developed during the early 1970s, provides an organized plan for breaking down the task of program development into manageable units called modules. In this context, discuss the different stages in PDLC.
2. Program designing tools help programmers to arrive at a logical flow. Compare the advantages and disadvantages of the various program designing tools.
3. Programming languages are of different types and they cater to different types of requirements. There are some programming languages like BASIC, COBOL, etc., which are very popularly used. Give a brief description of the popular programming languages.

7.11 Suggested Readings / Reference Material

1. Introduction to Information Technology, V. Rajaraman, PHI learning, 2018
2. Information Technology for Management, 2ed: Advancing Sustainable, Profitable BusinessGrowth, Turban, Volonino, Wood, O.P. Wali, Wiley India Pvt Limited, January 2021
3. Introduction to Information Systems - 6th edition, R. Kelly Rainer; John Wiley & Sons, Inc.2016
4. Information Technology: An Introduction for Today's Digital World , Richard Fox, Chapman and Hall/CRC; 2nd edition (August 21, 2020)
5. Information Technology for Management, Efraim Turban, Carol Pollard, Gregory Wood, Wiley, 2018.

Additional References:

1. Critchley, L., Where Nanotechnology, the IoT, and Industry 4.0 Meet., https://www.mouser.com/blog/where-nanotechnology-the-iot-and-industry-40-meet, 2019

2. Pan India implementation of HMIS over Indian Railways, Ministry of Railways., http://railministry.com/pan-india-implementation-of-hmis-over-indian-railways/ 2020
3. Vossler, C. How Long Does It Take To Order A New BMW?https:// www.bmwblog.com/2020/09/28/how-long-does-it-take-to-order-a-new-bmw/2020
4. Jay, A., 10 New ERP Trends & Forecasts for 2020/2021 – A Look Into What's Next. https://financesonline.com/erp-trends/2019
5. Gingiss, D., How Integrating Social Media Into The Rest Of The Business Will Increase Revenue., How Integrating Social Media Into The Rest Of The Business Will Increase Revenue (forbes.com), 2019

7.12 Answers to Check Your Progress Questions

Following are the answers to the Check Your Progress questions given in the unit.

1. **Programmers develop program designs** from program specifications. In the program designing stage, the first step involves identifying the main routine that is the program's major activity. Then, programmers break down various components of the main routine into smaller sub-routines or modules, until each sub-routine is highly focused and accomplishes only one single task.

2. **Defining the problem**

Defining the problem is the first step in the process of developing a software program. In this stage, the system analysts give the programmers their program specifications. These specifications define the input data, the processing that needs to be done, the output features, and the user interface design. Depending on the size of the job, the program development may then be handled by a single individual or a team of programmers.

3. Manuals are developed to enable users to understand how to operate a program. They are developed to explain thoroughly how the program works. These manuals provide an overview of the program's functionality, tutorials for new users, in-depth explanations of major program features, reference documentation of all program commands and a detailed description of error messages generated by the program.

4. The third stage in the Program Development Life Cycle (PDLC) is coding the program. Translating the algorithm into specific program language instructions is called coding. An appropriate programming language should be chosen and the program should be created by typing the code. Programmers must follow the rules of the language (syntax) carefully. These rules specify how certain operations should be expressed to arrive at appropriate results.

5. (b) Debugging

Testing and debugging of programs is done to detect and eliminate syntactical and logical errors in the program code. Syntactical errors arise if there are errors in the syntax while the program is being written. Logical errors arise when the programmer makes mistakes while designing the program. These errors are called bugs and the process of elimination is called debugging.

6. (a) Structured programming

In designing the program stage of the Program Development Life Cycle (PDLC), programmers develop program designs that elaborately specify the path of program execution. In this step, programmers identify the main goal of the program (output) and then break the program into manageable modules, each of which contributes to goal achievement. This approach is called structured programming or top-down program design.

7. v-ii-i-iv-iii

The Program Development Lifecycle (PDLC) provides an organized plan for breaking down the task of program development into manageable units called modules. Each module must be successfully completed before the programmer moves on to the next phase. Various steps involved in the PDLC are: defining the problem, designing the program, coding the program, testing and debugging the program, formalizing the solution, and maintaining the program.

8. Program design tools help develop a program. The various program design tools are flow charts, pseudocodes, decision tables, etc. A flow chart is a block diagram that outlines, step-by-step, the computational procedures necessary for accomplishing processing requirements of a computer program. Pseudocode is a type of structured English that is used for describing algorithms. It is the summarized form of a program written in a form that can easily be converted into real programming statements. Decision tables are used to outline a set of conditions that a computer program might encounter.

9. (b) Algorithm

Programmers develop algorithms to solve problems. An algorithm is a step-by-step description of a method for performing a task or arriving at a solution. It is a sequence of instructions designed so that if the instructions are executed in the specified sequence, the desired results will be obtained.

10. (b) The logic of a program flow chart should flow from bottom to top and right to left

The American National Standards Institute (ANSI) has developed a number of rules and guidelines to standardize the flow-charting process. According to ANSI, the logic of a program flow chart should flow from top to bottom and left to right, and not as mentioned in option (b).

11. (c) American National Standards Institute

The American National Standards Institute (ANSI) has developed a number of rules and guidelines to standardize the flow-charting process.

12. Decision tables are used to outline a set of conditions that a computer program might encounter. These tables specify the true or false status of a particular condition and exhibit the dependency of the condition on the values of certain attributes. They also indicate the action to be performed for each condition. Decision tables are used when large number of decisions have to be taken, when a lot of branching is involved, and when modeling of program logic is more complicated. They are also used in areas like programming, testing, business analysis, hardware designing, etc.

13. (c) Decision tables

A decision table consists of rows and columns divided into six parts or quadrants. The first part contains conditions. The condition stub describes conditions that exist in the program. The action stub contains action statements that correspond to respective conditions. The conditions and actions are also called causes and effects, respectively.

14. Iteration logic

Iteration logic is used for looping when one or more instructions are required to be executed several times on the basis of the same

condition. It uses two structures called DO … WHILE and REPEAT … UNTIL.

15. (d) Decision tables are better able to express total sequence of events needed to solve a problem than flow charts.

Decision tables are easier to draw and edit compared to flow charts. They provide more compact documentation. A small table can replace several pages of charts. It is easier to follow a particular path down one column than through several flow chart pages. Flow charts are better able to express the total sequence of events needed to solve a problem than decision tables. Flow charts are mostly preferred by programmers, especially beginners.

16. Machine language

Machine language consists of binary numbers (0s and 1s) that directly correspond to the computer's electrical pulses. It is also called the first generation programming language (1GL). The machine language is the only language that the computer can understand without translation. These languages are difficult to comprehend by humans as they are written in numbers.

17. (b) Source code, Object code

Assembly language is also called symbolic language. Assembly code is a program written in symbolic language using symbols instead of numbers. An assembler translates an assembly code into a machine code. A symbolic program written by a programmer in assembly language is called a source code. It is translated into a machine code called object code.

18. Query languages are languages that are used for getting information out of databases.

19. Assembly language; non-procedural languages

Assembly language is also called symbolic language. Assembly code is a program written in symbolic language. Here, symbols are used instead of numbers. An assembler translates an assembly code into a machine code. Fourth generation languages (4GLs) are called non-procedural languages as they do not require the programmer to specify procedures in detail.

20. (d) Application generators usually increase the time required for programming and testing.

An application generator is a type of language designed to facilitate development of a complete application. It includes defining input transactions, editing transactions, creating a database, updating files, generating reports and processing queries. They drastically reduce the time required for programming and testing.

21. (c) MNEMONIC

In assembly language programming, mnemonics are abbreviations used for program instructions.

22. Compiler

Language translators convert a source code into an object code. Compilers and interpreters are two types of translators. A compiler is a translation program that translates the entire source code into an object code and creates an object file. This file contains instructions in a specific machine language.

23. (a) The interpreter translates one line of source code at a time.

An interpreter is a program that executes the instructions that are written in a high-level language. It translates one line of the source code at a time and executes the translated instruction. As the program is executed line-by-line, the programmer can see exactly what each line does. Though interpreters run programs more slowly than compilers, they are helpful tools for learning and easier for debugging. Another advantage of interpreters is that they do not undergo the compilation phase. (During compilation, if any errors are found, the compiler identifies the error location and prints a program that highlights the location and likely cause of the error.)

24. Java is an object-oriented high-level programming language. It is the world's first truly cross-platform programming language, that makes possible the "write once, run anywhere" phenomenon. It is an all-purpose language with numerous features that makes it suitable for use on the World Wide Web. Java programs are designed to run on any type of commonly-used computer.

25. (c) Abstraction

Abstraction refers to the process of hiding the complexity and internal functionality of the objects.

26. Inheritance

Inheritance refers to the process by which one object acquires general properties such as the conditions and actions of other objects. For instance, all categories of vehicles like scooter, car, bus, etc share common features of the generic class of vehicles. All these vehicles can be used for transporting people and/or goods from one place to another.

27. (a) Encapsulation

Encapsulation binds together the code and data it manipulates. It helps prevent arbitrary data access.

Unit 8

Database Management Systems

Structure

8.1 Introduction

In the previous unit, we discussed program design, development life cycle, program design tools, object oriented programming. We have also discussed in the previous units, that storage of information is one of the most important operations carried out by a computer. Earlier, organizations stored data in a file-processing system, where records were stored in files. The file-processing system had certain problems related to data duplication, data access, and other security related issues. These difficulties led to the evolution of the Database Management System (DBMS). The main goal of a DBMS was to make easy the storage and retrieval of information, while providing security, preventing unauthorized access to information, and avoiding data loss due to a system failure.

In this unit, we introduce you to the database system, the various data models, and the database management approach. We would also be discussing about the entity-relationship and the relational database models.

Then, we would move on to discuss the concepts of data dictionary, data warehousing, data mining, and distributed database. Finally, we would be ending this unit with a discussion about the object-oriented databases and the object relational database system.

8.2 Objectives

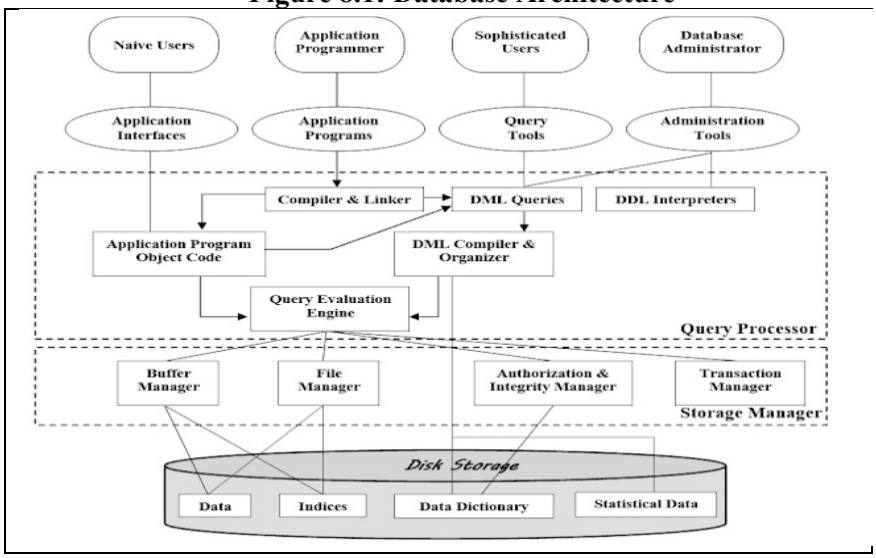
By the end of this unit, you should be able to:

- Name the components of a database system.
- Identify the types of data models.
- Recall the approaches for database management.
- Explain the entity-relationship model.
- Explain the relational database model.
- Define the need for data dictionary.
- Discuss the concept of data warehousing.
- Discuss the concept of data mining.
- Determine the need for distributed databases.
- Determine the various applications of object-oriented databases.
- Explain the structure of object relational database system.

8.3 Database System

A database is a set of data organized efficiently in a central location so that it can serve a number of information system applications. Figure 8.1 presents a database architecture.

Figure 8.1: Database Architecture



Source: ICFAI Research Center

A database is a data structure that stores organized information. Most databases contain multiple tables, which may each include several different fields. For example, a company database may include tables for

products, employees, and financial records. Each of these tables would have different fields that are relevant to the information stored in the table. A Database system is partitioned into modules that deal with each of the responsibilities of the overall system. The functional components of a database system can be broadly divided into the storage manager and query processor components. The database is used by different types of users like naive users, application programmers, sophisticated users and database administrators. Naive users are unsophisticated users who interact with the system by invoking one of the application programs that have been written previously. For example, a clerk in the university who needs to add a new instructor to a department might invoke a program called new hire. Application programmers are computer professionals who writes application programs. Application programmers can choose from many tools to develop applications which other users of the database can use to interact with the database for specific purposes. Sophisticated users interact with the system without writing programs. Instead, they form their requests either using a database query language or by using tools such as data analysis software. Sophisticated users may also write specialized database applications that do not fit into the traditional data-processing framework. One of the main reasons for using DBMSs is to have central control of both the data and the programs that access those data. A person who has such central control over the system is called a Database Administrator (DBA). The DBA is responsible for the everyday working, maintenance, and updating of the system. Hence, it is very important to understand the key roles and responsibilities of a Database Administrator.

The roles of a DBA include

- Schema definition
- Storage structure and access-method definition
- Schema and physical-organization modification
- Granting of authorization for data access
- Routine maintenance and activities, that includes
- Data backup
- Data recovery
- Database optimization
- Maintaining database users

The Query Processor is important because it helps the database system simplify and facilitate access to data. In a relational database system the query processor is the module responsible for executing database queries. The query processor receives as input queries in the form of SQL text,

parses and optimizes them, and completes their execution by employing specific data access methods and database operator implementations. The query processor communicates with the storage engine, which reads and writes data from the disk, manages records, controls concurrency, and maintains log files. High level views help to achieve this goal; with them, users of the system are not be burdened unnecessarily with the physical details of the implementation of the system. However, quick processing of updates and queries is important. It is the job of the database system to translate updates and queries written in a nonprocedural language, at the logical level, into an efficient sequence of operations at the physical level. The main components of the query processor are DDL interpreter which interprets DDL statements and records the definitions in the data dictionary, the DML compiler- DML translation and query optimization and the Query evaluation engine which executes low level instructions generated by DML compiler.

The function of the storage manager is to move data between disk storage and the main memory of computers. Databases typically require a large amount of storage space. Corporate databases usually require huge amounts of storage, all this cannot be accomplished only with the main memory. Secondary storage devices like hard disks have to be used. That is where the storage manager comes in, to quickly retrieve the required data from disk storage into main memory for processing and storing the processed files back into disk storage. Data are moved between disk storage and main memory as needed. Since the movement of data to and from disk is slow relative to the speed of the central processing unit, it is imperative that the database system structure the data so as to minimize the need to move data between disk and main memory. The storage manager is a program module that provides the interface between the low level data stored in the database and the application programs and queries submitted to the system. It is responsible for the interaction with the file manager. Raw data are stored in disks using the file system provided by the conventional operating systems. The storage manager is responsible for storing, retrieving and updating data in a database. The major components of the storage manager are the file manager, the buffer manager, authorization and integrity manager and transaction manager. The file manager is responsible for the allocation of space on disk storage and specifying the data structures that are used to represent data stored in disks. The buffer manager is responsible for fetching data from disk storage into main memory. It decided which data are to be cached in main memory. The buffer manager is a critical part of the database system. The buffer manager enables the database to handle data sizes that are much larger than the main memory size. Authorization and integrity manager checks for the

satisfaction of integrity constrains on the data. These constraints are put in place to avoid errors in the data. Authorization and integrity manager also checks to see whether the user has enough authority or privileges to access data. Transaction manager is responsible for ensuring that the database remains in a consistent and correct state, despite system failures. Transaction manager makes sure that concurrent transaction executions proceed without any conflicts. The storage manager implements several data structures as part of the physical system implementation like data files, data dictionary and indices. Data files store the different files in the database. The data dictionary stores metadata about the structure of the database. It specifies the particular schema of the database. Metadata can be thought of as data about data. Indices are used to provide fast access to data items. The location of every file stored in the database can be accessed from the index, which makes data retrieval fast and efficient.

A database is a set of data organized efficiently in a central location so that it can serve a number of information system applications.

8.3.1 Data Hierarchy

The data organized systematically in a sequential order is referred to as data hierarchy.

The data can be logically organized into characters, fields, records, files, and databases. They are discussed below:

- Character: A character can be a number, symbol or an alphabet and is the basic element that can be manipulated.
- Field: A group of characters form a field. For example, the grouping of the characters in a person's name forms a name field, and the grouping of numbers in a sales figure forms a sales amount field.
- Record: The related data fields are grouped together to form a record. For example, a labor record consists of name, hours worked, and pay. Records are of two types: fixed-length having limited data field size and variable-length having variable data field size.
- File: Related records constitute a file. For example, a labor record will have all information related to the laborers. Files may be temporary or permanent.
- Database: The logical collection of records and files is referred to as a database. For example, all employee details are grouped together to form an employee database.

8.3.2 Components of a Database System

The five major components are data, hardware, software, procedures, and users. They are discussed below:

8.3.2.1 Data: Data refers to raw facts and figures. For example, the labor details such as name, age, address, and salary are all raw facts or data. The data is stored in a repository called the database. Database is of two types: integrated and shared. Integrated database is a collection of related tables. If a table in a database can be used by other tables in the database, it is called as a shared database.

8.3.2.2 Hardware: The physical devices such as control units, hard disks, etc., and other related components form the hardware components of a system.

8.3.2.3 Software: Software is the interface between the hardware and the user. The software components include the applications software, operating system, utility software, and the DBMS. The user's requests for database access are handled by the DBMS as it acts as an interface between the user and the hardware.

Procedures: The series of steps that indicate ways to carry out tasks is referred to as a procedure. They are instructions and rules that govern the design and use of a database.

8.3.2.4 Users: The most important component of a system is a user. In a database system, users can be classified as application programmers, database administrators, and end users.

- **Application programmers:** They are given the task of writing application programs in languages such as COBOL, C, or Pascal. These application programs can retrieve, insert, and delete information in the database.
- **Database administrators:** The database administrators or DBAs help in implementing an organization's strategic and policy decisions by providing the necessary technical support. The DBAs create the database and are responsible for the performance of the system. They provide necessary data to the users and take care of database backup and recovery functions.
- **End users:** They interact with the database through an interface such as a software or online applications. For example, end users use Query by Example (QBE), a database query language, for interacting with the database system.

8.3.3 File Organization

A. The logical sequence of records organized in files is referred to as file organization. Files are organized based on parameters like storage efficiency, maintenance, reliability, ease of access, and update. Various types of file organization are discussed below:

- **Pile:** This is the simplest form of file organization, where the data is collected in the same order it arrived. There is no specified format or structure to save the records. It also has limitations such as data retrieval is not easy, it can be used only for temporary storage, and requires an exhaustive search in order to access the records.
- **Indexed:** Multiple indexes are used in an indexed file for different key fields. All index files are updated when a new record is added to the main file. The records in an indexed file may vary in length. Indexed files are used where timeliness of data is a critical issue.
- **Sequential:** In a sequential file organization, data is organized in the order in which data is entered into the file. All the files in a sequential file are of same length or fixed format. A unique key field, like a student ID, is used to identify a particular record. The main disadvantage of a sequential file is that it is difficult to enter new records. Sequential access (or serial access) is used for accessing databases in which the entire information has to be read. This is useful in batch processing, when most of the records in the file need to be processed at the same time.
- **Indexed-sequential:** The disadvantages of sequential file organization are overcome using an indexed-sequential file. The key field is used for organizing the records. An indexed-sequential file supports random access of records and has a search capability that searches the desired record.
- **Hashed:** Also called as direct file organization, the hash file contains records with a unique address called a hash value. A hashing algorithm is used to develop the hash value. The hash file contains a key field for each record on which hashing is done in order to get the location of the desired record. The hashed files are used in fixed length records and in files which require rapid access. File access from direct files is known as direct access, machine access, or random access.

Check Your Progress-1

1. What is a field?
2. Which of the following statements is **true**?
   a. A group of related records constitutes a field.
   b. Related files of data are grouped to form a record.
   c. A file is a collection of related databases.
   d. An identification field in a record is called a key field.

3. Which of the following is **true** regarding indexed-sequential file organization?
   i. Records are organized on the basis of a key field.
   ii. The index supports a random access of records.
   iii. The index provides a quick search capability that allows the user to reach the vicinity of the desired record rapidly.
   a. Only i and ii
   b. Only ii and iii
   c. Only i and iii
   d. i, ii, and iii
4. Which type of file organization is used when rapid access is required and the records are of fixed length?
5. A customer's name, account number, and account balance together constitute a:
   a. Character
   b. Foreign key
   c. Database
   d. Record

8.4 Data Models

Data models constitute a set of conceptual tools that describe data. These models specify the logical structure of a database. Data models are of three types. They are: record-based logical models, object-based logical models, and physical data models.

8.4.1 Record-Based Logical Models (RBLMs)

Data is described at the logical and view levels using RBLMs. The data stored in the database and its interrelationships are described using the logical level. In view level, a user sees a set of application programs, where the details of data types are hidden. The logical structure of the database and the description of its implementation are specified using an RBLM. It includes various models as follows:

- **Relational model:** A common key field is used to related date in several files in a relational model. The records in the file have the same key field with unique key field contents. The fields help in identifying the records. The relational model uses a two-dimensional table that can be used to represent data and relationships. The row in the table corresponds to a record while column corresponds to a field. A primary key, i.e., a unique key, is used to access records in a table. A relational database structure can link an employee file and a salary file by using a common field, such as an employee. The user can then request a report consisting of fields from both the employee and salary records.

- **Network model:** In this data model, data is represented by a collection of records and the relationships among data are represented by pointers or links. Pointers store the location or address of files and records. Network models may be suitable for handling voluminous transactions but not for handling ad hoc requests. Network databases enable many-to-many relationships.
- **Structured Query Language (SQL):** SQL (pronounced as Sequel) is a popular query language. It is a non-procedural query language. SQL has become a de facto standard for query languages because of its wide usage and recognition in the industry. It is a standard language used for relational databases. It includes both data definition and data manipulation operations.
- **Hierarchical model:** In a hierarchical model, the records are organized as collection of trees. This model handles large transaction-based applications such as inventory, check processing, file maintenance, sales-order processing, etc. An advantage of this model is that it can handle large number of structured transactions. These applications are large, have well-defined master file of information, and have a large number of well-structured transactions that are repeatedly processed.

8.4.2 Object-Based Logical Models (OBLMs)

In an OBLM, the user can specify the data constraints explicitly. Data can be described at the logical and the view levels. The entity-relationship (E-R) model and the object-oriented model fall under OBLMs. Other models include binary model, semantic data model, infological model, and the functional data model.

8.4.3 Physical Data Models

Data is described at the physical level using physical data in the form of bytes and words. These data models design the internal schema of a database. Schema describes the overall design of a database. The physical data models are also used for data normalization, identifying tables, columns and relationships, etc. Unifying model and frame memory model are some of the physical data models.

Check Your Progress-2

6. Which data model covers the Entity-Relationship (E-R) model that is based on the perception of a real world that consists of a collection of entities and their inter-relationships?
7. _____ is/are used in describing data at the logical level and at the view levels.
   i. Record-based logical models
   ii. Physical data models

iii. Object-based logical models

a. Only I and ii
b. Only ii and iii
c. Only I and iii
d. I, ii, and iii

8. What are record-based logical models (RBLMs)?
9. In a network model, _____ are used to store the location or address of files and records.
   a. Instances
   b. Schema
   c. Pointers
   d. Key fields
10. A tree-like structure of records in a database is seen in a _____ model while a tabular structure of records is present in a _____ model.
11. In a relational database, each column in a table represents
    a. Records/ relations
    b. Fields/ attributes
    c. Primary keys/ Foreign keys
    d. None of the above
12. What is SQL?
13. Which of the following is **false** about relational data model?
    a. In a relational database, data in several tables is related through the use of a common key field.
    b. In a relational database, a primary key is used to distinguish a record from other records in the table
    c. A relational DBMS is generally a three-dimensional table.
    d. As a relational DBMS can work with many files and relate all the data, many databases can be created and focused on different themes.

8.5 Database Management Approach

In a DBMS, the data records are grouped into a database for making it accessible by various application programs. A DBMS is an interface between a user and the database and allows users to create, store, retrieve, delete, organize, and maintain information in a database.

In order to reflect new transactions, the DBMS should be frequently updated and maintained. The DBMS allows the user to inquire about data stored in the database. This enables the user to obtain quick response to the required information and can also generate reports.

The DBMS helps in managing huge volumes of interrelated data effectively. DBMS prevents data redundancy. Different applications in banking, such as savings accounts, credit cards, cheque processing, and loan accounting are done on a regular basis. The customer data is consolidated into a common customer database. The creation, maintenance, and use of database systems in an organization are taken care of by the DBMS software. The main uses of database management software are database definition and application development, database maintenance, and database interrogation. These are discussed below:

- **Database definition and application development:** The database systems are maintained with the help of DBAs or database specialists. The most common databases are Microsoft SQL Server, Sybase, Oracle, and DB2 and use the *DDL (Data Definition Language)* commands that specify the modification to the database, relationship, data content, and structure. The information is stored in the data dictionary. The data dictionary consists of files present in the database, the number of records in the file, and the name and type of each field in the records. A data dictionary helps a database to access relevant files. A DBA requests the data dictionary with a query to retrieve details regarding files in the database. The active data dictionary (data dictionary currently in use) includes some standards for database operations. For example, the active data dictionary will not permit users to enter values that exceed the length specified for a field. The DBMS provides built-in software development tools or internal fourth generation programming languages to application developers for building custom applications. For example, application developers use DBMS to generate reports, forms, etc.
- **Database maintenance:** The database system needs to be maintained effectively. It includes procedures for adding, updating and deleting records. In order to store the latest and accurate data in the database, a database should be maintained. *DMLs (Data Manipulation Languages)* are used that help in modifying a database. Transaction processing programs and other end user packages, supported by DBMS help in maintaining the database.
- **Database interrogation:** Query languages and report generators are used for retrieving information from the database. Query languages support ad hoc querying of the database and are non-procedural in nature. *SQL (Structured Query Language)* is the most popular query language. Another query language used in database management packages is *QBE (Query By Example*).

- **Graphical and natural queries:** Databases can be accessed using simple *GUI (Graphical User Interface)*. It allows users to use features like point and click that are translated into queries by the software. This is easier to use compared to SQL.

8.5.1 Characteristics of a Well-designed Database System

A database should ensure efficiency, storage, retrieval, and ease of data access. In addition to this, a database system ensures data integrity, data independence, and data security. A well designed database should include the following features:

- **Data integrity:** The ability to consistently manipulate and navigate tables in the database is referred to as data integrity. The data should not be affected by data transmission errors, hardware malfunctions, and typographical errors. Certain programs that use data validation procedures should be used to notify the user in case of invalid input entry.
- **Data independence:** The data should be stored in such a way that a modification in the data structure or any change in the application program should not affect the data. This is referred to as data independence.
- **Prevention of data redundancy and inconsistency:** The database should be updated regularly without any redundancy. For example, an employee name may appear in two different databases in an organization. Any change in the name or address modified in one location should automatically be updated in another location. If updation does not take place, it leads to data inconsistency.
- **Data sharing and security:** The database should be designed in a way that makes sharing of data easy for users. The data should also support new applications in addition to the existing ones. To ensure data security, the data stored in the database should be accessible only to users authorized to access the system. The database should be protected from power or equipment failure by taking backup regularly.

8.5.2 Types of Databases

The databases are divided into several categories based on their use in an organization. Some of these databases are discussed below.

- **Operational databases:** The information related to the day-to-day operations in an organization is referred to as operational databases. They are also called subject area databases, transaction databases, or production databases. It stores information related to employees, inventory, customers, etc.

- **Analytical databases:** The data extracted from the operational and external databases are stored in analytical database. The information stored is vital for managers and the end users. They are also called as management databases, information databases, or multidimensional databases. Executive Information Systems, Online Analytical Processing, and Decision Support Systems use analytical databases.
- **Hypermedia databases:** The homepages and hyperlinked pages of multimedia are stored in a repository called hypermedia database. They store video, photo, graphics, and text. The HTML (Hyper Text Markup Language) files and GIF (Graphics Interchange Format) files stored in hypermedia databases are used by the World Wide Web.
- **Data warehousing:** The data extracted from an operational database is stored in a data warehouse. The data stored is edited, integrated, and updated regularly so that it can be used in market research, business analysis and decision support.
- **Distributed databases:** In a distributed database, the data is stored on many systems that are dispersed geographically and are connected through a network or a link. A distributed database may include information from operational or analytical databases, hypermedia or discussion databases, or any other type of database.

8.5.3 Database Development

Effective data management can be ensured by developing databases. Their creation involves a series of processes. Broadly, the database development process is divided into the following phases – database planning and modeling, database design, and finally, database implementation and maintenance.

Database planning and modeling: First, there is an investigation to identify the data types and quality for incorporation into a database. In the data planning process, an enterprise model is designed by the DBA. The model defines the primary business processes in the organization. During the process of database planning, both executives and users are consulted. In the next step, the key data elements in business processes and the relationships that exist between these entities are identified by the end users. An Entity Relationship (ER) diagram is then designed. A conceptual data model is then formed. This model accounts for all the basic facts and constraints under which the database will have to operate and it is particularly concerned with relationships between different entities and their attributes.

Database design: In this phase, a design for the DBMS is created, to be used for database implementation. This phase involves conceptual, logical, and physical design of the database. In the conceptual design, all information is represented in the form of high-level models. Logical designing transforms the conceptual model into a data model. The physical design determines the data storage structures and access methods.

Database implementation and maintenance: In database implementation, the hardware, software, the file sizes, and their organization, are decided upon. The actual data is also created in this phase. Database maintenance involves continuous monitoring and evaluation of the performance of the database, followed by enhancements to the earlier proposed design in order to fine-tune the performance.

8.5.4 Managing Data

The data in an organization is managed through a process called data resource management. Data resource management is also called as data management, or enterprise data management.

The Data Management Association (DAMA), an international organization, defines data resource management as, "the development and execution of architectures, policies, practices and procedures that properly manage the full data lifecycle needs of an enterprise."

With the help of database management, data resources can be properly managed. However, some problems may arise from the use of a DBMS such as huge installation costs, backup and recovery, hiring of specialized staff, etc. These are: (1) data administration, (2) data planning, and (3) database administration.

- **Data administration:** Data administration establishes and enforces policies and procedures for managing data. It manages, collects, stores and disseminates the data in away that makes that data as a standard resource that can be used by all users in the organization. Data administration controls data flow in an organization.
- **Data planning:** It is a corporate planning and analysis function that focuses on data resource management. For the data resources of a firm, data planning develops a data architecture. This works in conjunction with the strategic mission and plans of the organization.
- **Database administration:** This function ensures that the database management technology is put to effective use. It involves physical and logical design, developing and maintaining a data dictionary, monitoring and enhancing the performance of databases, and enforcing standards for database use and security.

8.5.5 Benefits and Limitations of DBMS

A DBMS has certain benefits and limitations. They are as follows:

8.5.5.1 Benefits: A DBMS reduces redundancy of data and improves concurrency. The data is integrated in order to make it accessible by multiple users. With the help of simple queries, information can be obtained from the DBMS. It reduces the maintenance costs and data security and reliability is ensured.

Limitations: A DBMS is expensive and complex in nature. The costs involved in data maintenance are low; however, DBMS requires frequent upgradation of the hardware components, thus resulting in increase in costs. The maintenance of centralized databases may lead to increase in the possibility of errors and failures.

8.5.5.2 Data Views

Data views allow users to access and modify files without knowing the physical location of the data. The types of data views include data abstraction and instances and schemas.

Data abstraction: The DBMS provides users an abstract view of the database system. Certain details are hidden from the users as to how the data is created, stored, and maintained. It does so by defining three levels of abstraction – the physical level, the logical level, and the view level.

- **Physical level:** The physical level is the lowest level of abstraction that describes the manner in which data is actually stored. In this level, complex low-level data structures are described in detail.
- **Logical level:** Also known as conceptual level, it is the next level of abstraction. The type of data stored and the interrelationships between them are described in this level. Here a record is described by a type definition and the interrelationships that exist between the record types are also described. Programmers and DBAs work at this level of abstraction.
- **View level:** This is the highest level of abstraction**.** At the view level, the users see a set of application programs that hide details of the data types. In addition to this, the views also provide a security mechanism to prevent users from accessing parts of the database.

Instances and schema: The information currently stored in the database is called an instance of the database. The overall design of the database is called schema. Database systems have several schemas, partitioned according to the levels of abstraction.

Check Your Progress-3

14. What are data manipulation languages (DMLs) used for?

15. Match the following:

i. Data integrity	p. Procedures for adding, updating and deleting records
ii. Data independence	q. Presence of correct and consistent data
iii. Database maintenance	r. The storage of data in such a way that it is not affected by changes in the application program

16. The information currently stored in the database is called _____ of the database.
    a. An instance
    b. The schema
    c. A model
    d. An abstraction

17. Match the following levels of abstraction in database management with their corresponding features.

i. Physical level	p. Simplified access for users to selected portions of the database
ii. Logical level	q. Conceptual description of the type of data and their interrelationships
iii. View level	r. Manner in which data is actually stored

18. Explain data independence.

19. The Database Management System (DBMS)
    i. Serves as an interface between users and the various databases
    ii. Helps the users to access data easily
    iii. Enables the programmers and administrators to create, delete, interrogate, and maintain databases to provide requisite information to the end-users and their organizations

a. Only i and ii
b. Only ii and iii
c. Only i and iii
d. i, ii, and iii

8.6 The Entity-Relationship Model

The Entity-Relationship (E-R) model was proposed by Peter Pin-Shan Chen in the year 1976. The E-R model was based on the perception of a real world that consists of a collection of basic objects called entities and their interrelationships (association). An entity may be a person, a book, or something abstract such as a loan.

The E-R model is used in the field of computing and for organizing data within databases or information systems. The E-R model consists of a component called the E-R diagram which is used for making a graphical or visual representation of the entities and their interrelationships.

8.6.1 Basic Elements in the E-R Model

The basic elements or constructs in the E-R model are entities, relationships, and attributes.

Entities: An entity consists of basic objects or concepts about which data is collected and stored. Employees, invoices, etc., are instances of entities. An entity can be independent or dependent. They are also referred to as strong and weak, respectively. An independent entity is one which does not depend upon another entity for identification, while a dependent entity depends upon another entity for identification. The occurrence of an entity, also known as the instance of an entity, is an individual occurrence of an entity. In the relational model, an entity is equivalent to a table, an occurrence is equivalent to a row, while an attribute is equivalent to a column.

There are also special entity types like associative entities and subtypes entities. Associative entities or intersection entities are used for associating or relating two or more entities to form a many-to-many relationship. Subtypes entities represent a subset of occurrences or instances of the parent entity which is called as a supertype or a generic entity. The subtypes entities have attributes or relationships that apply only to the subset.

8.6.2 Relationships: A relationship represents the association between two or more entities. For example, an organization consists of departments such as sales, marketing, production, manufacturing, etc. Each department consists of a number of employees, who are identified with an employee identification number. Relationships are classified on the basis of degree, connectivity and cardinality, direction, type, and existence of relationship.

Degree of relationship: It refers to the number of entities that are associated with the relationship. Binary relationships, ternary relationships, n-ary relationships, etc., are the different types of relationships classified based on degree. Binary relationships occur between two entities. When an entity is related to itself, such relationship is termed as a recursive binary relationship. Ternary relationships occur between three entities and are used to define situations when binary relationship becomes inadequate. N-ary relationship is the general form of relationship in degree n.

Connectivity and cardinality of relationship: It refers to the mapping of related entity occurrences in the relationship. These can be one or many. Cardinality refers to the actual number of associated instances for each of the two entities. Three types of relationships exist based on connectivity between entities. These are one-to-one, one-to-many, and many-to-many relationships.

- *One-to-one*: One-to-one or 1:1 relationships take place when one instance of an entity M is associated with one instance of another entity N. For example, the student name (M) is associated with the roll number (N).
- *One-to-many*: One-to-many or 1:N relationships take place when one instance of an entity M is associated with zero, one, or many instances of another entity N. However, entity N has only one instance of entity M. For example, a class (M) consists of a number of students (N). All these students (N) belong to only one class (M). This relationship is also called as many-to-one or N:1 relationship.
- *Many-to-many*: Many-to-many or M:N relationships take place when one instance of an entity M is associated with zero, one, or many instances of another entity M; and, one instance of entity N is associated with zero, one, or many instances of entity M. For example, in a class, a single project (M) can be assigned to many students (N) and at the same time, a single student (N) can be assigned to many projects (M). This relationship is also called as N:M or non-specific relationship.

Direction of Relationship: This indicates the flow of relationship from the originating entity to the terminating entity. In a relationship, the entity from which the relationship commences is called the parent entity and where it concludes is called the child entity. The connectivity of a relationship decides the direction of the relationship. For example, in a one-to-one relationship, the direction flows from an independent entity to a dependent entity. In one-to-many relationships, direction always flows from one side to many sides, only once at the parent entity. The direction will be an arbitrary one in situations where both the entities are independent or in case of many-to-many relationships.

Activity: Given below is a set of entities based on certain relationships. Identify the type of relationships that exists between these entities.

Student – Program

Student – Course

Faculty – Course

Head of Department – Department

Answer:

Identifying and non-identifying relationships: Relationships can be identifying or non-identifying relationships. An identifying relationship occurs when the primary key of the child entity contains all the attributes of the foreign key. In this relationship, the child entity can be identified only with the parent, i.e., the child entity is a dependent entity. A non-identifying relationship occurs when the primary key of the child entity does not contain all the attributes of a foreign key. In this relationship, both the parent and the child entities are independent, i.e., the child entity can be identified without the parent.

Attributes: Attributes are characteristics that describe an entity. They differentiate one entity instance from another. For example, attributes of a student include his/her roll number, first name, last name, class, marks, percentage, etc. Attributes can be classified as identifiers or descriptors.

Identifiers or keys are special attributes which are used for identifying the unique characteristics of a specific entity. For example, the student roll number is an identifier that uniquely identifies the employee. Descriptors, on the other hand, identify the non-unique characteristics of an entity.

8.6.3 Converting an E-R Model to a Relational Model

The E-R model is converted to a relational model so that it can be used in a database. The relational model consists of relations, tuples, and attributes. A relation is a two dimensional table, which is made up of tuples. A tuple in the relational model is a row in a database. It is made up of one or more attributes (or columns in the database model). Attribute is a characteristic or a property of the relation contained in a tuple.

Converting an entity into a relation: In an E-R model, entities are converted into relations. The attributes in the entity become the attributes in the relation. The identifier in the entity becomes a key in the relation. If a weak entity is not dependent on the identifier (key), then it should be treated like any other entity. But, if it is dependent, then the key or the identifier of the parent relation is copied into the dependent relation. The key is called as a foreign key.

Representing relationships: In case of 1:1 relationships, the key of one relation is stored in other relation. In case of 1:N relationships, parent entity lies on the '1' side of the relation while the child entity lies on the 'N' or 'Many' side of the relation. The key of the parent entity is copied into the child relation. M:N relationships cannot be directly represented in relations. To represent them, a third relation, called the intersection relation, should be introduced. The keys from the original two relations should be copied into the intersection relation.

8.6.3.1 Representing recursive relationships: They are also called as unary relationships. A recursive relationship exists between instances of a single type of entity. For example, various parts or components should be assembled to build a computer. Therefore, a computer is made up of one or many other parts or components. Also, a part or a component might be made up of one or more parts or components. This results in a recursive 1:N relationship. Other forms include 1:1 and N:M recursive relationships. If an employee trains one and only one other employee, it forms a recursive 1:1 relationship. If an employee trains one or many other employees and also that employee is trained by one or many other employees, then it forms a recursive N:M relationship. These can be represented directly in the form of a relationship by adding another copy of the key in the relation.

Representing ternary relationships: When binary or simultaneous relationships exist, a ternary relationship takes place. Let us take an example of a manufacturer who sends finished products to a warehouse. Here, there are three entities – the manufacturer, the finished products, and the warehouse. These entities are represented in a relationship. First, binary relationships between these entities are established such as between the manufacturer and the finished products, the manufacturer and the warehouse, and the warehouse and the finished products.

8.6.4 Representing subtypes entities:

Certain entities have subtypes related with them. For instance, a labor can be a supertype entity for which the subtype entities could be part-time labor, full-time labor, salaried labor, etc. The supertype entity should be converted directly into a relation by using the attributes in the entity and the key in the entity is noted down. Each subtype entity should be converted into relations by using only the attributes in the entity and the key of the supertype entity should be copied into each of the subtype entities.

8.6.4.1 E-R Structures

The E-R model is based on various structures. Following are some of the E-R structures:

Classic 1 to many structure: It is also called as the Master/Detail or Parent/Child structure. For example, Portfolio/Stocks, Delivery Manifests/Items are examples of such structures.

8.6.4.2 Tree structure: It is also called as a hierarchy. A tree structure is a series of 1:N relationships. In this structure, each node has only one parent and each parent has zero, one, or many children. The children with a common parent are called siblings. For example, a portfolio is a parent which consists of many mutual funds. All mutual funds are part of a single entity called the portfolio. They are together called as siblings. The tree structure can also be converted into a relationship using the same steps which were used for converting 1:N relationships.

8.6.4.3 Network structure: They are of two types, simple network structure and complex network structure. These networks can be converted into relationships using the same steps which were used for converting 1:N relationships. A simple network structure consists of 1:N relationships. But, in this structure, each node can have more than one parent. The relation would have two foreign keys as the entity has two parents. A complex network structure has only one M:N relationship. An intersection relation should also be introduced during the conversion.

8.6.5 Importance of the E-R Model

The E-R model has gained significant importance. For a database designer, an E-R model constructs such as entities and relationships can be easily be transformed into relational tables. The E-R model is simple and easy to understand by the user, requires minimal training, and can be used as a design plan for implementing a data model in database management software.

8.7 The Relational Database Model

In 1970, Edgar F. Codd proposed the relational model. The Relational Database Model, also known as Relational Database Management System (RDBMS) was based on the concept of tables or relations. In this model, the data is organized in the form of tables. A table consists of records and each record contains columns, also known as fields. The fields contain attributes like student roll number, student name, etc., while a row (also known as a tuple) identifies the details pertaining to a single instance like the details of a particular employee say, Steve. In relational databases, each row is distinctive and each column has a unique name. Each row has a key that identifies the row with the relation. Data records stored in tables are accessed using a primary key. The primary key is a field in a record which is used to distinguish a record from other records in the table. For example, employee ID will be the primary key in the employee table.

Relational databases allow designating certain fields as keys to facilitate specific field value search. At times, the fields will be referred with the same name in both the tables. For example, a table containing personal details and a table containing official details of the students would have a field for student roll number in both the tables. This student roll number field can be used to relate two tables. This facility offered by relational databases can also be extended to relate multiple tables.

Relational databases also have keys called foreign keys. These are primary keys of one table which can be included in another table for relating or joining the data. For example, we have two tables, a Customer table containing customer data and an Orders table containing customer orders. All orders can be associated with customers present in the orders table. Customer ID, the primary key of customer table, acts a foreign key in the orders table that associates the customers with the orders. Relational database management systems like DB2, Microsoft SQL Server, Oracle, etc are based on the relational model.

8.7.1 Structured Query Language

Query language is a specialized language which is used by users to request information from a database. Query languages can be procedural or non-procedural languages. A procedural language is one for which the user provides instructions to the system for performing a sequence of operations on the database.

The system takes these instructions and computes the desired information. In non-procedural languages, the user requests for the desired information but does not give a procedure for obtaining it. SQL, MDX, QBE, etc are examples of some query languages.

The Structured Query Language (SQL, pronounced as Sequel) is a popular query language. It has become a de facto standard for query languages because of its wide usage and recognition in the industry and is the standard language used for relational databases. SQL includes both data definition and data manipulation operations. SQL query consists of a command keyword through which users can create queries and extract the required information. The basic command consists of the following three subcommands:

SELECT: lists the data items to be retrieved.

FROM: lists the tables from which the data will be retrieved.

WHERE: specifies the conditions and parameters which will limit the search.

For instance, let us take an example of a sales_details table with data items: salesperson, city, state, and sales. If the marketing manager of a company wants to know the sales details in the city of Chennai for all salespersons, the command to be given would be:

SELECT salesperson, city, state, sales **FROM** sales_details **WHERE** city = 'Chennai'

Activity: A query has to be written to retrieve the sales details of salespersons who have sold one or more units of a product in the previous year. What are the SQL commands that should be used to generate the output with the customer name, city, state, and sales volume of salespersons who sold at least one unit?

Answer:

Exercise

A. The sales manager of a company wants to know the details of sales achieved by James, a salesperson. Write the SQL statement that can be used to find out the sales achieved by James. 'Salesdetails' is a table in which name, city, state, and quantity are the relevant fields.

8.7.2 Designing a Relational Database

In designing a relational database, the columns to be used in the table should be defined. The table contains the columns and enables the user to get information in any combination. The tables should be defined correctly in order to make additions and changes to the table.

Data normalization: The information consisting of data items on the forms and reports can be used for data collection, which is then organized into tables. This process referred to as data definition involves defining and organizing data within a structure (a table). Each table must obey certain rules called data normalization. Normalization is a technique used for designing relational databases. It refers to the optimization of tables to remove any redundancy and scalability issues that might arise when the amount of data is increased. It is a two step process which involves arranging the data in a tabular form after deleting the repeating groups and the duplicated data from relational tables.

The process of normalization is based on the concept of normal form. A relational table is said to be in a normal form if it satisfies a set of constraints. A database can be normalized upto six normal forms but most databases are normalized only up to the third normal form. To be in the first normal form, every table must have a primary key, which is a column (or a set of columns) that uniquely identifies each row. The table may contain one or two primary keys in some situations called as alternate keys. To be in second normal form, the table should be in the first normal form and there should not be any partial dependency, i.e., all the non-key fields (fields that are not primary) should be functionally dependent on the full primary key. To be in the third normal form, there should not be any transitivity. In other words, the non–primary fields in the table should be dependent on the primary key alone and there should not be any sort of dependency among the non-keys.

Activity: An airline company maintains a database of passengers who are their frequent fliers. This database is later used for calculating loyalty rewards. The following table lists the top frequent fliers of the company, the class under which they are categorized, and the gifts they have won based on the class. The various gifts offered by the airline

Contd. .....

company include free air travel, a 3-day stay in Australia, and Sony VAIOs. Do you think the table is complying with the third normal form? If not, what changes can be made in order to make the table comply with the third normal form?

Passenger ID	Passenger Name	Frequent Flier Class	Gift
231	Steve	Silver	Free Air Travel
658	Sameer	Gold	3 Day stay in Australia
961	Jack	Silver	Free Air Travel
57	Alicia	Platinum	Sony VAIO
125	Carlos	Gold	3 Day stay in Australia
225	Bill	Platinum	Sony VAIO
324	Sanjay	Platinum	Sony VAIO
25	Sarah	Silver	Free Air Travel
15	Katie	Gold	3 Day stay in Australia
685	Karthik	Silver	Free Air Travel

Answer:

8.7.3 Data entry: In a relational database, data is entered either directly or through input forms or input screens.

Direct entry of data: Direct entry of data into a relational database involves periodic collection of data and needs constant updation of tables. Most database systems automatically enter certain types of data, such as the current date, user's name, etc.

*Entry through input forms or input screens***:** Data is often entered into a database using input forms or data input screens. This makes the data entry process easier. Data can be entered using text files and data can also be imported from other software. Data input screens are similar to paper forms. Most input screens are empty screens, except for a menu line or some help message. Three types of information can be placed on an input screen: simple text, input blanks, and data retrieved from the database.

Reports: The formatted output derived from a query or by searching the database is called as a report. A report may include summaries of columns. Reports can be automatically produced using a tool called 'report writer.' A report writer defines the placement and format of columns, headings, footer information, column titles, and pagination in database reports. Such reports can be generated automatically with the help of application generators. Application generators are tools that combine various features from tables, forms, reports, and queries to generate an application. Such an application can be created without writing any programming commands.

Check Your Progress-5

20. Match the following cardinality relationships with their examples.

i. 1:1 relationship	p. One Indian citizen can have only one Permanent Account Number (PAN) issued by the Income Tax department.
ii. 1:N relationship	q. A customer can have several accounts with the same bank. A bank account can be jointly held by several customers.
iii. M:N relationship	r. Each state government employs thousands of Indians. A person cannot simultaneously be employed by multiple state governments.

21. What are input screens and what are they used for?

22. A table has four columns: Student ID, Course ID, Percentage Obtained, and Course Title. Student ID and Course ID together constitute the full primary key. Percentage Obtained depends on both Student ID and Course ID. Course Title depends only on Course ID. Which of the following statements is **true**?

    i. The table complies with the first normal form.
    ii. The table complies with the second normal form.
    iii. The table complies with the third normal form.

a. Only i
b. Only ii
c. Only i and ii
d. i, ii, and iii

23. What is a primary key?

24. A table has three columns: EmployeeID, Department, and HoD. EmployeeID is the primary key. Each Department has an HoD (Head of the Department). Which of the following statements is **true**?

    i. The table complies with the first normal form.
    ii. The table complies with the second normal form.
    iii. The table complies with the third normal form.

    a. Only i
    b. Only ii
    c. Only i and ii
    d. i, ii, and iii

25. The command, 'WHERE' in SQL is used to

26. What is the process of defining tables called?

27. _____ refers to the optimization of tables to remove any redundancy and scalability issues that might arise when the volume of data increases.

28. A/an _____ defines the placement and format of columns, headings, footer information, column titles, and pagination in database reports.

    a. Data loader
    b. Report writer
    c. Input form
    d. Application generator

8.8 MIS and RDBMS

RDBMS (Relational Database Management System) organizes, catalogs and controls data needed for management information systems. RDBMS plays a major role in implementation of MIS. RDBMS is the underlying database for any MIS. RDBMS ensures data integrity, information security and maintains data in the databases. RDBMS provides an orderly file structure for storing data. It stores data in the form of records in table. Each record consists of item or field or column values. RDBMS allows cataloging, access and partitioning of data.

Data security is the major feature in RDBMS. Some of the RDBMS even provide security at table or record level. They maintain security tables. These tables maintain details about user information, logins and passwords. RDBMS assigns rights and privileges to users. They have inbuilt security applications. When data is transmitted over Internet, data security should not be ignored. For this purpose, management information systems are kept under firewalls. MIS systems also have user access permissions and privileges.

For example, a Product table can have fields such as product name, product price, product dimensions, and product quantity, etc. A retail store can maintain an RDBMS with their product details. The MIS such as invoice systems, payment processing systems or customer billing systems make use of this RDBMS with the product details. The data in RDBMS can be processed either on batch mode or real-time model. For some of the applications, the banks do the batch processing at the end of the day. For some transaction, real-time update of data takes place.

RDBMS provides support for SQL (Structured Query Language). Using SQL, users can query the data, define the table structures, modify the table structures and update the data. Even some of the current big data system such as Hadoop even processes the unstructured data. Unstructured data is the data which cannot be represented in the form of relational database. The benefits of using RDBMS in management information systems include:

- Well organization of data in the form of records.
- Efficient data security
- Easy generation and customization of reports for efficient decision-making.
- Data integrity and consistency.
- Separation of business logic from data management responsibility
- Secured user authentication and access.
- Querying the data.

Database management systems other than RDBMS include indexed database management systems and sequential database management systems.

8.9 Data Dictionary

A data dictionary is a file that defines the way in which the database is organized. It contains details of all the files in the database, the records in each file, and the details of the fields in each record. A relational database system maintains data about the tables and other objects in the database. This information is called the data dictionary or the system catalog. The data dictionary stores names of the relations and its attributes, the domains and lengths of attributes, the names and definitions of views defined in the database, and the integrity or key constraints. A view is a virtual table that comprises the fields of one or more tables in the database or fields from other views. The reason a view is called a virtual table is because it does not store or contain data and is recreated each time it is accessed.

Database administrators query the status of a firm's data using data dictionaries. Data dictionaries enforce standard data element definitions whenever end users and application programs use a DBMS to access an organization's databases. For example, an active data dictionary would not allow a data entry operator to enter a non-standard definition of an employee record, nor it would allow him/her to enter a name that exceeds the defined size of that data element.

8.10 Data Warehousing

According to Bill Inmon, "A warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process." The data warehouse architecture represents the overall structure of data. It has three components; these are data source, data loaders and DBMS, and query and analysis tools. Data loaders are used for entering data into the DBMS and the query and analysis tools are used for extracting the desired data from the DBMS.

Data warehousing involves a number of techniques such as data cleansing (validating and correcting the data before inserting it into the warehouse), data mining and business intelligence, and metadata management. Metadata is referred to as 'data about data'. Metadata management enables companies to make an analysis of the impact of changes to database tables and track the owners of individual data elements. It also helps in building the data warehouse.

8.11 Data Mining

Data mining refers to extraction and analysis of data gathered from different sources and summarizing it into useful information. The main objective of data mining is to look for patterns in the data which can be used to formulate a business strategy or to look for unusual behavior. For example, a sudden rise in the sale of a particular brand of laptops.

Data mining applies statistical tools and techniques to find patterns in the large amounts of data and forecast future trends. In data mining, data is processed in a data warehouse to identify key factors and trends and historical patterns of business activity. Managers can use this information to make strategic changes in business operations to gain a competitive advantage in the market.

The need to identify patterns in large datasets requires intense effort & use of technology. Data mining, therefore, plays a very important role in utilizing data for extracting customer insights. Exhibit 8.1 presents data mining methodology

Exhibit 8.1: Data Mining

The process of utilizing data to acquire customer insights requires wrangling of large datasets followed by processing that reveals hitherto unnoticed patterns. Data mining, also referred to as knowledge discovery in data (KDD), is concerned with this challenging task of finding meaningful patterns in data and requires use of advanced technology and tools to accomplish it. This will have a direct impact on the data mining tool market that is likely to increase from a little over half a billion USD in 2018 to over $ 1.3 Billion by 2026. The CAGR between 2019 & 2026 works out to an impressive 11.42% and the contribution to this rapid growth comes from the demands for software datamining tools from SMEs as well as large enterprises.

It is clear that data mining will continue to be of great interest for years to come. Given below are some trends that this rapidly expanding sector is like to see:

1. The pharmaceutical & healthcare sector are likely to play a dominant role for data mining;
2. Data mining will continue to move from coding by specialists to use of automated tools;
3. While automation will certainly increase usage, various enterprise software will see increasing use of embedded data mining to make data mining an integral part of the decision making process;
4. Besides the pharmaceutical & health care sector, spatial & geographic data mining will also contribute immensely to demand for data mining tools.
5. With demand shooting up, one should expect product and vendor consolidation in the space to produce few large vendors from the rather fragmented market today.

Source: https://www.datamation.com/big-data/data-mining-trends/

Check Your Progress-6

29. Which of the following components of data warehousing enables companies to make an analysis of the impact of changes to database tables, tracking the owners of individual data elements, and is also required to build the warehouse?
    a. Data cleansing
    b. Data application
    c. Business intelligence
    d. Metadata management

30. What is a data warehouse? What are its components?

31. What is data mining?

32. Which of the following is/are **true** regarding data mining?
    a. Data mining refers to analyzing data from different sources and perspectives and summarizing it into useful information.
    b. Data mining applies statistical techniques to find patterns in large amounts of data.
    c. Data mining processes data in a data warehouse to identify key factors, trends, and historical patterns of business activity.
    d. All of the above

33. A retail chain analyzed the buying patterns of 20-25 year old women. It discovered that the women belonging to this category shopped every fortnight for cosmetics and during this time they also tended to buy music and/or movie CDs. This information enabled the retail chain to move the music and movie CDs display closer to the cosmetics. Which process/technique would have helped the management of the retail chain in gaining such information about the purchasing patterns of the consumers?

8.12 Distributed Databases

Many organizations replicate and distribute copies or parts of databases to network servers at various sites. The distributed databases reside on network servers on the World Wide Web, on corporate intranets or extranets, or on other company networks.

A distributed database is a database that comprises of two or more data files which are placed at various sites on a computer network. It is a single logical database that is physically distributed across computers on a network to make it available to different users for access. The users feel that the entire database is at their location and are unaware of the replication of the data that happens behind the scenes.

The computers in a distributed database system communicate with one another through various communication media, such as high-speed networks and telephone lines. The computers called sites/nodes in a distributed database system may vary in size and function, ranging from workstations to mainframe systems. Each site is an independent database system. The distributed databases are controlled by a centralized distributed database management system (DDBMS), wherein the data stored in database is periodically synchronized. It ensures that any modifications made to the data from a particular location are automatically reflected in the data stored at other locations on the computer network. A distributed database has many advantages like ease of data sharing, degree of control over data, etc.

8.13 Objected-Oriented Databases

Object-oriented databases provide a mechanism to store complex data such as images, audio and video, etc. An object-oriented database (also referred to as object-oriented database management system or OODBMS) is a set of objects. In these databases, the data is modeled and created as objects.

Malcolm Atkinson and his fellow associates defined OODBMS in the following manner, “An object-oriented database system must satisfy two criteria: it should be a DBMS, and it should be an object-oriented system, i.e. to the extent possible, it should be consistent with the current crop of object-oriented programming languages. The first criterion translates into five features: persistence, secondary storage management, concurrency, recovery and an ad hoc query facility. The second one translates into eight features: complex objects, object identity, encapsulation, types or classes, inheritance, overriding combined with late binding, extensibility and computational completeness.”

ODBMS encapsulates the data and the code that operates on that data in an object. To create an object-oriented database, one can either add concepts of object orientation to the existing database languages, or extend existing object-oriented languages to deal with databases by adding database concepts. Orion, Jasmine, Matisse, ObjectStore, etc are examples of object-oriented databases.

The object-oriented data model deals with various applications as follows:

- **Computer-Aided Design (CAD):** A CAD database is a document management system that automatically archives drawings and retrieves them when necessary.
- **Computer-Aided Software Engineering (CASE):** A CASE database stores data required by software developers pertaining to source code, dependencies among software modules, definitions and uses of variables, and the development history of the software system.
- **Multimedia databases:** They are high-capacity/high-performance DBMSs that support multimedia data types as well as other basic alphanumeric data types, and handle large volumes of multimedia information.
- **Office Information Systems (OIS):** They include workstation-based tools for document creation and document retrieval and the maintenance of appointment calendars.
- **Hypertext databases:** Hypermedia databases have links that point to other documents and may also be structured in specific ways so that they can be indexed. Hypertext databases support the ability to retrieve documents based on links, and to query documents based on their structure.

Check Your Progress-7

34. _____ is a single logical database that is physically distributed across computers on a network.
    a. Relational database
    b. Data warehouse
    c. Distributed database
    d. Object-oriented database

35. What is CASE? Explain its importance.

36. _____ are high-capacity and high-performance database management systems that support text, images, and graphic objects.
    a. Relational databases
    b. Multimedia databases
    c. Analytical databases
    d. Distributed databases

37. What are hypertext databases?

8.14 Object Relational Database System

Relational databases face certain limitations such as the structure is not reusable, a specific interface cannot be determined, database cannot be maintained on a continuous basis and data is drawn from a limited set of data types. These limitations can be overcome with the help of Object Relational Database Management System (ORDBMS).

ORDBMS is a relational database that is object-oriented on the front end. Applications interface with this database as if data is stored in the form of objects and the system converts these objects into data tables containing rows and columns. The data is handled in the similar manner as in a relational database. After the data is retrieved, it is reassembled back into complex objects. The ORDBMS was developed with an objective to enable the software developers to view problems from an enhanced level of data abstraction.

ORDBMS enables the software developers to incorporate their customized data types and methods into the database. These databases support both object and relational models, thus, enabling the reusability of the structure. Data modeling and querying also becomes easier as ORDBMS uses complex data types. The main advantage of ORDBMS is that the programmer need not write any code for converting object data between RDBMS and object database as software for conversion is available. PostgreSQL, UniSQL, Valentina and VMDS are examples of ORDBMS.

8.15 Summary

- A database is a collection of related data that is organized to make access easier. The components of a database system include data, hardware, software, procedures, and users.
- Data is logically organized into characters, fields, records, files, and databases in an information system. The most widely accepted record-based data models are the hierarchical, network, and relational models.
- The database management system (DBMS) serves as an interface between users and the various databases. DBMS is software that enables the programmer and administrators to create, delete, store, retrieve, organize, interrogate, and maintain databases to provide requisite information to the end users and their organizations.
- The Entity-Relationship (E-R) model, proposed by Peter Pin-Shan Chen in 1976 was based on the perception of a real world that consists of entities and their inter-relationships. The model was proposed as a basis for unifying the different views of data which include the network and the relational views.

- A Relational Database Management System stores data in the form of tables and then links or relates those tables to provide answer to the diverse queries of users.
- Data dictionaries contain details of all the files in the database, the records in each file, and the details of the fields in each record.
- According to Bill Inmon, "A warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process." The data warehouse architecture represents the overall structure of data and it consists of three components; these are data source, data loaders and DBMS, and query and analysis tools.
- Data mining refers to extracting and analyzing the data from different sources and summarizing it into useful information. The main objective of data mining is to look for (hidden) patterns in the data that can be used to formulate a business strategy or to look for unusual behavior.
- Distributed databases are used by many organizations to improve database performance and security. The database is stored on several computers that communicate with one another through various communication media.
- The concept of Object Relational Database System (ORDBMS) overcomes the limitations of a relational database system. It enables the reusability of the structure, regular and periodic maintenance of the database, and a clear-cut interface.

8.16 Glossary

- **Analytical databases:** Databases that store information that is extracted from the operational and external databases. They contain information that is vital to the managers and the end users.
- **Class:** In object-oriented programming, a class constitutes a set of objects that share a common structure and behavior.
- **Computer-Aided Software Engineering (CASE):** A set of tools to help application developers complete the software development process more quickly and more accurately.
- **Data dictionary:** A file that defines the way in which the database is organized. It contains details of all the files in the database, the records in each file, and the details of the fields in each record.

- **Data hierarchy:** It refers to organizing data systematically and in a proper sequence. From the user's perspective, data can be logically organized into characters, fields, records, files, and databases.
- **Data independence:** The storage of data in such a way that it is not affected by any modification in the data structure or changes in the application program. Data that is logically independent is not affected by a change in the application program.
- **Data input screens:** Data is rarely entered directly into database tables. Instead, input forms or data input screens are used for entering the data to make the data entry process easier. Most input screens are empty screens, except for a menu line or some help message.
- **Data integrity:** Data integrity is said to exist if the data is correct and has no inconsistencies. To ensure data integrity, programs should use data validation procedures that define acceptable input ranges for each field in the record.
- **Data Management Association (DAMA):** An international organization that works for developing advanced concepts and practices of data resource management.
- **Data mining:** Refers to extracting and analyzing the data from different sources and summarizing it into useful information.
- **Data models:** A set of conceptual tools that describe data. These models are mainly used to specify the overall logical structure of a database and to provide a higher-level description of the implementation.
- **Data normalization:** This is a technique used for designing relational databases. Data normalization refers to the optimization of tables to remove any redundancy and scalability issues that might arise when the amount of data is increased.
- **Data warehousing:** It comprises data that is extracted from the operational database. Data is collected and stored in a data warehouse over several years.
- **Database Administrators (DBAs):** These are specialists who provide the necessary technical support for implementing the strategic and policy decisions related to the data in an organization.
- **Database Management System (DBMS):** A collection of software programs that allows users to organize and store data in a uniform way and to access day from more than one file.

- **Database:** A repository of stored data. A database consists of a collection of related data that is organized in a logical form to make access easier.
- **Distributed database:** A database where the data is not stored at a single physical location but is distributed over many computers. These computers are geographically dispersed and connected by a network or communication link.
- **Entity:** An entity consists of basic objects or concepts about which data is collected and stored. These may be places, people, objects, events, customers, orders, or something abstract like a loan or a holiday. Employees, invoices, projects, etc. are examples of entities.
- **Entity-Relationship (E-R) model:** It is a basis for unifying the different views of data which include the network and the relational views. The E-R model was based on the perception of a real world that consists of a collection of basic objects called entities and their interrelationships (association).
- **File:** A group of related records constitute a data file. Files are classified based on their permanence. That is, some files are temporary in nature as they undergo updation from time to time, and others are permanent.
- **File-processing system:** In this system, records are stored in various files. In order to extract, add, or delete records in a particular file, various application programs need to be written.
- **Hierarchical model:** In this data model, the records are organized as a collection of trees rather than as arbitrary graphs. The hierarchical model is designed to handle large transaction-based applications like sales-order processing, check processing, inventory updating, patient billing, insurance file maintenance, etc.
- **Hypertext/hypermedia databases:** These databases have links that point to other documents. Hypertext databases support the ability to retrieve documents based on links, and to query documents based on their structure.
- **Logical level:** This level of abstraction describes the type of data stored in the database and their interrelationships. Logical level is also known as the conceptual level.
- **Metadata management:** Metadata is referred to as 'data about data'. Metadata management helps companies to make an analysis of the impact of changes to database tables and track the owners of individual

data elements; metadata management also helps in building the data warehouse.

- **Network model:** In this data model, data is represented by a collection of records and the relationships among data are represented by pointers or links. Pointers store the location or address of files and records.
- **Object Relational Database System (ORDBMS):** A relational database that is object oriented on the front end. Applications interface with this database as if data is stored in the form of objects.
- **Object:** In object-oriented programming, an object plays a key role with respect to a request for an operation. The request invokes the operation that defines some action to be performed. Objects have their own state and behavior, and these in turn define the relationship.
- **Object-Based Logical Models (OBLMs):** These data models allow the user to specify data constraints explicitly. They are also used for describing data at the logical and the view levels.
- **Object-Oriented Database Management System (OODBMS):** This database management system encapsulates the data and the code that operates on that data in an object. To create an object-oriented database, one can either add concepts of object orientation to the existing database languages, or extend existing object-oriented languages to deal with databases by adding database concepts.
- **Object-oriented databases:** The state, behavior, and relationships of these objects are characterized in line with the object-oriented data model. In other words, in these databases, the data is modeled and created as objects.
- **Office Information Systems (OIS):** These information systems allow queries about schedules, documents, and contents of documents. They include workstation-based tools for document creation and document retrieval and the maintenance of appointment calendars.
- **Optimization analysis:** A complex extension of goal-seek analysis in which the target is not fixed. In this analysis, one or more variables are changed after taking the constraints into account until the best alternative or the optimal value is found.
- **Physical data models:** These data models are used to describe data at the physical or the lowest level, i.e., in the form of bytes and words. Physical data models are used for identifying tables, identifying columns, identifying relationships, data normalization, etc.

- **Physical level:** The lowest level of abstraction that describes the manner in which data is actually stored. At the physical level, complex low-level data structures are described in detail.
- **Pile file organization:** In this type of file organization, data is collected and saved in the order in which it arrives. There is no specified structure or format in which the records are saved.
- **Primary key:** A field in a record that is used to distinguish a record from other records in the table. A primary key uniquely identifies a record. It may also be a reference key in another table.
- **Query language:** A specialized language used to request information from a database. The Structured Query Language is a popular query language. It is a non-procedural query language.
- **Record:** Related fields of data are grouped to form a record.
- **Record-Based Logical Models (RBLMs):** These data models describe data at the logical and view levels. In RBLMs, the database is structured into numerous types of records that are of fixed format, i.e., each record type defines a fixed number of fields or attributes and each field is usually of a fixed length.
- **Relational Database Management System (RDBMS):** A database model which is based on the concept of tables or relations. In a RDBMS, the data is organized in the form of tables.
- **Relational databases:** These databases allow the designating of certain fields as keys to facilitate specific field value search.
- **Relational Model:** In this model, data in several files is related through the use of a common key field. Each record in the file has the same key field but unique key field contents. The field can thus be used to identify a record.
- **Relationships:** It can be defined as the way in which data is shared between entities. A relationship represents the association between two or more entities.
- **Schemas:** The overall design of the database is referred to as the schema.
- **Subtype entities:** Certain entities have subtypes related with them. For instance, an employee can be a supertype entity for which the subtype entities could be part-time employee, full-time employee, salaried employee, consultant, etc.

- **View level:** The highest level of abstraction. Despite the use of simpler structures at the logical level, some complexity remains, because of the large size of the database. Interaction with the system is simplified by the DBMS software and users need to access only a part of the database.

8.17 Self-Assessment Test

1. A database management system consists of five components - data, hardware, software, procedures, and people. Explain the role of these components in a database management system.
2. Files are organized based on parameters like rapid access, storage efficiency, and reliability. In this context, explain the types of file organization.
3. Data models specify the logical structure of a database. What are the different data models? Explain.
4. Database systems should ensure data integrity, data independence, and data security. Describe the characteristics of a well-defined database system.
5. The E-R model is based on the perception of a real world that consists of a collection of basic objects called entities and their interrelationships (association). Describe the basic elements in the E-R model.
6. The E-R model is used for organizing data within databases or information systems. What are the various E-R structures? Explain.
7. Normalization is a technique used for designing relational databases. Explain the process of normalization with examples.
8. Explain the following concepts:
   - Data dictionary
   - Data warehousing
   - Data mining
9. An object-oriented database supports various applications. Explain with examples.

8.18 Suggested Readings / Reference Material

1. Introduction to Information Technology , V. Rajaraman , PHI learning, 2018
2. Information Technology for Management, 2ed: Advancing Sustainable, Profitable Business Growth , Turban, Volonine, Wood , O.P. Wali, Wiley India Pvt Limited, January 2021
3. Introduction to Information Systems - 6th edition, R. Kelly Rainer; John Wiley & Sons, Inc.2016

4. Information Technology: An Introduction for Today's Digital World , Richard Fox, Chapman and Hall/CRC; 2nd edition (August 21, 2020)
5. Information Technology for Management, Efraim Turban, Carol Pollard, Gregory Wood, Wiley, 2018.

Additional References:

1. Critchley, L., Where Nanotechnology, the IoT, and Industry 4.0 Meet., https://www.mouser.com/blog/where-nanotechnology-the-iot-and-industry-40-meet, 2019
2. Pan India implementation of HMIS over Indian Railways, Ministry of Railways., http://railministry.com/pan-india-implementation-of-hmis-over-indian-railways/ 2020
3. Vossler, C. How Long Does It Take To Order A New BMW? https://www.bmwblog.com/2020/09/28/how-long-does-it-take-to-order-a-new-bmw/2020
4. Jay, A., 10 New ERP Trends & Forecasts for 2020/2021 – A Look Into What's Next. https://financesonline.com/erp-trends/2019
5. Gingiss, D., How Integrating Social Media Into The Rest Of The Business Will Increase Revenue., How Integrating Social Media Into The Rest Of The Business Will Increase Revenue (forbes.com), 2019

8.19 Answers to Check Your Progress Questions

Following are the answers to the Check Your Progress questions given in the unit.

1. A field consists of a group of characters. For example, the grouping of the characters in a person's name forms a name field, and the grouping of numbers in a sales figure forms a sales amount field.

2. (d) An identification field in a record is called a key field.

A character is the most basic element of data that consists of a single alphabet, number, or a symbol. A field consists of a group of characters. Related fields of data are grouped to form a record. A group of related records constitutes a data file. A database is a collection of logically related records and files. Data present in several files is related through the use of a common key field. A particular field is defined as the key field, which uniquely identifies a record. Records are organized based on the key field. Each record in the file has the same key field but is unique in its contents.

3. (d) i, ii, and iii

In indexed-sequential file organization, records are organized on the basis of a key field. An index supports the random access of records. It provides a quick search capability that allows the user to reach the vicinity of the desired record rapidly. The index also contains the key field and a pointer to the main file. When a query is executed, the index is searched to find the highest key value which is equal to or less than the desired key value. The search continues in the main file at the location indicated by the pointer. New records in the indexed-sequential file are added to the overflow file. Then the record and index is updated in the main file.

4. Hashed file organization

In hashed file organization, hashing is done on the key value to get the location of the record. This type of organization is used when rapid access is required and fixed length records are used.

5. (d) Record

Related fields of data when grouped together form a record. For example, a pay-roll record consists of data fields like name, salary, etc. A customer's name, account number, and account balance constitute a record.

6. Object-based logical model

Object-based logical models (OBLM) are used for describing data at the logical and view levels. The Entity-Relationship (E-R) model falls under OBLM. It is based on the perception of a real world that consists of a collection of basic objects called entities and their inter-relationships (association). An entity may be a person, a book or something abstract such as a loan or a holiday.

7. (c) Only i and iii

Both record-based and object-based logical models are used for describing data at logical and view levels. Physical data models are used to describe data at the physical or the lowest level, i.e., in the form of bytes and words.

8. Record-based logical models (RBLM) are used for describing data at the logical and view levels. These models are used both to specify the overall logical structure of the database and to provide higher level description of the implementation.

Relational, network, and hierarchical models are widely accepted record-based models.

9. (c) Pointers

In the network model, data is represented by collections of records, and relationships among data are represented by pointers. Pointers are used to store the location or address of files and records.

10. Hierarchical; relational

In the hierarchical model, the records are organized as a collection of trees rather than as arbitrary graphs. Also, the records are not represented by pointers but on the basis of the values that they have. In the relational model, a two-dimensional table is generally used for representing data and relationships. Each row corresponds to a record, while each column corresponds to a field. Unique or primary keys are used to access records in a table.

11. (b) Fields/ attributes

A relational database consists of tables, each of which is assigned a unique name. A table consists of rows and columns. Each row in a table represents a record and each column represents attributes or fields.

12. Structured Query Language (SQL) is a popular query language that includes both data definition and data manipulation operations. It consists of a command keyword followed by additional information. By combining keywords, users can create queries that extract just the information they are looking for.

13. (c) A relational DBMS is generally a three-dimensional table

In a relational database, data in several files is related through the use of a common key field. A relational DBMS is generally a two-dimensional table. Each row corresponds to a record, while each column corresponds to a field. In the relational database model, unique keys or primary keys are used to access records in a table. A primary key is used to distinguish a record from other records in the table. As a relational DBMS can work with many files and relate all the data, many databases can be created and focused on different themes.

14. Database maintenance entails modifying the database to store the latest and most accurate data in the database. DMLs (Data Manipulation Languages) are used for modifying the data in database tables.

15. **i/q, ii/r, iii/p**

An ideal database is one that provides fast and easy access to data and ensures efficiency in its storage and retrieval. Database systems should also ensure data integrity, data independence and data maintenance. Data integrity is said to exist if the data is correct and has no inconsistencies. Data independence refers to the storage of data in such a way that it is not affected by modification in data structure or changes in the application program. Database maintenance implies that a good database system should have a system in place for data maintenance. This includes procedures for adding, updating, and deleting records.

16. (a) **An instance**

Databases undergo a lot of change over time as information is constantly being inserted and deleted. The information currently stored in the database is called an instance of the database. The overall design of the database is called schema. Database systems have several schemas partitioned according to the levels of abstraction.

17. **i/r, ii/q, iii/p**

The physical level is the lowest level of abstraction; it describes the manner in which data is actually stored. At the physical level, complex low-level data structures are described in detail. The logical level is the next level of abstraction that describes the type of data stored in the database, and their interrelationships. This is also known as the conceptual level. The entire database is described in terms of a small number of relatively simple structures. The view level is the highest level of abstraction. Despite the use of simpler structures at the logical level, some complexity remains, because of the large size of the database. Interaction with the system is simplified by the DBMS software and users need to access only a part of the database. The system provides many views for the same database.

18. Data independence refers to the method of data storage in which data is not affected by modification in data structure or changes in the application program. Data independence can be either physical or logical. When we say that data is physically

independent, it means that data is not affected by a change in the data structure. Data that is logically independent is not affected by a change in the application program.

19. (d) i, ii, and iii

In the database management approach, data records are consolidated into databases that can be accessed by different application programs. The Database Management System (DBMS) serves as an interface between users and the various databases. This helps the users to access data easily and enables the programmers and administrators to create, delete, interrogate, and maintain databases to provide requisite information to the end-users and their organizations.

20. i/p, ii/r, iii/q

One-to-one or 1:1 relationships take place when one instance of an entity P is associated with one instance of another entity Q. One Indian citizen can have only one Permanent Account Number (PAN) and vice versa. One-to-many or 1:N relationships take place when one instance of an entity P is associated with zero, one, or many instances of another entity Q. However, entity Q has an association with only one instance of entity P. Every state government can have many employees. The same employee cannot simultaneously be employed by multiple state governments. Many-to-many or M:N relationships take place when one instance of an entity P is associated with zero, one, or many instances of another entity Q. And, one instance of entity Q is associated with zero, one, or many instances of entity P. One customer of a bank can have one or more accounts. And the same account can be jointly held by multiple customers also.

21. Data is rarely entered directly into database tables; instead, input forms are used. This makes the data entry process easier as input screens are similar to paper forms. Most input screens are empty screens, except for a menu line or some help message.

22. (a) Only i

o be in the first normal form, every table must have a primary key, which is a column (or set of columns) that uniquely identifies each row. In this table, Student ID + Course ID is the primary key. So the table complies with the first normal form. To be in second normal form, the table should be in the first normal form and there should not be any partial dependency, i.e., all the non-key fields (fields that are not primary) should be functionally dependent on the full primary key. Here, the course title functionally depends

only on the course ID (part of the primary key) rather than on the full primary key. Therefore, this is a situation of a partial dependency. So the table does not comply with the second or higher normal forms.

23. In relational databases, unique keys or primary keys are used to access records (rows) in a table. The primary key is a field in a record that distinguishes a record from other records in the table. For example, a social security number, which is unique to a particular individual, is used to access data pertaining to different individuals. A primary key of a table can be a reference key in another table.

24. (c) Only i and ii

To be in the first normal form, every table must have a primary key, which is a column (or set of columns) that uniquely identifies each row. In this table, EmployeeID is the primary key. So the table complies with the first normal form. Since there is only one key column, there is no possibility of a non-key field being dependent on a part of the primary key. So, the table is also in the second normal form. To be in third normal form, the table should be in the second normal form and there should not be any transitivity, i.e., the non-primary key fields (fields that are not primary) should be dependent on the primary key alone. There should not be any sort of dependency among the non-keys. Here, the field HoD is dependent on the field Department. Therefore, this is a situation of transitivity. So, the table does not comply with the third normal form.

25. Specify the conditions and parameters which will limit the search

The command, 'WHERE' in Structured Query Language (SQL) is used to specify the conditions and parameters which will limit the search. For example, a 'WHERE' command, can be used to find out the salespersons who have achieved a sales target of more than 2,000 units.

26. Data definition

Data definition refers to a process which involves defining and organizing data within a structure (a table).

27. Data normalization

Data definition involves defining and organizing data within a structure (tables). The information consisting of the data items on the forms and reports can be used for data collection. The data items

displayed on these forms and reports should then be organized into tables. Each table must obey certain rules called data normalization. Normalization refers to the optimization of tables to remove any redundancy and scalability issues that might arise when the amount of data is increased.

28. (b) Report writer

Tools like 'report writer' automatically produce reports. A report writer defines the placement and format of columns, headings, footer information, column titles, and pagination in database reports. Such reports can be generated automatically with the help of application generators.

29. (d) Metadata management

Metadata management takes place during the entire process of identifying, acquiring, and querying the data. Metadata is referred to as 'data about data'. It helps companies to make an analysis of the impact of changes to database tables, tracking the owners of individual data elements, and is also required to build the warehouse.

30. Data warehouse is a system that stores, retrieves, manages, or otherwise manipulates massive amounts of data that may be from the organization's databases or from external sources. It consists of three components: data source, data loaders and DBMS, and query and analysis tools.

31. Data mining is the process that extracts and analyzes data from different sources and summarizes it into information. Statistical tools and techniques are used to discern patterns in large amounts of data thus enabling forecast of future trends and actions. This helps businesses to become proactive and arrive at decisions that are based on knowledge.

32. (d) All of the above

Data mining refers to the analysis of data from different sources and perspectives and summarizing it into useful information. It applies statistical techniques to unearth patterns in large amounts of data. It processes data in a data warehouse to identify key factors, trends and historical patterns of business activity. This information can be used by managers to make strategic changes in business operations to gain a competitive advantage in the market.

33. Data mining

Data mining refers to the process of analyzing data from different sources and perspectives and summarizing it into information. The aim is to look for patterns in the data which can be used to formulate a business strategy or to look for something unusual. This involves processing the data in a data warehouse to identify key factors, trends and historical patterns of business activity. In the given context, the management of the retail chain must have used data mining to make changes in its business operations to gain a competitive advantage in the market.

34. (c) Distributed database

Distributed database is a single logical database that is physically distributed across computers on a network. A distributed database comprises two or more data files which are placed at various sites on a computer network.

35. CASE stands for Computer Aided Software Engineering. A CASE database is an application of the object-oriented data model. It stores data required by software developers such as source code, dependencies among software modules, definitions and use of variables, and the history of the software system.

36. (b) Multimedia databases

Multimedia databases are an application of the object-oriented data model. They are high-capacity/high-performance DBMS that support multimedia data types such as text, images, graphic objects, video and audio. They also support basic alphanumeric data types, and handle large volumes of multimedia information.

37. Hypertext databases are an application of the object-oriented data model. They have links that point to other documents and may also be structured in specific ways so that they can be indexed. Hypertext databases support the ability to retrieve documents based on links, and to query documents based on their structure.

8.20 Answers to Exercises

Following are the answers to the Exercises given in the unit.

A. SELECT name, city, state, quantity FROM salesdetails WHERE name = 'James'

In structured query language (SQL), the command SELECT is used to list the data items (fields) to be retrieved. The command FROM is used to list

the tables from which the data will be retrieved. The command WHERE specifies the conditions and parameters which will limit the search. In the given situation, the sales manager wants to find out the sales details of the salesperson James. 'Salesdetails' is the table with fields name, city, state, and quantity. Therefore, the statement to be used for finding out the quantity achieved by James is: SELECT name, city, state, quantity FROM salesdetails WHERE name = 'James'.

Unit 9

Computer Networks

Structure

9.1 Introduction

In the previous unit we discussed about database management systems and also realized that, computers help in converting raw data into useful information. But, how can one share this information with another person? This is possible through computer networks. In this unit, we introduce you to computer networks.

Computers are said to be interconnected if they are able to exchange information. In simple words, a network links two or more computers and enables sharing of information between them.

Many developments have taken place in the field of computing and communication technologies. Information is being shared and disseminated through computer networks due to advances in communication technologies. A network is a data communication system that links multiple computers for sharing of information. They may be connected through a copper wire, fiber optic cables, microwaves, or communication satellites.

In this unit, we would first discuss about the basics of computer networks and the classification of networks. We shall then move on to discuss about the peer-to-peer and client/server networks. Finally, we would be discussing the various network topologies, network infrastructure, and the network architectures and protocols.

9.2 Objectives

By the end of this unit, you should be able to:

- Define the basics of computer networks.
- Reproduce the classification of networks.
- Compare peer-to-peer and client/server networks.
- Recall the various types of network topologies.
- Explain the network infrastructure.
- Explain the network architecture and its protocols.

9.3 Basics of Computer Networks

A computer network is a data communication system that links two or more computers and enables sharing of information between them. Computers are connected through cables, radio waves, telephone lines, infrared light beams, or satellites. Depending on the usage of the network, they are classified into Local Area Network (LAN), Metropolitan Area Network (MAN), and Wide Area Network (WAN).

9.3.1 Elements of a Network

The elements of a network include network interface cards (NICs), cable, hub, and protocol.

Network Interface Card: NICs allow a computer to communicate over a network.

Cable: A cable is a physical medium that connects two network devices for enabling communication for the purpose of data transfer.

Hub: A hub is the central connecting point that connects all the network devices through network cables.

Protocol: A protocol is a set of rules regarding communication that ensures that all the computers in the network communicate in the same language.

9.3.2 How Does a Network Work?

Consider a simple network connecting several computers to a hub using cables. To send a message from Computer A to Computer B, the NIC of Computer A first translates the file from the binary form into electric pulses

based on the protocol. The hub receives the pulses through the cable, and sends them on to the other computers. The pulses are translated by the NIC of Computer B into binary data that forms the file. For enabling this data transfer, both the computers A and B should have the same protocol

9.3.3 Computer Network Applications

Computer networks have revolutionized many aspects of our life. The benefits that can be derived through the use of computer networks depend on the type of the user (Business or Individual).

Use of computer networks by individuals: Individuals can access information through the Internet. Networks can also be used for managing bank accounts and paying bills electronically. In addition to this, people gain access to the World Wide Web (WWW), online catalogs of various companies, e-mail, video conferencing, etc enabling efficient gathering and sharing of information.

Activity: A large retailer of books wants to make the best use of information technology to market and distribute its products. In what all ways can computer networks help the company do this? Explain.

Answer:

_____

_____

_____

_____

_____

.

Use of computer networks by businesses: Business organizations consider computer networks as an asset as it allows storage and retrieval of information. The computers enable organizations to maintain records, manage inventory, monitor productivity, and generate reports. The shared network enables employees in an organization to transfer data, share input/output devices, etc. Networks also facilitate exchange of information between people through intranets or extranets. Intranets facilitate sharing of internal organizational information among the employees of a company. An extranet is a private network that acts as a link to various internet channels. It is extended to external users of the company like the suppliers, distributors, customers, business partners, etc., through the internet.

9.4 Components of Unified Communication

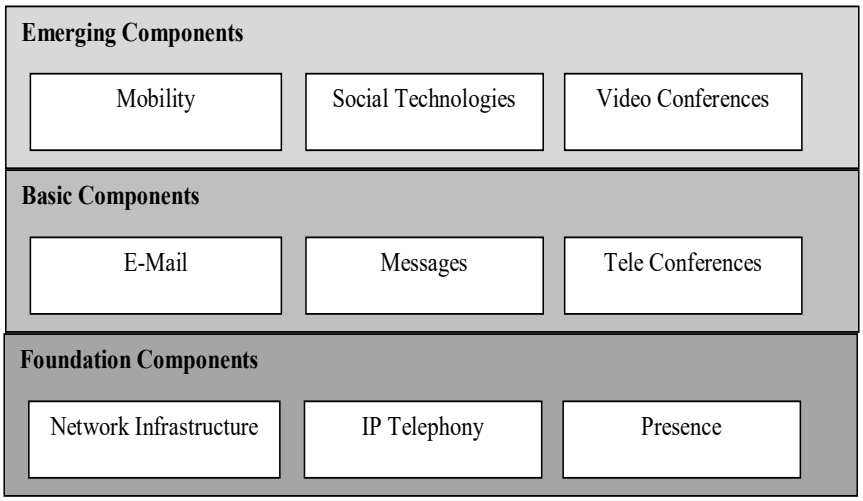
Unified communication integrates enterprise information systems in the organization with the organizational communications infrastructure. Unified communications makes use of complex technologies to work together in order to meet organizational purpose. The organization's decision to go for unified communications should be based on Return on Investment (RoI) and its adaptability to changes. Unified communications is an evolving field and it has future scope. The organization should have roll out plan for having unified communication.

The components of unified communication include foundation components, basic components and emerging components. Foundation components are the core components of the system such as network infrastructure, IP telephony (Voice over IP) and Presence, which has the capacity to check whether the lines are occupied or not.

Basic components include services and applications needed for day to day business operations such as e-mail, messaging, instant messaging, faxing, voicemail and conferences including audio and web conferences.

Emerging components include current technologies for mobile communications, social media technologies and video conferencing to reduce travel expenses for the organization.

Figure 9.1: Components of Unified Communication



Sources: Cisco (2013), "Unified Communications Components: Understanding Your True Unified Communications Needs", Available online at http://www.firewall.cx/cisco-technical-knowledgebase/cisco-services-tech/938-cisco-unified-communications-components.html.

9.5 Classification of Networks

Networks are classified based on transmission and scale. Based on transmission, they are classified into point-to-point networks and broadcast networks.

Point-to-point networks connect individual pairs of machines, where the message visits one or more machines to travel from source to destination. These networks are also called unicast or one-to-one networks and cover large areas. On the other hand, broadcast networks have a single communication channel which is shared by all the machines on the network. These networks are used for sending messages to all the computers on the network. Broadcast networks usually cover smaller areas.

Based on the scale or size/area of networking, networks are classified into Local Area Networks (LANs), Metropolitan Area Networks (MANs), and Wide Area Networks (WANs). Virtual private networks and wireless networks are other categorization of networks.

9.5.1 Local Area Network (LAN)

LAN is a computer network that covers a limited area and provides networking capabilities to a group of computers. They link different departments or workgroups within an organization but can also be used for connecting several offices or a group of buildings.

LANs allow sharing of resources like hardware, software or data between computers. LANs are constructed using hardware such as NICs, cables, and hubs.

9.5.1.1 Advantages of LANs

The advantages of LAN include:

- LANs facilitate sharing of software and data among computer systems on the network.
- LANs provide centralized backup facilities by allowing users to store information as a backup in servers. This enables users to retrieve information in case of a system failure or accidental loss of data.
- LANs have the ability to share the hardware resources like printers and other devices. They also facilitate sharing of data and other resources among the computer systems on the network.
- Other advantages of LAN include enhancing communication from one computer to another and thereby reducing cost and time required in disseminating messages. LANs are limited to a building or an institution thereby providing better control, security, and flexibility to the users on the network.

9.5.1.2 Disadvantages of LANs

- LANs may cause difficulty in resource sharing if numerous computers are attached to a LAN.
- The threat of virus is major problem in LAN as all the computers are linked.
- The sharing of information can also be a matter of concern in LAN as unknown users may gain unauthorized access to company data by knowing user ids and passwords.
- All the network devices in a LAN are managed by servers. If a server breaks down, then all the systems will be affected.
- LANs have more complex architecture than a set of individual standalone computers. Therefore, they require maintenance, and also, usually someone to look after and run the system.

9.5.2 Metropolitan Area Network (MAN)

A MAN is an interconnection of various LANs and spans a larger area. MANs have medium-sized networks and cover a radius of 2.5 to 25 kms. A MAN is owned by a group of users or a single organization that provides network services to the users. These networks can also be owned as public utilities. MANs have high speed network capabilities and allow sharing of regional resources. A MAN is larger than a LAN and smaller than a WAN.

9.5.3 Wide Area Network (WAN)

A WAN is a computer network covering a larger geographical area spanning a country or a continent. These networks can be used to connect two or more LANs through public networks like telephone or through satellites or leased lines. WANs cover areas that span 1,000 to 10,000 kms. However, these networks also suffer from security related issues as they are vulnerable to hacking as they are accessed by users located worldwide. The Internet is a well-known example of a WAN.

9.5.4 Virtual Private Networks (VPNs)

These are private communication networks that use the Internet to connect employees of a company located at remote areas through a virtual connection. Businesses and government related services use VPN as it provides security. The biggest advantage of a VPN is that a company can establish secure intranets between its offices and manufacturing plants.

VPNs enable computers in different geographical areas to interact with the help of a LAN. VPNs transfer information over the Internet by encrypting and encapsulating the data in packets thereby ensuring security. VPNs mostly are maintained by Internet Service Providers (ISPs).

9.5.5 Wireless Networks

A wireless network is a computer network or a telephone network that enables users to connect through radio waves. Wireless networks applications are found in mobiles, notebook computers and personal digital assistants (PDAs). Wireless networks help people send and receive mails, make phone calls, access documents, etc., at any time or from anywhere.

Wireless networks use three major technologies:

(1) **Wireless Application Protocol (WAP):** WAP, an international standard for wireless communication. It is a protocol that connects mobile devices to the Internet. WAP allows internet to be accessed through a mobile phone or a PDA.

(2) **Short Messaging Service (SMS):** SMS allows users to send text messages using mobile phones or pocket PCs facilitating sharing of information.

(3) **Global System for Mobile (GSM) Communications:** GSM, a standard for mobile communications is a mobile telephone protocol that enables international roaming possible for mobile phone users.

Check Your Progress*1

1. Define a computer network? What are the advantages of having a computer network?
2. Match the following:

i. Wireless Application Protocol	p. Allows sending text messages to mobile phones
ii. Short Messaging Service	q. Mobile telephone protocol
iii. Global System for Mobile	r. Allows mobiles to connect to the Internet

3. What is a protocol?
4. In terms of size, which of the following gives the correct increasing order of different types of networks?

a. LAN<WAN<MAN
b. LAN>MAN>WAN
c. LAN<MAN<WAN
d. LAN>WAN>MAN

5. A communications network in an office or a building is a _____
6. What are Virtual Private Networks (VPNs)? Explain their uses.
7. _____ are privately-owned networks which are used in a single building or an office to connect various workstations and share other resources such as printers, files, etc.
8. Given below are statements pertaining to networks. Indicate true/false.
   a. VPNs transfer information by encrypting and encapsulating it into data packets and sending the packets over the Internet.
      True/False
   b. LANs are built and maintained by Internet Service Providers (ISPs).
      True/False
   c. VPNs allow a company to establish secure Intranets over the Internet.
      True/False
   d. In point-to-point network a message can directly reach the destination system.
      True/False
   e. WANs are vulnerable to hacking.
      True/False
   f. The Internet is an example of a LAN.
      True/False
   g. Larger networks use point-to-point technology.
      True/False
   h. LANs facilitate sharing of data and software among the computer systems on the network.
      True/False
   i. Point-to-point networks have a single communication channel shared by all the machines on the network.
      True/False
   j. WANs are limited only upto a building, office, campus, etc.
      True/False
   k. LANs facilitate sharing of peripherals like printers and other devices.
      True/False

l. A VPN is a computer network that covers a large geographical area.
True/False

m. LANs provide better control, security, and flexibility to the users on the network.
True/False

9.6 Peer-to-Peer and Client/Server Networks

Network operating systems are divided into peer-to-peer network operating systems and client/server network operating systems.

9.6.1 Peer-to-Peer

A peer-to-peer network runs on the computers of individual users allowing users to share files and resources on their computers and access resources on other computers. They do not have a file server. The resources available on the network are controlled by the users. Peer-to peer networks are designed for small to medium LANs. Internet music sharing services like Napster and Gnutella are examples of peer-to-peer networking systems.

9.6.2 Client/Server

In a client/server network, one computer acts as a server that runs the software enabling it to perform jobs requested by other computers called clients. Servers perform functions like resource sharing, access control, and communication. Various types of servers include file servers, application servers, e-mail servers, communication servers, back up and archive servers, etc. Novell Netware and Windows 2000 are examples of client/server networking systems.

9.7 Network Topologies

Topology refers to the physical or logical layout of a network. The topology of a network is a geometric representation of all devices connecting to each other in the network. Mesh, star, bus, ring, and tree are the basic types of network topologies or structures.

9.7.1 Mesh Topology

In mesh topology devices are connected to each other using a dedicated point-to-point links. The dedicated link ensures that the traffic is carried only between the two connected devices. This helps avoid traffic problems. A mesh topology is robust meaning that if a link stops functioning, it will not affect the entire system thereby ensuring privacy and security. A mesh topology requires huge cabling, number of I/O ports and its installation and reconfiguration is difficult.

9.7.2 Star Topology

In star topology, devices are connected to a centrally located hub through a point-to-point link. The devices are not directly linked. Star topology does not permit direct traffic between the devices. The hub acts as an intermediary between the devices. For example, if a computer wants to send data to another, it first sends the data to the hub, which in turn transmits the data to other devices. Devices in mesh topology require only one link and one I/O port in order to connect with others and it is easy to install and reconfigure. Star topology is also robust but requires huge cabling. StarLAN is an example of a star topology.

9.7.3 Bus Topology

A bus topology has a multipoint configuration and the devices are linked through a long cable. The nodes of the devices are connected to the cable using drop lines and taps. A drop line is a connection or link that runs between the main cable and the device, while a tap is a connector which joins the main cable or penetrates into the covering of the cable to make contact with the metallic core. When a signal travels along the backbone, some energy gets transformed into heat resulting in the signal becoming weaker. A bus topology is easy to install, requires less cabling but is difficult to reconfigure. Any puncture or break in the bus topology stops transmission. ARCnet and Ethernet are examples of bus topology.

9.7.4 Ring Topology

In a ring topology, each device is linked to two devices on either side of it through a dedicated point-to-point connection. Signals in a ring topology are passed in one direction from device to device till it reaches its destination. Each device has a repeater that regenerates the bits and passes them whenever a device receives a signal which is meant for another device. It is easy to install and reconfigure a ring topology. Ring topologies are subject to traffic and media constraints. A rupture in the ring disables the whole network. Fiber Distributed Data Interface (FDDI) is an example of ring topology.

9.7.5 Tree Topology

In a tree topology, the nodes of a few devices are linked to a centrally located hub that controls the traffic in the network. The other nodes are indirectly connected to the hub through a secondary hub, which in turn, is connected to the main hub. The central hub called the active hub uses a repeater which regenerates the bit patterns before sending them out. This strengthens the transmission and increases the traveling distance of a signal. The secondary hubs can be active or passive. A passive hub gives a simple connection to the devices. The secondary hub allows more devices

to be connected to the active hub thereby increasing the traveling distance between the devices.

9.7.6 Hybrid Topologies

In hybrid topologies, different topologies are used to connect a single network. For example, different departments in an organization can use different topologies.

Activity: XYZ Ltd., a management consulting firm, is evaluating wireless and wired networking options, for the computer systems at its head office in Mumbai. What are the factors that the company must consider, before choosing between a wired network and a wireless network? Explain.

Answer:

Check Your Progress-2

9. Which of the following statements is **false** regarding peer-to-peer operating systems?
   a. Peer-to-peer operating system runs on the computer of individual users
   b. Users can share files and resources on their computers and access the resources of other computers
   c. They have a centralized management source or a file server
   d. They are designed for small to medium local area networks

10. What is a network topology?

11. Which of the following is **not** a LAN topology?
   a. Ring
   b. Star
   c. Hub
   d. Bus

9.8 Network Infrastructure

The network infrastructure determines the efficiency and reliability of any network. The networks should be properly managed and controlled to enable efficient use of the available bandwidth. Several devices are used for achieving this. Some of them are discussed below:

9.8.1 Network Interface Cards

Network Interface Cards (NICs) are adapters installed in a computer providing physical connection to a network. NIC provides a transparent interface enabling communication between the computer and the network. Systems supported by LANs are provided with an NIC designed exclusively for the LAN. They provide a dedicated connection to a network and control the information flow between the user and the network. The NIC breaks the data into packets and then transmits them. ISA, PCI and PCMCIA cards are examples of NIC cards.

9.8.1.1 Type of LAN Protocols

Three types of LAN protocols are Ethernet, FDDI and the Token Ring. Ethernet is a widely used LAN technology developed by Xerox. It uses a coaxial cable or a special grade of twisted pair wires to transmit data. Token Ring is another LAN technology, developed by IBM. Ethernet and Token Ring networks use different NIC cards. A FDDI is used to connect two or more LANs over large distances. An FDDI operates over a fiber optic cable at a speed of 100 Mbps.

9.8.1.2 Type of Cabling

The type of cabling determines the type of NIC used. The cabling system must meet current and future needs for data transmission, physical layout, and communication requirements. The commonly used cabling systems are coaxial cable and thin Ethernet. A repeater may be used for extending the distance between network segments. The repeater amplifies the signal received from one segment of a network and sends the amplified signal to other segments.

The Institute of Electrical and Electronic Engineers (IEEE) 802 Standards define the cable access method as a protocol that determines the order, priority, and amount of time that a computer can transmit on a shared cable. The IEEE Standards committee also defines addresses for NICs so that no two cards share the same address.

9.8.2 Repeater

A repeater is a networking device that transmits a weak or low-level signal at a higher level or at higher power. It amplifies or regenerates the signal received from one segment of a network and sends or retransmits the amplified signal to the other segments. A repeater ensures that the signal does not lose its strength while traveling over a network.

9.8.3 Bridge

A bridge is a device that is used for connecting multiple network segments. The increase in the network length may result in slowing down of the system due to increased access to the system. Therefore, the traffic can be reduced by splitting into smaller LANs that are connected through a bridge.

9.8.4 Routers

A router is a network device that routes data packets between two networks. A router interconnects individual computers or LANs, and WANs in order to make internetworking possible. It can take up decisions related to finding the best path for routing the data packet. It can also act as a firewall preventing unwanted packets from entering and leaving the network.

9.8.5 Hubs

Various LAN segments are connected through centralized points called hubs. Hubs are used to build structured cabling systems. A hub joins multiple computers together to form a network within which they communicate. A hub has a series of ports to which network cables are added. High-end hubs can support a variety of networks such as token ring and optical networks. Token Ring MAUs (Multistation Access Unit) and Ethernet hubs are examples of hubs.

9.8.6 Switch

A switch is a device that connects networks together. Switches have certain advantages compared to bridges related to their speed, price and superior performance. Bridge is used only to increase the length of the network where as a switch is used to distribute traffic from a network segment to the destinations which connect the segment to another network segment.

Check Your Progress-3

12. Which of the following statements is **false** regarding Network Interface Cards (NICs)?
    a. NICs are adapters installed in a computer that enable the physical connection to a network
    b. An NIC ensures a dedicated, full-time connection to the network

c. The NIC takes the data to be sent, breaks it into packets, and transmits the packets

d. NICs cannot be connected to hubs

13. A computer network infrastructure consists of components like network interface cards, cable, hub, and protocol. What are network interface cards?

14. Routers, Hubs, and Bridges are _____.

a. Network software

b. Software structuring techniques

c. Wireless networks

d. Networking devices

15. _____ can also act as firewalls preventing unwanted packets from entering and leaving the network.

16. _____ uses a coaxial cable or a special grade of twisted pair wires to transmit data.

a. FDDI

b. Bluetooth

c. Infrared LAN

d. Ethernet

9.9 Network Architectures and Protocols

The hardware, software and, communication channels require standards for the interfaces between them for the purpose of compatibility. To support the development of advanced data communication networks, computer manufacturers and international organizations developed standards called *protocols* and *network architectures*. A protocol can be defined as a standard set of rules and procedures for the control of communications in a network. The two important network architectures are the Open Systems Interconnection (OSI) reference model and the Transmission Control Protocol/Internet Protocol (TCP/IP) reference model.

9.9.1 The OSI Reference Model

The Open Systems Interconnection (OSI) reference model defines a framework of networking for the implementation of various protocols. In this model, the controls are passed on from one layer to the next. The model establishes communications between different devices. The OSI reference model is based on a proposal developed by the International Standards

Organization (ISO) as a first step towards the international standardization of protocols used in the various layers for communication between two end users in a network. The model is called the *OSI reference model* as it connects open systems (systems that are open for communication with other systems). There are seven layers in the OSI reference model. They are discussed below:

9.9.1.1 The Physical Layer

The physical layer helps in data transmission on the telecommunication media in the network. This layer takes care of the physical medium used between devices and provides the hardware required for sending and receiving data.

9.9.1.2 The Data Link Layer

The data link layer establishes an error-free communication path between computers over the physical channel. The data packets are split into frames and are transmitted sequentially by the sender. Using multiplexing, the data link layer simultaneously transmits two or more messages or signals over the same channel. The data packets are encoded and are decoded into bits in this layer.

9.9.1.3 The Network Layer

The network layer provides routing for establishing connections among network links. A logical path called as a virtual circuit is established between the source and the destination computers for sending data from one node to another. The routing and forwarding decisions are taken at the network layer. This layer ensures that the data packet is securely delivered to the destination.

9.9.1.4 The Transport Layer

The transport layer provides control standards for a communication session to enable the two processes at both ends to exchange data reliably and sequentially. This layer splits the message into smaller units and appends a header to each unit. The message flow between the sender and the receiver is also controlled. The transport layer ensures the complete transfer of data, error recovery, and flow control between the systems.

9.9.1.5 The Session Layer

The session layer provides means for establishing, maintaining, and terminating a dialogue or session between the two end users. The dialogue type is specified, initiated and the direction of message flow is regulated at the session layer.

9.9.1.6 The Presentation Layer

The presentation layer or syntax layer, provides appropriate data transmission formats and codes. It performs text compression, code conversion, or security encryption on the message. This layer tries to convert the data into a form that can be accepted and understood by the application layer.

9.9.1.7 The Application Layer

The application layer provides communication services for end-user applications. This layer is application-specific and user-oriented and provides services that directly support the end user of the network. It offers services like file transfers, remote file access, database management, etc. It also provides network transparency to the end users so that they are ignorant of the physical distribution of the various network resources being utilized by them.

9.9.2 TCP/IP Reference Model

The TCP/IP reference model is used as a de facto standard for transmitting data over the networks by the Internet. The four important TCP/IP layers are discussed below.

9.9.2.1 Network Interface Layer

Also known as the network access layer, it is the lowest layer in the TCP/IP reference model. This layer sends and receives the TCP/IP packets over a network medium. This way, TCP/IP connects different types of network. The network interface layer is similar to the data link and physical layers of the OSI reference model.

9.9.2.2 The Internet Layer

This layer manages the addresses of each message so that it reaches the right destination. The address is verified by the gateway and the packet is forwarded. The whole architecture of the TCP/IP reference model is held by the internet layer. It allows hosts to inject data packets into any network and make them travel independently to the destination.

The Internet layer defines an official packet format and protocol called IP (Internet Protocol). This layer delivers IP packets to their destination. This internet layer functions in a manner similar to OSI's network layer as packet routing is a major activity in both the layers.

9.9.2.3 The Transport Layer

It is the next layer in the TCP/IP reference model. Here the message is assembled into small data packets and are sent over the Internet to their destination. These packets are converted back into the original message at the TCP layer. This layer, like OSI's transport layer, allows peer entities

on the source and destination hosts to carry on the conversation. TCP and UDP are examples of two end-to-end protocols.

9.9.2.4 The Application Layer

This layer is just above the transport layer and contains all higher level protocols such as virtual terminal protocols (TELNET), file transfer protocols (FTP), etc. The virtual terminal protocol allows a user on one computer to log onto a distant machine and work there. The FTP enables efficient file transfer from one computer to another. Simple Mail Transfer Protocol is a specialized protocol for file transfer.

9.9.3 Switching Techniques

The various ways in which data can be transmitted over a network are called as switching techniques. Following are the three switching techniques used for building networks:

9.9.3.1 Circuit Switching

Circuit switching refers to a process in which a switch opens a circuit to establish a link between a sender and a receiver. The link remains open until the communication session is completed. This technique of switching is the most commonly used technique for building a communications network. Integrated Services Digital Network (ISDN) is an example of circuit switching.

9.9.3.2 Message Switching

In message switching, data is routed or transmitted from one switching device to another. This technique is also known as store-and-forward technique as the message to be transmitted is stored temporarily and then sent to its destination at a later stage. The message switching technique allows the routing of messages over the networks that are not accessible at all times. E-mail delivery is an example of message switching.

9.9.3.3 Packet Switching

Packet switching uses short messages. A message that exceeds the maximum length stipulated by the network is broken down or subdivided into fixed or variable groups called packets. These packets are then individually transmitted through the network to their destinations. Frame Relay packet switching protocol is used by large companies for their WANs. Asynchronous Transfer Mode (ATM) is an emerging packet switching technology that has high capacities and the capability to carry video, voice and data. Switched Multimegabit Data Services (SMDS) and Frame Relay are other examples of packet switching.

Network strategies have undergone a lot of changes in past few decades. A shift has been observed in the focus of networking strategy in organisations as newer and more advanced technologies are being introduced. Exhibit 9.1 presents the network architecture of 2021.

Exhibit 9.1: Network Architecture 2021

The evolution in network technologies over the past three decades can be broadly grouped under three key phases. These are characterised by functional imperatives viz. connectivity, bandwidths and application fingerprinting that have driven network strategies in these phases.

Connectivity

Network strategies during 1990s and early 2000s were focused on network connectivity. Protocol such as frame Network strategies during 1990s and early 2000s were focused on network connectivity. L2 networking protocols such as frame relay, ATM and ISDN were the primary focus for routers. Overall connectivity was based on complex devices and often led to poor experience. However, situations have improved with time. Today network connectivity is faster, reliable and much cheaper than it used to be, even in case of personal connectivity with the help of smartphones and consumer broadbands.

Bandwidth

The second phase during 2000s-2010s was characterised by advancements in network bandwidth. Telecommunication companies were now able to build better connections with the help of Optical and Dense Wavelength Division Multiplexing (DWDM) networks that led to increased bandwidths. Higher bandwidths encouraged more consumption of networking services which eventually resulted in mass production and lower costs.

Application networking

The third era in development of network architecture was mainly focused on application networking. Connectivity and bandwidth are more or less resolved issues, as most consumers today have easy access to network connectivity and sufficient bandwidths. Application networking emerged with the use of Software Defined WANs (SD-WAN) technologies and is enabled primarily through overlay networking. Users can now connect to applications from anywhere through networks, using any available bandwidth.

Source: https://packetpushers.net/network-architecture-2021-old-network-technologies-remain-relevant/

Example: Intel File Transfer System

Intel has an Intel File Transfer System (FTS) for the purpose of securely routing binary files between business groups owned by Intel and other trading entities (TEs). TEs use a protocol for sending files to any Intel

Contd. .....

business group. The file along with the metadata is routed to destination that is already configured within Intel. Intel business groups can also send files routed through the FTS to the TEs.

Intel has designed the architecture in a way that an Outsource Provider (OP) called Hubspan, a leading provider of Integration-as-a-service. Hubspan acts as an intermediary between the TEs using protocols such as HTTP, SMTP or FTP. Between Intel and OP, a single interface is maintained. This model provides flexibility to TEs for choosing a method for meeting their business needs. At the same time, the methods choosen by the TEs meet the Intel Information Security guidelines and requirements.

Adapted from "Intel File Transfer System," https://supplier2.intel.com/filetransfer/.

Check Your Progress-4

17. Two end-to-end protocols, TCP and UDP, are defined in the _____ of the TCP/IP reference model.
    a. Internet layer
    b. Transport layer
    c. Application layer
    d. Network layer

18. In the Open System Interconnection (OSI) reference model, which layer is responsible for ensuring error-free communication?

19. Explain the functions of the session layer in the Open System Interconnection (OSI) reference model.

20. Switching techniques refer to the various ways in which data can be transmitted over a network. Which among the following switching alternatives is most commonly used for building a communications network?
    a. Circuit switching
    b. Packet switching
    c. Message switching
    d. Hybrid switching

21. In the Transmission Control Protocol/Internet Protocol (TCP/IP) reference model, file transfer protocols, virtual terminal protocols, and electronic mail protocols are present in the _____ layer.

22. Which among the following statements is **false** regarding message switching?
    a. Data is routed or transmitted one block at a time from one switching device to another
    b. The message to be transmitted is sent to its destination soon after without being temporarily stored
    c. It allows routing of message over the networks that are not accessible at all times
    d. It is also known as store-and-forward technique

23. Which layer of the OSI Reference Model deals with splitting a message into smaller units?

24. Which among the following standard protocols is most widely used by networks like the Internet?
    a. HTTP
    b. TCP/IP
    c. SMTP
    d. SLIP

25. What is Asynchronous Transfer Mode (ATM)?

26. Organizations which are large in size and have wide area networks use
    a. Message switching
    b. Circuit switching
    c. Frame relay packet switching
    d. Asynchronous transfer mode

27. _____ layer in the Open System Interconnection (OSI) reference model is also called the syntax layer.

28. The application layer in the TCP/IP reference model contains protocols like virtual terminal protocols, file transfer protocols, and electronic mail protocols. What is the function of the File Transfer Protocol (FTP)?

9.10 Summary

- A computer network links two or more computers and enables sharing of information between them.

- The elements of a network include network interface cards (NICs), cable, hub, and protocol.
- Based on transmission, networks are classified into point-to-point networks and broadcast networks.
- Based on scale or size/area of networking, networks are mainly classified into Local Area Networks (LAN), Metropolitan Area Networks (MAN), and Wide Area Networks (WAN).
- Network operating systems are divided into peer-to-peer network operating systems and client/server network operating systems.
- The important network topologies are mesh, star, bus, ring, tree, and hybrid.
- Network infrastructure consists of various devices like Network Interface Cards (NICs), buses, repeaters, bridge, routers, hubs, and switches.
- Most networks support protocol hierarchies. The two important network protocols are OSI and TCP/IP protocols.
- Data can transmitted over a network through various switching techniques like circuit switching, message switching, and packet switching.

9.11 Glossary

- **Asynchronous Transfer Mode (ATM):** An emerging packet switching technology. These networks have high capacities, and are capable of carrying video, voice, and data.
- **Bandwidth:** The speed and capacity of telecommunication networks can be classified by their bandwidth. Bandwidth displays the frequency range of a channel and determines its maximum transmission rate and measured in bits per second (bps) and sometimes as bytes per second (Bps).
- **Bridge:** A device that is used for connecting multiple network segments. As a network grows larger, more resources need to be shared, and messaging within the network increases, which will slow down the network.
- **Broadcast networks:** These networks have a single communication channel which is shared by all the machines on the network. These networks are used for sending messages to all the computers on the network.
- **Bus topology:** A network topology that has a multipoint configuration and in which all the devices are linked through a long cable. This topology has a multipoint configuration.

- **Cable access method:** A protocol that determines the order, priority, and amount of time that a computer can transmit on a shared cable.
- **Circuit Switching:** Circuit switching refers to a process in which a switch opens a circuit to establish a link between a sender and a receiver. The link remains open until the communication session is completed.
- **Client/server network:** In this network, at least one computer acts as a server. A server is a computer that runs the software which enables it to perform jobs requested by other computers, called clients.
- **Computer network:** A computer network is a data communication system where two or more computers are linked in order to exchange data (documents, sheets, etc...) and share resources like (CD-ROM, Printers, Storage Devices).
- **Ethernet:** A widely used LAN technology developed by Xerox. It uses a coaxial cable or a special grade of twisted pair wires to transmit data.
- **GSM:** Stands for Global System for Mobile Communications. GSM is a standard for mobile phones that makes international roaming possible between users of mobile phones.
- **Hub:** Centralized points to which various segments in a LAN are connected. Hubs are used to build structured cabling systems. A hub joins multiple devices together to form a network within which they communicate.
- **Hybrid topologies:** A network topology in which two or more network topologies are used for connecting single network.
- **Local Area Network (LAN):** LANs are usually privately owned networks. Based on the requirements these types of networks can be limited to a single office, building, or campus. They are network computers that are in one location, or in nearby offices and buildings.
- **Mesh topology:** In mesh topology, every device is connected with every other device through a point-to-point link. The link is a dedicated one, i.e., the link carries traffic only between the two devices that are connected.
- **Message Switching:** In this, data is routed or transmitted one by one, from one switching device to another. This technique is also known as the store-and-forward technique as the message to be transmitted is stored temporarily and then sent to its destination at a later stage.

- **Metropolitan Area Network (MAN):** A larger version of LAN that interconnects several LANs located in a metropolitan area or even a small country. MANs have medium-sized networks and usually cover a radius of 2.5 to 25 kms.
- **Multiplexing:** A process through which a system is designated for transmitting two or more messages or signals simultaneously over the same channel.
- **Network Interface Cards (NICs):** Network Interface Cards (NICs) are adapters installed in a computer that enable its physical connection to a network. These cards provide a transparent interface between the computer and the network.
- **Open System Interconnection (OSI) reference model:** A framework of networking for implementing the various protocols. The controls in the OSI model are passed on from one layer to the next to through all the seven layers.
- **Packet Switching:** A switching technique that uses short messages. A message that exceeds the maximum length stipulated by the network is broken down or subdivided into fixed or variable groups called packets.
- **Peer-to-peer network:** An operating system that runs on the computers of individual users. The users can share files and resources on their computers and access the resources of other computers. They do not have a centralized management source or a file server.
- **Point-to-point networks:** These networks connect individual pairs of machines. In point-to-point network, the message has to first visit one or more intermediate machines to travel from source to destination. These networks are also called unicast or one-to-one networks.
- **Protocol:** A protocol is a set of rules regarding communication that ensures that all the computers in the network communicate in the same language.
- **Repeater:** A networking device which is used to retransmit a weak or low-level signal at a higher level or at higher power. Whenever a signal has to travel over large networks (or several small networks) the signal gradually loses its strength and may fail and data may be lost.
- **Ring topology:** In this type of network topology, each device is linked to two devices on either side of it through a dedicated point-to-point connection. Each device is physically linked to its immediate neighbors.

- **Routers:** A network device that routes data packets between two networks. Routers make internetworking possible by interconnecting both individual computers or LANs, and WANs.
- **SMS:** It stands for Short Messaging Service. It is a service available on mobile phones and other such devices that facilitates sending of short messages to others. Generally, SMS allows users to send text messages that are up to 160 characters long to mobile phones.
- **Star topology:** In this network topology, each device is connected to a centrally located hub through a point-to-point link. There is no direct link between the devices. Star topology does not permit direct traffic between the devices. The hub acts as an intermediary between the devices.
- **Switch:** A switch is a device that connects networks together. The superior speed and reasonable price of switches have made them more popular than bridges. A switch is used to distribute traffic from a source network segment to a destination network segment
- **Transmission Control Protocol/Internet Protocol (TCP/IP) Reference Model:** A network architecture that was developed to support multiple networks. The TCP/IP reference model was used as a de facto standard for the Internet to transmit data over the networks.
- **Tree Topology:** In this network topology, the nodes of some devices are linked to a centrally located hub. The other nodes are indirectly connected to the hub through a secondary hub, which in turn, is connected to the main hub. The central hub is called the active hub while the secondary hubs can be active or passive.
- **VPNs:** In order to provide security for business and corporates and government related services on the Internet a technology called Virtual Private Networks (VPNs) has been implemented.
- **Wide Area Network (WAN):** A computer network that covers a large geographical area, generally a country or a continent. These networks are used for connecting two or more LANs so that the users and computers of one location can communicate with those located at other places.
- **Wireless Application Protocol (WAP):** An international standard for wireless communication applications. WAP is a special protocol that allows mobiles to connect to the Internet. This standard allows access to the Internet from a mobile phone or a PDA.

- **Wireless network:** A computer or telephone network that connects users using radio waves. Wireless networks are the fastest growing networks and do not require wires to make the connection.

9.12 Self-Assessment Test

1. A computer network is a data communication system that links two or more computers and enables sharing of information between them. Describe the components and applications of computer networks.
2. Based on the scale of networking, networks are classified into local area networks, metropolitan area networks, and wide area networks. Describe these networks.
3. The basic types of network topologies are mesh, star, bus, ring, and tree. Explain these topologies.
4. The efficiency and reliability of any network is determined by its network infrastructure. Explain the devices used for achieving this.
5. The two important network architectures are the Open Systems Interconnection (OSI) reference model and the Transmission Control Protocol/Internet Protocol (TCP/IP) reference model. Explain these models in detail.

9.13 Suggested Readings / Reference Material

1. Introduction to Information Technology, V. Rajaraman, PHI learning, 2018
2. Information Technology for Management, 2ed: Advancing Sustainable, Profitable Business Growth, Turban, Volonino, Wood, O.P. Wali, Wiley India Pvt Limited, January 2021
3. Introduction to Information Systems - 6th edition, R. Kelly Rainer; John Wiley & Sons, Inc.2016
4. Information Technology: An Introduction for Today's Digital World, Richard Fox, Chapman and Hall/CRC; 2nd edition (August 21, 2020)
5. Information Technology for Management, Efraim Turban, Carol Pollard, Gregory Wood, Wiley, 2018

Additional References:

1. Critchley, L., Where Nanotechnology, the IoT, and Industry 4.0 Meet., https://www.mouser.com/blog/where-nanotechnology-the-iot-and-industry-40-meet, 2019
2. Pan India implementation of HMIS over Indian Railways, Ministry of Railways., http://railministry.com/pan-india-implementation-of-hmis-over-indian-railways/ 2020
3. Vossler, C. How Long Does It Take To Order A New BMW?

   https://www.bmwblog.com/2020/09/28/how-long-does-it-take-to-order-a-new-bmw/2020
4. Jay, A., 10 New ERP Trends & Forecasts for 2020/2021 – A Look Into What's Next. https://financesonline.com/erp-trends/2019
5. Gingiss, D., How Integrating Social Media Into The Rest Of The Business Will Increase Revenue., How Integrating Social Media Into The Rest Of The Business Will Increase Revenue (forbes.com), 2019

9.14 Answers to Check Your Progress Questions

Following are the answers to the Check Your Progress questions given in the unit.

1. A computer network can be defined as an inter-connected collection of autonomous computers that usually links two or more computers and enables sharing of information. Fundamentally, computer networks enable sharing of resources like documents, storage space, and printers. At more complex levels, these networks are used to exchange audio and/or video data, for instance, audio conferencing, video conferencing, etc.

2. **i/r, ii/p, iii/q**

Wireless networks are the fastest growing trend in networks – no wires are necessary to make the connection. Wireless networks use three types of protocols: Wireless Application Protocol (WAP), Short Messaging Service (SMS), and Global System for Mobile (GSM). While WAP is a special protocol that allows mobiles to connect to the Internet, SMS is a service of sending text messages of up to 160 characters length to mobile phones and GSM is the most widely used mobile telephone protocol world over.

3. A protocol is a standard set of rules and procedures for the control of communications in a communications network. It consists of a

set of rules regarding communication that ensures that all the computers in the network communicate in the same language.

4. (c) LAN<MAN<WAN

Local Area Network (LAN) covers a limited area such as office, classroom, manufacturing plant, etc. A Metropolitan Area Network (MAN) is a bigger version of a LAN and normally uses similar technology. It evolved primarily to permit LANs to interconnect across a metropolis or a city. A Wide Area Network (WAN) covers a large geographical area, generally a country or a continent.

5. Local Area Network

LANs are usually privately owned networks. Based on the requirements, this type of networks can be limited to a single office, a building, or campus. They network computers that are in one location, or in nearby offices and buildings.

6. Virtual Private Networks (VPNs) use public networks like the Internet to connect employees located at distant offices or remote areas. These networks use virtual connections which are routed through the Internet from the company's network to the remote area. VPNs enable a geographically dispersed group of hosts to interact with the help of an extended LAN in a highly secure virtual environment.

7. Local Area Networks

Networks are of different types and serve different purposes. These are local area network (LAN), metropolitan area network (MAN), and wide area network (WAN). LANs are privately-owned networks which are used in a single building or an office to connect various workstations and share other resources such as printers, files, etc.

8. Statements (a), (c), (e), (g), (h), (k), and (m) are true while statements (b), (d), (f), (i), (j), and (l) are false.

VPNs are built and maintained by Internet Service Providers (ISPs). In point-to-point network, to travel from source to destination, a packet may have to first visit one or more intermediate machines. These networks consist of many connections between individual pairs of machines. Larger networks use point-to-point technology. A WAN is a computer

network that covers a large geographical area, generally a country or a continent. The Internet is an example of a WAN.

9. (c) They have a centralized management source or a file server

A peer-to-peer operating system runs on the computers of individual users. The users can share files and resources on their computers and access the resources of others. However, they do not have a centralized management source or a file server. These networks are designed for small to medium local area networks.

10. Topology refers to the physical or logical layout of a network. The topology of a network is a geometric representation of all devices connecting to each other in the network. Mesh, star, bus, ring, and tree are the basic types of network topologies or structures.

11. (c) Hub

The various network topologies are ring, star, bus, mesh, tree, and hybrid topologies. Hub is the hardware that controls traffic in a network.

12. (d) NICs cannot be connected to hubs

NIC are adapters installed in a computer that enable its physical connection to a network hub through a cable.

13. NIC are adapters installed in a computer that enable its physical connection to a network hub through a cable. They provide a transparent interface between the computer and the network and enable the former to communicate with the latter. They control the flow of information between the user and the network and are used for sending and receiving messages from other computers.

14. (d) Networking devices

The efficiency and reliability of any network depends on the network infrastructure. The networks have to be managed and controlled well to enable the efficient use of available bandwidth. For this purpose, various kinds of devices called networking devices are used. These are network interface cards, buses, repeaters, bridges, routers, hubs, gateway, and switches.

15. Routers

Routers build inter-networks, in which networks that use different topologies and protocols can be connected. They determine the best path for delivery on the network and can act as a firewall preventing unwanted packets from entering and leaving the network.

16. (d) Ethernet

Networks are broadly divided into Ethernet and Token Ring. Ethernet, developed by Xerox is a widely used LAN technology. It uses a coaxial cable or a special grade of twisted pair wires to transmit data.

17. (b) Transport layer

The transport layer of the TCP/IP reference model allows peer entities on the source and destination hosts to carry on a conversation. This layer defines two end-to-end protocols: transmission control protocol (TCP), and user datagram protocol (UDP).

18. Data link layer

In the OSI reference model, the data link layer is responsible for establishing an error-free communications path between computers over the physical channel.

19. In the Open System Interconnection (OSI) reference model, the session layer provides the means for establishing, maintaining, and terminating a dialogue or session between two end users. It specifies the dialogue type (one way, two way alternative, or two way simultaneous), initiates a dialogue, and even regulates the direction of message flow.

20. (a) Circuit switching

Circuit switching refers to a process in which a switch opens a circuit to establish a link between a sender and a receiver. The link remains open until the communication session is completed. A regular landline telephone service works on circuit switching.

21. Application

The most important TCP/IP layers are the Internet layer, the transport layer, the network layer and the application layer. The application layer is just above the transport layer. It contains all the higher level protocols like the virtual terminal protocols, file transfer protocols, and electronic mail protocols.

22. (b) The message to be transmitted is sent to its destination soon after without being temporarily stored

In message switching, data is routed or transmitted one block at a time from one switching device to another. This switching technique is also known as store-and-forward technique as the message to be transmitted is stored temporarily and then sent to its destination soon after.

23. Transport Layer

In the OSI reference model, the transport layer provides control standards for a communication session to enable two processes at both ends to exchange data sequentially in a reliable manner. This layer splits the message into smaller units and appends a header to each unit. It also controls message flow between the sender and the receiver so that a 'fast' sender does not overwhelm a 'slow' receiver with too much data.

24. (b) TCP/IP

Transmission Control Protocol/Internet Protocol (TCP/IP) reference model was used in Advanced Research Projects Administration Networks (ARPANET) and is presently used by the Internet, as well as intranets, and extranets. The TCP/IP reference model is used as a de facto standard for transmitting data over networks via the Internet. It uses several protocols, the major ones being TCP and IP.

25. Asynchronous Transfer Mode (ATM) is an emerging, high-capacity cell switching technology. An ATM switch breaks voice, video, and other data into fixed cells of 53 bytes and routes it to its next destination on the network. Companies which require fast, high-capacity multimedia capabilities for voice, video, and data communication are developing ATM networks.

26. (c) Frame relay packet switching

Frame relay handles the heavy telecommunication traffic of interconnected local area networks within a company's wide area network. It is often used by large companies for their WANs.

27. Presentation

The presentation layer provides appropriate data transmission formats and codes. It performs text compression, code conversion, or security encryption on the message. This layer tries to convert

the data into a form that can be accepted and understood by the application layer.

28. It enables users to move data efficiently from one computer to another

The application layer is situated above the transport layer in the TCP/IP reference model. The various protocols that the layer supports are: virtual terminal protocols, file transfer protocols, and electronic mail protocols. The file transfer protocol enables users to move data efficiently from one computer to another.

Unit 10

Telecommunication Networks

Structure

10.1 Introduction

In the previous unit, we have discussed that it is possible to share information if the computers are interconnected. However, mere interconnectivity would not facilitate transfer of information. The computer systems need telecommunication networks to aid such a transfer. In this unit, we introduce you to telecommunication networks.

Telecommunication involves transmission of signals for the purpose of facilitating communication. A telecommunication system obtains information from the transmitter and converts it into a signal for transmission using a medium; this signal is received by the receiver at the other end. A telecommunication network consists of links through which messages are transferred from one part of the network to another.

Over the years, there has been a rapid change in the telecommunication technology that brought about a significant change in long distance communication. With the advent of the internet, intranet, VSAT (Very Small Aperture Terminal), ISDN (Integrated Services Digital Network),

ADSL (Asymmetric Digital Subscriber Line), cellular telephony, video conferencing, etc., distance has not been a barrier to communication.

In this unit, we provide you with a basic understanding of various aspects of telecommunications like enterprise-wide collaboration, electronic and mobile commerce, and internal business systems. Apart from this, we would also explain you about the various media used in information transmission and the manner in which telecommunication helps build internetworked enterprises.

10.2 Objectives

By the end of this unit, you should be able to:

- Explain the telecommunication system, its service providers, the transmission media, and the regulatory authorities.
- Determine the business applications of telecommunications.
- Summarize the applications scope, technology scope, and industry scope of telecommunication networks.
- List the key telecommunication network components.
- Identify the telecommunications equipment used in WAN.
- Discuss about network convergence and its advantages.

10.3 The Telecommunication System

The first telecom network was the Public Switched Telephone Network (PSTN). The PSTN used analog technology. In the PSTN, the data to be transmitted is carried by copper wires in the form of electrical pulses from one end to another. The services provided by a PSTN is sometimes referred to as the Plain Old Telephone Service (POTS).

With the advent of telecommunication networks like ISDN and Fiber Distributed Data Interface (FDDI), there has been a shift from analog to digital technology. The telecommunication system comprises of several elements like service providers, the transmission media, and standards and regulations/regulatory authorities.

10.3.1 Service Providers

Wireline telephone companies can be divided into two categories such as the local exchange carriers (LECs) and the long distance carriers, also known as the interexchange carriers (IXCs).

10.3.1.1 Local Exchange Carriers

Local Exchange Carriers (LECs) provide wired services for subscribers. Lines that are drawn from homes and business organizations end at local exchanges. These local exchanges are connected to other local exchanges within a network of local connections called local access and transport area (LATA), or to long distance carriers. Without accessing a long-distance carrier, a subscriber can place a call within a LATA.

A LATA is a network of local connections, where the subscriber can place a call without accessing a long-distance carrier. LATAs are linked to Subscriber Loop Carriers (SLCs) by means of copper wires called twisted pair cables. The area served by an SLC is called the local loop.

10.3.1.2 IntereXchange Carriers (IXCs)

An IntereXchange Carrier (IXC), also known as the long distance carrier, provides long distance communication services ouside the LATA. For example, in the United States, Verizon and Sprint are IXCs.

10.3.2 Transmission Media

Transmission media includes all the physical media such as twisted-pair wires, coaxial cables, optical fibres, terrestrial microwaves, communication satellites, cellular phone systems, and wireless LANs through which the data is transmitted between the computer systems. A technique called multiplexing has been developed for long distance transmission. Multiplexing is a process through which a system is designated for transmitting two or more messages or signals simultaneously over the same channel.

10.3.3 Standards and Regulations/Regulatory Authority

The telecommunication standards are governed by the International Telecommunication Union (ITU), previously called as the Consultative Committee for International Telephony and Telegraphy (CCITT). The ITU was established on May 17, 1865 in Paris and is headquartered at Geneva, Switzerland. The ITU is divided into three sectors. Telecommunication Standardization or ITU-T, Radio communication or ITU-R, and Telecommunication Development or ITU-D.

ITU is an international organization, where the government and private companies coordinate with each other to facilitate global telecommunication networks and services and develop communications technology. It also standardizes, regulates, and coordinates the various radio and telecommunication standards to make communication easy across various countries. The regulatory authority for telecommunications in India is the Telecom Regulatory Authority of India (TRAI).

Check Your Progress-1

1. IntereXchange Carrier (IXC) comes into picture when
   a. The user places a call inside the local transport area
   b. The user places a call outside the local transport area
   c. The user places a call inside or outside the local transport area
   d. None of the above
2. _____ provide wired services to the subscribers.
3. Expand PSTN.
4. Plain Old Telephone Service (POTS) refers to
   a. The technology used in ISDN
   b. The technology used in FDDI
   c. The service provided by the Public Switched Telephone Network (PSTN)
   d. The name of the service provider which offered PSTN connections
5. The telecommunication standards are governed by the Consultative Committee for International Telephone and Telegraphy (CCITT). CCITT is currently known as
6. What is local loop?
7. Explain the role played by the International Telecommunications Union (ITU). What are the various sectors in the ITU?
8. Which is the regulatory authority for the telecommunications industry in India?
9. ISDN stands for
   a. Integrated Services Digital Network
   b. Indian Services Digital Network
   c. Integrated Services Dialup Network
   d. Indian Services Dialup Network
10. Multiplexing is
    a. The ability to execute more than one task i.e., a program at the same time
    b. A mechanism by which the output of one command is passed on as input to another command
    c. The ability of hardware or software to handle the processing demands of a wide range of end users, transactions, queries, and other information processing requirements

d. The process of designating a system for the simultaneous transmission of two or more messages or signals over the same channel

10.4 Telecommunications in Business

Telecommunications help companies overcome geographic, time, cost, and structural barriers. For example, intranets, extranets, the Internet and organizational networks can be used for transmitting relevant information. They also provide better customer service by reducing the delay in fulfilling orders and speeding up the billing process. Moreover, they minimize the cost incurred by traditional means of communication by eliminating the need for expensive business trips and simplifying the business processes. Teleconferencing allows customers, suppliers, and employees to participate in meetings and collaborate on joint projects. Telecommunications can be used in internal business systems, enterprise collaboration systems, electronic commerce, and mobile commerce systems.

10.4.1 Internal Business Systems

Internal business applications of telecommunications include internal transaction processing, process control, inquiry processing, activity monitoring, workflow systems and management support systems. For example, a company can link systems through local area networks (LAN) enabling managers to query the corporate database and generate reports.

10.4.2 Enterprise Collaboration Systems

Enterprise collaboration systems are information systems using a variety of information technologies enabling people to work together. They facilitate communication, coordination, and collaboration among the members of an organization and workgroups. It uses tools such as groupware, intranets, extranets, and the Internet for enhancing communication in the organization. E-mail and video conferencing are examples of enterprise collaboration systems.

10.4.3 Electronic Commerce

Electronic commerce applications support the buying and selling of products and services, electronic data interchange, electronic banking, interactive marketing, and supply chain management.

10.4.4 Mobile Commerce Systems

Mobile commerce systems, also known as M-Commerce, is a form of electronic commerce that allows buying and selling of products/services through wireless handheld devices like cell phones or personal digital

assistants using the Internet. The M-Commerce is facilitated with the help of Wireless Application Protocol (WAP). Users can view news services, stock and banking applications with the help of m-commerce.

Activity: ABC Ltd., a company that manufactures and sells refrigerators and air conditioners is looking for new ways to sell its products. How useful would a mobile commerce system prove to be, for the company? Explain.

Answer:

Activity: XYZ Ltd., an Internet Service Provider is facing problems related to customer service. Describe the ways in which modern telecommunications technologies can help the company improve its customer service.

Answer:

10.5 Scope of Telecommunication Networks

Telecommunication networks provide various communication tools to managers that help them take timely decisions. The scope of telecommunications can be categorized into applications scope, technology scope, and industry scope.

10.5.1 Applications Scope

Telecommunication networks are used in e-commerce, enterprise collaboration and also business functions such as operations and distribution. An organization's telecommunication network is used for office communication and as local and global computer networks that help to minimize costs, share resources, look for new customers and suppliers, and develop new products and services.

10.5.2 Technology Scope

Technology in the telecommunication industry is witnessing rapid change. The analog technology used earlier was slow and prone to errors and noises, but the new digital technology is reliable, accurate, provides significantly greater speed of transmission, transfers huge amounts of information, and also transmits data, audio, and video.

10.5.3 Industry Scope

With the advent of the Internet, telecommunication industry grew at a rapid rate. They provide various other services like video telephony, videoconferencing, interactive TV, video-on-demand, cable TV, etc. Liberalization and opening up of various economies in the early 1990s was also responsible for the growth of the industry. The industry became fiercely competitive with the influx of new players. Globally, the industry is moving towards a greater number of competitive vendors, carrier alliances and network services. The opening up of cellular and basic telephone services to private players also led to revolutionary changes in the Indian telecom sector.

The telecommunication industry continues to be at the forefront of innovation, expansion, and new prospects, ensuring that it remains a highly competitive market for all. Exhibit 10.1 details the changing trends of telecom in 2021.

Exhibit 10.1: Game-Changing Telecommunication Trends to Follow in 2021

The telecommunications sector has undergone lot of transformations in the year 2021 and is no longer what it was prior to the pandemic. In order to continue their position in the market and recover from the global pandemic, telecom companies are expanding, re-building policies, and developing better and advanced technologies. Given below are the key telecommunication trends to follow in 2021:

1. Expansion of 5G: Apart from offering ultrafast speed, 5G also empowers users, both business and consumers with reliable, low latency communications at lower costs. It will also act as a key enabler for all IoT deployment that requires fast machine-type communications.
2. Machine Learning and Artificial Intelligence (AI): Adoption of machine learning and AI in the telecom sector will continue to grow. Predictive maintenance, network optimization, chatbots, virtual assistants, and conversational AI to improve customer experience will be the key focus areas for these technologies.

Contd. .....

3. Edge computing: Edge computing refers to moving computation from data centres to the network's edge. This is made possible with the help of intelligent objects and network gateways that provide dependable and effective cloud services. Telecoms can use edge computing to accelerate data processing in real time via linked devices.
4. DevOps: DevOps is a set of practices that focus on software development and IT operations. Telecom companies are adopting DevOps for better network control, no manual configuration, real-time orchestration, shorter time to market for new applications, and more efficiently implementing highly scalable solutions.
5. Internet of Things(IoT): As more and more businesses shift towards intelligent factories and and IoT devices, telecom companies will have to adopt mechanisms that can enable devices to connect people, data, and processes with the help of technologies like 5G.
6. High resolution content transfer: There is an increased usage of better quality and large size content today, due to availability of better bandwidths. Telecom networks are now building capacities for high-speed and low-latency transmissions of such data which includes media such as videos, music, images as well as augmented reality and virtual reality content.

Source: https://www.business2community.com/communications/game-changing-telecommunication-trends-to-follow-in-2021-02407459

Check Your Progress-2

11. What are mobile commerce systems?
12. Process control, inquiry processing, activity monitoring, workflow systems, and management support systems are all part of the
    a. Mobile commerce systems
    b. Enterprise collaboration systems
    c. Internal business systems
    d. All of the above
13. What are the application areas of telecommunication networks?

10.6 Telecommunication Network Components

A telecommunications network is network that allows transfer of messages from one part of the network to another through various other links and

nodes. The key telecommunication network components are (a) Media (b) Terminals (c) Computers (d) Channels (e) Processors and (f) Software.

10.6.1 Telecommunication Media

The data is transmitted and received in a telecommunications network through telecommunication media. It is classified as Guided Media and Unguided Media.

10.6.1.1 Guided Media

Guided media, also called as bounded media, consists of a transmission line through which the waves are guided. The guided media consists of a cabling system that guides the signals through a specific path. These include open wires, twisted-pair wires, coaxial cables, and fiber optic cables, which provide physical link to the devices in a network. In guided media, transmission of data depends on the medium of transmission used.

Open wires: Open wires (similar to electrical wires hung between poles) do not have protection or shielding from noise or other kinds of interferences.

Twisted-pair wires: Twisted-pair wires are mostly used in star and hub networks and transmit both voice and data. A twisted-pair wire consists of copper wires that are twisted into a pair in order to reduce noise. Twisted-pair wires provide high-speed data transmission, when they are twisted to form cables.

Coaxial cables: Coaxial cables are mostly used for LAN and bus networks. A coaxial cable consists of a copper or aluminum wire wrapped with spacers to insulate for minimizing interference from other signals and protect the cable. Coaxial cables provide high speed data transmission and can be used as an alternative to twisted-pair wires in high-service metropolitan areas for cable TV systems.

Optical fibres: Optical fibres use cables consisting of one or more hair-thin filaments of glass fiber wrapped in a protective jacket. They can conduct light pulses generated by lasers at transmission rates as high as 30 billion bits per second. Optical fibres have the highest bit rates among all the telecommunications media.

10.6.1.2 Unguided Media

Unguided media, also called as unbound media or wireless media, is a medium through which the transmission and reception takes place through an antenna. Though they have a means for the signals to travel, the signals are not guided through a specific path. Here, the data signals are not bound by a cabling system as in the guided media. Unguided media include terrestrial microwaves, communication satellites, cellular phone systems, and wireless LANs, which use microwave and other radio waves.

Terrestrial microwaves: Terrestrial microwaves cover long distances and involve earthbound microwave systems that transmit high-speed radio signals between the relay stations spaced approximately at 30 miles. Microwaves provide enough bandwidth to support voice channels and a few television channels.

Communication satellites: Communication satellites are special applications of microwaves. They use microwave radio as a telecommunication medium and are placed in stationary geosynchronous orbits. These are powered by solar panels and transmit microwave signals at a rate of several hundred million bits per second. Communication satellites are mainly used for voice and video transmission where high speed transmission of large volumes of data is required. Some large corporations have developed satellite networks, using satellite dish antennas known as VSAT (Very Small Aperture Terminal) in order to connect their distant work areas.

Cellular phone systems: Cellular phone systems use a number of radio communication technologies and enable placing of calls through the wireless communication systems. The transmitted signals are coordinated and controlled by the computers and other communication processors as the users move from one place to another.

Wireless LANs: A wireless LAN is a network that links two or more computer systems. Wireless LANs may involve a low frequency radio technology called spread spectrum or a high frequency radio technology similar to a digital cellular technology. Digital cellular technology refers to the use of digital technology in cellular networks. Spread spectrum is the technology which is widely used in the wireless LANs and is based on radio waves that facilitate communication between computer systems or devices within a limited area.

Infrared is the other wireless LAN technology used in wireless LANs. This technology uses infrared rays for establishing network links between various components of a LAN.

Infrared systems: The infrared systems use the infrared rays for transmitting and receiving data.

10.6.2 Terminals

An input/output device using a telecommunication network to transmit or receive data is called a terminal. The devices used include telephones, office equipment, etc.

10.6.3 Computers

In order to enhance information processing, computers are connected to the telecommunication network. A mainframe may serve as a host computer in large networks and a micro computer can serve as a small server in small networks.

10.6.4 Communication Channels

A communication channel refers to a path taken by data to accomplish the communication task. The channel carries signals from a sender to the receiver through predefined routes. The data transmission is performed using either digital data signals or analog data signals. Switching techniques and bandwidth are important aspects of communication channels.

10.6.5 Transmission Systems

There are two types of transmission systems – analog transmission systems and digital transmission systems. In analog transmission systems, signals are transmitted through a medium as continuously varying electromagnetic waves. In a digital transmission system, the signals are transmitted as discrete pulses that are measured in bits per second and are similar to those transmitted in computer systems. This speeds up the process of transmission and thus enables the movement of larger amounts of information with greater economy compared to analog systems. Digital transmissions encounter problems related to attenuation, i.e., the signal strength decreases. Other problems include high power requirement.

10.6.5.1 Switching Techniques

Switching techniques are important in telecommunication networks. They are of three types namely, circuit switching, message switching, and packet switching.

10.6.5.2 Bandwidth

The bandwidth specifies the speed and capacity of telecommunication networks, displays the frequency range of a channel, and determines its maximum transmission rate. Bandwidth refers to the amount of data that

can be transmitted within a fixed period of time. It is usually measured in bits per second (bps) and sometime as bytes per second (Bps).

10.6.6 Telecommunication Processors

Telecommunication processors help in data transmission and reception between the terminals and computers. The functions of processors include converting data from digital to analog and back, coding and decoding of data, and increasing speed, accuracy, and efficiency of the communications flow. Modems and multiplexers, switches, routers, hubs, and gateways are examples of telecommunication processors.

10.6.6.1 Modems

The word 'modem' is derived from modulation and demodulation. A modem is a device that converts digital signals from input/output devices into appropriate frequencies at a transmission terminal, and converts them back into digital signals at a receiving terminal. A modem is used for sending digital data over a telephone network. The modulator converts the data into a signal suiting the telephone line while the demodulator re-converts the signal back into digital data. Other functions supported by modem include transmission error control, automatic dialing and faxing, etc.

10.6.6.2 Multiplexers

A multiplexer enables a single communication channel to carry simultaneous data transmissions from numerous terminals. It merges transmission signals received from various terminals at one end of the communications channel and a similar device called demultiplexer separates the individual transmissions at the receiving end. Multiplexing includes frequency division multiplexing, where a high speed channel is divided into multiple slow-speed channels, and time division multiplexing, where the multiplexer divides the time during which each terminal can use the high-speed line into very short time slots or time frames.

10.6.6.3 Internetworked Processors

The efficiency and reliability of an internetwork depends on the network infrastructure like switches, hubs, routers, etc. An internetwork is a collection of interconnected networks. These networks connect two or more dissimilar telecommunication networks into an internetwork using specialized processors called internetworked processors. Switches, routers, hubs, and gateways are the various internetworked processors.

10.6.7 Telecommunication Software

The telecommunication software includes programs for controlling telecommunication activities and managing the functions of the telecommunications network. It facilitates communication between diverse networks. The software packages provide various communication support services like network and security management; these packages control transmission, access, and errors in the process.

a) Network and security management: The telecommunications networks provide protection to the communications network from unauthorized access. Data transmissions are protected by encryption techniques. Data is scrambled into a coded form before transmission and decoded upon arrival.

b) Controlling transmission, access, and errors: The telecommunications software allows users to send and receive messages, commands, data, etc. The software works with the communication processor to establish communication parameters like mode, speed, and direction of the transmission. This also involves automatic telephone dialing and redialing, logging on and off with appropriate account numbers and security codes, and automatic answering of telephone calls from another computer.

   The distortions in communication channels may result in errors like noise and power surges. The telecommunications software detects and corrects the error using the process of retransmissions. Retransmission is a process in which signal is sent back to the computer so that it can retransmit those data units that are error prone.

Check Your Progress-3

14. An example of a wireless LAN technology is _____.
    a. Ultraviolet LAN
    b. Infrared LAN
    c. ISDN
    d. Both (a) and (b)

15. _____ are transmitted as continuously varying electromagnetic signals

16. Spread spectrum is a
    a. Guided Media Technology
    b. Low frequency radio technology
    c. Wireless LAN technology
    d. Both (b) & (c)

17. What is a modem? Explain its functions.

18. Which of the following telecommunication media can conduct light pulses generated by lasers at transmission rates as high as 30 billion bits per second?
    a. Coaxial cables
    b. Fiber optics
    c. Twisted-pair wires
    d. Territorial microwave

19. A communications medium that uses pulses of laser light in glass fibers is referred to as _____.

20. The various media over which data is transmitted and received are called telecommunication media. Which among the following has the highest bit rates among all the telecommunications media?
    a. Open wires
    b. Twisted-pair wires
    c. Coaxial cables
    d. Optical fibers

21. What is a terminal?

22. _____ signals are transmitted as discrete pulses.

23. Which of the following telecommunication media is mainly used for voice and video transmission where high speed transmission of large volumes of data is required?
    a. Communications satellites
    b. Twisted pair cables
    c. Wireless LANs
    d. Cellular phone systems

24. Bandwidth is measured in

25. Telecommunication media refers to the various media over which data is transmitted and received. Which among the following telecommunication media is mostly used in star networks and hub networks?

    a. Open wires
    b. Twisted-pair wires
    c. Coaxial cables
    d. Optical fibers

26. Explain the functions of telecommunication processors.

27. What is guided media and unguided media?

28. Which of the following uses a microwave radio as a telecommunication medium and is placed in stationary geosynchronous orbits?

    a. Territorial microwave
    b. Communications satellites
    c. Cellular phone systems
    d. Wireless LANs

29. Explain briefly about the various internetworked processors.

10.7 Telecommunications Equipment for WAN

Wide Area Network (WAN) refers to a computer network that spans a region, nation, or a global area. These networks connect the local area networks or the metropolitan area networks. Multinational companies use WANs to connect their offices in different countries and states to facilitate communication and information. The most popular WAN is the Internet. WANs have a local access number, called point of presence (POP), and long distance trunk lines, called backbones.

10.7.1 Point of Presence (POP)

A Point of Presence refers to a city or a location that connects one network with another. It is a point at which a line from the IXC is connected with the line from the local telephone company (or to the user in the absence of a telephone company).

10.7.2 Backbones

The backbone carries the traffic on the communications network and is also used for joining the LANs together. LANs are connected to the backbone through bridges and/or routers. The backbone serves as a communication highway for LAN-to-LAN traffic.

10.7.3 Uses of Wan

10.7.3.1 LAN-to-LAN Connections

A LAN cannot link computers at diverse geographical locations. Instead, a WAN can be used for connecting LANs at two or more geographically separated locations.

10.7.3.2 Acquiring Transactions

The information about a transaction can be instantly transmitted to control computers using a WAN. When a transaction occurs, the point of sale terminal is linked to a data communication network and the data acquired is collected for recording purposes.

10.7.3.3 Electronic Data Interchange (EDI)

Electronic Data Interchange (EDI) sets standards for B2B transactions and speeds up the process of ordering, invoicing, and payments. For example, the set of interchanges between a buyer and a seller like Request for Quotation (RFQs), purchase order, etc.

Check Your Progress-4

30. What are backbones?
31. Which of the following transactions are supported by electronic data interchange?
    a. Mobile commerce
    b. Business-to-business
    c. Business-to-consumer
    d. Consumer-to-consumer
32. What is a point of presence (POP)?

10.8 Network Convergence

Network convergence refers to an integration of data, voice, and video into a single network. It is also called as converged communications, data and

voice integration, voice and data convergence, etc. Planning plays a vital role in achieving network convergence. Companies can merge their separate networks into a single network infrastructure for transporting and connecting to the entire communications traffic.

Converged networks provide multiple services that are especially useful for companies that run multimedia applications like video collaboration, distance learning, and unified messaging systems.

Converged networks enable the centralized management of the network, and lower the costs involved in maintenance and support. Converged networks make use of a variety of wired and wireless devices. These devices enable creation of a new level of customer care and service that is integrated in nature. Converged networks help in enhancing the productivity of the organization. These networks also provide scope for collaboration within and outside the organization.

For an organization that uses converged networks, it is much easier to make changes in the functions (i.e., change, add, delete) when there is a change in the staff (i.e., relocated, appointed, downsized). Converged networks provide various options to the organization to make use of new equipment by replacing the existing ones that have become obsolete or those which are on lease. In converged networks, a single communication network infrastructure is used. This reduces the costs involved in the network bandwidth. This network infrastructure also reduces the administration costs and the cost of integrating applications. All these in turn enable reduction of the total ownership cost.

10.9 Summary

- Telecommunication refers to any transmission, emission, or reception of signals, through wires, radio, optical, or other electromagnetic systems.
- The telephone companies in the US include the local exchange carriers (LECs) and the interexchange carriers (IXCs).
- The telecommunication industry is regulated by International Telecommunications Union (ITU), a division of the United Nations, headquartered at Geneva, Switzerland.
- The business applications of telecommunication include Enterprise Collaboration Systems, Electronic Commerce Systems, and Internal Business Systems.
- Telecommunication networks provide various communication tools to managers that help them take timely decisions. The scope of

telecommunications can be categorized into applications scope, technology scope, and industry scope.

- The basic components of a telecommunication network are media, terminals, computers, channels, processors, and software.
- Wide Area Network (WAN) refers to a computer network that spans a region, nation, or a global area.
- Network convergence combines the communication channels of data, voice, and video and integrates into a single network infrastructure. This enables centralized network management and in turn reduces the maintenance and support costs.

10.10 Glossary

- **Analog transmission systems:** These systems use signals which are replicas of sound waves or the images that are being transmitted. These signals are transmitted through a medium as continuously fluctuating electromagnetic waves.
- **Backbone:** It carries the most traffic on the communications network. It is also used for joining the LANs together (either within a building or across a city or a country). The backbone serves as a communication highway for LAN-to-LAN traffic.
- **Coaxial cables:** Coaxial cables are mostly used for local area networks and in bus networks. A coaxial cable consists of a copper or aluminum wire wrapped with spacers to insulate and protect the cable.
- **Communication channel:** The path taken by data to accomplish the given communication task. Communication channels carry signals from sending stations to receiving stations along predefined routes. The data transmission is performed using either digital data signals or analog data signals.
- **Communication satellites:** These are special applications of microwaves. Satellites use microwave radio as a telecommunication medium but are placed in stationary geosynchronous orbits. Communication satellites are mainly used for voice and video transmission where high-speed transmission of large volumes of data is required.
- **Digital cellular technology:** This technology refers to the use of digital technology in cellular networks. Digital cellular technology enables wireless Internet applications, provides better service quality, and creates more security for the customer.

- **Guided media:** Also called bounded media, it consists of a solid medium called a transmission line through which the waves are guided. For guided media, transmission of data depends largely upon the transmission medium used.
- **Infrared technology:** A wireless LAN technology used in wireless LANs. The infrared technology uses infrared rays to establish network links between the various LAN components.
- **IntereXchange Carriers (IXCs):** Also known as the long distance carrier, IXCs come into play when a user places a call outside the local transport area. These IXCs provide long distance communication services within the LATA.
- **International Telecommunication Union (ITU):** Previously called as the Consultative Committee for International Telephony and Telegraphy (CCITT), the ITU governs the telecommunication standards. It ensures that there is a smooth transmission of calls placed by subscribers of two different companies and to help in creating a global communications network. It standardizes, regulates, and coordinates the various radio and telecommunication standards so that people belonging to one country can communicate with those living in another.
- **Local Access Transport Area (LATA):** A network of local connections. Within LATA, the subscriber can place a call without accessing a long-distance carrier.
- **Local Exchange Carriers (LECs):** These exchanges provide wired services for subscribers. A local exchange is the place at which the lines are drawn from homes and business organizations.
- **Mobile Commerce (M-Commerce):** A form of electronic commerce in which products and/or services are bought and sold through wireless handheld devices like cell phones or personal digital assistants using the Internet. This e-commerce application of telecommunication enables users to access the Internet without cables or wires, through their mobile devices.
- **Modem/Modulation and demodulation:** A device that converts digital signals from input/output devices into appropriate frequencies at a transmission terminal, and converts them back into digital signals at a recurring terminal. A modem is generally used for sending digital data over a telephone network.

- **Multiplexers:** A communication processor that enables a single communication channel to carry out simultaneous data transmissions from numerous terminals. It combines or merges the transmission signals that are received from various terminals at one end of the communications channel.
- **Network convergence:** It refers to an integration of data, voice, and video into a single network. Network convergence is also called as converged communications, data and voice integration, voice and data convergence, etc.
- **Point of Presence (POP):** It refers to a city or a location where one network can be connected to another. It is a point at which a line from the intereXchange carrier is connected with the line from the local telephone company or to the user in the absence of a telephone company.
- **Public Switched Telephone Network (PSTN):** The first telecom network that was developed. The PSTN telephone lines used analog technology. In this network system, copper wires carry the transmitted data in the form of electrical pulses from one end to another.
- **Retransmission:** A process in which a signal is sent back to the computer so that it can retransmit those data units that are error prone. Most of the error correction methods use the process of retransmissions.
- **Spread spectrum:** A technology widely used in wireless LANs. This technology is based on radio waves which facilitate communication between computer systems or devices within a limited area.
- **Telecom Regulatory Authority of India (TRAI):** The regulatory authority for telecommunications, including cable services in India. TRAI also ensures that consumer interests are protected.
- **Telecommunication media:** The various media over which data is transmitted and received in a telecommunications network.
- **Telecommunication processors:** These support data transmission and reception between the terminals and computers. Processors perform a variety of control and support functions in a telecommunication network.
- **Telecommunication software:** It consists of programs that control telecommunication activities and manage the functions of the telecommunications network.

- **Telecommunication system:** It consists of a transmitter that obtains the information to be transmitted and converts it into a signal for transmission, a medium of transmission through which the signal is transmitted, and a receiver that receives the signal and converts it back into the original information.
- **Telecommunication:** It involves the transmission of signals over a distance for communication purposes.
- **Telecommunications network:** A network in which the links and nodes are arranged in such a manner that it allows transfer of messages from one part of the network to another through various other links and nodes.
- **Terminals:** Any input/output device that uses telecommunication networks to transmit or receive data. The input/output devices could include telephones, office equipment, and transaction terminals.
- **Transmission media:** All the physical media through which the data is transmitted between the computer systems. These include the open wires, twisted-pair wires, coaxial cables, optical fibers, terrestrial microwaves, communication satellites, cellular phone systems, and wireless LANs.
- **Twisted-pair wires:** It consists of copper wires twisted into a pair. The wires are twisted together to reduce noise. Such networks can transmit both voice and data. Twisted-pair wires are more flexible than coaxial cables and optical fibers and it is very easy to install them.
- **Unguided media/unbound media/wireless media:** Media through which transmission and reception takes place through an antenna. This type of a medium has a means for the signals to travel.
- **Unified Messaging Systems (UMS):** These systems facilitate network convergence. Unified Messaging Systems deliver voice mail, e-mail, and faxes from a single system.
- **Wireless LAN:** It is a network which is used for linking two or more computer systems without the use of wires. A wireless LAN eliminates or greatly reduces the need for wires and cables.

10.11 Self-Assessment Test

1. The telecommunication system is comprised of the service providers, the transmission media and the regulatory authorities. Explain in detail.
2. Telecommunication systems can be used in various business applications. What are their advantages?

3. The key telecommunication network components are Media, Terminals, Computers, Channels, Processors, and Software. What are the functions of each of these components?
4. The various media over which data is transmitted and received in a telecommunications network are called telecommunication media. Describe the types of telecommunication media.
5. Telecommunication processors support data transmission and reception between the terminals and computers. In this context, explain the various telecommunication processors.
6. Network convergence refers to an integration of data, voice, and video into a single network. In this context, explain the advantages of network convergence.

10.12 Suggested Readings/Reference Material

1. Introduction to Information Technology, V. Rajaraman, PHI learning, 2018
2. Information Technology for Management, 2ed: Advancing Sustainable, Profitable BusinessGrowth, Turban, Volonino, wood, O.P. Wali, Wiley India Pvt Limited, January 2021
3. Introduction to Information Systems - 6th edition, R. Kelly Rainer; John Wiley & Sons, Inc.2016
4. Information Technology: An Introduction for Today's Digital World, Richard Fox, Chapman and Hall/CRC; 2nd edition (August 21, 2020)
5. Information Technology for Management, Efraim Turban, Carol Pollard, Gregory Wood, Wiley, 2018

Additional References:

1. Critchley, L., Where Nanotechnology, the IoT, and Industry 4.0 Meet., https://www.mouser.com/blog/where-nanotechnology-the-iot-and-industry-40-meet, 2019
2. Pan India implementation of HMIS over Indian Railways, Ministry of Railways., http://railministry.com/pan-india-implementation-of-hmis-over-indian-railways/ 2020
3. Vossler, C. How Long Does It Take To Order A New BMW? https://www.bmwblog.com/2020/09/28/how-long-does-it-take-to-order-a-new-bmw/2020
4. Jay, A., 10 New ERP Trends & Forecasts for 2020/2021 – A Look Into What's Next. https://financesonline.com/erp-trends/2019

5. Gingiss, D., How Integrating Social Media Into The Rest Of The Business Will Increase Revenue., How Integrating Social Media Into The Rest Of The Business Will Increase Revenue (forbes.com), 2019

10.13 Answers to Check Your Progress Questions

Following are the answers to the Check Your Progress questions given in the unit.

1. (b) The user places a call outside the local transport area

An Inter Exchange carrier (IXC) comes into play when a user places a call outside the local transport area. Transmission media like copper wires, fiber-optic cables, microwaves, and satellites are used to provide regional and long-distance services between local telephone exchanges.

2. Local Exchange Carriers

Local Exchange Carriers (LECs) provide wired services for subscribers. Local Access and Transport Area (LATA) is a network of local connections. A subscriber can place a call without accessing a long-distance carrier within LATA.

3. Public Switched Telephone Network

The Public Switched Telephone Network or PSTN was the first telecom network. The PSTN telephone lines used analog technology.

4. (c) The service provided by the Public Switched Telephone Network (PSTN)

The Public Switched Telephone Network or PSTN was the first telecom network that was developed. The telephone lines provided by PSTN used analog technology. The service provided by the PSTN is called the plain old telephone service (POTS).

5. International Telecommunications Union

Telecommunication standards are governed by the International Telecommunications Union (ITU), previously called the Consultative Committee for International Telephony and Telegraphy (CCITT). The ITU was established on May 17, 1865 in Paris. It is currently a specialized division of the United Nations and is headquartered at Geneva, Switzerland.

6. A LATA is a network of local connections. LATAs are linked to Subscriber Loop Carriers (SLCs) by means of copper wires called

twisted pair cables. The area served by an SLC is called the local loop.

7. The International Telecommunications Union (ITU) was established on May 17, 1865 in Paris. It is responsible for governing telecommunication standards. The ITU is divided into three sectors. These are: Telecommunication Standardization or ITU-T, Radio communication or ITU-R, and Telecommunication Development or ITU-D.

8. **TRAI**

In India, the Telecom Regulatory Authority of India (TRAI) is the regulatory authority for telecommunications, including cable services.

9. **(a) Integrated Services Digital Network**

ISDN or Integrated Services Digital Network is a telephone network designed to allow digital transmission of voice and data over ordinary telephone copper wires.

10. **(d) The process of designating a system for the simultaneous transmission of two or more messages or signals over the same channel**

Multiplexing is the process of designating a system for the simultaneous transmission of two or more messages or signals over the same channel. Multiplexing is done in several ways: frequency division multiplexing, time division multiplexing, and advanced statistical time division multiplexing.

11. Mobile commerce systems, also known as M-Commerce, is a form of electronic commerce in which products and/or services are bought and sold through wireless handheld devices like cell phones or personal digital assistants using the Internet. This e-commerce application of telecommunication enables users to access the Internet without cables or wires, through their mobile devices.

12. **(c) Internal business systems**

Internal business applications of telecommunications include internal transaction processing, process control, inquiry processing, activity monitoring, workflow systems and management support systems. For instance, a company can link wide area and local area networks, so that managers can make inquiries and generate reports from corporate databases stored on distant network servers.

13. The scope of telecommunications can be categorized into Industry scope, Technology scope and Application scope. Telecommunication networks are being used for carrying out electronic commerce, enterprise collaboration, and for supporting business functions of operations and distribution.

14. (b) Infrared LAN

A wireless LAN enables users to move around with a computer or a device and stay connected to the network. Infrared LAN is a wireless LAN that makes use of infrared rays to establish network links between the various LAN components.

15. Analog signals

In analog transmission systems, signals are transmitted through a medium as continuously varying electromagnetic waves. Twisted-pair wires, coaxial cables, optical fibers, atmosphere, water, or space can be used as the transmission media for analog transmission systems.

16. (d) Both (b) & (c)

A wireless LAN eliminates or greatly reduces the need for wires and cables. LAN radio may involve a high frequency radio technology similar to digital cellular or a low frequency radio technology called spread spectrum.

17. Modem is a device that converts digital signals from input/output devices into appropriate frequencies at a transmission terminal, and re-converts them into digital signals at a receiving terminal. They support functions like transmission error control, automatic dialing and faxing.

18. (b) Fiber optics

Fiber optics use cables consisting of one or more hair-thin filaments of glass fiber wrapped in a protective jacket. They can conduct light pulses generated by lasers at transmission rates as high as 30 billion bits per second. The size and weight of these cables is substantially lesser than conventional cables. These also work at a greater speed and have larger carrying capacity.

19. (c) Fiber optic cables

Optical fibers or fiber optic cables use cables consisting of one or more hair-thin filaments of glass fiber wrapped in a protective jacket. These have the highest bit rates among all telecommunications media. The size and weight of these cables is

substantially lesser than conventional cables and they also work at a greater speed and have larger carrying capacity.

20. (d) Optical fibers

Optical fibers use cables consisting of one or more hair-thin filaments of glass fiber wrapped in a protective jacket. Optical fibers have the highest bit rates among all the telecommunications media. The size and weight of these cables is substantially lesser than conventional cables; they also work at a greater speed and have greater carrying capacity.

21. A terminal is one among the five basic components that supports an organization's telecommunication activities. Any input/output device that uses telecommunication networks to transmit or receive data is called a terminal. This includes telephones, office equipment, and transaction terminals.

22. Digital

In a digital transmission system, the signals are transmitted as discrete pulses. These pulses are measured in bits per second and are similar to those transmitted in computer systems.

23. (a) Communications satellites

Communications satellites use radio microwave as the telecommunication medium, but they are placed in stationary geosynchronous orbits. These are mainly used for voice and video transmission where high speed transmission of large data volume is required.

24. Bits per second

Bandwidth displays the frequency range of a channel and determines its maximum transmission rate. It is measured in bits per second.

25. (b) Twisted-pair wires

Twisted-pair wires are the transmission medium that is mostly used in star and hub networks. A twisted-pair wire consists of copper wires twisted into a pair. The wires are twisted together to reduce noise. Such networks can transmit both voice and data.

26. Telecommunication processors perform a variety of control and support functions in a telecommunication network. The functions of processors include converting data from digital to analog and back; coding and decoding of data; and controlling various activities like speed, accuracy, and efficiency of the communications flow between computers and terminals.

27. The various media over which data is transmitted and received in a telecommunications network are called telecommunication media. They are basically classified as guided media and unguided media. Guided media consists of a cabling system. The cabling system guides the signals through a specific path. These include open wires, twisted-pair wires, coaxial cables, and fiber optic cables, which provide physical link to the devices in a network. Unguided media is a medium through which the transmission and reception takes place through an antenna. These signals are not guided through a specific path as data signals are not bound by a cabling system as in guided media. Transmission of data for unguided media largely depends upon the bandwidth produced by the antenna.

28. (b) Communications satellites

Communication satellites use radio microwave as a telecommunication medium, but are placed in stationary geosynchronous orbits. These are powered by solar panels and transmit microwave signals at a rate of several hundred million bits per second.

29. Switches, routers, hubs, and gateways are internetworked processors. A switch connects different telecommunication circuits in a network for a message to reach its intended destination. A router interconnects networks based on different rules or protocols, so the message can be routed to its destination. Hubs provide automatic switching among connections called ports for shared access to a network resource. Gateways interconnect networks that use different communication architectures.

30. Backbones are long distance trunk lines that carry most of the traffic on the communications network. A backbone is used for joining the LANs together either within a building or across a city or a country. LANs are connected to the backbone through bridges and/or routers.

31. (b) Business-to-business

Electronic Data Interchange (EDI) is used for setting standards for B2B transactions and to speed up the process of ordering, invoicing, and payments. Two companies having compatible systems can establish connections using an EDI.

32. WANs have a local access number called point of presence (POP). A point of presence (POP) means a city or a location where a network can be connected to another one. POP allows customers to access WAN by means of a local analog telephone call or a direct digital hookup.

IT & Systems

Course Components

BLOCK I	**Introduction to Information Technology and Systems**
Unit 1	Computer Systems – An Overview
Unit 2	Operating Systems
Unit 3	Fundamentals of Information Systems
BLOCK II	**Applications of Information Technology in Business**
Unit 4	Personal Productivity Software
Unit 5	Enterprise Collaboration Systems
Unit 6	Management Information Systems
BLOCK III	**Software and Database Concepts, and Networks**
Unit 7	Program Design and Programming Languages
Unit 8	Database Management
Unit 9	*Computer Networks*
Unit 10	*Telecommunication Networks*
BLOCK IV	**Management of MIS**
Unit 11	MIS – Planning and Design
Unit 12	MIS – Implementation, Evaluation, and Maintenance
Unit 13	Information Resources Management and IT Governance
Unit 14	Global IT Management
Unit 15	MIS in Specialized Areas
BLOCK V	**Enterprise Functions and E-Business**
Unit 16	Basics of E-Business and Enterprise Application Integration
Unit 17	Supply Chain Management and E-Business
Unit 18	Enterprise Resource Planning
Unit 19	CRM and E-Business
BLOCK VI	**Advanced Topics in IT**
Unit 20	Cloud Computing
Unit 21	Business Intelligence and Big Data
Unit 22	Current Trends in Software Design and Architecture
Unit 23	Mobile and Social Technologies
Unit 24	IT and Business Process Management