Chapter 1

The Nature and Characteristics of GIS

GEOGRAPHIC Information System (GIS) is a broad term that is applicable to a system that uses computer facilities for handling data, referenced in the spatial domain, with the help of appropriate data sets, for carrying out spatial analysis, and for presenting the results in the most convenient mode of display. In a GIS environment, maps are used to visualise spatial data to reveal and understand relations between them. Maps are no longer just the final products they used to be, when they functioned as a medium for both storing and presenting spatial data. The introduction of digital data has brought about a split between the two functions of storage and presentation [Krack 1996].

In modern GIS, the map is replaced by a database accessed through a software system; the software simply reproduces graphic products that look like traditional maps, but with added visual representations. At this level, the technology only provides a replacement for the map, often at a much higher cost. The next level of technology requires more sophisticated software that can reorganise the raw data to discover new relationships. It is at this level that the real advantages of a GIS become apparent.

1.1. DEFINING A GEOGRAPHIC INFORMATION SYSTEM

THERE are many definitions for the term geographic information system, each developed from a different perspective or disciplinary origin. See Appendix A. Some
focus on the map connection, some stress the database or the software, and still others emphasise applications such as decision support. One of the most general definitions, developed by consensus among 30 specialists, is as follows:

Geographic information system is a system of hardware, software, data, people, organisations, and institutional arrangements for collecting, storing, analysing, and disseminating information about areas of the earth. (Cited in [Chrisman 1997] page 7)

Unimpressive, as it may sound initially, this definition reveals all the crucial characteristics of a GIS, as the terms, system, data, information, organisations, et cetera, are expanded to their intended meanings. It also gives an idea of the sweeping range and scope for the applications of GIS.

In the digital perspective, a GIS could be considered to provide users not only with an array of tools for managing and linking attribute and spatial data, but also with advanced modelling functions, tools for designing and planning, and advanced imaging capabilities. While many of these capabilities also exist in other types of systems such as visualisation and virtual reality systems, modern GIS are unique because of their emphasis on providing users with a representation of objects in a cartographically-accurate spatial system and on supporting analysis and decision making [Mennecke 1997].

A GIS, therefore, must fulfil the following functions:

Provide tools for the creation of digital representation of the spatial phenomena, also known as data acquisition and encoding.
Handle and secure these encodings efficiently, by providing tools for editing, updating, managing and storing; for reorganisation or conversion of data from one form to another, and for verifying and validating those data.
Facilities for information browsing, querying, summarising, and simulation, enabling the user to gain additional insight into theoretical or applied problems.
Chapter 1. The Nature and Characteristics of GIS

Assist the task of spatial reasoning by providing for efficient retrieval of data for complex queries. Create human compatible output in varied forms of printed tables, plotted maps, pictures, sound, scientific graph, or other suitable multimedia modes.

1.2. THE STRUCTURE OF A GEOGRAPHIC INFORMATION SYSTEM

IN the light of the above criteria, a GIS could be considered as a software that has several subsystems and which communicates with the various hardware devices in its environment. It includes a database (DBMS) subsystem for storing and managing data, and a graphics subsystem for displaying the relevant spatial data. These two subsystems are controlled by the computer operating system. They communicate, through the graphics workstations, with the users by means of a user interface and command language interpreters.

The devices employed in a typical GIS are high capacity disks for storing data in various formats (text, graphic, and image), digitising tablets, and scanners to enter graphic data, high resolution printers/plotters for presenting the results. Connecting the GIS computer to a computer network, will allow for the exchange of data with other users (Figure 1.1)

The focus of GIS design is slowly turning towards incorporating them with multimedia data types. Pictures, sound, animated sequences, and unstructured texts may form an integral part of geographically located entities in information systems. According to Kemp, multimedia’s technical issues, such as storage, retrieval, display, and transmission, impose considerable additional demands on the GIS. Video/audio data are several orders of magnitude greater than even bitmaps or rasters, and they often require special storage devices. Storing and managing such data make GIS data
management, already designed to handle large volumes of spatial data, even more complex [Kemp 1995]. Kemp cites other characteristics of multimedia data types that have negative implications in integrating them with GISs, based on the current state of multimedia technology.

As noted in 1.1, a GIS ideally must include the tools to keep track of data, bring information into the database, produce output, provide security, and allow for a range of spatially as well as non-spatially-oriented processing.
Data acquisition involves the direct capture of data, either through remotely sensed images, or close at hand, by still video cameras used on a field trip. This could also be accomplished through digitisation of accurate maps, tables of data from socio-economic surveys, *et cetra*, all of which have involved some kind of prior processing. Direct digital encoding via satellite is followed by an interpretation step.

Tools for output may be many, depending in part on the kind of information in the system. Analytically-oriented systems will most likely provide capabilities for the creation of physical documents in the form of maps, graphs, presentation graphics, tables, or narrative text. Query-oriented systems will focus more on the preparation of virtual output in the form of quickly produced maps or graphs, the contents of relevant segment of the database, or derived statistics, all made available immediately in response to a user request.

The processing functions are the tools for performing the spatial analysis. Again, these depend on many factors. A census data may require aggregation and classification. A query-oriented system is more likely to emphasise functions of search and retrieval.

1. 3. CHARACTERISTICS OF SPATIAL DATA

MAPS have been used as media for revealing details of shapes of entities as well as position and variations in attributes. While maps can clearly depict objects, simple or complex, unitary or compound, the cartographical techniques employed for effectiveness and the high information content are usually a challenge to the imagination of the map designer. A map model, nevertheless, remains an alternative
form of looking at reality so much so, that the designer of a GIS should be aware of
the cartographical conventions of generalisation, feature displacement, and
symbolisation. The map model has limitations and may distort the perception of space.
It is based on a two-dimensional view of the world, and is only one influence on
conceptualising spatial objects and their variations.

The computer-based organisation of digitised data into tables or other formats,
is quite different from the map model. **Data model**, the term used for a comprehensive
set of conceptual tools for organising data with a semantic basis, addresses the
computing issues of minimising the storage requirements and enhancing performance.

Spatial data are different from the usual data stored in databases, as they
involve an infinite number of points in space, instead of a fixed number of entities like
“parts-id”, or number of people in a city. They are present in both intensional and
extensional\(^1\) forms of representation and hence there will be varying degrees of
precision in the representations [Laurini & Thompson 1994].

Spatial entities can be viewed from several perspectives regarding
dimensionality and precision. For some applications of GIS, the reduction in
dimensionality associated with space-filling curves is important. The broad extension
of this concept of precision of representation forces us to recognise that for many
phenomena, boundary lines are artefacts, and that a field-oriented view can be more
appropriate than an entity-orientation. In short, spatial data are more complex than

\(^1\) In the extensional mode of representing a road (or a river), the end points of a series of segments are stored. A point
between the end points of an arbitrary segment will not exist in the database. In the case of a non-spatial entity, like an
amount or a number, data will always be extensional as every digit of an invoice amount will be important.
When the road is stored in intensional mode, the end points as well as the equation of an arc (or other methods for
constructing the arc) that could approximate the road are stored. Therefore any point between the end points of the road
can be generated.
non-spatial data. They are special with regard to continuity in space, anisotropicity, sphericity, multi-dimensionality, duality (attribute and spatial) of access and intensionality [Laurini & Thompson 1994].

1.4. THE EVOLUTION OF GEOGRAPHIC INFORMATION SYSTEMS

GEOGRAPHIC Information Processing have their roots in the history of thematic cartography. Modern GISs evolved from the combinations of increased capabilities of computers, better analytical techniques, and a renewed interest in environmental/social responsibilities. The earliest systems such as the Canadian Geographic Information Systems (CGIS) [Tomlinson 1976], Storage and Retrieval of Water Quality Control (STORET) [Green 1964], Management Information Assembly and Display System (MIADS) [Amidon 1964], were all consequences of the above factors and application requirements.

The GIS-user community is a diverse group drawn from a wide spectrum of disciplines. Foresters, geographers, city administrators, public utilities, academic researchers—all exploring the same territory, but from different directions and with different goals. As soon as the cartographers succeeded in the production of fairly accurate base maps, attempts were initiated for making thematic mapping wherein, different attributes of a given geographic region are appropriately defined and displayed. The forerunner to a geographic information system came into being when maps portraying magnetic variations with isolines, and wind directions with arrows were produced [Robinson 1982].
The transportation modelling study, commissioned in 1955 by the Detroit Metropolitan Area, might well have been the first GIS application. By early 1960's academicians were beginning to undertake research along these lines. If transportation studies were the first automated geographic analysis systems, with statistical capabilities, resource inventories gave widespread applications to the emerging field. Governments in Canada, and the United States began to see the immensity of opportunities that could be realised by the electronic storage and retrieval of land-based data. In 1962, Roger Tomlinson of the Canada Land Inventory, developed the Canadian Geographic Information System (CGIS) [Tomlinson 1976]. CGIS was designed for more than just one specific application. Designed as a polygon-based system, it had many shortcomings in terms of real-time graphic editing, thereby preventing it from being an interactive system. Besides possessing storage and retrieval capabilities, CGIS could reclassify attributes, change scales, merge and create new polygons, and create lists and reports.

At the same time, the United States Forest Service was developing its own land Management Information Assembly and Display Systems (MIADS). This was a much more advanced system, one that could not only store and retrieve attribute of a given cell, but also perform simple overlay functions and mathematical calculations and prepare simulations over time. The outputs were generated on a line printer, but arranged topologically to produce cartographic images [Amidon 1964]. MIADS, therefore, could be considered as the first full-service GIS in the natural resource environment.
Along with the development of environment-based GIS models of the 1960's, there were similar attempts on large-scale, geo-based urban models. One such model was developed in Australia in 1970. TOPAZ (Techniques for the Optimum Placement Activities in Zones) is a general planning technique that has been applied at the regional, urban, and facility planning levels [Brotchie et al. 1980].

One of the early systems that has had a great impact in the applications-sector of GIS was the Dual Independent Map Encoded (DIME) file system of the United States Census Bureau. The DIME files represented a dual encoded map of streets and intersections along with geographic codes and addresses. Each street segment was coded twice, once as an edge to one block (right), and again as an edge to the other block (left). A block is a unique area defined by a boundary of street segments, and uniquely numbered. Geographic coordinates were specified for each intersection with each node and block. The computer could then construct two independent networks and then match them to ensure that all areas are accounted for and that the network is complete [Totschek et al. 1969]. Although technically the DIME file is only a data file, it has probably been used in more applications than any other system.

1.5. THE IMPORTANCE OF GIS

THERE has been a profound change in the processing of geographic information since the emergence of information technology. Even though a number of new journals and new technical societies in this field are being spawned, and several universities in the U. S. are instituting academic programmes in GIS at undergraduate and graduate levels, the importance of this technology is not well-understood, and it remains a grossly under-utilised technology.
In April 1994, President Clinton of the United States issued an Executive Order under the title “Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure,” in which he justified the order by noting that

"Geographic information is critical to promote economic development, improve our stewardship of natural resources, and protect the environment. Modern technology now permits improved acquisition, distribution, and utilization of geographic (or geospatial) data and mapping. The National Performance Review has recommended that the executive branch develop, in cooperation with State, local, and tribal governments, and the private sector, a coordinated National Spatial Data Infrastructure to support public and private sector applications of geospatial data in such areas as transportation, community development, agriculture, emergency response, environmental management, and information technology."  
(Cited in [Guptill 1994])

The above order puts the importance of GIS in the current perspective.

Urban planners and cadastral agencies need detailed information about the distribution of land and resources in the area under their jurisdiction. Civil engineers need to plan the layout of roads and canals and to estimate construction costs, including those of cutting away of hill sides and filling in valleys. When a highway is planned to pass through a heavily populated state such as Kerala, its alignment at every zone could become controversial. This often results in re-alignment, entailing cost over-runs and waste of manpower. If a properly designed GIS is available for the entire state, the initial alignment can be drafted accurately and presented for public discussion. Any revisions resulting from these discussions can be made within hours, and the final alignment can be published within a matter of weeks, instead of many years. The police department need to know the spatial distribution of various kinds of crimes and accidents; medical institutions are interested in the spatial distribution of
sickness and disease; and the commercial interests are eager to know the distribution of consumer preference.

The enormous infrastructure of water supply lines, electricity distribution, telephone circuits, sewerage lines—all need to be recorded and manipulated in map-form, if the civic authorities are to react timely and efficiently to the breakdown in any of these public utilities. A robust GIS for a given town can display at different layers, the pertinent information of the pipeline layout, location of the valves, alternate supply route; type and route of power supply lines, interlinkings, location of transformers, location and types of circuit breakers, _et cetera_. This will enable the utility authorities to despatch the right technicians to the right place with the right tools, instead of perambulating across the town from one corner to another in search of the fault as is the current style of fault repairing.

A GIS is a tool box as well as a database. As a tool box, it can allow planners to perform spatial analysis using its geoprocessing and cartographic modelling functions such as data retrieval, map overlay, and connectivity [Berry 1987]. Map overlay is probably the most useful tool for planners who have a long tradition of using map overlays in the analyses of land suitability. As a database, spatial and textual data can be linked by a geo-relational model for data query and retrieval. Planners can also extract data from the GIS database, use them as input for other modelling and analysis programme, or merge them with data amassed from spatially conducted surveys, for making planning decisions.

The inventory, analysis, and mapping capabilities of GIS have wide applications in urban and regional planning, ranging from data retrieval and site selection to project
monitoring and implementation, development control, mapping and zoning analysis [Anthony 1991]. A GIS is most useful at the analysis, forecasting, and plan-formulation stages. Through the use of map overlays, problem areas such as urban redevelopment districts and environmentally vulnerable areas can be identified. At the plan formulation stage, land suitability maps can be generated using GIS for the preparation of master plans and land-use plans.

Remote sensing provides one of the latest and cheapest source of data of the terrestrial surface (vide page 5). Images recorded by sensors onboard geosynchronous satellites are being beamed at regular intervals at favourable cost per unit area. Such images can be processed to produce basic data sets, combined with other types of spatially referenced data, and integrate with a GIS to generate value-added products. Various spatial information pertaining to urban areas can be extracted from digital remotely sensed images using new analytical tools such as pixel-based spatial re-classification procedure. The latter technique also allows surface and volume descriptions of buildings. Further, the object-based stereo digitising system allows heights and volumes of buildings to be aerially determined. The output of this can be imported into a GIS system [Barnsley et al. 1993]. In the area of soil conservation, approaches along the following lines are bound to be fruitful:

A hydrological model can be linked to a GIS with special techniques that allow on-line data handling with time series grids, and the linked up combo can be used for Catchment run off forecasting and evaluation of the spatial distribution of run off production. Evaluation of spatial distribution of evapotranspiration and soil moisture content on a catchment scale. Distributed prediction of flow in different river reaches for flood or pollution control.
GISs can be designed, incorporating capabilities for the analysis of hazards. Hazards from surface instabilities such as land slides and subsidence can be modelled and embedded in a GIS [Wadge et al. 1993]. In fact, many countries are moving towards making evaluation of the visual impacts of any new development proposals, a statutory requirement, as part of the environmental assessment.

It is the developing nations like India and the heavily-populated regions like the state of Kerala that should derive the maximum utility from tools like GIS. The uncontrolled and indiscriminate growth of every settlement—panchayat, municipal, or corporation areas alike—is turning the towns and cities fast into huge uninhabitable "slums." It is ironic that GIS remains an unknown entity in these over-crowded parts of the world, where it is needed the most.

It would be fool-hardiness to think of GIS as mere means of coding, storing, and retrieving data about aspects of the earth's surface. In a real sense, the data in a GIS should be considered to represent a model of the real world domain, because these data can be accessed and transformed interactively, and they can serve as a test bed for studying the environmental process or for simulating the possible long-term effects of planning decisions. Such studies can preclude the planners' decisions turning into death-traps for future generations.

1.6. MAJOR APPLICATION AREAS OF GIS

Apart from the applications of GIS in the solution of environmental and social problems, as discussed in Section 1.5, there are many other areas where GIS can be
put to productive use. These are categorised and discussed in the following Subsections.

1. 6. 1. Automatic Mapping

One of the first applications of GIS was that of capturing spatial data to automatically generate maps [Coppock & Rhind 1991]. Software designed to support automated mapping (AM) represents a powerful tool for business applications because it enables managers to generate spatial data in-house. It did not prove itself to be very popular because data capture was problematic for its users. There were other problems due to positional accuracy, attribute accuracy, completeness (are all the relevant objects included in the map?) and interpretive issues (is this a tree or a bush?), et cetra [Goodchild 1992]. The training of personnel was another constraint that dissuaded organisations from whole-heartedly accepting the idea of in-house map-making.

Although error propagation and training continue as problems, potential opportunities still exist for business to make use of AM tools. Remote sensing and global positioning systems (GPS) allow more accurate map production by removing paper map as a data source [Goodchild 1992]. In the United States, Electric companies lead the rest of the business in the use of AM, having begun early in the 1970s to use commercial GIS technology, for managing facilities data like the transmission line routing, locations of switching yards, locations of transformers, et cetra, customer records, order processing, and network analysis [Mennecke 1997].
Petroleum companies are some of the largest users of AM operations in the world. They have adopted GIS and digital mapping for supporting their operational and exploratory activities like managing well locations, lease information, and seismic information.

As with many technologies, AM applications have been caught in the Internet. InfoNow, a company based in Aurora, Colorado in the United States, has an on-line service called FindNow which enables a subscribing company to provide customers with information and maps showing the location of facilities, service centres, or retail locations. Visa Plus uses this service to develop an ATM (automatic teller machine) Location Service that allows Visa customers to locate the three nearest ATM machines to a specified street address or intersection.

1.6.2. Facilities Management (FM)

GIS fulfills “an ever-increasing demand for information pertaining to the location, condition, and performance of the utilities infrastructure” [Rector 1993]. The demand for information is not confined to utilities. Many organisations in other sectors must manage and control facilities such as manufacturing plants, distribution centres, retail outlets, and other departments of the organisations.

The key functions used in FM are the spatial visualisation, and database management functions. Most FM applications use historical or transaction (real-time) data for managing or monitoring facilities [Schek & Wolf 1993]. They rely heavily on the imaging capabilities of GIS. The AM functions of GIS (defined on Page 14) are often combined with FM functions to provide organisations with a system for
generating, managing, and utilising maps and other spatial data, in the management of the organisation’s physical plant [Mennecke 1997].

1.6.3. Transportation and Logistics

Transportation systems use tools and algorithms such as transportation network models and material flow models that came from disciplines such as operations research and production management. Specific GIS tools that fit into this category include vehicle routing and navigation system, intelligent vehicle highway systems, dispatch systems, production control systems, and inventory systems [White 1991]. Each of these technologies represent useful applications that managers can use to develop tactics to reduce waste, reduce personnel and fuel costs, and provide better customer service [Lapalme et al. 1992].

Logistical problems are common to many industry segments. Electric utility companies use GIS to produce location maps so that meter readers can plan their daily routes in advance; American Automobile Association uses GIS to support routing analysis and travel information reporting to its members. Car rental firms are increasingly including navigation systems in their rental vehicles. General Motors uses GIS-related technology to provide vehicle navigation systems. Railway companies use GIS for managing rail maintenance history by route and milepost down to each individual track. Federal Express uses GIS for tracking packages along their routes.

1.6.4. Spatial Decision Support Systems

Much of what managers do in business relates to planning and making decisions. Strategic decision making generally involves decisions that are broad in
scope, unstructured, and focussed on long time frames. Information systems
developed to provide strategic decision support to managers have generally been
designed to provide access to critical data, analytic and modelling tools, and
communication support. Many of these systems, however, represent spatial data and
information inadequately [Mennecke 1997]. The term spatial decision support system
(SDSS) has been proposed to represent easy-to-use systems that have capabilities for
manipulating and analysing spatial data [Densham 1991]. SDSSs provide capabilities
to input and output spatial data and information. They allow representation of
complex spatial structures, and they include analytical tools for spatial, geographical,
and statistical analyses. As such, SDSSs are an important class of GIS designed for
use by middle- and upper-level managers. GIS can be effectively employed in strategic
endeavours such as corporate downsizing, organisational restructuring, site selection
(as employed by McDonald’s), and competitive analyses (as employed by Time Warner
Communications to determine cellular phone market potential).

A software tool, called Spatial Group Choice, developed under a collaborative
project of the University of Idaho (Dr Piotr Jankowski) and the University of
Washington (Dr Timothy Nyerges), and funded by the National Science Foundation,
combines interactive multimedia map visualisation, with multiple criteria decision
models¹ This Collaborative Spatial Decision-Making (CSDM) system helps problem
solving in a group setting, comprised of elected officials, interest group
representatives, and the public, enabling the group members to develop consensus-
based solutions to decision problems. Several experimental groups are testing the

¹ http://www.idaho.edu/~piotrj/
http://weber.u.washington.edu/~nyerges/
http://weber.u.washington.edu/~ljmoore/csdm.html
software on a real spatial decision problem of habitat restoration and economic development along the Duwamish waterway, near Seattle, WA, USA.

1.6.5. Design and Engineering

Computer aided drafting (CAD) systems have been routinely used by engineering firms to develop archive architectural drawings. GISs have the necessary features which, with suitable modifications, can perform the functions of a CAD system. While CAD systems have rudimentary links to databases and deal with relatively small quantities of data, they do not allow users to assign symbols based on user-defined criteria. They are also constrained by limited capabilities. Organisations that are involved in engineering and design (e.g., Electric utilities, communication engineering, highway construction, et cetera) would benefit immensely by coupling their GIS usage with CAD technologies.

ARC/INFO is an extensive GIS with vector and raster capabilities including import, georeferencing, editing, analysis and output. [Mayall & Hall 1994] found that by integrating ARC/INFO\(^1\) with AutoCAD\(^2\), and in conjunction with other models, landscape changes could be simulated or predicted. They used ARC/INFO Macro Language to output GDS command files from the graphic and attribute data stored in the GIS. This procedure allows an end-user to obtain quite realistic 2D and 3D digital representations of current landscapes and to produce visual simulations of landscape change. It also allows numerous planning applications, such as site location,

---

1. ARC/INFO is registered trade mark of Environmental Systems Research Institute, Inc., USA.
2. AutoCAD is registered trade mark of Autodesk, Inc, USA.
environmental impact analyses, transportation routing, and resource allocation, to be undertaken more efficiently than is possible using the traditional, manual methods.

1. 7. CURRENT ISSUES IN GIS

It is difficult to draw up a comprehensive list of topics which are consuming the interests of those involved in GIS activities. However, they can be grouped into broad areas and the topics in each area be identified as in the sub-sections to follow [Morrison 1994]. The technology being relatively young, it should not be surprising to note that the areas are not disjoint, but with perceptible overlapping among the areas.

1. 7. 1. Data Related Area

The topics that could be linked to this category are

- Spatial data integration
- Data access methods
- Data conversion
- Object-oriented approach

1. 7. 1. A. Spatial Data Integration

The majority of GISs currently in use are based on relational databases, wherein attribute data and their spatial references are organized and handled separately. In ARC/INFO, for example, the attributes are stored in conventional DBMS, but spatial data are manipulated using conventional file handling techniques. The gain in flexibility attained through this approach is countervailed by the fact that coordinate data are not subject to the same rigorous checkings as in the case of attribute data so that, in these schemes, spatial data suffer degradation in quality control in terms of security and integrity of data. The Starburst project [Lohman et
Chapter 1. The Nature and Characteristics of GIS

al. 1991], [Widom 1996] at the IBM Research Centre at Alamaden, California, and the POSTGRES project [Stonebraker et al. 1990], [Stonebraker 1995] at the University of California, Berkeley, are developing extensions to standard relational database systems.

There is a great deal of interest in the United States in the creation of large spatial databases that would constitute a national geospatial network. This framework database is likely to be a federation of discrete geographic databases, each maintained by the respective participatory agency. The major technical challenge is to synchronize these discrete databases into a federated database system so that they appear as one unified database [Guptill 1994].

1. 7. 1. B. Data Access Methods

In alphanumerical databases, retrievals are normally attribute-based; but in spatial databases we have to retrieve data based on location. Retrievals in spatial databases are made faster through the use of spatial indexing. The subject of spatial indexing in a database context is one of the most difficult problems in GIS.

There are several methods for producing a spatial index, using space-filling curves, cell trees, quadtrees, R-trees, R*-trees, or sphere trees, but none emerges as easy or tractable. Only a few commercial GISs today provide spatial indexing capabilities [Ooi 1990], [Laurini & Thompson 1994].

Spatial databases often are split into several thematic overlays, each concerning a particular aspect, for example, the street layout, sewerage system, network of canals,
et cetra. For handling these layers, it becomes necessary to create as many indices as there are layers. When a given area involves several thematic layers, the processing of countless indices becomes time-consuming.

1.7.1. C. Data Conversion

The Global Positioning System (GPS) of the navigational satellites give access to positional information in near and real time, thereby enabling accurate and rapid surveys. Position data from GPS will therefore find increasing usage along with existing GISs, for new applications or updating an existing GIS. The integration of GPS data with spatial databases, however, are not easy tasks [Dodson & Haines-Young 1993]. Inconsistencies occur due to discrepancies in the definition of map data, or due to internal mismatch within many of the older data used for the existing GIS.

1.7.1. D. Object-Oriented Approach

As noted in 1.7.1.A., the general-purpose databases are inadequate when dealing with spatial information. The existing GISs are incapable of storing the earlier states of the databases, previous states are discarded as soon as a transaction is committed. A database associated with a GIS should incorporate time as part of the semantics of the system so that, not only spatial, but also the temporal, information of the relevant geographic objects could be elicited.

Object-oriented (OO) approach persuades one to view objects as basic constructs for modelling the problem domain, rather than taking an implementation-oriented approach to the problem. The Edinburgh group has shown that the static, atemporal limitation of a GIS can be alleviated through a data-centred model, named
Chapter I. The Nature and Characteristics of GIS

TCObject data model, which provides a user-oriented data structuring, along with temporal and spatial semantics [Ramachandran et al. 1994]. Notable, among others, who have reported their activities involving object-oriented data modelling or OO databases, are [Worboys et al. 1990], [Worboys et al. 1993], [Ervin 1993], [Hamre 1994], [Shaw et al. 1996], [Sendler 1996], and [Shekhar et al. 1997].

The major advantage of OO Technology (OOT) is that it can encapsulate old applications, making the transition to the new system less painful. In practice, however, it has been found that OOT introduces new ways of developing software, requiring new tools, new programming paradigms, and new metrics, thereby masking the anticipated advantages.

While the field-based models of GIS see the world as a continuous layer of surface over which features (e.g., elevations, road crossings, et cetera) vary, the object-oriented models treat the world as a collection of recognisable objects corresponding to different user-views. Some investigators have been attempting to couple an OO database containing actual data for class descriptions, semantic relationships, spatial and aspatial attributes, et cetera, with a deductive database containing rules and procedures for deriving spatial and aspatial relations [Smith et al. 1987], [Jones & Luo 1994], [Ku et al. 1994]

1.7.2. Technology Related Areas

There are areas besides those covered by 1.7.1, that are receiving intensive attention of researchers involved with geographic information systems, and some of them are listed below
1.7.2. A. Knowledge-Based Approaches

The design of a GIS faces great challenge in the storage, retrieval, and processing of many large and complex data. Knowledge representation in traditional programming is procedural and only complete and structured knowledge, restricted to certain variety, can be represented appropriately. In knowledge-based systems, however, domain knowledge is formalized in a declarative form and allows for a natural representation of a large variety of knowledge types (fragmentary, unstructured, uncertain, incomplete, approximate, as well as structured, complete, certain, et cetera). For these and other reasons, knowledge-based systems have a reputation of being very flexible and general, and hence suitable for many applications.

For some GIS tasks it will be appropriate if the knowledge is embedded within the procedural code. For example, an algorithm to perform an overlay operation on two sets of spatial data could contain the knowledge of how to perform the operation. In many other tasks, it is more appropriate to keep the knowledge separate, within a knowledge base. A GIS for selecting the site for locating an airport, or a petrochemical complex, would rely on separate knowledge bases relating to different areas of expertise (aviation expert, meteorologist, petrochemical expert, spatial modeller, earth scientist, town planner, land developer) to fill their roles in an integrated spatial decision support system for users with different backgrounds. The approaches and methods currently available for KBS development, however, are partial or non-systematic [Roman 1990], [Guida & Mauri 1993], [O'Neal & Edwards...
In spite of the efforts for formulating methodologies for the design of knowledge-based systems along the patterns adopted for software engineering, there is a scarcity of concepts and methods regarding the important aspects of development of knowledge-based systems like, formal requirement specifications, evaluation of performance, complexity, and quality. A major reason for this is that the life-cycle of knowledge base design is different from that of software engineering.

1.7.2. B. Spatial Analysis and Modelling

Statistical analysis techniques for spatial data are of considerable value in increasing the capability of a GIS in spatial information management and analysis. Even though the need for integrating GIS with spatial analysis has been recognised during the past several years, it is hard to find a commercial GIS today that is equipped with spatial analysis tools. The most important reasons for this lacuna are, the obscurity of the field of spatial analysis, the lack of a comprehensive spatial analysis package, and the complexity of the spatial analysis techniques [Goodchild et al. 1992]. One group at the Canada Centre for Mapping has developed a prototype of a package for statistical analysis of spatial data [Zhang et al. 1994]. Spatial autocorrelation measures the degree to which a spatial phenomenon is correlated to itself in space. Investigations by Chou [1993] indicate that the effects of spatial weighting functions on statistics of spatial autocorrelation can be identified from correlograms that show higher-order spatial relationships, these were very difficult tasks earlier.
1.7.2. C. Temporal Models

Space and time—where it happened and when—are fundamental to decision making and prediction. Most research into spatio-temporal representations has focussed on extending relational database models, treating time as a point or interval. [Raafat et al. 1991] introduced an extended time-domain relational algebra into the DBMS, to manage image-processing histories and analyse remotely sensed data over time. More recently, spatiotemporal models have been proposed for recording spatial changes over time, relating to specific geographic objects instead of their locations [Pequet & Wentz 1994], [Worboys 1994], and [Ramachandran et al. 1994]

1.7.2. D. Multimedia Integration With GIS

Thematic map is a powerful visual medium for presenting geographic data. With the advent of digital maps, many new methods were opened for presenting temporal geographic data, more as a phenomenon, than as a series of flat maps that correspond to data at different points in time. This is further augmented with the arrival of multimedia (MM).

Multimedia allows for interactive integration of sound, animations, text, and video images. While a conventional GIS can work with coordinates, pixels, their attributes, and spatial relations, MM technology offers links to all kinds of other geographic information. These could include text describing a parcel of land, photograph of objects that exist in the GIS database, or a video showing the environment of the area. The image of an old map could conveniently be linked to the system via hyperlinks, thereby obviating the laborious process of encoding the map.
Individual MM elements (video/still images, sound, *et cetera*) can be linked to maps to enhance geographic information for visual exploration, analysis, and presentation. While some of these applications have been realized as stand-alone development projects, they have not been incorporated effectively into full GIS [Kemp 1995], [Krack 1996]. A major hurdle for integrating MM and GIS is the lack of a logical data model for spatial and aspatial data that can subsume MM.

In conventional databases, data items are represented by a set of attributes that are of basic data types. Such schemas are not suitable for representing MM objects. MM database systems should manage voice, image, video, text, and numeric data, uniformly. Because information consists of several types of data, these databases should, in some cases, evaluate a query condition by referring to several data that are of different types. The type of the query condition given by a user, may differ from the target data to be retrieved. To illustrate, consider the changes in landscape of a given area, recorded over a period of time, and stored in a MM database. The attribute is represented by a series of video clips, showing the images of landscape at different points on time. The particular video clip for retrieval and display will depend on the period of time indicated in the query.

In other applications, the query may be similarity-based, intended to "find something similar to 'that one'". Instead of *key-based* retrieval, what is needed in MM databases, therefore, are *similarity-based* queries and *content-based* retrievals. In large databases, indexes are crucial for data retrieval at reasonable speed. But conventional indexing techniques like B-trees and inverted files are not effective with similarity-based queries and content-based retrievals. Researchers from different
fields, ranging from GIS [Shepherd 1994] to medical imaging [Petrakis & Orphanoudakis 1993], are trying to find solutions to these problems, using techniques of neural networks [Wu 1997], and object-oriented data model [Yoshitaka et al. 1994].

The methodology of management of MM databases are still in a fluid state. While in DBMS, the focus is on search and indexing techniques, in VDBMS (video database management system), efficient retrieval operations of video in a terabyte database remains a great challenge [Yeo & Yeung 1997]. Commercial object-relational database systems¹, which are the state-of-the-art for implementing MM database systems have been unsatisfactory in intuitive query environments that are characteristics of the MM usage [Grosky 1997].

1. 8. GIS AND HUMAN FACTORS

Because of the visual nature of GIS, issues related to the nature of the task, the visual layout and presentation of the display, and the cognitive effects that these issues produce in the user, are all critical considerations in the use of GIS. Not surprisingly, therefore, extensive research on the human factors vis-à-vis GIS is under way, under the broad area of computer application known as Human Computer Interfaces (HCI) [Adam et al. 1997], [Gentner & Grudin 1996], [Vetter et al. 1995], [Nyerges 1993], [Turk 1993], and [Benbasat 1986].

¹ DB2 Universal Database (IBM): http://www.ibm.com
Oracle Media Server (Oracle): http://www.oracle.com
HCI is less mature as a technology but has lower entry barriers than system architecture. Challenges related to HCI include how to serve a much more diverse and less expert user community than that of professional programmers and how to support a wider variety of applications in an environment of increasingly powerful and sophisticated technology [Adam et al. 1997]. HCI can enhance the “convenience-of-use” factor of applications by providing better interfaces and by conducting studies to determine the effectiveness of alternative approaches to interface design.

Fundamental to this area of inquiry is the consideration of the physical characteristics of the system. Issues such as the layout of features on the screen, the colour and saturation of display objects, the number and type of display objects used, the nature of the input and output devices, and the arrangement of the physical components of the system—all have important impacts on the way people interact with the technology [Benbasat et al. 1986], [Turk 1993]. With the wide range and scope of applications that GIS may be used for, it is also important to consider task characteristics when studying human factors in GIS. A task framework is needed to provide a better understanding where GIS should be used and how it should be applied to specific applications [Nyerges 1993]. Another important issue in human factors research relates to the cognitive characteristics of GIS users [Turk 1993], [Traynor & Williams 1997]. It has been found that differences in individual spatial cognitive abilities had important impacts on decision maker effectiveness and efficiency [Crossland et al. 1993]. A better understanding of cognitive skills and styles, influence of individual cognitive characteristics on user effectiveness — all these factors have impact on a wider acceptance of GIS.
1.9. APPLICATIONS AND DEVELOPMENT OF GIS IN INDIA

1.9.1. Activities at WWF-India

Indira Gandhi Conservation Monitoring Centre (IGCMC) was established in New Delhi, India, by the WWF\(^1\)-India, with support from the Government of India. Using a GIS, built from data acquired via satellites, IGCMC has prepared several databases for ten major protected areas, viz.,

- Buxa
- Ranthambhore
- Periyar
- Palamau
- Pench
- Kalakad
- Mundanthurai & Simlipal Tiger Reserve
- Nagarhole (Renamed as Rajiv Gandhi) National Park
- Gir National Park
- Great Himalayan National Park

These databases will be used to study the forest cover changes, contours, reserve-forest boundaries, and the distribution of endangered plant species. They will also assist in conducting diagnostic studies to identify the frailties of these areas and to launch corrective measures and sustainable developments.

In the case of the protected areas, a project covers wildlife distribution, human settlements, forest-type distribution, river networks, *et cetera*. For a second project on endangered plant species, IGCMC is preparing spatial distribution maps, to be used for assessing conservation threats at regional scale using other spatial data such as forest cover, population pressure, forest contiguity, and the legal status of forest areas\(^2\).

---


2 http://www.igccpc.org/wr/wr-98-99wr98indi.html
1. 9. 2. Activities at NRSA

At the National Remote Sensing Agency (NRSA), Hyderabad, India, attempts are being made to build an expert system that could be used as an interface to ARC/INFO GIS [Hebbar 1991].

NRSA has developed two systems — one under a project for Drought Monitoring and Prevention, and the other under a project for Land and Water Resources Monitoring. The former system became operational in 1989. The main input data source for both these systems are satellite, and the output for both are in the form of tabular data.

National Informatics Centre (NIC) of the Planning Commission of India has an active project for designing a GIS for administrative and planning purposes. The input data sources used for this project are satellite data, census data, and field data. The outputs will be maps, tables, and summary reports. These are available over NICNET (NIC Network), but with access limited to government officials. The system has been operational since 1990, unfortunately, these websites do not get updated—some are of several months "down-to-date"!

1. 10. MOTIVATION FOR THE PRESENT WORK

It was pointed out earlier that the majority of the currently popular GISs were built around relational databases (vide page 19). Relational databases are severely handicapped when it comes to handling spatial data. The situation is compounded when it comes to formulating queries even for the simplest of spatial operations [Orenstein & Manola 1988]. Information in a relational data model is scattered.
Implementation of the normalisation requirements tends to allocate data to small relations, disseminating the attributes of a particular entity into several separate but connectable tables. These factors make queries in relational databases cumbersome.

Because the relational database requires that records must be of fixed lengths, spatial data have to be stored in external files, and this degrades the access times of GISs associated with relational databases. Moreover, the relational data model lacks the ability to specify explicitly the semantic information about relationships [Mohan and Kashyap 1988].

While conventional databases have been successful as data transaction record processing in the business world, the domains of geometric and multimedia GISs require more flexible and expressive approaches than that provided by the relational model. Therefore, object-oriented approaches, developed as a result of the grafting of several roots including artificial intelligence concepts such as 'objects', 'frames', 'abstract data types', 'instances', 'rules', 'inferencing', et cetera — collectively known as knowledge-based approach — offer greater promise for the future of GIS. In such an approach, a feature, simple or complex, is reached not via attributes but directly; and an entity has an existence separate from the descriptive thematic data.

The progress in computer visualisation and multimedia technology is liberating the field of cartography from the domain of maps, into an entirely new process of communication [Fraser 1994]. With the advent of data acquisition via remote sensing satellites, the need to convert increasing volumes of data into usable information is increasing. The computer is becoming a direct source of proliferation of data as well as new ways of visualising it. The driving force in all these changes has been the need
for increasing the interactivity between the user and the computer — much more than the new technologies. The urge for increasing the interactivity between a GIS and its user, has also been the principal motivation for the present work.

For realising this objective, a GIS has been designed, invoking the recent techniques in the representation and manipulation of knowledge, while treating the geographic data as "minute pieces of knowledge" that are inter-related within the domain of a map. The knowledge contained in the points, lines, and textual descriptions, are stored in memory as rules, using OPS5, a production system. OPS5 has the facility for analysing the code during execution. It can start the execution with a set of overly specific rules, and generate new rules that are more general and abstract. When the system over-generalises, it would backtrack or specialise some of the rules that are too general.

The major advantages of using a production system in the design of a GIS are the following.

A production system is modularised at the level of the working memory elements. According to [Harmon 1995], it is only when a developer uses both objects and production rules that he has really modularized an application in the most systematic manner. In formulating the rules, it is not required to make a distinction at the top level between spatial and aspatial components. Data will be accessible in an application-independent manner to many users. All the data in a production system are brought under one umbrella, with consequent advantages in terms of integrity, reliability, and security.

The data model conceived for the present research was based on a three-pronged strategy. Firstly, it would be a vector representation strategy because
Chapter 1. The Nature and Characteristics of GIS

geographic data entities are stored as a vector of coordinate points that define the geometry and topology of these entities. Secondly, it would be a logic-based structure because two geographic entities, the polygon and the line, are represented as predicate symbols describing certain relationship between a series of objects that are attributes of space [Jayaprakash and Menon 1995]. Thirdly, it is a relational structure because the polygon and the line predicate structures can be interpreted as implementation of two relations (tables).

1.11. SUMMARY

THIS chapter has provided an overview of the concepts and issues associated with the GIS technology. The peculiarities of geographic information processing, vis-à-vis general information processing, have been briefly outlined, and the problems and difficulties in building a GIS, highlighted. The enormous potentials of GIS, brought about by the rapid growth in database, visualisation, multimedia, and communication, et cetera, have been pointed out. The importance of GIS has reached such a level that the United States has launched the massive National Spatial Data Infrastructure for coordinating geographic data acquisition and access. Major applications of GIS include automated mapping, design and engineering, demographic and market analysis, environmental monitoring, facilities management, natural resource management, strategic planning and decision making, and transportation and logistics. Current issues that surround GIS, human factors that have a bearing on the quality of GIS, have also been discussed. In the end, the features of the GIS implemented under the present research are described briefly.