UNIT 7

DATABASE STRUCTURES

Structure

7.1 Introduction

Expected Learning Outcomes

7.2 Overview of Data in a Computer

Digital Representation of Data and Geospatial data

Related Technical Issues

7.3 Coding the Basic Models for Input to the Computer

7.4 Spatial Data Structure

7.5 Vector Data Structure

7.6 Raster Data Structure

7.7 Relative Merits of Raster and Vector Data Structures

7.8 Data Structures and Data Analysis

7.9 Concept of Database

7.10 Concept of Geodatabase

7.11 Database Management System

7.12 Summary

7.13 Activity

7.14 Terminal Questions

7.15 References

7.16 Further/Suggested Readings

7.17 Answers

7.1 INTRODUCTION

Spatial data is an important component in day to day life as whatever we see and do is associated with the spatial component. The geospatial data represent both spatial or geographical and non-spatial or descriptive or statistical information about the spatial entity. For example, we have a vehicle traversing from point A to point B covering a distance of 10 km. Here, the spatial data like a road and its associated information like road length, width, road condition, number of coffee shops, hotels, etc. form a descriptive information or non-spatial data. To further describe, and know how spatial relation exists between different variables different models are used. To run these models the spatial data structures are helpful that furnish required information to the computer in a GIS environment. Further, the pattern and relationship between the spatial data can be understood through spatial analysis.

In the previous unit you have read about concepts of data models and their types such as raster and vector data models, comparison of vector and raster models and their advantages and disadvantages in GIS. In this unit, we will discuss overview of data in a computer, coding the basic models for input to the computer, spatial data structure, and vector and raster data structures. In addition, we will also discuss relative merits of vector and raster structures, concept of database and geodatabase and database management system.

Expected Learning Outcomes

After studying this unit, you should be able to:

- ❖ discuss the overview of data in a computer with emphasis on digital representation of
- ❖ data and geospatial data, and related technical issues;
- ❖ describe coding the basic models for input to the computer;
- ❖ illustratespatial data structure, vector and raster data structures;
- ❖ compare and know the merits ofraster and vector data structures; and
- ❖ explain the concept of database and geodatabase, and database management system.

7.2 OVERVIEW OF DATA IN A COMPUTER

From layman's language, data refers to range of possible information which can be collected from various sources. There are different techniques or methods by which a data can be gathered. For example, to know the extent of deforestation in remote forest ranges, we can gather data using satellite images, ground based data, photographs, survey techniques for boundary information and much more. In modern ways, processing the data using computer is much easier than traditional and manual methods which were very tedious, time consuming and error prone.

Thereforecomputer supported data can be in different formatssuch as numerical data, images, coding, notes, and voice based, animation, and much more combinations. Further, data can be categorised into analog data and digital data. Analog data is data transmitted in a continuous manner. On the other hand, digital data is the electronic format or language that machines can read and understand through script or programming. In simple terms, digital data is binaryinformation converted into a machine-readable or digital format (Fig. 7.1).Hence, computerswith data arecapable of performing computations and making logical decisions at a speed of millions, and even billions of times faster than human can think and act. Computers process data under the control of sets of instructions called computer programs.

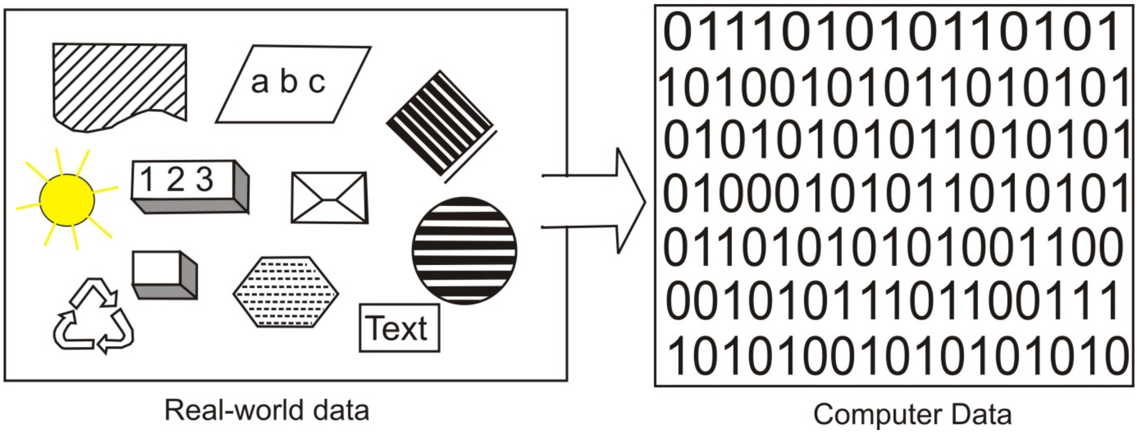

Fig.7.1: Representation of real world data in a computer. (Source: https://homepage.cs.uri.edu/faculty/wolfe/book/Readings/Reading02.htm)

7.2.1 Digital Representation of Data and Geospatial Data

The representation or the reproduction of data in digital form which is supported by computersis nothing but the digital representation of any data. All information is represented as sequence of 1's and 0's(Fig. 7.1). This information can represent *numbers, text, instructions, sound and image/audio.* They are represented as numbers in a computed and manipulated or modified per requirements.

A single 1 or 0 is known as a bit. To represent more than one value, we need to use the combination or sequence as a group of 8-bits. A byte is defined as a group of 8-bits.Therefore, a word varies in size but is usually a 32-bits. Computers store bits in fixed-sized units (8-bits, 16-bits, 32-bits and 64-bits).

Geospatial data is the representation of data at specified point of time linking to a specific location on the earth. It provides valuable inputs to relationship between geographic variables with specific patterns and trends. It combines location information (geographic coordinates represented as latitudes and longitudes) with attribute or descriptive information about an object, event or phenomena under investigation with temporal information (representing time at which the location and its attributes linked).

Geospatial data can comprise of large sets of spatial data obtained from many diverse sources in varying data formats. This can include census data, satellite imagery, rainfall data, mobile data, hand drawn maps and social media data. Geospatial data has its maximum benefits when it can be shared, and investigated in combination with other business data and resource data. Geospatial data consists of vector data and raster data.Vector data is represented by means of points, lines and polygons (electric pole, manhole, cities, roads, mountains and water bodies).Raster data is pixelor cell or grid based which are recorded as rows and columns. Raster data consists of digital imagery which is more complex, such as photographs and satellite images.

7.2.2 Related Technical Issues

Handling geospatial data is always a technical challenge due to its large data sets. This is one of the major hindrances to get full benefits of the geospatial data. Data volume is too high to handle even for more than a week. For example, a typical weather data is estimated to be around 100 TB for a single day of operations. This needs considerable storage and limits the access capability of many departments.

Secondly, the data available in different file format and makes it not accessible and ready for critical use in times of need. Furthermore, data accuracy varies with different mapping procedures and acquisition methods.Therefore data preparation is one of the most challenging initial phases of an organisation. After data acquisition, processing the raw data and making it compatible and ready to use needs specialised skill and application knowledge on concepts of geography and mathematics to perform specific tasks like data alignment, use of right geographic parameters for specific analysis.

Geospatial data acquired by different departments may vary in map scale, accuracy and method of processing. This can potentially impact the final output if the data need integration of various real time data to handle some disaster situation where resource mobilisation needs accurate input.

7.3 CODING THE BASIC MODELS FOR INPUT TO THE COMPUTER

You have read about overview of data in a computer in the previous section. Now let us discuss the coding of basic models for input to the computer. When we refer data entry into a computer to perform some specific tasks, be it a table query, complex analysis, or designing an application, input and output acts as the communication between a computer program and its end user. Input is driven by the user who gives specific command to the program; output is something given to the user by the computer.In this way, data is entered into a program in a computer, either by the developer or digitally, are known as inputs. Interestingly, inputs are stored in variables and used to run the intended program. To understand the benefits of a program which is being run in a computer, developer would include output mechanism to see the end result which is more computer or user friendly, quick in processing and benefit the larger audience.

From GIS perspective, computer takes many forms – spatial, non-spatial information or combined to be fed into a system for data manipulation, storage, retrieval, analysis and final output. Input can be in many forms like satellite images, aerial photos, GPS readings, drone data, LiDAR and ground collected data. The computer system with GIS functionality can handle large volumesof data efficiently and produce desired output in many forms useful and easily readable even by general public.

7.4 SPATIAL DATA STRUCTURE

You have read about coding the basic models for input to the computer in the previous section. Now let us discuss the spatial data structure in detail.

Structures that provide information required for computer to process spatial data model in digital form are defined as spatial data structure. On the other handspatial data structures are the set of rules that are applied to characterise the geographic data in a GIS. Spatial data structures are fundamental aspects of any GIS that influence its performance and functioning capabilities. Hence

Contributor: Dr. O. M. Murali

spatial data structures are important in analysis ofa geographic data. As far as GIS is concerned, spatial data structures are categorised as vector data structures and raster data structures.

When we consider the internal processes in a computer, data structure is a systematic way of storing and organising data and when invoked through command prompt, can be used efficiently. If we think about the spatial data structures, they are known for representations of data for storage, retrieval, manipulation and visualisation through analytical ability that provide the end result for key decisions. The unlimited power of a GIS purely depends on the enormity of the data fed into it.

It is clear that having the right spatial data structure is very important prerequisites. It should accurately model or depict the real world situation like flood inundation, forest fire, tracking of endangered wildlife species,tracking pollution level and much more which varies from the scope of the applications. Secondly, it should be capable of retrieving the data as requested by the end user or the developer. It should have standard procedures for scalable implementation through simple and complex code or script driven application.Above all, it should be able to process efficiently the large volume of both spatial and non-spatial data in a meaningful way.

For this reason data structures are developed that progressively access and process large volumes of data. These data structures are hierarchical, network and relational depending on the specific purpose and application needs (Fig. 7.2 to 7.4). The data structures vary and differ for each application and depend on nature of the collected spatial and descriptive data.These data structures will be discussed in detail in Unit 11 of this Course.

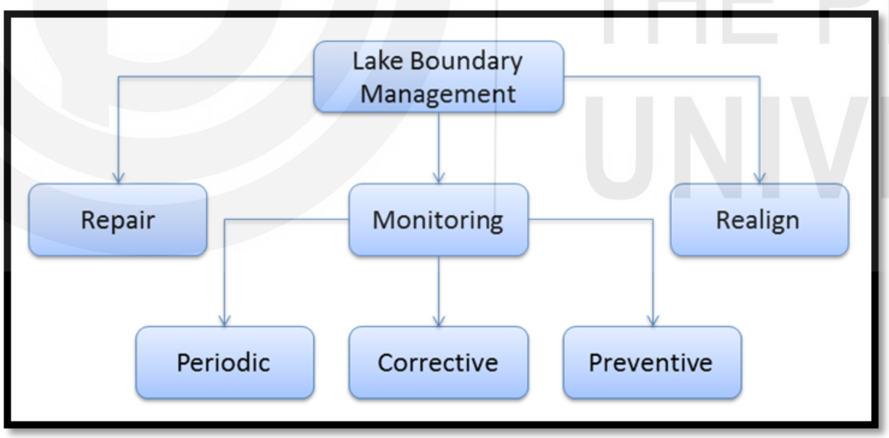

Fig. 7.2: Hierarchical data structure.

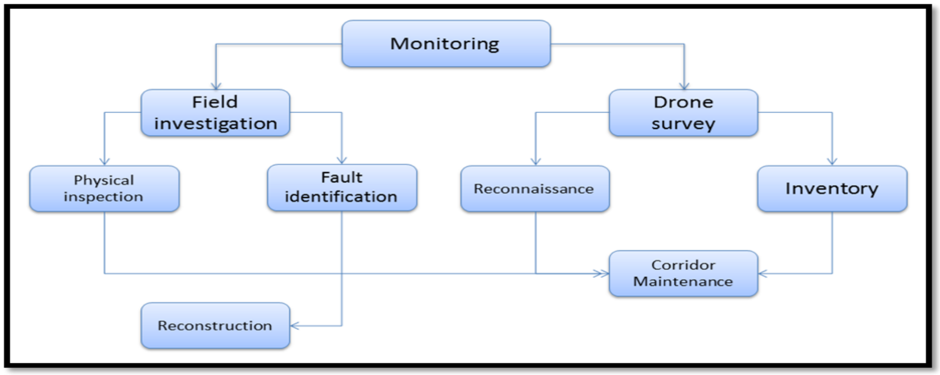

Fig. 7.3: Network data structure.

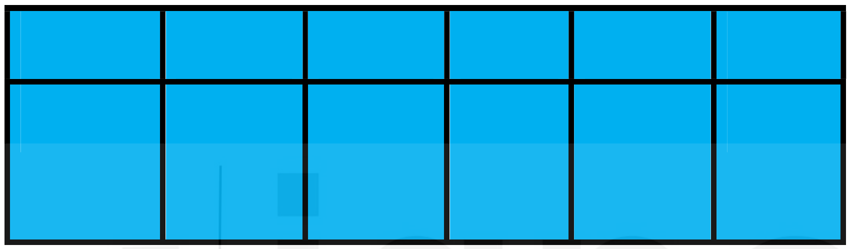

Fig. 7.4: Relational database structure.

7.5 VECTOR DATA STRUCTURE

In the above section you have read that vector data structures are part of spatial data structure. Let us discuss the vector data structure in detail.Vector data structure or model is a very interesting real world representation where different combination of *point, line, area or polygon* is represented to depict their area of interest in a meaningful way so that anybody can understand thematic representation (Fig. 7.5).

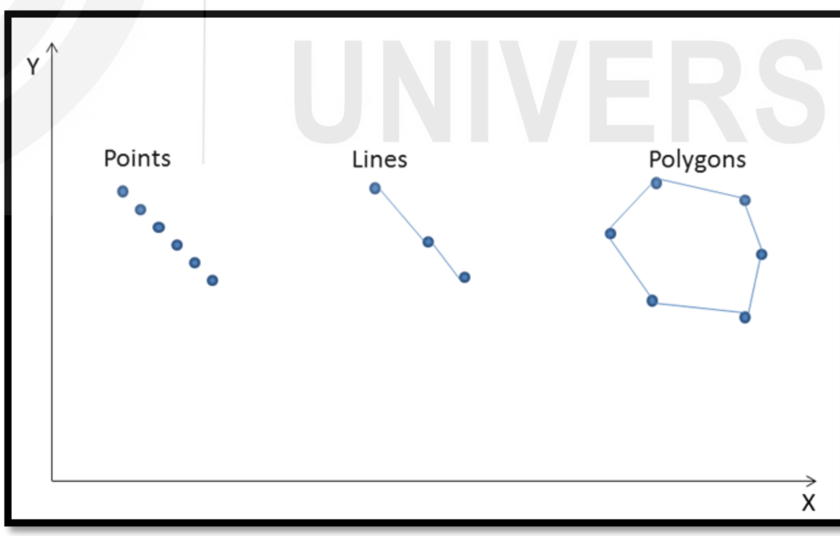

Fig. 7.5: Representation of vector data.

Vector data structures are the data structures that generally store the geometries of the geographic objects that are represented. They store the geometries using a series of points and lines known as vectors. In these data structures, geometries are characterised with different dimensions (Fig. 7.6). A point is represented by zero geometry with x and y coordinates (Fig. 7.6a). A

Contributor: Dr. O. M. Murali

line also known as arc or a series of line segments occasionally termed a polyline (Fig. 7.6b). A polygon is symbolised with two-dimensional geometry with same start and end point where sequential arrangement of one-dimensional geometries are established (Fig. 7.6c). Polygons are known to be simple or complex. The simple polygons are characterised by one polygon and complex polygons are aggregates of more than one polygon. The complex polygons may possibly have holes, with island geometriesinside the holes. A three-dimensional geometry is known as solid. The solid is represented by assemblage of two-dimensional geometries and a third dimension called as Z-coordinate (Fig. 7.6d). The Z-coordinate generally represents height in relation to a reference point. Similar to the complex polygon the solid may be an aggregate of numerous solids and may contain holes.

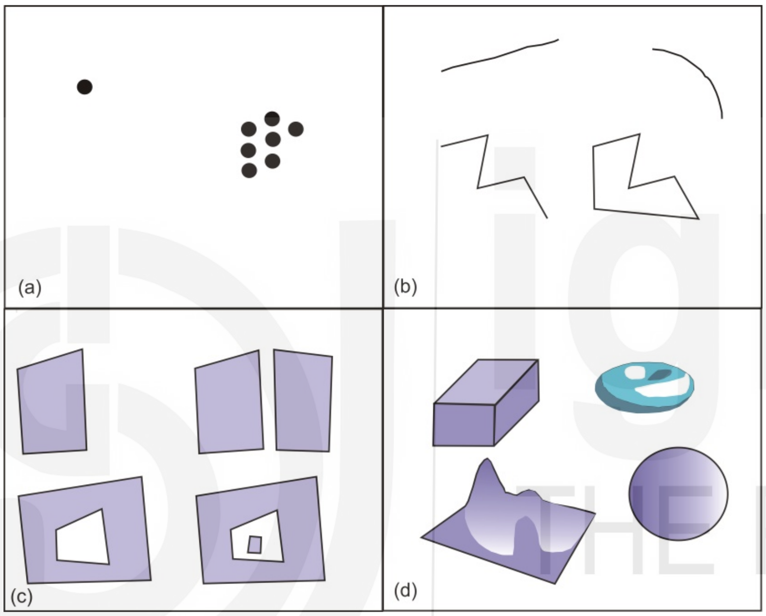

Fig. 7.6: Vector data structures with different geometries: a) 0-dimensional; b) 1-dimensional; c) 2- dimensional; and d) 3-dimensional. (Source: https://www.academia.edu/3045247/Spatial_Data_Structures)

7.6 RASTER DATA STRUCTURE

You have read about raster data structures in the previous section. Just as the vector data structure raster data structure is also part of spatial data structure. Let us discuss the raster data structure in detail.

Raster or grid data structure refers to the storage of the raster data for data processing and analysis by the computer. There are mainly three commonly used data structures such as cell-by-cell encoding, run-length encoding, and quad tree (Chang, 2010).

Cell-By-Cell Encoding Data Structure

This is the simplest raster data structure and is characterised by subdividing a geographic space into grid cells. Each pixel or grid cell contains a value. A grid matrix and its cell values for a raster are arranged into a file by row and column. Fig. 7.7 shows the cell-by-cell encoding data structure. Digital Elevation Models (DEMs) are the best examples for this method of data structure. In DEMs, the neighbouring elevation values are more or less the same. In Fig. 7.6, value 1 represents the grey cells and 0 has no data. This cell-by-cell encoding method can also be used for storage of data in satellite images. Most of satellite images consist of multispectral bands and each pixel in a satellite image has more than one value. Mainly three formats such as Band Sequential (.bsq), Band Interleaved by Lines (.bil), and Band Interleaved by Pixels (.bip) are used to store data in a multiband/multispectral imagery.

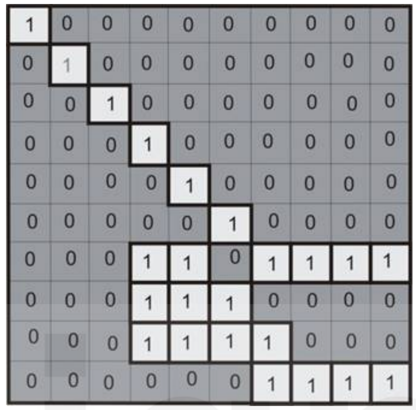

Fig. 7.7: Cell-by-cell encoding data structure.

Run-Length Encoding Data Structure

Run-length encoding (RLE) algorithm was developed to handle the problem that a grid often contains redundant or missing data. When the raster data contains more missing data, the cell-by-cell encoding method cannot be suggested. In RLE method adjacent cells along a row with the same value are treated as a group called as***run***. If a whole row has only one class, it is stored as the class and the same attributes are kept without change. Instead of repeatedly storing the same value for each cell, the value is stored once together with the number of the cells that makes the run. Fig. 7.8 explains the run-length encoding structure of a polygon. In the figure, the starting cell and the end cell of the each row denote the length of group and are generally called as run. RLE data compression method is used in many GIS packages and in standard image formats.

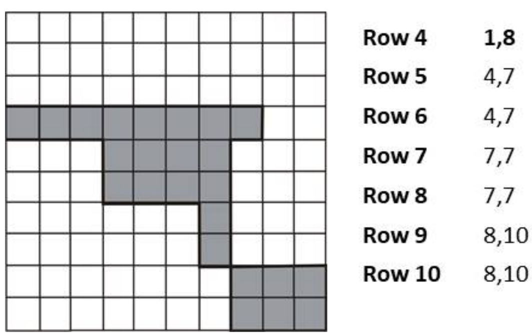

Fig. 7.8: Run-length encoding data structure.

Quad Tree Data Structure

To compress the data as well as to save the space in original grid, quad tree data structure can be used (Fig. 7.9). A quad tree works by dividing a grid into four quadrants for the available data. The available data quadrant is again split into four half-size quadrants and so on until the individual pixel is reached. The

Contributor: Dr. O. M. Murali

attribute data for all the pixels of the quadrant remains the same even if it is divided.

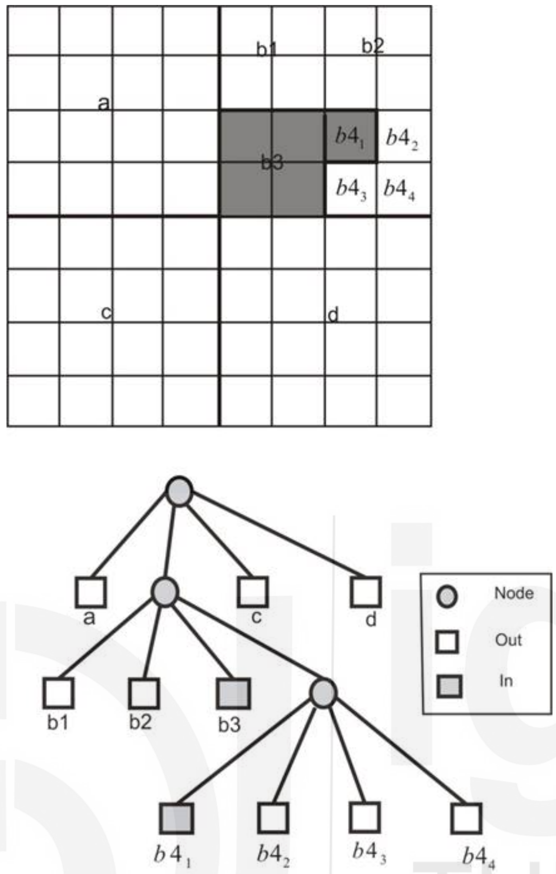

Fig. 7.9: Quad tree data structure.

SAQ II

a) Data can be characterised into _________ and __________ data.

b) The computer system with GIS ____________ can handle large volumes of data _________ and produce desired output.

c) Spatial data structures are the set of rules that are applied to _________the geographic __________in a GIS.

d) In Run-length encoding method neighbouring ___________along a row with the same value are treated as a group known as__________.

7.7 RELATIVE MERITS OF RASTER AND VECTOR DATA STRUCTURES

To store geospatial data in an organised manner computer followsthe logical phenomena known as data structures for representing both raster and vector data. The link between data representation and analysis in GIScan be understoodby knowing the merits anddemerits of raster and vector structures. Both thesestructures have advantages and disadvantages for storing geospatial data. Let us discuss themerits and challenges of raster and vector data structures in detail.

Merits of Raster Data Structure

- Data structure is simple and grid or cell based which facilitates easy processing.
- Widely used for visualising contiguous surfaces, especially suited for weather related analysis.
- Low-cost technology
- Easy to manipulate location specific attribute data.
- Different types of spatial analysis and filtering can be performed.
- Multi-layer stacking and related operations are quite fast to perform.
- Data sources are multiple that can be from satellite images, aerial photography, even drone and field data.
- The data obtained using satellite imageries are near real time and quite useful for tracking of natural hazards and even monitoring of forest fire, oil spills, and environmental pollution.

Challenges of Raster Data Structure

- Difficult to store large volumes of data.
- Inaccurate representation of point, line and area features.
- Decreased spatial resolution while using large grid cells to reduce data volumes. Thus resulting in loss of information and inability to recognise remarkably defined structures.
- Basic raster maps are appealingly less agreeable though graphic representation is acceptablewith a reduced problem.
- Coordinate transformation is challenging and takes much time if no specific logarithms and hardware are used. Consequently resulting in information loss and grid cell shape distortion.

Merits of Vector Data Structure

- Compact data structure.
- Entity data models can be notably represented.
- Topology can be described clearly which makes it best for proximity and network based analysis.
- Coordinate transformation and rubber sheeting techniques are easy to perform.
- Graphic representation is precise at all scales.
- Retrieval, updating, and generalisation of graphics and attributes are possible.

Challenges of Vector Data Structure

- Complex data structure.
- Joining various polygons for intersection and overlay analysis is very challenging and needs high end computers.

- Data preparation is quite expensive as it needs to be obtained from primary data sources like satellite images and ground truthing.
- Preparation of high quality cartographic output is expensive, time consuming and person dependent.
- Highly skilled resources required for processing.
- Spatial analysis and filtering in elementary components like polygons is difficult without additional database as they are known to be homogenous internally.

7.8 DATA STRUCTURES AND DATA ANALYSIS

In the previous section we have discussed about merits and challenges of raster and vector data structures. In this section we will discuss the data structures in relation to data analysis.In GIS, spatial data structures are characterised as vector and raster data structures. Each is having unique representation of geographic features. With raster, it is depicted as grid cells such as tessalations, run length compression, quad tree, BSQ/BIL/BIP, DBMS and file formats. In this data structure, attribute is represented as a single DN fora particular cell. Input data comes from remote sensing satellite images like LANDSAT, SPOT, IRS series, scanned images, elevation data from different sources. This data structure is a best representation for continuous features – elevation, temperature, rainfall, soil type and land use inventory.

Some of the concepts of using raster data structure are given below:

Grid often has its origin in the upper left, but

- State plane and UTM – lower left
- Lat\long and Cartesian – centre

Single value is associated with each cell

- 8 bits assigned to values therefore 256 possible values (0-255)

Rules needed to assign value of cell if object does not cover entire cell

- Majority of the area
- Value at cell centre
- Touches cell

Raster orientation: angle between true north and direction defined by raster columns

Class: set of cells with same value (For example: alluvial soil)

Zone: set of contiguous cells with same value

Neighbourhood: set of cells adjacent to a target cell in some systematic way

On the other hand, vector data is represented through geographic coordinates as point, line, polygon, TIN and file formats. In this data structure, attribute is referenced through unique ID number to tables. Much of the data comesfrom census and as wellas derived from satellite images. It is the best for representing discrete features – property lines, political boundaries and transportation.In this structure all the data is stored in one file,coordinates/boundaries for adjacent polygons are stored twice. Features in the coverage have unique identifiers – point ID, polygon ID, arc ID. The common identifiers provide link to coordinates table for where condition and attribute table for what condition. But, concepts are associated with relational data base which is a prerequisite for the vector data structure. The file formats include – coverage, shape file and geodatabase.

7.9 CONCEPT OF DATABASE

In the previous section we have discussedthe data structures in relation to data analysis. In this section we will discuss the concept of database.

In our daily transaction, we directly or indirectly with knowledge or without our knowledge use large database to help us achieve our desired action. But we don't even pay attention to this hidden database of our personal details which is in use by many external people. When we go to a retail shop, shop keeper stores our purchasing behaviour through entry of our phone number. Likewise, when we use ATM to withdraw money, machine use the database from our card details which in turn is stored. In short, database is the organised collection of data. It supports electronic storage, manipulation, retrieval and display of data. At the click of a button, different department use a database to store data of people, their purchasing details, address and other details.

Data Base Management System (DBMS) is a software package that helps us to create, manipulate and manage adatabase. By this way, DBMS provides tools to create, insert, delete, modify or update data in the database. Examples of DBMS includelicensed packages like ORACLE, SQL Server and Open source packages like MySQL, PostgreSQL, etc.

Earlier, file based systems had certain limitations like separation, isolation of data, duplication, inconsistent data, incompatible file format and fixed queries. To overcome the above said limitations, database approach emerged. It is a persistent collection of logically related data. Database helped in reduced redundancy, shared data, data independence, improved integrity, efficient data access, multi-user interfaces, improved security, backup, recovery and concurrent transactions,

7.10 CONCEPT OF GEODATABASE

In the previous section we have discussedthe concept of database. Similarly in this section we will discuss the concept of geodatabase.

At the basic level, geodatabase storesspatial and attribute data and the relationships that exist among them. Geodatabase, which is a vector data format, features and their associated attributes can be structured to work together as an integrated system using rules, relationships, and topological associations. In other words, it allows us to model the real world in simple or complex way as per our needs. Geodatabase can be created, edited, and managed using the standard menus and tools in ArcCatalog™ and ArcMap™ of ArcGIS software.

Contributor: Dr. O. M. Murali

Types of Geodatabase

We have two types of geodatabases: a) Personal and b) Enterprise.

a) Personal geodatabase

A personal geodatabase has the .mdb file extension (a format used by Microsoft Access) and can be read by multiple people at the same time, but one person to edit at a time (Fig. 7.10). A personal geodatabase has a maximum size of 2 gigabytes (GB). Vector data is stored in the database, while raster data is referenced.

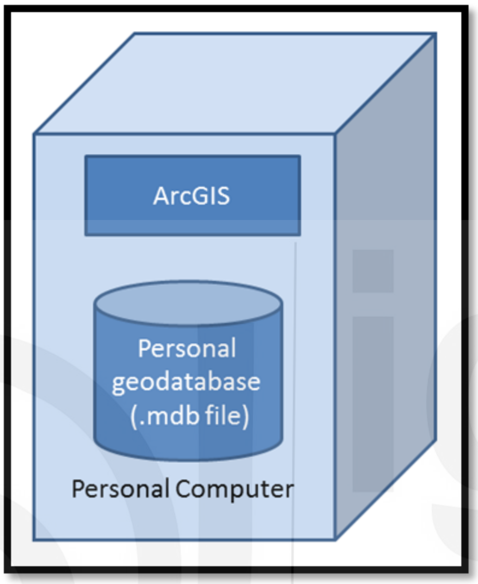

Fig. 7.10: Personal geodatabase.

b) Enterprise geodatabase

Enterprise geodatabase is sometimes called as ArcSDE® geodatabase. It is suitable for large workgroups and enterprise GIS implementations. It can be read and edited by multiple users at the same time, and can store both vector and raster data (Fig.7.11).Enterprise geodatabase requires ArcSDE software and a DBMS (Database Management System) such as IBM DB2, Informix, Oracle, or Microsoft SQL Server. The spatial data is stored in the DBMS and ArcSDE allows you to view and work with the data from your GIS software applications. With some database management systems, you can directly connect to your data from ArcCatalog or ArcMap. For others, ArcSDE must manage the connection.

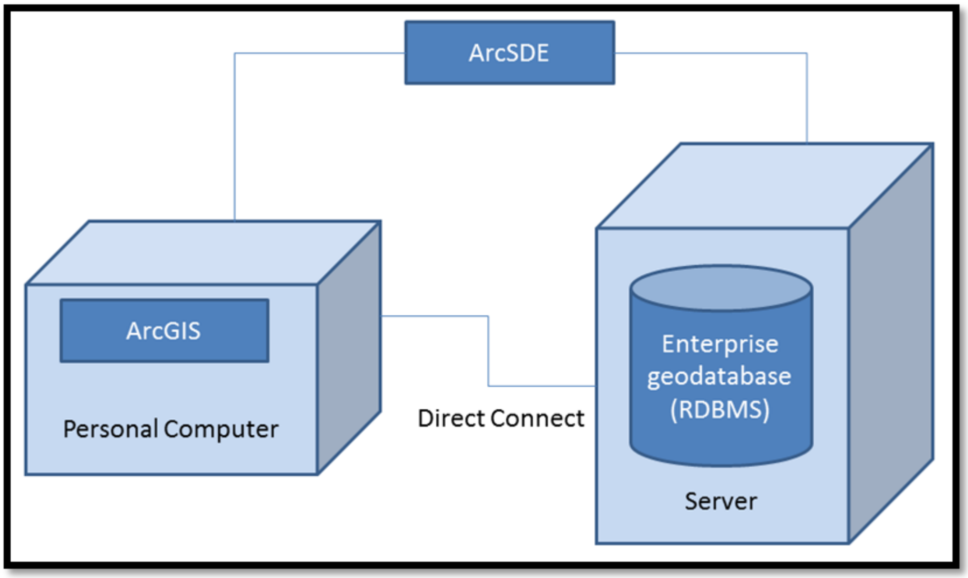

Fig. 7.11: Enterprise geodatabase.

7.11 DATABASE MANAGEMENT SYSTEM

In the previous section we have discussed the concept of geodatabase. In this section we will discuss the concept of database management system.

Real world entities or objects can be stored into computer compatible files which are simpleor complex and well organised. The data record in a file contains all the requiredinformation for each entity. Interestingly, data records may be of same or variable lengths.Likewise, spatial databases contain many files with data on related aspects of the samefeature or other connected or linked together. The prime concern is that the proper organisation is importantfor storing these files in an appropriate way to model real world phenomena andensure efficient storage and retrieval of data.

A software package designed tofacilitate creation and maintenance of computerised database is called as aDatabase Management System (DBMS). There are many procedures forefficient storage and retrieval of the data and to model logically the real worldfeatures i.e. hierarchical, network, relational, multidimensional and object-basedwhich are used in GIS that will be discussed in detail in Unit 11 of this Course. The fundamental objective of the DBMS isto make data quickly available to different users. Three primary database integrity stepsare given below.

- Protecting the database against deletion and corruption.
- Maintaining the integrity of the database to facilitate the addition, removal, and updating, if necessary.
- Privacy of the database.

A database is the collection of data records, files, and other database objects. There aremany types of DBMS, from small systems that run on personal computers to huge systems that run on mainframes. The given set of rules for building a query is known as a querylanguage. Different DBMSs support different query languages. But there isa semi-standardised query language called Structured Query Language (SQL). Many GIS utilise the DBMS as part of the system because it has the capability to handle data explicitly. As discussed earlier the hierarchical, network, relational, multidimensional and object oriented data structures (non-spatial data structure database models) are often used in spatial information system.

SAQ II

a) Data structure is simple and _________based which facilitates easy _________of data.

b) In __________data structure, attribute is referenced through _________number to tables.

c) _________provides tools to create, insert, delete, modify or update data in the database.

d) ________stores spatial and attribute data and the relationships that exist between them.

e) Semi-standardised query language used in DBMS for querying is known as _________.

7.12 SUMMARY

Let us summarise what you have studied in this unit.

- Data refers to range of possible information which can be collected from various sources. Computer supported data can be in different formats such as numerical data, images, code, notes, and voice based, animation, and much more combinations
- Geospatial data is the representation of data at specified point of time linking to a specific location on the earth. It comprise of large sets of spatial data obtained from many diverse sources in varying data formats.
- Managing geospatial data is always a technical challenge due to its large data sets. Data accuracy varies with different mapping procedures and acquisition methods. Therefore data preparation is one of the most challenging initial phases of an organisation.
- When we refer data entry into a computer to perform some specific tasks, be it a table query, complex analysis, or designing an application, input and output acts as the communication between a computer program and its end user.
- Spatial data structures are fundamental aspects of any GIS that influence its performance and functioning capabilities and helps in analysis of a geographic data. The capacity of a GIS purely depends on the enormity of the data fed into it.
- Vector data structures are the data structures that generally store the geometries of the geographic objects that are represented as points and lines known as vectors.

- Raster data structure refers to the storage of the raster data for data processing and analysis by the computer. The three commonly used data structures are cell-by-cell encoding, run-length encoding, and quad tree.
- The association between data representation and analysis in GIS can be identified by knowing the merits and demerits of raster and vector structures. Both these structures have advantages and disadvantages for storing geospatial data.
- Raster data structures are depicted as grid cells such as tessalations, run length compression, quad tree, BSQ/BIL/BIP, DBMS and file formats. Whereas vector data is represented through geographic coordinates as point, line, polygon, TIN and file formats.
- Database is the organised collection of data. It supports electronic storage, manipulation, retrieval and display of data. Geodatabase stores spatial and attribute data and the relationships that exist between them.
- A software package designed to facilitate creation and maintenance of computerised database is called as a Database Management System (DBMS). Different models for efficient storage and retrieval of data are hierarchical, network, relational, multidimensional and object-based.

7.13 ACTIVITY

- You have learnt about the raster and vector data structures. Now download satellite images from the websites. Open an image using image processing software and observe the clarity of pixel while you zoom the image. And also try to identify various spatial features.

7.14 TERMINAL QUESTIONS

1. Discuss the digital representation of data and geospatial data and their related technical issues.
2. What is spatial data structure? Explain vector data structure.
3. Explain the relative merits and demerits of raster and vector data structures.
4. Explain the concept of geodatabase.

7.15 REFERENCES

- Burrough, P. A. and McDonnell, R. A., (1998), Principles of Geographical Information Systems, Oxford University Press, New York
- Chang, K.-t., (2010), Introduction to Geographic Information Systems, Tata McGraw-Hill, New Delhi.
- https://homepage.cs.uri.edu/faculty/wolfe/book/Readings/Reading02.htm
- https://www.academia.edu/3045247/Spatial_Data_Structures
- https://courseware.cutm.ac.in/wp-content/uploads/2020/05/lecture-5.pdf

(The above websites were retrieved between 9/11/2022 and 12/11/2022)

7.16 FURTHER/SUGGESTED READINGS

Contributor: Dr. O. M. Murali

- Burrough, P. A. and McDonnell, R. A., (1998), Principles of Geographical Information Systems, Oxford University Press, New York
- Chang, K.-t., (2010), Introduction to Geographic Information Systems, Tata McGraw-Hill, New Delhi.

7.17 ANSWERS

SAQ I

a) analog, digital

b) functionality, efficiently

c) characterise, data

d) cells, run

SAQ II

a) grid or cell, processing

b) vector, unique ID

c) DBMS

d) Geodatabase

e) Structured Query Language (SQL)

Terminal Questions

1. Please refer to subsection 7.2.1 and 7.2.2.
2. Please refer to section 7.4 and 7.5.
3. Please refer to section 7.7.
4. Please refer to section 7.10.

Contributor: Dr. O. M. Murali

GLOSSARY

Area	:	A closed two-dimensional shape defined by its boundary.
Coverage	:	A topological vector data format
Data	:	Any collection of related facts arranged in a particular format; often, the basic elements of information that are produced, stored, or processed by a computer.
Database	:	One or more structured sets of persistent data, managed and stored as a unit and generally associated with software to update and query the data. A simple database might be a single file with many records, each of which references the same set of fields.
Data structure	:	The format of the data as stored and manipulated on the computer.
DBMS (Database Management System)	:	Software designed to access and structure a database.
Line	:	A shape having length and direction but no area, connecting at least two x, y coordinates. Lines represent geographic features too narrow to be displayed as an area at a given scale, such as contours, street centerlines, or streams, or linear feature, with no area, such as state and district boundary lines.
Models	:	the abstract representation of the real world in various forms like pictorial/graphical/sculpture.
Network database	:	A database that is based on the built-in connections across tables.
Object-oriented data model	:	A data model that uses objects to organize spatial data. An object is an entity such as a land parcel that has a set of properties and can perform operations upon requests.
Point	:	A single x, y coordinate that represents a geographic feature too small to be displayed as a line or area at that scale.
Raster data model	:	It is a computer based representation of spatial or geographical data which are saved as grid cells evenly spaced as rows and columns.
Spaghetti data	:	It is a non-topological vector data model, where spatial relations are not possible.
Spatial data	:	Information about the locations and shapes of geographic features, and the relationships between them; usually stored

		as coordinates and topology.
SQL (Structured Query Language)	:	A standard language interface to relational database management systems.
Topology	:	It can be defined as the organisation of spatial relationships between features in a GIS.
Vector data model	:	It is a computer based representation of spatial or geographical data which are saved as coordinates where every feature is a discrete object.

ABBREVATIONS

ASCII	:	American Standard Code for Information Interchange
CAD	:	Computer Aided Design
CGIS	:	Canadian Geographic Information Systems
CPU	:	Central Processing Unit
DBMS	:	Database Management System
DDL	:	Data Definition Language
DEM	:	Digital Elevation Model
DGPS	:	Differential Global Positioning System
DIME	:	Dual Independent Map Encoding
DML	:	Data Manipulation Language
ESRI	:	Environmental Systems Research Institute
GBF	:	Geographic Base File
GIRAS	:	Geographical Information Retrieval and Analysis System
GIS	:	Geographical Information System
GPS	:	Global Positioning System
GRASS	:	Geographic Resources Analysis Support System
GUI	:	Graphical User Interface
IGIS	:	Integrated GIS and Image Processing Software
IMS	:	Information Management System
ISRO	:	Indian Space Research Organisation
IIRS	:	Indian Institute of Remote Sensing
IIT	:	Indian Institute of Technology
JIS	:	Joint Information System
LAMIS	:	Local Authority Management Information System
NRSC	:	National Remote Sensing Centre
OGC	:	Open Geospatial Consortium
OODB	:	Object-Oriented Database
OOPL	:	Object-Oriented Programming Language
PC	:	Personal Computer
RLE	:	Run-Length Encoding
SAC	:	Space Applications Centre
SLDB	:	Swedish Land Data Bank
SQL	:	Structured Query Language

WWW : World Wide Web