

T.Y. B.Sc. (IT) : Sem. VI
Principles of Geographic Information Systems



Time : 2½ Hrs.]

Prelim Question Paper Solution

[Marks : 75]

Q.1 Attempt any THREE question of the following : **[15]**

Q.1(a) Define GIS. **[5]**

Ans.: A GIS is a computer-based system that provides the following four sets of capabilities to handle georeferenced data:

1. Data capture and preparation
2. Data management, including storage and maintenance
3. Data manipulation and analysis
4. Data presentation

This implies that a GIS user can expect support from the system to enter (georeferenced) data, to analyse it in various ways, and to produce presentations (including maps and other types) from the data. This would include support for various kinds of coordinate systems and transformations between them, options for analysis of the georeferenced data, and obviously a large degree of freedom of choice in the way this information is presented (such as colour scheme, symbol set, and medium used).

Q.1(b) Explain the terms 'GISystems', 'GIScience' and 'GIS Application' **[5]**

Ans.: **GISystems:**

A geographic information system is a computerized system that facilitates the phases of data entry, data management, data analysis and data presentation specifically for dealing with georeferenced data. In the 'wider' sense, a functioning GIS requires both hardware and software, and also people such as the database creators or administrators, analysts who work with the software, and the users of the end product.

GIScience:

The discipline that deals with all aspects of the handling of spatial data and geoinformation is called geographic information science (often abbreviated to geoinformation science or just GIScience).

Geo-Information Science is the scientific field that attempts to integrate different disciplines studying the methods and techniques of handling spatial information.

Related terms include geoinformatics, geomatics, and spatial information science. These are all similar terms which have much the same meaning.

GIS Application: A GIS Application is the methodology of using a GIS Software (such as QGIS, ArcGIS etc.) in a specific field of application (such as studying the average temperature of an area or studying the decline in fisheries of a river etc.). This term is distinct from the term 'GIS Software', since a GIS Software can do many GIS Applications within it, whereas a GIS Application is the use of a GIS Software for a specific purpose.

Q.1(c) Explain the term 'database'. **[5]**

Ans.: A database is a repository for storing large amounts of data. It comes with a number of useful functions:

1. A database can be used by multiple users at the same time i.e. it allows concurrent use,
2. A database offers a number of techniques for storing data and allows the use of the most efficient one i.e. it supports storage optimization,
3. A database allows the imposition of rules on the stored data; rules that will be automatically checked after each update to the data i.e. it supports data integrity,
4. A database offers an easy to use data manipulation language, which allows the execution of all sorts of data extraction and data updates i.e. it has a query facility,

5. A database will try to execute each query in the data manipulation language in the most efficient way i.e. it offers query optimization.

Databases can store almost any kind of data. Modern database systems organize the stored data in tabular format. A database may have many such tables, each of which stores data of a certain kind. It is not uncommon for a table to have many thousands of data rows, sometimes even hundreds of thousands.

Q.1(d) Define Geographic phenomena. [5]

Ans.: A GIS operates under the assumption that the relevant spatial phenomena occur in a two- or three-dimensional Euclidean space, unless otherwise specified. Euclidean space can be informally defined as a model of space in which locations are represented by coordinates—(x, y) in 2D; (x, y, z) in 3D—and distance and direction can be defined with geometric formulas. In the 2D case, this is known as the Euclidean plane, which is the most common Euclidean space in GIS use.

In order to be able to represent relevant aspects of real world phenomena inside a GIS, we first need to define what it is we are referring to. We might define a geographic phenomenon as a manifestation of an entity or process of interest that:

- Can be named or described,
- Can be georeferenced, and
- Can be assigned a time (interval) at which it is/was present.

The relevant phenomena for a given application depends entirely on one's objectives. For instance, in water management, the objects of study might be river basins, agro-ecologic units, measurements of actual evapotranspiration, meteorological data, ground water levels, irrigation levels, water budgets and measurements of total water use. Note that all of these can be named or described, geo-referenced and provided with a time interval at which each exists.

Q.1(e) Explain the concept of geographic objects. [5]

Ans.: When a geographic phenomenon is not present everywhere in the study area, but somehow 'sparsely' populates it, we look at it as a collection of geographic objects. Such objects are usually easily distinguished and named, and their position in space is determined by a combination of one or more of the following parameters:

- Location (where is it?),
- Shape (what form is it?),
- Size (how big is it?), and
- Orientation (in which direction is it facing?).

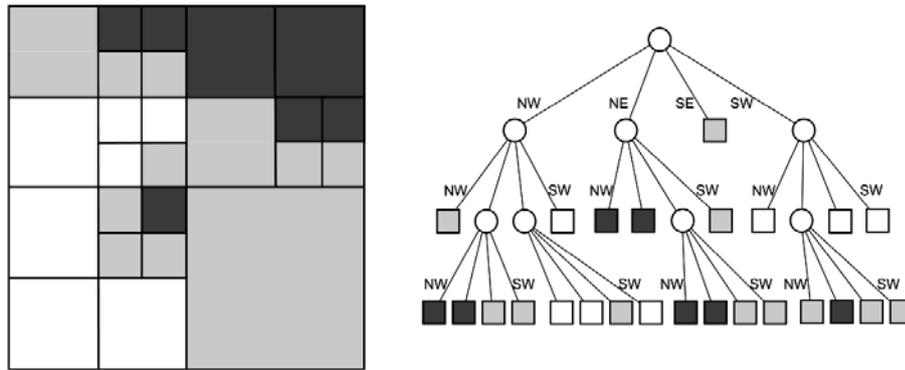
How we want to use the information about a geographic object determines which of the four above parameters is required to represent it. For instance, in an in-car navigation system, all that matters about geographic objects like petrol stations is where they are. Thus, location alone is enough to describe them in this particular context, and shape, size and orientation are not necessarily relevant. Shape is usually important because one of its factors is dimension. This relates to whether an object is perceived as a point feature, or a linear, area or volume feature.

Q.1(f) Explain the concept of irregular tessellations. [5]

Ans.: Regular tessellations provide simple structures with straightforward algorithms, which are, however, not adaptive to the phenomena they represent. Essentially this means they might not represent the phenomena in the most efficient way. For this reason, substantial research effort has also been put into irregular tessellations. Again, these are partitions of space into mutually disjoint cells, but now the cells may vary in size and shape, allowing them to adapt to the spatial phenomena that they represent. We discuss here only one type, namely the region quadtree, but we point out that many more structures have been proposed in the literature, and have also been implemented.

Irregular tessellations are more complex than the regular ones, but they are also more adaptive, which typically leads to a reduction in the amount of memory used to store the data. A well-known data structure in this family—upon which many more variations have been based—is the region

quadtree. It is based on a regular tessellation of square cells, but takes advantage of cases where neighboring cells have the same field value, so that they can together be represented as one bigger cell. A simple illustration is provided in Figure



Q.2 Attempt any THREE question of the following : [15]

Q.2(a) Explain the hardware and software trends in computers in the past few decades. [5]

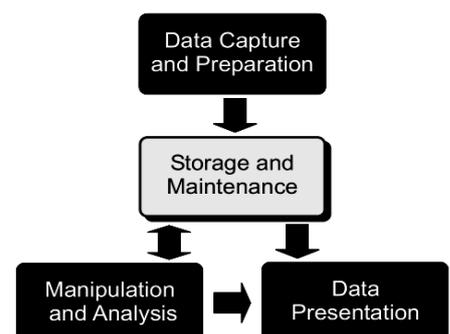
Ans.: Advances in computer hardware seem to take place at an ever-increasing rate. Every several months, a faster, more powerful processor generation replaces the previous one. Computers are also becoming increasingly portable, while offering this increased performance. The computing power that we have available in today's handheld computers is a multiple of the performance that the first PC had when it was introduced in the early 1980's. In fact, current PCs have orders of magnitude more memory and storage capacity than the so-called minicomputers of 25 years ago. To illustrate this trend: compare a typical early 1980's PC with a 2 MHz CPU, 128 kbytes of main memory, and a 10 MByte hard disk to the current generation of desktop PC's.

Computers are also becoming increasingly affordable. Hand-held computers are now commonplace in business and personal use, equipping field surveyors with powerful tools, complete with GPS capabilities for instantaneous georeferencing. To support these hardware trends, software providers continue to produce application programs and operating systems that, while providing a lot more functionality, also consume significantly more memory. In general, software technology has developed somewhat slower and often cannot fully utilise the possibilities offered by the exponentially growing hardware capabilities. Existing software obviously performs better when run on faster computers.

Alongside these trends, there have also been significant developments in computer networks. In essence, today almost any computer on Earth can connect to some network, and contact computers virtually anywhere else, allowing fast and reliable exchange of (spatial) data. Mobile phones are more and more frequently being used to connect to computers on the Internet. The UMTS protocol (Universal Mobile Telecommunications System), allows digital communication of text, audio, and video at a rate of approximately 2 Mbps. The new HSDPA protocol offers up to 10 times this speed. Looking at these developments it is clear that the combination of a GPS receiver, a portable computer and mobile phone has already dramatically changed our world, certainly so for out-of-office activities of Earth science professionals.

Q.2(b) Write a note on 'GIS Architecture and Functionality'. [5]

Ans.: A GIS consists of several *functional components*-components which support key GIS functions. These are data capture and preparation, data storage, data analysis, and presentation of spatial data. Figure shows a diagram of these components, with arrows indicating the data flow in the system.



For a particular GIS, each of these components may provide many or only a few functions. Arguably, the system should not be called a geographic information system if any one of these components is missing. It is important to note however, that the same function may be offered by different components of the GIS: for instance, data capture and data storage may have functions in common, and the same holds for data preparation and data analysis.

Q.2(c) Define and explain the concept of Spatial Data Infrastructure (SDI). [5]

Ans.: For reasons that include efficiency and legislation, many organizations are forced to work in a cooperative setting in which geographic information is obtained from, and provided to, partner organizations and the general public. The sharing of spatial data between the various GISs in those organizations is of key importance and aspects of data dissemination, security, copyright and pricing require special attention. The design and maintenance of a Spatial Data Infrastructure (SDI) deals with these issues.

An SDI is defined as "the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data". Fundamental to those arrangements are—in a wider sense—the agreements between organizations and in the narrow sense, the agreements between software systems on how to share the geographic information. In SDI, standards are often the starting point for those agreements. Standards exist for all facets of GIS, ranging from data capture to data presentation. They are developed by different organizations, of which the most prominent are the International Organization for Standardisation (ISO) and the Open Geospatial Consortium (OGC).

Typically, an SDI provides its users with different facilities for finding, viewing, downloading and processing data. Because the organizations in an SDI are normally widely distributed over space, computer networks are used as the means of communication. With the development of the internet, the functional components of GIS have been gradually become available as web-based applications.

Q.2(d) Explain the stage of Spatial Data Presentation. [5]

Ans.: The presentation of spatial data, whether in print or on-screen, in maps or in tabular displays, or as 'raw data', is closely related to the disciplines of cartography, printing and publishing. The presentation may either be an endproduct, for example as a printed atlas, or an intermediate product, as in spatial data made available through the internet.

Method	Devices
Hard copy	<ul style="list-style-type: none"> • printer • plotter (pen plotter, ink-jet printer, thermal transfer printer, electrostatic plotter) • film writer
Soft copy	<ul style="list-style-type: none"> • computer screen
Output of digital data sets	<ul style="list-style-type: none"> • magnetic tape • CD-ROM or DVD • the Internet

Cartography and scientific visualization make use of these methods and devices to produce their products.

Q.2(e) Explain the various reasons for using a DBMS. [5]

Ans.: There are various reasons why one would want to use a DBMS for data storage and processing.

- A DBMS supports the storage and manipulation of very large data sets.
- A DBMS can be instructed to guard over data correctness.
- A DBMS supports the concurrent use of the same data set by many users.
- A DBMS provides a high-level, declarative query language.
- A DBMS supports the use of a data model. A data model is a language with which one can define a database structure and manipulate the data stored in it.
- A DBMS includes data backup and recovery functions to ensure data availability at all times.
- A DBMS allows the control of data redundancy.

Q.2(f) Explain Spatial Database Functionality.**[5]**

Ans.: DBMS vendors have over the last 20 years recognized the need for storing more complex data, like spatial data. The main problem was that there is additional functionality needed by DBMS in order to process and manage spatial data. As the capabilities of our hardware to process information has increased, so too has the desire for better ways to represent and manage spatial data. During the 1990's, object-oriented and object-relational data models were developed for just this purpose. These extend standard relational models with support for objects, including 'spatial' objects.

Currently, GIS software packages are able to store spatial data using a range of commercial and open source DBMSs such as Oracle, Informix, IBM DB2, Sybase, and PostgreSQL, with the help of spatial extensions. Some GIS software have integrated database 'engines', and therefore do not need these extensions. ESRI's ArcGIS, for example, has the main components of the MS Access database software built-in. This means that the designer of a GIS application can choose whether to store the application data in the GIS or in the DBMS. Spatial databases, also known as geodatabases, are implemented directly on existing DBMSs, using extension software to allow them to handle spatial objects.

A spatial database allows users to store, query and manipulate collections of spatial data.

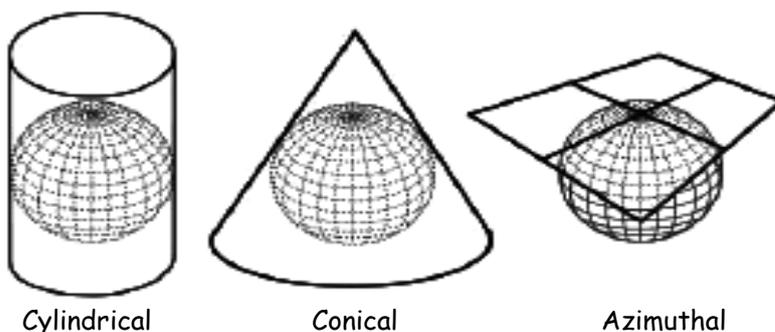
A geodatabase allows a wide variety of users to access large data sets (both geographic and alphanumeric), and the management of their relations, guaranteeing their integrity. The Open Geospatial Consortium (OGC) has released a series of standards relating to geodatabases that (amongst other things), define:

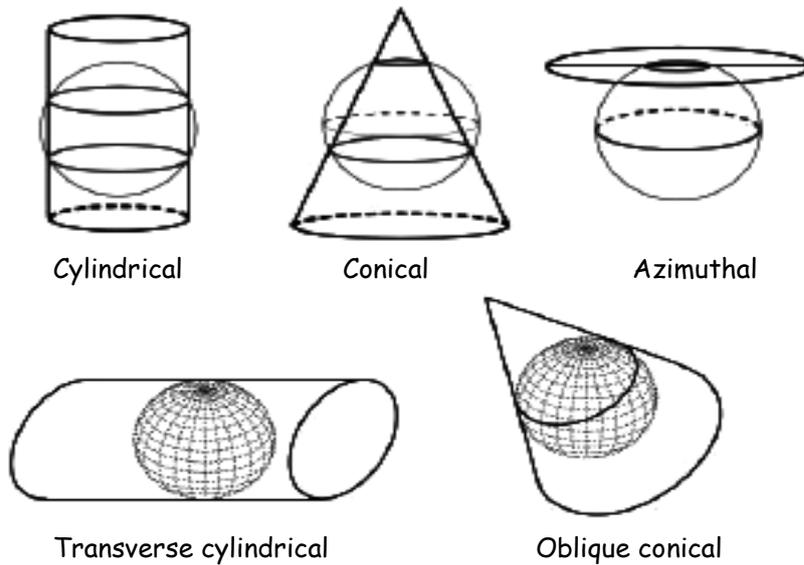
- Which tables must be present in a spatial database (i.e. geometry columns table and spatial reference system table)
- The data formats, called 'Simple Features' (i.e. point, line, polygon, etc.)
- A set of SQL-like instructions for geographic analysis.

Q.3 Attempt any THREE question of the following :**[15]****Q.3(a) Explain how map projections are classified.****[5]**

Ans.: Hundreds of map projections have been developed, each with its own specific qualities. These qualities in turn make resulting maps useful for certain purposes- By definition, any map projection is associated with scale distortions-There is simply no way to flatten out a piece of ellipsoidal or spherical surface without stretching some parts of the surface more than others. The amount and which kind of distortions a map will have depends on the type of the map projection that has been selected.

Some map projections can be visualized as true geometric projections directly onto the mapping plane, in which case we call it an azimuthal projection, or onto an intermediate surface, which is then rolled out into the mapping plane. Typical choices for such intermediate surfaces are cones and cylinders. Such map projections are then called conical, and cylindrical, respectively. Figure shows the surfaces involved in these three classes of projections.





Q.3(b) Explain 2D Polar to 2D Cartesian Coordinate Transformations. [5]

Ans.: The transformation of polar coordinates (α, d) , into Cartesian map coordinates (x, y) is done when field measurements, angular and distance measurements are transformed into map coordinates. The equation for this transformation is:

$$x = d(\sin(\alpha))$$

$$y = d(\cos(\alpha))$$

The inverse equation is:

$$\alpha = \tan^{-1}\left(\frac{x}{y}\right)$$

$$d^2 = x^2 + y^2$$

Q.3(c) Explain the concept of Satellite Based positioning. [5]

Ans.: Satellite-based positioning was developed and implemented to address military needs, somewhat analogously to the early development of the internet. The technology is now widely available for civilians use. The requirements for the development of the positioning system were:

- Suitability for all kinds of military use: ground troops and vehicles, aircraft and missiles, ships;
- Requiring only low-cost equipment with low energy consumption at the receiver end;
- Provision of results in real time for an unlimited number of users concurrently;
- Support for different levels of accuracy (military versus civilian);
- Around-the-clock and weather-proof availability;
- Use of a single geodetic datum;
- Protection against intentional and unintentional disturbance, for instance, through a design allowing for redundancy,

A satellite-based positioning system set-up involves implementation of three hardware segments:

1. The space segment, i.e. the satellites that orbit the Earth, and the radio signals that they emit,
2. The control segment, i.e. the ground stations that monitor and maintain the space segment components, and
3. The user segment, i.e. the users with their hard- and software to conduct positioning.

Q.3(d) Explain how to obtain spatial data by 'Direct Capture'. [5]

Ans.: One way to obtain spatial data is by direct observation of the relevant geographic phenomena. This can be done through ground-based field surveys, or by using remote sensors in satellites or airplanes. Many Earth sciences have developed their own survey techniques, as ground-based techniques remain the most important source for reliable data in many cases.

Data which is captured directly from the environment is known as primary data.

With primary data the core concern in knowing its properties is to know the process by which it was captured, the parameters of any instruments used and the rigour with which quality requirements were observed.

Remotely sensed imagery is usually not fit for immediate use, as various sources of error and distortion may have been present, and the imagery should first be freed from these. This is the domain of remote sensing.

An image refers to raw data produced by an electronic sensor, which are not pictorial, but arrays of digital numbers related to some property of an object or scene, such as the amount of reflected light.

In practice, it is not always feasible to obtain spatial data by direct spatial data capture. Factors of cost and available time may be a hindrance, or previous projects sometimes have acquired data that may fit the current project's purpose.

Q.3(e) Explain the need to combine spatial data from various sources. [5]

Ans.: A GIS project usually involves multiple data sets, so the next step addresses the issue of how these multiple sets relate to each other. There are four fundamental cases to be considered in the combination of data from different sources:

1. They may be about the same area, but differ in accuracy,
2. They may be about the same area, but differ in choice of representation,
3. They may be about adjacent areas, and have to be merged into a single data set.
4. They may be about the same or adjacent areas, but referenced in different coordinate systems.

Q.3(f) Describe the process of Point Data Transformation. [5]

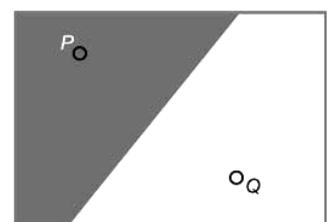
Ans.: We may want to transform our points into other representations in order to facilitate interpretation and/or integration with other data. Examples include defining homogeneous areas (polygons) from our point data, or deriving contour lines. This is generally referred to as interpolation, i.e. the calculation of a value from 'surrounding' observations. The principle of spatial autocorrelation plays a central part in the process of interpolation.

In order to predict the value of a point for a given (x, y) location, we could simply find the 'nearest' known value to the point, and assign that value. This is the simplest form of interpolation, known as nearest-neighbour interpolation. We might instead choose to use the distance that points are away from (x, y) to weight their importance in our calculation.

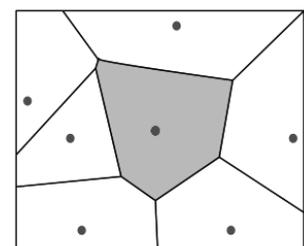
Following are some important techniques for interpolation:

(i) Interpolating discrete data:

If we are dealing with discrete (nominal, categorical or ordinal) data, we are effectively restricted to using nearest-neighbour interpolation. This is the situation shown in Figure, though usually we would have many more points.



If the desired output was a polygon layer, we could construct Thiessen polygons around the points of measurement. The boundaries of such polygons, by definition, are the locations for which more than one point of measurement is the closest point. An illustration is provided in Figure:



(ii) Interpolating continuous data:

Interpolation of values from continuous measurements is significantly more complex.

Since the data are continuous, we can make use of measured values for interpolation. There are many continuous geographic fields—elevation, temperature and ground water salinity are just a few examples. Commonly, continuous fields are represented as rasters, and we will almost by default assume that they are.

Q.4 Attempt any THREE question of the following :

[15]

Q.4(a) Explain the classification of Analytical GIS Capabilities.

[5]

Ans.: There are many ways to classify the analytical functions of a GIS.

1. **Classification, retrieval, and measurement functions:** All functions in this category are performed on a single (vector or raster) data layer, often using the associated attribute data.
 - Classification allows the assignment of features to a class on the basis of attribute values or attribute ranges (definition of data patterns). On the basis of reflectance characteristics found in a raster, pixels may be classified as representing different crops, such as potato and maize.
 - Retrieval functions allow the selective search of data. We might thus retrieve all agricultural fields where potato is grown.
 - Generalization is a function that joins different classes of objects with common characteristics to a higher level (generalized) class. For example, we might generalize fields where potato or maize, and possibly other crops, are grown as 'food produce fields'.
 - Measurement functions allow the calculation of distances, lengths, or areas.

2. **Overlay functions:** These belong to the most frequently used functions in a GIS application. They allow the combination of two (or more) spatial data layers comparing them position by position, and treating areas of overlap—and of non-overlap—in distinct ways. Many GISs support overlays through an algebraic language, expressing an overlay function as a formula in which the data layers are the arguments. In this way, we can find
 - The potato fields on clay soils (select the 'potato' cover in the crop data layer and the 'clay' cover in the soil data layer and perform an intersection of the two areas found).
 - The fields where potato or maize is the crop (select both areas of 'potato' and 'maize' cover in the crop data layer and take their union).
 - The potato fields not on clay soils (perform a difference operator of areas with 'potato' cover with the areas having clay soil).
 - The fields that do not have potato as crop (take the complement of the potato areas).

3. **Neighbourhood functions:** Whereas overlays combine features at the same location, neighbourhood functions evaluate the characteristics of an area surrounding a feature's location. A neighbourhood function 'scans' the neighbourhood of the given feature(s), and performs a computation on it.
 - **Search functions** allow the retrieval of features that fall within a given search window. This window may be a rectangle, circle, or polygon.
 - **Buffer zone generation** (or buffering) is one of the best known neighbourhood functions. It determines a spatial envelope (buffer) around (a) given feature(s). The created buffer may have a fixed width, or a variable width that depends on characteristics of the area.
 - **Interpolation functions** predict unknown values using the known values at nearby locations. This typically occurs for continuous fields, like elevation, when the data actually stored does not provide the direct answer for the location(s) of interest.
 - **Topographic functions** determine characteristics of an area by looking at the immediate neighbourhood as well. Typical examples are slope computations on digital terrain models (i.e. continuous spatial fields). The slope in a location is defined as the plane tangent to the topography in that location. Various computations can be performed, such as:
 - determination of slope angle
 - determination of slope aspect

- determination of slope length
 - determination of contour lines, These are lines that connect points with the same value (for elevation, depth, temperature, barometric pressure, water salinity etc).
4. **Connectivity functions:** These functions work on the basis of networks, including road networks, water courses in coastal zones, and communication lines in mobile telephony. These networks represent spatial linkages between features. Main functions of this type include:
- **Contiguity functions** evaluate a characteristic of a set of connected spatial units. One can think of the search for a contiguous area of forest of certain size and shape in a satellite image.
 - **Network analytic functions** are used to compute over connected line features that make up a network. The network may consist of roads, public transport routes, high voltage lines or other forms of transportation infrastructure. Analysis of such networks may entail shortest path computations (In terms of distance or travel time) between two points in a network for routing purposes. Other forms are to find all points reachable within a given distance or duration from a start point for allocation purposes, or determination of the capacity of the network for transportation between an indicated source location and sink location.
 - **Visibility functions** also fit in this list as they are used to compute the points visible from a given location (viewshed modelling or viewshed mapping) using a digital terrain model.

Q.4(b) Explain the various spatial selection queries.

[5]

Ans.: Following are the spatial selection query types:

1. **Interactive Spatial Selection:**

In interactive spatial selection, one defines the selection condition by pointing at or drawing spatial objects on the screen display, after having indicated the spatial data layer(s) from which to select features. The interactively defined objects are called the selection objects; they can be points, lines, or polygons. The GIS then selects the features in the indicated data layer(s) that overlap (i.e. intersect, meet, contain, or are contained in) with the selection objects. These become the selected objects.

2. **Spatial Selection by Attribute Conditions:**

It is also possible to select features by using selection conditions on feature attributes. These conditions are formulated in SQL if the attribute data reside in a geodatabase. This type of selection answers questions like "where are the features with?"

3. **Combining Attribute Conditions:**

When multiple criteria have to be used for selection, we need to carefully express all of these in a single composite condition. The tools for this come from a field of mathematical logic, known as propositional calculus.

Above, we have seen simple, atomic conditions such as $\text{Area} < 400000$, and $\text{LandUse} = 80$. Atomic conditions use a predicate symbol, such as $<$ (less than) or $=$ (equals). Other possibilities are \leq (less than or equal), $>$ (greater than), \geq (greater than or equal) and \neq (does not equal). Any of these symbols is combined with an expression on the left and one on the right. For instance, $\text{LandUse} \neq 80$ can be used to select all areas with a land use class different from 80. Expressions are either constants like 400000 and 80, attribute names like Area and LandUse, or possibly composite arithmetic expressions like $0.15 \times \text{Area}$, which would compute 15% of the area size.

Atomic conditions can be combined into composite conditions using logical connectives. The most important ones are AND, OR, NOT and the bracket pair (...). If we write a composite condition like

$\text{Area} < 400000 \text{ AND } \text{LandUse} = 80$

we can use it to select areas for which both atomic conditions hold true. This is the meaning of the AND connective.

Q.4(c) Explain the technique named 'Classification'.

[5]

Ans.: **Classification** is a technique of purposefully removing detail from an input data set, in the hope of revealing important patterns (of spatial distribution). In the process, we produce an output data set, so that the input set can be left intact. We do so by assigning a characteristic value to each element in the input set, which is usually a collection of spatial features that can be raster cells or points, lines or polygons. If the number of characteristic values is small in comparison to the size of the input set, we have classified the input set.

The pattern that we look for may be the distribution of household income in a city. Household income is called the classification parameter. If we know for each ward in the city the associated average income, we have many different values. Subsequently, we could define five different categories (or: classes) of income: 'low', 'below average', 'average', 'above average' and 'high', and provide value ranges for each category. If these five categories are mapped in a sensible colour scheme, this may reveal interesting information. This has been done for Dar es Salaam in Figure in two ways.



Fig.: Two classifications of average annual household income per ward in Dar es Salaam, Tanzania. Higher income areas in darker greens. Five categories were identified, (a) with original polygons left intact; (b) with original polygons merged when in same category. The data used for this illustration are not (b) factual.

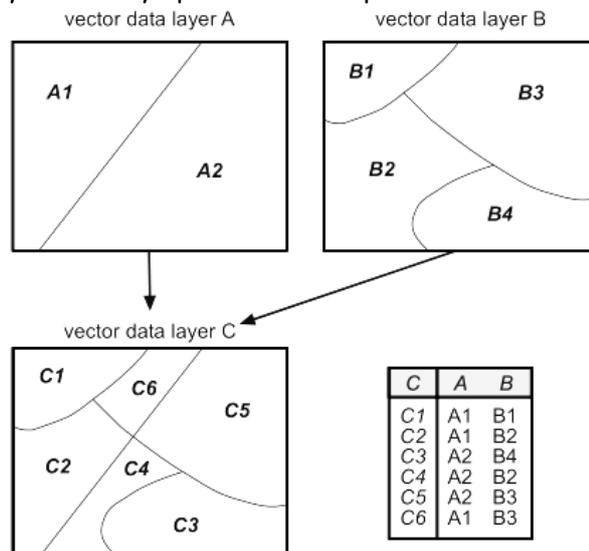
Q.4(d) Explain Vector and Raster Overlay Functions.

[5]

Ans.: Overlay is the process of taking two spatial layers and combining them into one. It is mainly of two types:

1. Vector overlay:

In the vector domain, overlay is computationally more demanding than in the raster domain. Here we will only discuss overlays from polygon data layers, but we note that most of the ideas also apply to overlay operations with point or line data layers.



2. Raster Overlay:

Vector overlay operators are useful, but geometrically complicated, and this sometimes results in poor operator performance. Raster overlays do not suffer from this disadvantage, as most of them perform their computations cell by cell and thus they are fast.

GISs that support raster processing—as most do—usually have a language to express operations on rasters. These languages are generally referred to as map algebra, or sometimes raster calculus. They allow a GIS to compute new rasters from existing ones, using a range of functions and operators. Unfortunately, not all implementations of map algebra offer the same functionality.

The syntax often differs for different GIS software packages.

When producing a new raster we must provide a name for it, and define how it is computed.

This is done in an assignment statement of the following format:

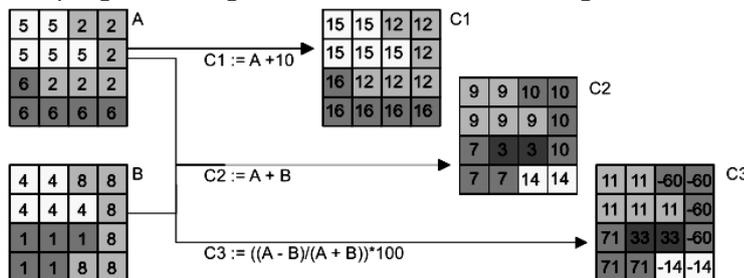
Output_raster_name := Map_algebra_expression

Q.4(e) Explain arithmetic operators on raster data.

[5]

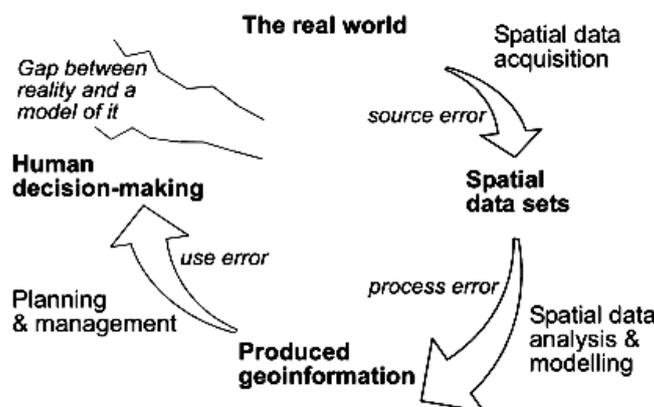
Ans.: Various arithmetic operators are supported. The standard ones are multiplication (×), division (/), subtraction (−) and addition (+). Obviously, those arithmetic operators should only be used on appropriate data values, and for instance, not on classification values.

Some simple map algebra assignments are illustrated in Figure:



Q.4(f) Explain how errors propagate in spatial data processing.

Ans.: A number of sources of error may be present in source data. It is important to note that the acquisition of base data to a high standard of quality still does not guarantee that the results of further, complex processing can be treated with certainty. As the number of processing steps increases, it becomes difficult to predict the behaviour of error propagation. These various errors may affect the outcome of spatial data manipulations. In addition, further errors may be introduced during the various processing steps, as illustrated in Figure.



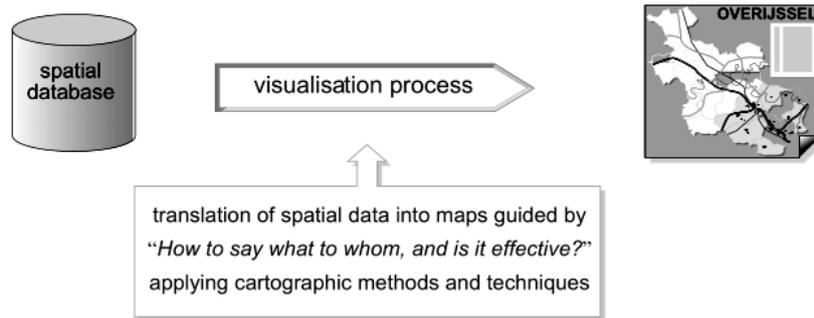
Q.5 Attempt any THREE question of the following :

[15]

Q.5(a) Explain the visualization process in brief.

Ans.: The characteristic of maps and their function in relation to the spatial data handling process was explained in the previous section. In this context the cartographic visualization process is considered to be the translation or conversion of spatial data from a database into graphics. These

are predominantly maplike products. During the visualization process, cartographic methods and techniques are applied. These can be considered to form a kind of grammar that allows for the optimal design and production for the use of maps, depending on the application (see Figure).



Q.5(b) Describe the various visualization strategies.

[5]

Ans.: There are two main strategies for visualization:

- **Present:**

Traditionally the cartographer's main task was the creation of good cartographic products. This is still true today. The main function of maps is to communicate geographic information, i.e. to inform the map user about location and nature of geographic phenomena and spatial patterns. This has been the map's function throughout history. Well-trained cartographers are designing and producing maps, supported by a whole set of cartographic tools and theory as described in cartographic textbooks.

- **Explore:**

Exploration means to search for spatial, temporal or spatio-temporal patterns, relationships between patterns, or trends. In case of a search for patterns, a domain expert may be interested in aspects like the distribution of a phenomenon, the occurrence of anomalies, the sequence of appearances and disappearances. A search for relationships between patterns could include: changes in vegetation indices and climatic parameters, location of deprived urban areas and their distance to educational facilities. A search for trends could, for example, focus on the development in distribution and frequency of landslides. Maps not only enable these types of searches, findings may also trigger new questions, and lead to new visual exploration (or analytical) acts.

Q.5(c) Explain the types of cartographic data that one can represent using cartographic symbols.

[5]

Ans.: Data will be of a qualitative or quantitative nature. Qualitative data is also called nominal or categorical data. This data exists as discrete, named values without a natural order amongst the values. Examples are the different languages (e.g. English, Swahili, Dutch), the different soil types (e.g. sand, clay, peat) or the different land use categories (e.g. arable land, pasture). In the map, qualitative data are classified according to disciplinary insights such as a soil classification system represented as basic geographic units: homogeneous areas associated with a single soil type, recognized by the soil classification.

Quantitative data can be measured, either along an interval or ratio scale. For data measured on an interval scale, the exact distance between values is known, but there is no absolute zero on the scale. Temperature is an example: 40 °C is not twice as warm as 20 °C, and 0 °C is not an absolute zero. Quantitative data with a ratio scale does have a known absolute zero. An example is income: someone earning \$100 earns twice as much as someone with an income of \$50. In order to generate maps, quantitative data are often classified into categories according to some mathematical method.

In between qualitative and quantitative data, one can distinguish ordinal data. These data are measured along a relative scale, based on hierarchies. For instance, one knows that one value is 'more' than another value, such as 'warm' versus 'cool'. Another example is a hierarchy of road types:

'highway', 'main road', 'secondary road' and 'track'. The different types of data are summarized in Table.

Measurement Scale	Nature of data
Nominal, categorical	Data of different nature/identity of things (qualitative)
Ordinal	Data with a clear element of order, though not quantitatively determined (ordered).
Interval	Quantitative information with arbitrary zero.
Ratio	Quantitative data with absolute zero.

Q.5(d) Explain how to map various types of data.

[5]

Ans.: Basic elements of a map, irrespective of the medium on which it is displayed, are point symbols, line symbols, area symbols, and text. The appearance of point, line, and area symbols can vary depending on their nature. Most maps in this book show symbols in different size, shape and colour. Points can vary in form or colour to represent the location of shops or they can vary in size to represent aggregated values (like number of inhabitants) for an administrative area. Lines can vary in colour to distinguish between administrative boundaries and rivers, or vary in shape to show the difference between railroads and roads. Areas follow the same principles: difference in colour distinguishes between different vegetation stands.

1. How to Map qualitative data:

If, after a long fieldwork period, one has finally delineated the boundaries of a province's watersheds, one likely is interested in a map showing these areas. The geographic units in the map will have to represent the individual watersheds. In such a map, each of the watersheds should get equal attention, and none should stand out above the others.



2. How to Map Quantitative Data:

When, after executing a census, one would for instance like to create a map with the number of people living in each municipality, one deals with absolute quantitative data. The geographic units will logically be the municipalities. The final map should allow the user to determine the amount per municipality and also offer an overview of the geographic distribution of the phenomenon. To reach this objective, the symbols used should have quantitative perception properties. Symbols varying in size fulfil this demand.



3. How to Map Terrain Elevation:

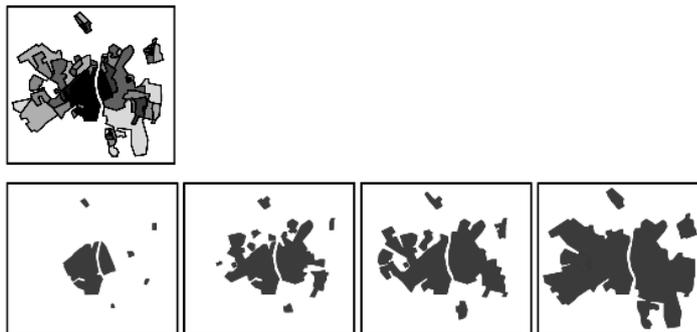
Terrain elevation can be mapped using different methods. Often, one will have collected an elevation data set for individual points like peaks, or other characteristic points in the terrain. Obviously, one can map the individual points and add the height information as text. However, a contour map, in which the lines connect points of equal elevation, is generally used. To visually improve the information content of such a map the space between the contour lines can be filled with colour and value information following a convention, e.g. green for low elevation and brown for high elevation areas. This technique is known as hypsometric or layer tinting. Even

more advanced is the addition of shaded relief. This will improve the impression of the three-dimensional relief.

4. How to Map Time Series:

Advances in spatial data handling have not only made the third dimension part of GIS routines. Nowadays, the handling of time-dependent data is also part of these routines. This has been caused by the increasing availability of data captured at different periods in time. Next to this data abundance, the GIS community wants to analyse changes caused by real world processes. To that end, single time slice data are no longer sufficient, and the visualization of these processes cannot be supported with only static paper maps.

Mapping time means mapping change. This may be change in a feature's geometry, in its attributes or both. Examples of changing geometry are the evolving coastline of the Netherlands (as displayed in Figure), the location of Europe's national boundaries, or the position of weather fronts.



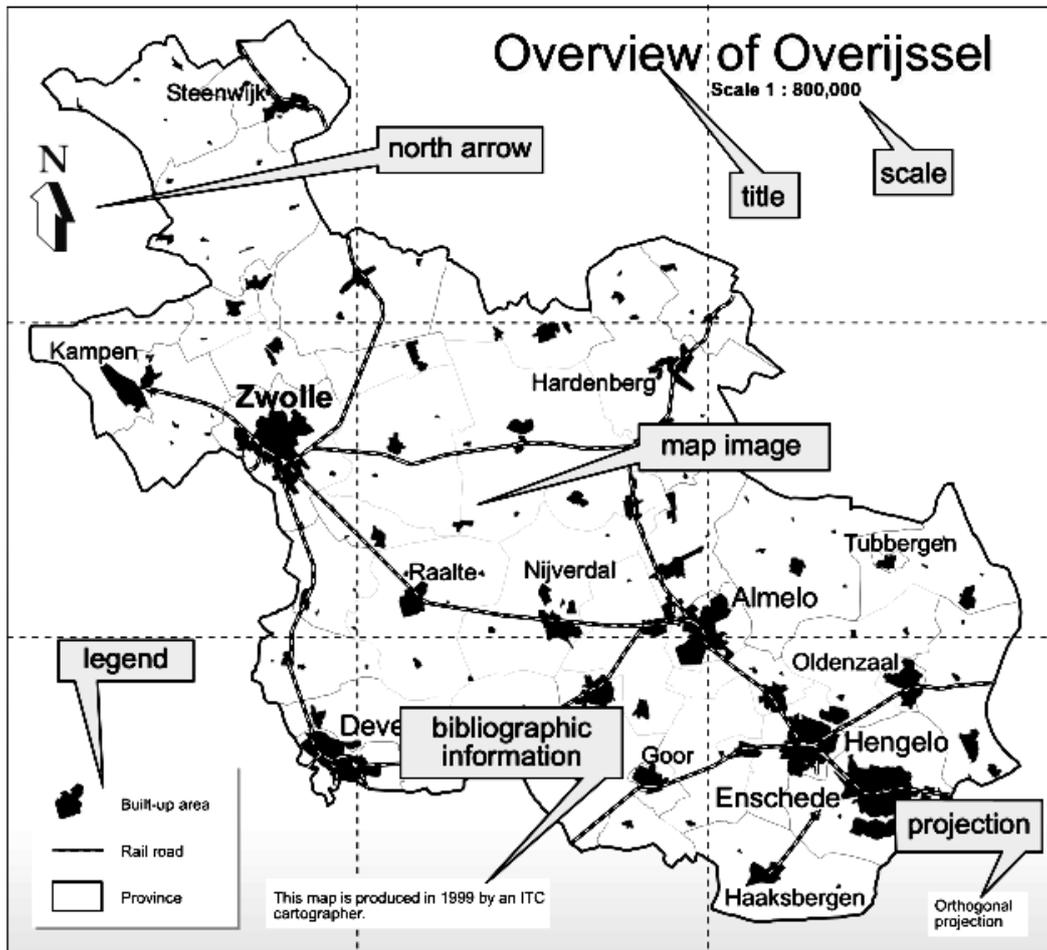
Q.5(e) Explain the concept of Map Cosmetics.

[5]

Ans.: Most maps in this chapter are correct from a cartographic grammar perspective. However, many of them lack the additional information needed to be fully understood that is usually placed in the margin of printed maps. Each map should have, next to the map image, a title, informing the user about the topic visualized. A legend is necessary to understand how the topic is depicted. Additional marginal information to be found on a map is a scale indicator, a north arrow for orientation, the map datum and map projection used, and some lineage information, (such as data sources, dates of data collection, methods used, etc.).

Further information can be added that indicates when the map was issued, and by whom (author / publisher). All this information allows the user to obtain an impression of the quality of the map, and is comparable with metadata describing the contents of a database or data layer.

Figure illustrates these map elements. On paper maps, these elements (if all relevant) have to appear next to the map face itself. Maps presented on screen often go without marginal information, partly because of space constraints. However, on-screen maps are often interactive, and clicking on a map element may reveal additional information from the database. Legends and titles are often available on demand as well.



Q.5(f) Explain the techniques for Map Dissemination. [5]

Ans. : The map design will not only be influenced by the nature of the data to be mapped or the intended audience (the 'what' and 'whom' from "How do I say What to Whom, and is it Effective"), the output medium also plays a role. Traditionally, maps were produced on paper, and many still are.

Currently, most maps are presented on screen, for a quick view, for an internal presentation or for presentation on the WWW. Compared to maps on paper, onscreen maps have to be smaller, and therefore their contents should be carefully selected. This might seem a disadvantage, but presenting maps on-screen offers very interesting alternatives. In one of the previous paragraphs, we discussed that the legend only needs to be a mouse click away. A mouse click could also open the link to a database, and reveal much more information than a paper map could ever offer. Links to other than tabular or map data could also be made available.

Maps and multimedia (photography, sound, video, animation) can be integrated. Some of today's electronic atlases, such as the Encarta World Atlas are good examples of how multimedia elements can be integrated with the map. Pointing to a country on a world map starts the national anthem of the country or shows its flag. It can be used to explore a country's language; moving the mouse would start a short sentence in the region's dialects.

The World Wide Web is nowadays a common medium used to present and disseminate spatial data.

