THE DEVELOPMENT OF A
GEOGRAPHICAL INFORMATION SYSTEM

By

Capt. Ajmal Ijaz Tarrar
Capt. Khalid Qureshi
Capt. Aqeel Anwar
Capt. Mohsin shaheen

Submitted to the Faculty of Computer Sciences,
National University of Sciences and Technology, Rawalpindi,
in partial fulfillment for the requirement of an BE Degree.

September 2000

Dedicated to

our beloved parents

Abstract

The history of using computer for mapping and spatial analysis shows that there has been a parallel development in automated data capture, data analysis, and presentation in broadly related fields. This multiplicity in efforts in several initially separate but closely related fields is now resulting in possibility of linking many kinds of spatial data processing together into truly general purpose geographical information system, as the technical and conceptual problem are overcome. The tremendous development in the fields of GIS has paved ways for new systems to be developed. These system should typically be equipped with features like various methods to input the spatial data (Data capturing methods), provision of different and relevant view to the user of stored data, some new database design for keeping attributes attached to spatial data, new searching and retrieval methods, and three dimensional representation of maps.

The aim of research was to develop a prototype Geographical Information System. This new system is not intended to replace the existing commercially available systems. But rather is an exercise for students to implement and adopt new ideas. The project was an effort to implement some of the above mentioned features.

Acknowledgements

We are humbly thankful to "Almighty Allah" for his blessings on us, and giving us the wisdom, Knowledge and understanding to complete this work.

We wish to express our deepest gratitude and appreciation to our project advisors Dr Dill Muhammad Akbar, Lt Col Hamid (Engrs) and Dr Saeed Murtaza for their supervision, guidance, dedication and commitment to this project.

We also Express our thanks to the officers of C4I directorate (GHQ) who helped us and encouraged us. Our very special thanks to officers of Army Survey and Photogramatic unit, and specially to Major Abrar (Engrs) who gave us his very precious time and ideas for our project.

We are thankful to Capt. Faisal Karim (BESE-2) for all his technical support, which he rendered to us

Our sincere gratitude goes to National University of Sciences and Technology and Military College of Signals for the provision of facilities for this research.

Contents

CHAPTER 4

IMPLEMENTATION

CHAPTER 5

CONCLUSION AND FUTURE WORK

BIBLIOGRAPHY

Chapter 1

Introduction

Introduction.

This chapter is about the introduction of the proposed research area, the goal of the thesis, its general background and the outline of the thesis. First section presents the general outline of the area of research and it covers only recent trends in Geographical Information Systems. The second section covers the major goal of the thesis and sets the objectives for the research. The third section reviews the GIS background and trends in brief. The last section presents the outline of the thesis.

Area of Research

Geographical Information Systems are a special case of information systems with a capability to integrate spatial and non-spatial data. As Geographical Information Systems have gained importance in the mainstream decision making applications as well as in the area of efficient resources management. The tremendous development in this field has paved ways for new systems to develop. These systems should typically be equipped with tools for data input, transforming the data into different projection and coordinate systems, should have the capability to perform complex analysis based on the stored spatial and non spatial data. Such systems should also provide features like provision of different and relevant view to the users of stored data, highly flexible database design for keeping attributes attached to spatial data, new methods for search and retrieval of data. The idea of this project is to develop a prototype Geographical Information System which implements most of above mentioned features.

Goal of the Thesis

The major goal of this research is to develop a fully operational Geographical Information Systems equipped with a graphical user Interface and tools to digitize a map and then add non spatial attributes to objects on the map. The system will integrate data from multiple sources like satellite imagery, aerial photography etc and then perform necessary transformations to bring the data in a common projection/coordinate systems. Our other goal is to design an efficient database structure, which should facilitate the task of not only accessing any required information but also gives accurate results to the complex queries that may be made by a planning or decision support system. The database will store the geometry as well topology of the objects.

Background of GIS

In the past 10 years or so Geographic information System are gaining importance for storing and organizing spatial information in a computer

[Lauterbatch]. GIS empowered by dropping costs of workstations and increasing power of graphics, are beginning to make their way from urban planning, natural resource management, and cartographic production shops into main stream business [Murphy 95].

Geographic Information System are characterized by a wide variety of architectures or platforms, applications, and processing need due to a heterogeneous universe of users including geographic, natural resource management environment sciences and archeology. For example in the comprehensive GIS reference yet [Merger 91b], the authors suggested some eleven definitions of GIS. If there is a distinctive element in these definitional offering, it is that GIS offer the manipulation of a spatial dimension referenced to the earth. For example one of the definitions states that,

> *A Geographic Information System (GIS) is a computer program for storing, retrieving, analyzing, and displaying cartographic data.*

For an efficient and flexible use of these systems it is necessary to combine data acquisition, development of an evaluation scheme and GIS in an integrated concept [Liberman 91]. Appropriate support for acquisition of and operation of GIS may be different from other type of systems. GIS may enter an organization as a shrink-wrapped end user computing application or as a specialized function demanding customized hardware (e.g. digitizer) as well as software. In attempting to define the functionality of GIS, there is a need to review the functions they perform.

While there is a large range of the sophistication in GIS, a typical GIS has to perform following few functions [Maguire 91a].

- Data Capture
- Data Transfer
- Validate & Edit the data
- Store & Structure

- Generalize
- Transform
- Query
- Analyze
- Output

These functions, when considered without reference to spatial versus attribute distinction, are very generic, and could be used to describe high level functionality of many typical GIS. Some other approaches distinguish GIS on the basis of tools, ability to manipulate diverse data types and data sources, and the use of modeling [Bracken 89]. Some experts observe that GIS is developing two traditions: spatial information and spatial analysis. The spatial information stresses large databases representing inventories of geographically important features such as land use, timber reserves, or vegetation cover. For this tradition, geography can be defined as “access mechanism”. Spatial analysis demands a greater range of data models and additional functionality in which geography plays a more significant role in analysis and modeling.

One of the most-cited reasons for adopting GIS is improvement in decision making and problem solving [Attenucci 91]. While there are application in business where the creation of maps may be and end in itself, use of GIS as a decision support tool is likely to be more wide spread. As these systems move into broader use, international standard organizations are coming in contact with this new type of system. In parallel to the advancement in this technology, the quantity of application has increased. From high quality cartography to land planning, natural resource management, environment assessment and planning, ecological research etc, Geographical Information System promise to be one of the most extensive computer application ever to emerge.

Outline of the Thesis

This thesis is comprised of four chapters. In this chapter the covered topics include the general overview of the area of research, the goal of the thesis which include the major objective of research and the brief background and key concepts of the GIS.

The second chapter presents the history of maps and cartography. It discusses the concepts, foundation and architecture of the Geographical Information Systems.

The third chapter gives the Concept of the project. and discusses topics like thematic mapping, Digital Elevation Model, Geo-Referencing. The Final Chapter gives the details of implementation.

Chapter 2

BACKGROUND AND LITERATURE REVIEW

History of GIS

The prosperity if the developed world is dependent on it's information society. Today there are more number of people employed in the collection, storage, manipulation and output of information than in the Industrial, Services or Agriculture industries. Information is required to make decision. With the advent of computer and satellite the quantity and availability of information has greatly improved.

The importance of land to a society can not be emphasized enough. Through out history and even today people as individual or nation fight for ownership of land. It is a source, which can significantly enhance the wealth of nations [Yang 93]. The collection of data about the spatial distribution of significant properties of the earth's surface has long been an important part of the activities of organized societies. From the earliest civilization to modern times, spatial data has been collected by Navigators, Geographers and surveyors and rendered into pictorial form by the mapmakers or cartographers. Originally, maps were used to describe far-off places, as an aid for navigation and military strategist [IIodgkiss 81]. In Roman times, the agrimensores, or land surveyors, were an important part of the government, and the result of their work may still be seen in digital form in the landscapes of Europe today [Dike 71]. The decline of the Roman Empire led to the decline of surveying and map making. Only in the eighteen century did European civilization once again reached a state of organization such that many governments realized the value of systematic mapping of their lands. National government bodies were commissioned to produce topographical maps of whole countries. These highly disciplined institutes have continued to this day to render spatial distribution of the earth's surface, or topography, into map form. During the last 200 years many individual styles of map have been developed, but there has been a long, unbroken tradition of high cartographic standards that gas continued until the present.

As the European powers increased their influence over the globe, they spread their ideas and method of map making to the countries that fell under their sway. As scientific study of the earth advanced, so did the new material needed to be mapped. The development in the assessment and understanding of the natural resources, geology, geomorphology, soil sciences and ecology, that began in the nineteenth century and have continued to this day, provided new material to be mapped. Whereas topographical maps can be regarded as general purpose because they do not set out to fulfill any specific aim (e.g. they can be interpreted for many different purpose), maps for the distribution of rock types, soil series or land use are made for more limited purpose. These specific purpose maps are often referred to as *thematic maps* because they contain information about a single subject or theme. To make the thematic data easy to understand, thematic maps are commonly drawn over a simplified topographic base by which users can orient themselves.

In the twentieth century, the demand for map of the topographic and specific themes of the earth's surface, such as natural resource has accelerated greatly. Stereo aerial photography and remotely sensed imagery have allowed photogramists to map large areas with great accuracy [Raafat 91]. The same technology has also given the earth resource scientist, the geologist, the soil scientist, the ecologist, and the land specialist enormous advantages for reconnaissance and semi-detailed mapping. The study of land evaluation arose through the need to match the land requirements for producing food and supporting populations to the resources of eliminate, soil water, and available technology.

The need for spatial data and spatial analysis has not been restricted to earth scientist. Urban planners and cadastral agencies need detailed information about the distribution of land and resources in towns and cities. Civil engineers need to plan the routes of r oads and canals and to estimate construction costs, including those of cutting of hillsides and filling in valleys. Police department need to know the spatial distribution of various kind of crime, medical organizations, the distribution of sickness and disease, commercial interests the distribution of sales outlets, and potential markets. The enormous infrastructure of what are collectively known as utilities i.e. water, gas electricity, telephone lines, sewerage system, all need to be recorded and manipulated in map form.

Until computers were applied to mapping, all kinds of mapping had one point in common. The spatial database was a drawing on a piece of paper of film. The information was encoded in the form of points, lines and areas. These basic geographical entities were displayed using various artifacts, such as diverse symbolism or color or text codes, the meaning of which is explained in a legend.

Where more information was available than could be printed in the legend on the map, then it was given in an accompanying memoir, because the paper map, and its accompanying memoir, was the database, there were several very important consequence for the collection, coding and use of the information it contained.

- The original data has greatly reduced in volume, or classified, in order to make them understandable and representable; consequently, many local details were often filtered away and lost.
- The map had to be drawn extremely accurately and the presentation, particularly of compels thermos, had to be very clear.
- The sheer volume of information meant that areas that are large with respect to the map scale could only be represented by a number of map sheets. It is a common experience that one's area of interest is frequently near the junction of two, if not more map sheets.
- Once data had been put into a map, it was not cheap or easy to retrieve them in order to combine them with other spatial data.
- The printed map is static, qualitative document. It is extremely difficult to attempt quantitative spatial analysis within the units delineated on a thematic map without resorting to collecting new information for the specific purpose in hand.

The collection and compilation of data and the publication of printed map is a costly and time-consuming business. Consequently, the extraction of single theme from a general-purpose map can be prohibitively expensive if the map must be redrawn by hand [Matwin 95]. It was not important that initial mapping costs were large when a map could be thought of as being relevant for a period of 20 years or more. But there is now such a need for information about how the earth's surface is changing that conventional map making techniques are totally inadequate. For example, for some kinds of mapping, such as weather charts or the distribution net of a telephone company, there can be a daily, or even hourly need for the spatial database to be brought up to date, which is just simply not possible by hand. Essentially, hand drawn map or the map in a resource inventory is a snapshot of the situation seen through the particular filter of a given surveyor in given discipline at a certain moment in time. More recently, the aerial photography, but more specially the satellite image, have it possible to see how landscapes changes over period of time, to follow the slow march of desertification or erosion or the swifter progress of forest fires, floods, locust swarms or weather systems. But the products of airborne and space sensors are not maps, in original meaning of the word, but photographic images or streams of data on magnetic tapes. The digital data is not in the familiar form of points, lines and areas representing the already recognized and classified features of the earth's surface, but are coded in picture elements pixels cells in two dimensional matrix that contain merely a number indicating the strength of reflected electromagnetic radiation in a given band. New tools were need to turn these streams of numbers into pictures and to identify meaningful patterns.

The cartographers, initially, did not posses the skill to use these new tools and so the fledgling science of remote sensing, image analysis, and pattern recognition were nursed into being, not by the traditional custodian of spatial data, but by mathematicians, physicists and computer scientists (with, it must be said, much support from military authorities). These new practitioners of the art of making images of the earth have taken very different approach to that of the conventional field scientist, surveyor and cartographers. In the beginning, they often made exaggerated claims about the abilities of remote sensing and image analysis to recognize and map the properties of the earth's surface without expensive ground surveys. Gradually it has come to be realized that the often very striking images produced from remotely sensed data only have a real value if they can be linked to ground truth, a certain amount of field survey is essential for proper interpretation. And to facilitate calibration, the images have to be located properly with respect to a proper geodetic grid, otherwise the information cannot be related to a definite place. The need for a marriage between remote sensing, earthbound survey, and cartography arose, which has been made possible by the class of mapping tools known as 'GEOGRAPHICAL INFORMATION SYSTEMS' or GIS.

THE GEOGRAPHICAL INFORMATION SYSTEMS

The history of using computer for mapping and spatial analysis shows that there has been a parallel development in several broadly related fields of automated data capture, data analysis and presentation [Burrough 92]. These fields are cadastral and topographical mapping, thematic cartography, civil engineering, geography, mathematical studies of spatial variation, soil science, surveying and photogrammetry, rural and urban planning, utility network, and remote sensing and image analysis. Military application have overlapped and even dominated several of these monodisciplinary fields. Consequently there has been much duplication of effort and a multiplication of discipline specific jargon for different application in different lands. This multiplicity of effort in several initially separate but closely related field is now resulting in the possibility of linking many kinds of spatial data processing together into truly general purpose geographical information systems, as technical and conceptual problems are overcome.

Essentially, all these disciplines are attempting the same sort of operation, namely to develop a set of powerful tools for collecting, storing, retrieving at will, transforming, and displaying spatial data from the real world for a particular set of purposes. This set of tools constitutes a 'Geographical Information System' [Cadastrl Surveying 95]. Geographical data describe the objects from the real world in following terms;-

- Their position with respect to a known coordinate system.
- Their attributes that are unrelated to position (such as color, cost, pH, incidence of disease, etc).
- Their spatial interrelations with each other (topographical relations), which describe how they are linked together or how one can travel between them.

Difference between Computer Graphics and GIS. Geographical information systems differ from computer graphics because the latter are largely concerned with the display and manipulation of visible material. Computer graphics systems do not pay much attention to the non-graphic attributes that the visible entities might or might not have, and which are useful data for analysis. Good computer graphics package is by itself not sufficient for performing the tasks expected, nor are such drawing packages necessarily a good basis for developing such a system.

Difference between CAD and GIS. Geographical information systems do have a lot in common with computer-aided design (CAD) systems used for drafting a wide range of technical objects, from a plane to the layout a microchip. Both GIS and CAD systems need to be able to relate objects to a frame of reference, both need to handle non-graphic attributes and both need to be able to describe topological relations. This major differences between GIS and CAD systems are the much greater volume and diversity of the data input to GIS system and the specialized nature of the analysis method used. These differences can be so large that an efficient system for CAD may be quite unsuitable for GIS and vice versa.

Geographical information systems should be thought of as being very much more than simply the means of coding, storing and retrieving data about aspects of the earth's surface. In a very real sense the data in a geographical information system, whether they are coded on the surface of a piece of paper or as invisible marks on the surface of a magnetic tape, should be thought of as representing a model of the real world. Because this data can be accessed, transformed and manipulated interactively in a geographical information system, it can serve as a test bed for studying environmental processes or for analyzing the results, or of anticipating the possible result of planning decisions. By using the GIS in a similar way that a trainee pilot uses a flight simulator, it is, in principal, possible for planners and decision-makers to explore a range of possible scenarios and to obtain an idea of the consequence of a course of action before the mistakes have been irrecoverably made in the landscapes itself.

Components of GIS

Geographical information system has following two components,

- Computer hardware.
- Sets of application software.

Computer hardware for GIS

The computer or central processing unit is linked to a disk drive storage unit, which provides space frosting data programs. A digitizer or another device is used to convert data from maps and documents into digital form and send them to computer. A plotter or other kind of display device is used to present the result of the data or program on magnetic tape, or for communicating with other systems. Inter computer communication can also take place via a network system over special data lines, or over telephone lines using a Modem. The user controls the computer and peripherals (a general term for plotters, printers, digitizer and other apparatus linked to the computer) via a visual display unit (VDU), known as terminal. The user's terminal might itself be a microcomputer, or it might incorporate special hardware to allow maps to be displayed quickly. There is a very wide range of devices that can be used to fill these general hardware requirements.

GIS software component

- Data input module.
- Data storage and database management.
- Data output and presentation.
- Data transformation.
- Interaction with user.

Data Input

Data input covers all aspects of transforming data captured in the form of existing maps, field observation and sensors (including aerial photography, satellites and recording interments) into a compatible digital form. A wide range of computer tools are available for this purpose, including the interactive terminal or visual display unit (VDU), the digitizer, lists of data in text files, scanners (n satellites or airplane for direct recording of data or for covering maps and photographic images) already written on a magnetic media such as tapes, drums and disks. Data input and the verification of data needed to build a geographical database.

Data Storage

Topology, and attributes of geographical elements (points, lines and areas representing objects on the earth's surface) are structured and organized, both with respect to the way they are handled in the computer and how they are perceived by the users of the system. The computer program used to organized the database is known as a Database Management System (DBMS).

Data output and presentation

Data output and presentation concerns the ways the data are displayed and the results of analysis are reported to the users. Data may be presented as maps, table and figures (graphs and charts) in a variety of ways, ranging from the ephemeral image on cathode ray tube (CRT) through hard copy output drawn on printer or plotter to information recorded on magnetic media in digital form.

Data transformation

Data transformation embraces two classes of operation,

- Transformation needed to remove errors from the data or to bring them up to date or to match them to other data sets.
- The large array of analysis methods that can be applied to the data in order to achieve answers to the questions asked to the GIS.

Transformation can operate on the spatial and the non-spatial aspects of the data, either separately or in combination. Many of these transformation , such as those associated with scale-changing, fitting data to new projections, logical retrieval of data, and calculation of area and perimeters, are of such a general nature that one should not expect to find them in every kind of GIS in one form or another. Other kinds of manipulation may be extremely application specific, and their incorporation into a particular GIS may be only to satisfy the particular users of that system. The kinds of transformation method available, their optimum use and misuse, the ways in which sets of simple transformations can be combined in order achieve certain types of geographical or spatial modeling.

Interaction with the users

This module is concerned with the interaction with the user, query input which are essential for the acceptance of the use of nay geographical information system. Certainly it is an aspect that until relatively recently has received less attention than it deserves. It is only in the last few years that the average user has received less attention than it deserves. It is only in the last few years that the average user has been able to make direct contact with computer other than via the impersonal and unforgiving media of punched paper tapes and cad handed into the computing center. The widespread introduction of the personnel computer and of programs that are operated by commands chosen from a menu (a list), or that are initiated by a response to requests in an English like command language of verbs, nouns , and modifiers has broken down the barriers that once frightened many a would be computer user away for life.

Definition of a Map

A map is a set of points, lines, and areas that are defined both by their location in space with reference to a coordinate system and by their non spatial attributes. A map is usually represented in two dimension but there is no reason to exclude higher dimension except through the difficulty of portraying them on a flat piece of paper.

The map legend is the key linking the non-spatial entities. Non spatial attributes may be indicated visually by colors, symbols or shading, the meaning of which is defined in the legend. For Geographical Information Systems. A region is a set of pixels, areas or polygons that are described by a single legend unit. A region may be made up of several discrete occurrences, which may be uniform or which may contain polygons belonging to regions of another kind. Although the eye can easily distinguish the topological relationships between regions, these relationship must be explicitly built into any digital presentation.

Geometric classification of spatial information

In addition to organizing spatial data by themes, mapped information is also structured as points, lines and polygons.

Point Data Examples of point data include location of wells, post office, manholes, stream gauges, bird nesting sites or control points.

Line Data Examples of line data include road networks, utility lines, stream drainage, and fault lines.

Polygon Data Examples of polygon data include land use, vegetation cover, electoral districts, soil types, and zoning.

Textual data base

Besides the spatial information in a map, the GIS can usually store non-spatial information that is related to the spatial entities. For instance an urban GIS database may have a map theme of property boundaries. Attached to each parcel will be a textual database that might store the name of the owner, the address, the assessed value of the property, or the type of services and utilities on the site.

Geographical data in the computer

When geographical data is entered into a computer the user will be at most ease if geographical information system can accept the phenomenological data structures that he has always been accustomed to using. But computers are not recognized like human minds and must be programmed to represent phenomenological structures appropriately [Mufti 83]. Moreover, the way the geographical data are visualized by the user is frequently not the most efficient way to structure the computer database. Finally, the data have to be written and stored on magnetic devices that need to be addressed in a specific way. We can represent theses four stages as follows:-

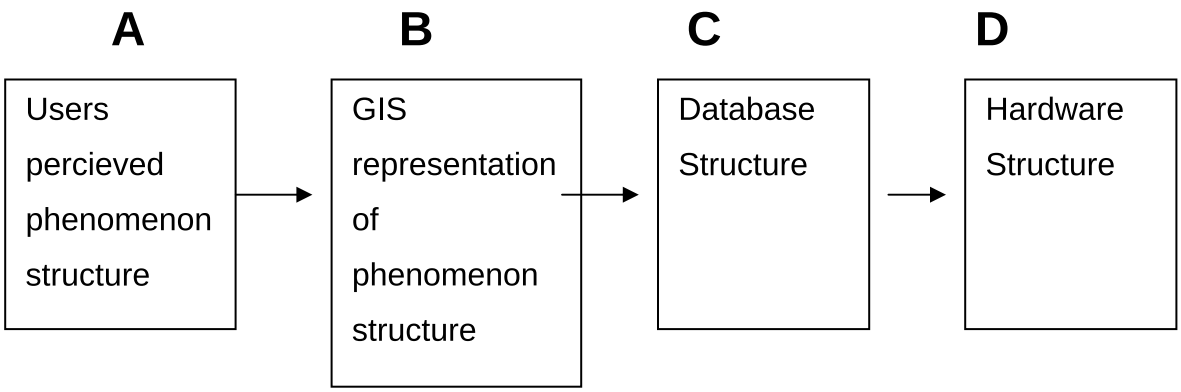


Fig 2.4 The stages followed during the Database development

The essential feature of any data storage system is that it should be able to allow data to be accessed and cross referenced quickly. There are several ways of achieving this, some of which are more efficient than others. Unfortunately, there seems to be no one best method that can be used for all situations. This explains in part the massive investment in labor and money in effective Database Management Systems, which are the computer programs that control data input, output, storage and retrieval from a digital database.

Data/Knowledge Structure and Computer representation of geographical data

Unlike many other kinds of data handled by modern information systems, geographical data processing is complicated by the fact that it must include information about position, possible topological connections, and attributes of the objects recorded. The topological and spatial aspects of geographical data processing distinguish systems designed for graphics and mapping from those other modern data systems such as those used for banking, library searches, airline bookings or medical records [Cheng 93].

Geographical data is referenced to locations on the earth's surface by using a standard system of coordinates. The coordinates system may be purely local, as in the case of a study of limited area, or it may be that of a national grid or an internationally accepted projection such as the 'Universal Transverse Macerator Coordinate System' or (UTM). Geographical data is very often recognized and described in term of well-established geographical 'objects', or phenomena [Giloi 78]. All geographical studies have used phenomenological concepts such as town, river, flood plain, ecotype, soil association as fundamental building blocks for analyzing and synthesizing complex information. These phenomenological building blocks are very often grouped or the hierarchically defined taxonomies, for example the hierarchy of country province town district, or the hierarchy of most soil classification systems or of plants and animals.

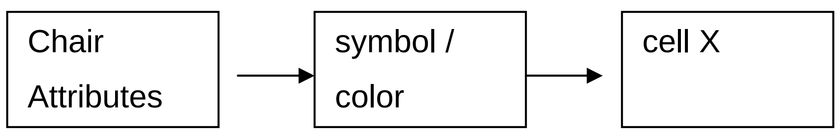
The human eye is highly efficient at recognizing shapes and forms, but the computer needs to be instructed exactly how spatial should be handled and displayed. Essentially there are two contrasting, but complementary ways of representing spatial data in the computer that we shall refer to as explicit and implicit ways of describing spatial entities.

Vector and Raster GIS

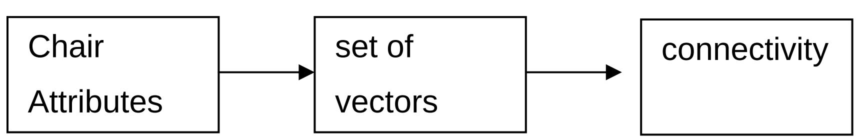
There are two major methods of storing mapped information. Geographic Information Systems which store map features in vector format store points, lines and polygons with high accuracy. They are preferred in urban applications where legal boundaries and the analysis of networks are important.

Raster Geographic Information Systems, which store map features in raster or grid format, generalize the location of features to a regular matrix of cells. Raster GIS data structures are preferred for digital elevation modeling, statistical analysis, remotely sensed data, simulation modeling and natural resource applications.

There are two different ways in which a chair can be **explicitly** or **implicitly** represented in a computer. Explicit representation means that the form of the chair is built up from a set of points on a grid or “Raster”. So that the computer knows that this set of points represents a chair and not a table, each cell is given the same code value ‘C’. In practice, the C’s would not be themselves displayed but would be represented by a numerical value or a color or gray scale. We would then have the. following simple data structure for the chair:



The implicit representation makes use of a set of lines, defined by starting and end points and some form of connectivity. The starting and end points of the lines define ‘Vectors’ that represent the form of the chair; pointers between the lines indicate to the computer how the lines link together to form the chair, the data structure is:



The implicit representation requires fewer numbers, implying fewer storage spaces, to store the information about the chair (the vector representation uses XY pairs and connecting pointers and the raster representation uses cells). Second, the vector representation is aesthetically more pleasing than the raster image. The raster image would need to be based on a 0.5 mm grid to produce the equivalent resolution, thereby requiring XY pairs. Third, the connectivity information allows directed spatial searches to be made over the chair. On the other hand if the shape or size of the chair has to be changed, this can be done much quicker and easier in the raster representation than in the vector. In a raster representation data update merely involves deleting certain values and writing the new ones. In vector representation, not only must the coordinates be updated but the connectivity must also be rebuilt.

Vector Geographical Information System

The fundamental primitive of a vector GIS is point. Joining points with line (usually straight but in some system areas and curves) creates objects. Sets of lines define areas which are commonly referred to as polygons. Vectors are usually used for transportation and infrastructure. They are used when relating to the cadaster2 and the DCDB3 is a Vector GIS.

The point represents an object e.g. a manhole or a depending upon the scale. It is used to represent 2 polygons where the sides have not been recorded. Identification, feature code, coordinates and a pointer represent a point to any attribute data sets. When a polygon does not exist then its centroid is used to attach the attributes.

Early Vector GIS's created spaghetti models where all points lines and polygons were recorded as separate entities. Each model created was independent of adjoining models and common points and lines were duplicated. This is non-topological model. A topological model only requires points and lines to be stored once. This minimizes the storage requirements and all allows for efficient manipulation of data for spatial queries. It also maintains the spatial relationship between entities. Although many lines may be shared between two polygons, all are input and coded twice and merged and then only one version is stored. All points and resulting lines are stored as coordinates.

There are four main methods of inputting vector data into system.

- Direct keyboard entry, bearing and distances or coordinates.
- Conversion from other digital sources.
- Scanning
- Digitizing

Scanning is often used for raster while digitizing is more common for vector.

The digitizing of a map or plan will require at least three points to be identified, and the plan or map to be related to the coordinate system. The more control points there are the plan or map to be related to the coordinate system. The more control points there are the better the digitized product, because as most maps were not drafted for the purpose of digitizing, and paper maps are unstable and may stretch or shrink errors will occur. These errors are automatically entered into the GIS database at the same level as the source maps.

Raster Geographical Information System

The raster model divides the entire study area into rectangle grid, where each grid square is known as a cell. Each cell contains a single value and it is the spatial entity or basic spatial unit (BSU) to which attribute data may be stored.

The smaller the grid cell the greater the resolution and the greater the quantity of data to be stored. Each layer in the system is capable of storing one item of information.

For example for a topographical map

- Elevation or contours 1 layer
- Creeks 1 layer

- Roads 1 layer
- Fences 1 layer
- Building 1 layer

A total of 5 layers for this information.

The matrix of columns and rows is suitable for input to most computers. All values are related to their matrix position. To be able to correlate data from different sources it must be all related to the same coordinate systems. The same origin and the same orientation.

The basic principle for setting up a raster GIS is best explained by following points:

- What information is needed to support this decision e.g soil type, rainfall, frost, slop, etc
- Define the limit of the area. What items of geographic information need to be considered.
- What are the sources of this information, existing maps, aerial photographs etc
- Does anyone hold this data in digital form.

- The analysis consist of a series of operations which must be carried out in a sequence:
- Using one or more layers, create a new layer or text information (the GIS analysis depend upon both the software and database).
- Analysis using all the information to determine the most suitable area.

One of the major problems with Raster GIS is the quantity of data that has to be stored to obtain high resolution e.g one LANDSAT image has 10,000 pixels or cells.

Grid cell data could be added to the system by manually coding in ASCII. The GIS program can reformat the data to its special processing needs. The information could be collected by entering each cell, obviously this is expensive and time consuming. For some data entry in a victor form is more efficient. The extent of a cell type e.g edge of soil A is digitized in Vector form by capturing the location of points along the boundary. It is then assumed that these points are joined by straight lines. This vector representation then has to be converted to raster by an operation known as vector raster conversion.

‘A USA launched satellite used for weather forecasting etc.

Digitizing is the collection of spatial data rather than non spatial data. The spatial entity will indicate:

- A feature code.
- A point, line, polygon identification number.

To ensure that maximum use of digitized data, it must be fully described:

- The digitizing system used.
- The source of digitized data e.g 1:2000, paper maps 1976.

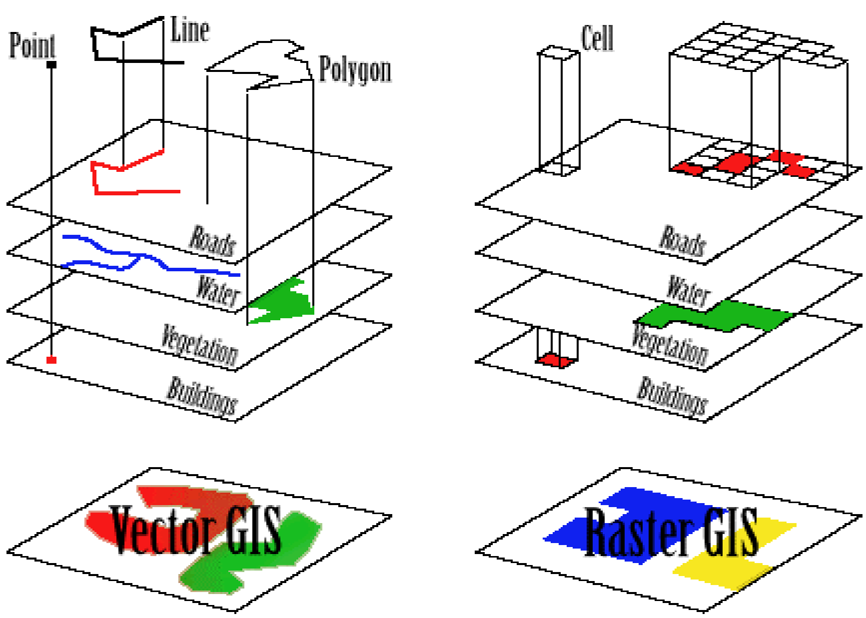
From this information the expected accuracy and reliability may be determined.

Chapter 3

Conceptualization

Thematic Mapping

Maps in Geographic Information Systems are represented thematically. A standard topographic map will show roads, rivers, contour elevations, vegetation, human settlement patterns and other features on a single map sheet. In a GIS these features are categorized separately and stored in different map themes or overlays. For example, roads will be stored in a separate overlay. Likewise, rivers and streams will each be stored as a separate theme. This way of organizing data in the GIS makes maps much more flexible to use since these themes can be combined in any manner that is useful. The following illustration shows conceptually how maps are stored as themes in a GIS.



How Information is stored in a GIS

Each different theme is stored on a separate overlay. The overlays on the left represent a vector based GIS, where the information is stored as a series of points, lines and polygons. The overlays on the right represent a raster based GIS, where the information is stored as a series of discrete units called cells.

Querying the GIS

The GIS stores both spatial and non-spatial data in a database system which links the two types of data to provide flexible and powerful ways of querying or asking questions about the data. An example of a spatial query might look like this:

> ‘Locate and display all playgrounds downstream of landfills within 100 year floodplain’

This type of query is answered by a set of commands to the GIS that then generates a map display of all sites meeting the criteria expressed in the query. The user may also query the GIS by the textual attributes in the tabular database and then display the map features that correspond to these attributes. An example of this type of query is as follows:

> ‘Display all water mains installed before 1950 with a diameter less than 12 inches’

This query results in a map display of the water mains in the study area with the specific mains in the query highlighted. Alternatively a report could be generated which lists the complete information on each segment of water mains which meet the criteria in the query.

GIS data integration

Many Geographic Information Systems handle both vector and raster data from a wide variety of sources including satellite imagery, cadastral information, hand digitized maps and scanned images.

Geo-referencing

In order to ensure that all maps in a GIS database overlay accurately, the data set is 'geo-referenced' to a common coordinate system. In many countries the Universal Transverse Macerator (UTM) projection is commonly used to define coordinates in the GIS.

Spatial analysis

Spatial analysis is a set of analytical procedures applied to GIS data to describe, predict, or assess environmental or social issues. Spatial analysis techniques include methods for:

- Reclassifying map overlay features
- Measuring distance, and area
- Interpolating values
- Identifying the co-occurrence of values on different map themes (overlay analysis)

Digital Elevation Modeling (DEM)

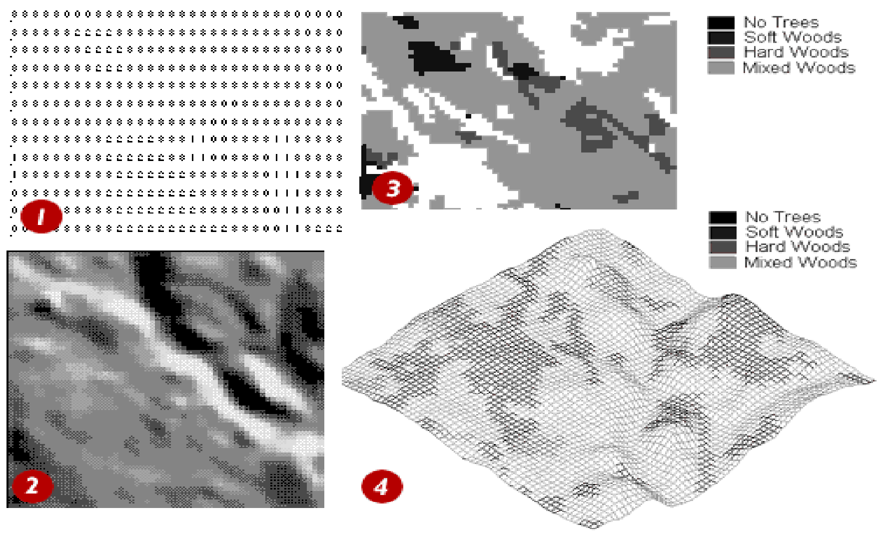
There are also specific methods for analyzing terrain including:

- Calculating slope
- Solar aspect
- View Shed analysis (calculating visibility)
- Runoff analysis

Automated cartography and data visualization

Geographic Information Systems also have sophisticated graphic capabilities for map overlay production and data visualization in plan and perspective. Conventional techniques for producing cartography are automated in Geographic Information Systems. In addition, most systems will provide methods for displaying maps in three dimensions. In the illustration below, the same region is depicted as:

- a set of numbers
- a shaded relief map overlay
- a hatched drawing in plan view
- a perspective drawing draped over an elevation map.



There are also other techniques for representing spatial data, including charts histograms and statistical tables.

GIS APPLICATIONS

GIS are now used extensively in government, business, and research for a wide range of applications including environmental resource analysis, land use planning, location analysis, tax appraisal, utility and infrastructure planning, real estate analysis, marketing and demographic analysis, habitat studies, archaeological analysis, and military planning.

Natural resources management

This is one of the first major areas of application it includes management of

- Wildlife habitat.
- Wild and scenic rivers.
- Recreation resources.
- Floodplains.
- Wetlands.
- Agricultural lands.
- Aquifers.
- Forests.

Facilities management

One of the largest areas of application has been in facilities management. Uses for GIS in this area have included

- Locating underground pipes and cables.
- Balancing loads in electrical networks.
- Planning facility maintenance.
- Tracking energy use.

Land management

Local, state, and federal governments have found GIS particularly useful in land management. GIS has been commonly applied in areas like

- Zoning and subdivision planning.
- Land acquisition.
- Environmental impact policy.
- Water quality management.
- Maintenance of ownership.

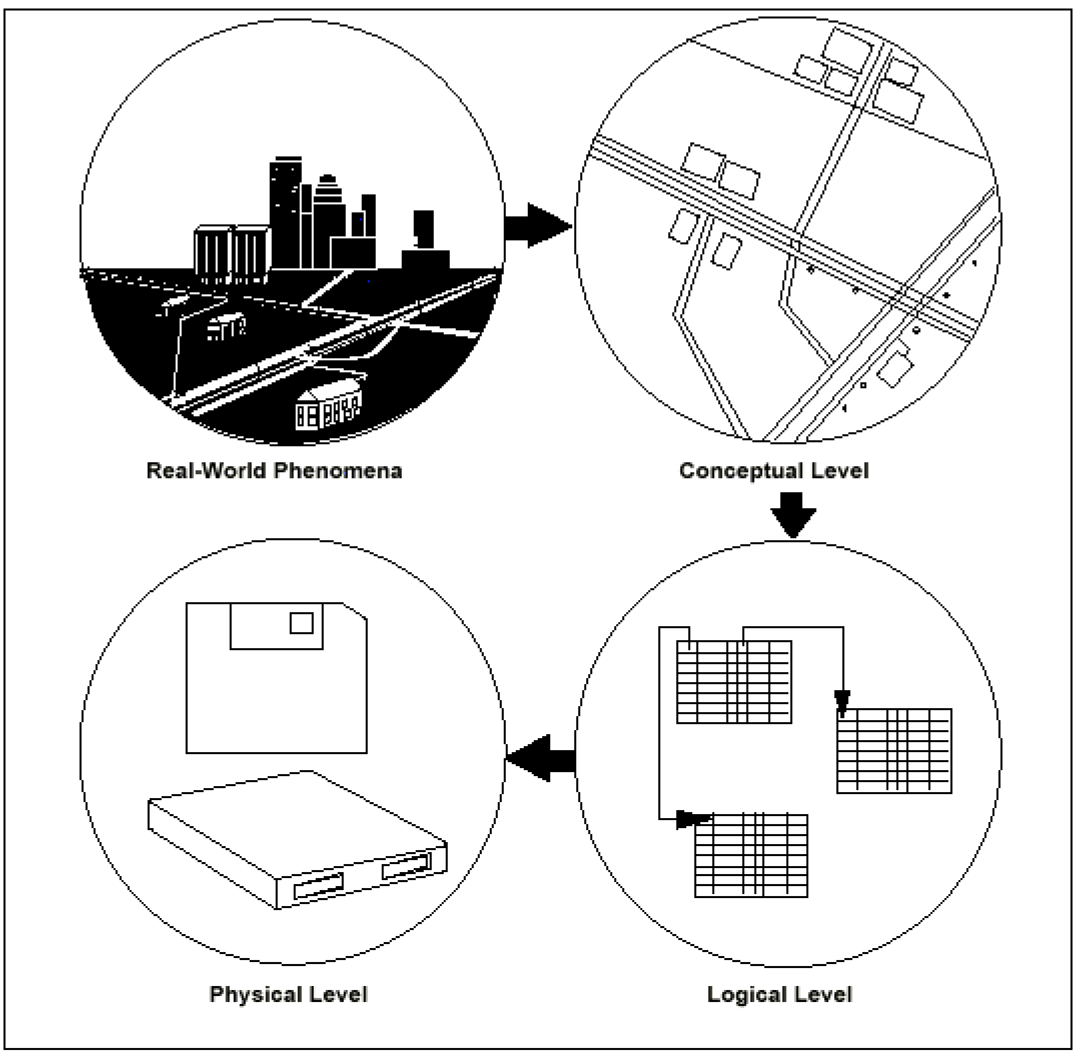
Street networks

More recent and innovative uses of GIS have used information based on street-networks. GIS has been found to be particularly useful in

- Address matching.
- Location analysis or site selection.
- Development of evacuation plans.

Spatial Data Model.

We have defined the Model of Spatial data at three "levels," from the real world to the **physical encoding** of the data (see Figure 1). The **conceptual level** describes a way to represent real-world entities, including their geometric and topological characteristics and relationships. **The logical level** presents a data model for identifying and encoding information for digital storage. We have also defined the physical level with rules and specific formats for encoding data in a



RDBMS.

Figure 1: Overview of Spatial Data Model

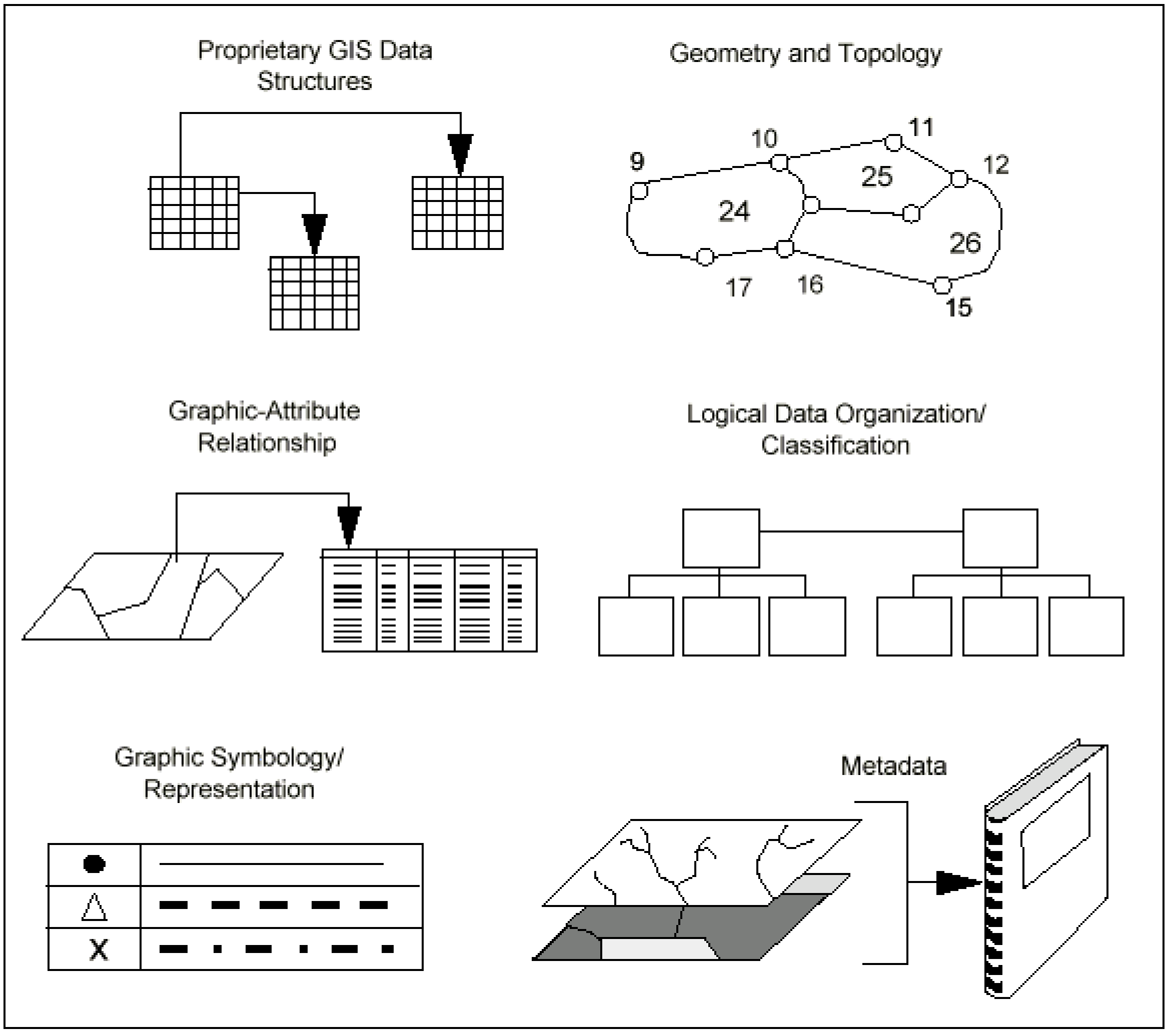
Challenges in Spatial information representation.

As illustrated in Figure 2, many characteristics of spatial databases must be taken into account in a fully effective spatial modeling process:

- *Geographic information system* and *automated mapping* software packages use different proprietary structures for storing graphic data. As a result, batch restructuring of data is necessary if it is transferred to and used with another software package.
- Spatial databases represent *spatial features* in terms of their *geometry* (graphic representation) as well as their topological relationships, which explicitly define the connectivity and adjacency of features.
- Spatial databases store not only the graphic representations of features but non-graphic attributes associated with those features. These attributes should be included in the data and their link with specific *spatial features* should be maintained.
- Spatial databases are built using specific approaches for grouping and classifying *spatial features* and attributes.
- GISs and *automated mapping* software packages allow for the selection of specific symbology (e.g., line styles, point symbols) for the graphic display of spatial features .
- It is becoming a standard practice of GIS user organizations to build and maintain metadatabases that contain information about the content, quality, and characteristics of a spatial database.

A complete and fully effective spatial data representation mechanism should provide for the encoding and exchange of all of these characteristics and components of spatial databases among disparate computing platforms. Our model has been designed to accomplish this goal without loss or corruption of information

Figure 2: Fundamental Issues in Spatial Data Modeling



Spatial Data Model (Conceptual Level)

This Model is based on a *conceptual model* of spatial data that defines the characteristics of *objects*—the building blocks for a digital representation of a spatial entity like a river, building, utility line, or water well. Objects within this model may be *simple objects* (the most basic representative elements like points or line segments) or *aggregate objects* (which combine multiple simple objects into a larger whole, e.g., a data layer).

The spatial model defines both the *geometry* (graphic depiction) and the *topology* (connectivity and spatial relationships) of map features as shown in Figure 3. These map features may be graphically represented as points (zero-dimensional vector objects), as lines (one-dimensional vector objects), as areas (two-dimensional vector objects), or in gridded or raster form. Figure 4 illustrates how spatial entities may be represented as zero-, one-, or two-dimensional objects.

Figure 3: Geometry vs. Topology

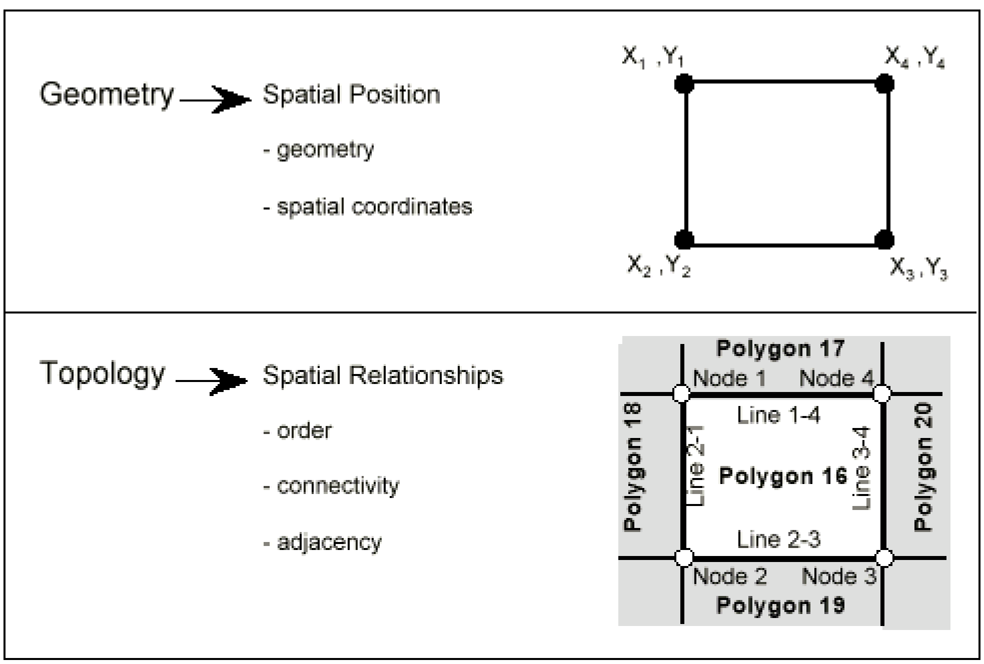


Figure 4: Map Feature Representations

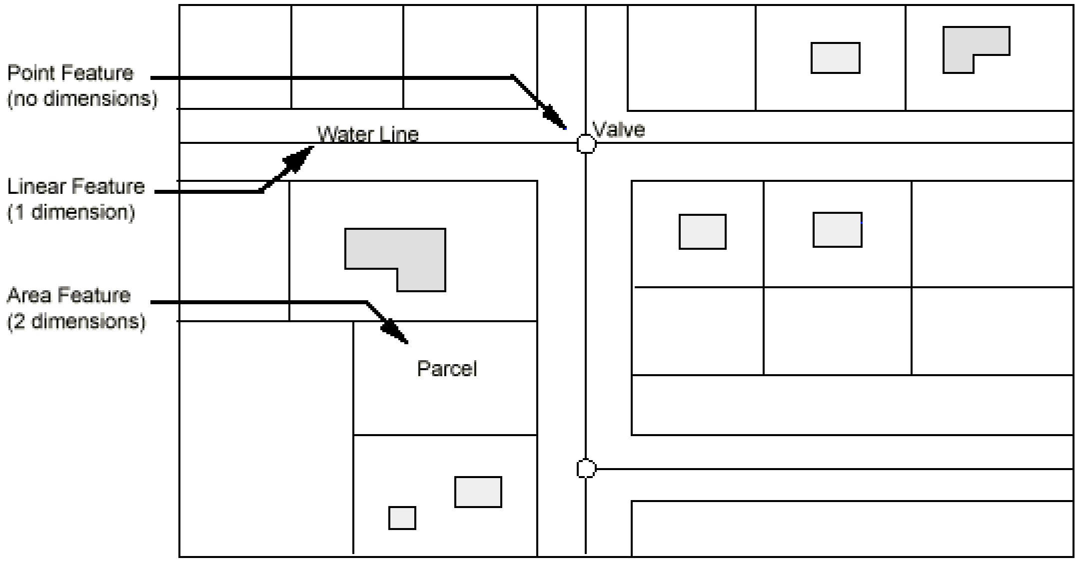


Table 1 identifies the different simple object types included in the conceptual model. Depending upon the particular type of data and requirements of particular users, a map feature could be digitally represented by “geometry only” objects or “geometry-topology” objects. For instance, using the example in Figure 7, a water valve could be represented as a point (geometry-only) object or as a node (geometry-topology) object if it is being treated as part of a linear topological network of the water distribution system. Whether or not a map feature is represented by geometry only or by topological representation depends in part on the software being used and on the specific applications of the user.

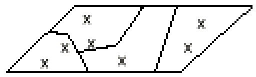
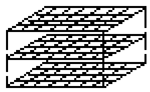
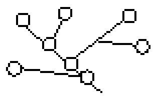
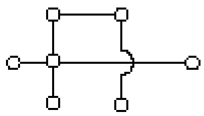
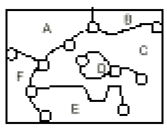
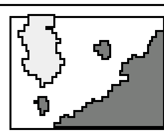
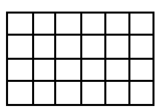
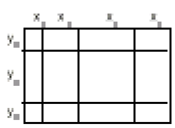
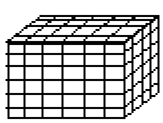
Table 1: SDTS Simple Objects

Feature Types	Simple Objects Geometry Only	Simple Objects Geometry/Topology
Point Features (Zero-dimensional)	Point (includes subtypes of Entity, Area, and Label Points)	Node (Planar or Network)
Linear Features (One-dimensional)	Line String, Arc, G-Ring	Link, Chain, GT-Ring
Area Features (Two-dimensional)	G-Ring , G-Polygon	GT-Polygon, Universe Polygon, Void Polygon
Raster Surfaces (Two-dimensional)	Pixel, Grid Cell, Labeled Grid Cell	N/A
Raster Surfaces (Three-dimensional)	Voxel , Labeled Voxel	N/A

Several types of *aggregate objects* are defined in this model because they are effective in providing a context for use of *simple objects* in a digital representation. With simple objects as the building blocks, aggregate objects denote collections of simple objects that represent real-world phenomena and, therefore, they provide a basis for defining a specific spatial phenomenon. For example, the aggregate object, planar graph, may represent a road network that is concisely defined in terms of its component simple objects, thereby facilitating a consistent data representation. Aggregate objects are explained in Table 2.

The *composite object* is a specially-defined object type that is any aggregation of simple objects or other composite objects. This object type is useful because it allows the flexibility to define an object for transfer that consists of any collection of other objects.

Table 2: Aggregate Spatial Objects

	**Layer**	General term describing a collection of instances (occurrences) of spatial features in a single theme
	**Raster**	One or more overlapping layers for the same grid, labeled grid, voxel space, or other raster data
	**Planar Graph**	Linear objects inter-connected on a two-dimensional surface with a node at each intersection (e.g., a topologically structured stream network or street network)
	**Network**	A graph of linear objects inter-connected topologically, which, when projected onto a two-dimensional surface, allow multiple nodes at a single location or no nodes at intersection points (e.g., a gas distribution pipeline with overlapping pipes in three-dimensional space)
	**Two-Dimensional Manifold**	A planar graph and its associated polygons which totally exhaust a surface (e.g., land cover map, parcel map)
	**Digital Image**	Two-dimensional array of regularly spaced pixels (e.g., unclassified satellite image, digital orthophotograph)
	**Grid**	Matrix of cells forming a mesh with repeating pattern (e.g., grid map, digital terrain model)
	**Labeled Grid**[1]	A two-dimensional set of labeled grid cells [1] forming an irregular rectangular pattern; each labeled grid cell is identified by a spatial label
	**Voxel Space, Labeled Voxel Space**[1]	Three-dimensional grids in which each voxel represents a three-dimensional volumetric unit (the three-dimensional equivalent to the grid or labeled grid)

Spatial Data Model (Logical Level)

Our design also includes the logical format for representation of spatial data. This logical format defines the content and basic format of a series of *modules*, each of which contains a specific category of information. Figure 5 shows the five major categories of modules into which other individual modules are grouped.

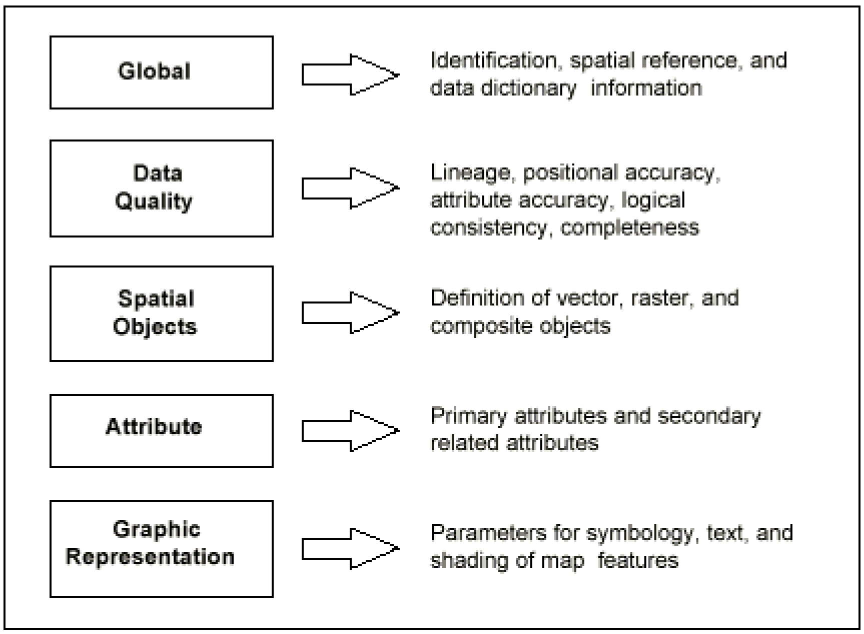


Figure 5: Categories of Spatial Data Modules

Spatial Data Model (Physical Level)

The model does not include any special constraints at physical level and it can be implemented in any RDBMS and on any platform however some of the differences across various platforms can effect the required results , therefore we have also specified some implementation details at physical level (for example integer data type is required to be 32 bit).

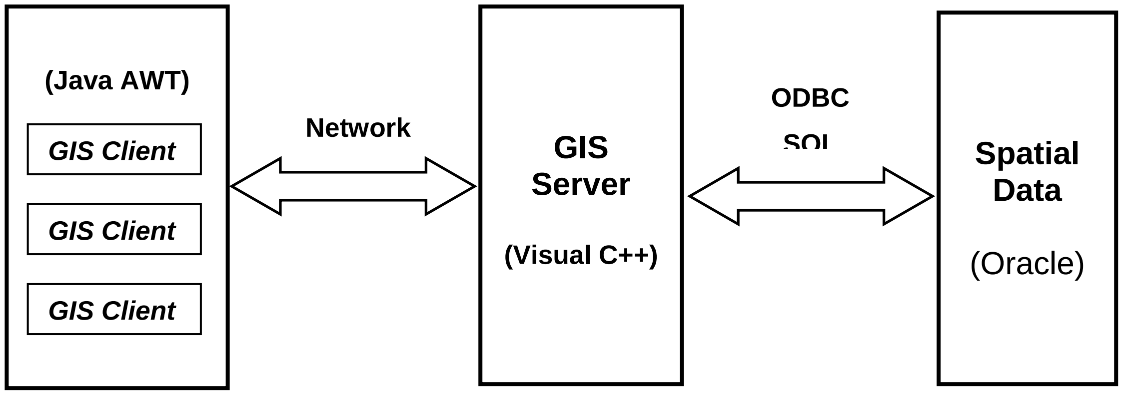
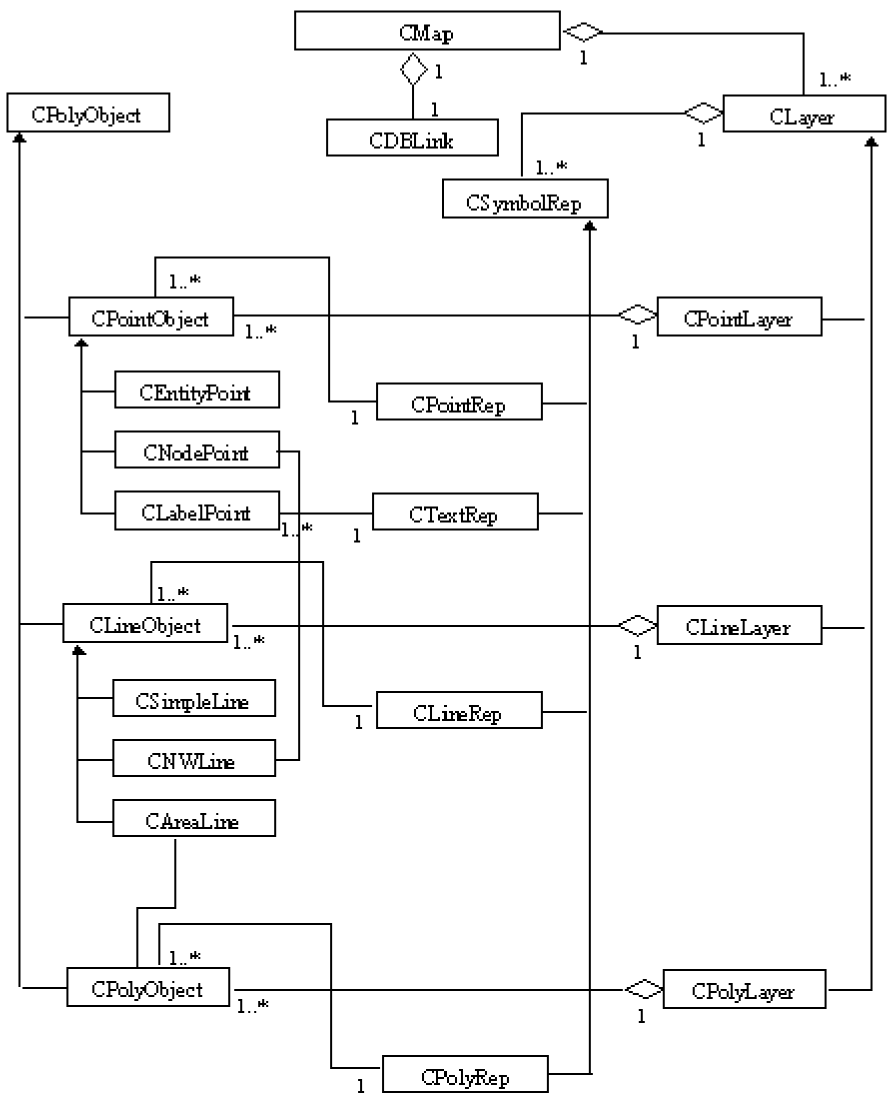


Figure 0-1 Outline Architecture

Final Design.

The design is mainly derived from the data model. The main classes and their relationship is shown in Figure-1.



The main classes and their responsibilities are as under:-

CGObject

CGObject is the generalization of all classes that represent a real world entity. Objects are grouped into three types base on similarity of their geometry. Each object keep its spatial address, its topology, and can access its non Geo-Spatial attributes. Each object can also serialize itself and represent itself graphically.

CLayer

CLayer and its derived classes are containers for Object classes. Each Layer class can contain one or many Objects of similar type.

CSmbolRep

CSymbolRep and its derived classes determines how the objects should be portrayed within fixed scale ranges.

CMap

CMap is the container class of different layers. It represent layers in a common domain. CMap also provides interfaces to external applications for handling spatial queries.

CDbLink

CdbLink class provides persistence to CMap and all other contained objects.

features. The architure of the system should be built in a modular way to allow the future modification and interfacing possible.