Geographic information systems and cartography

M. F. Goodchild

The GIS phenomenon shows no sign of weakening, but cannot last for ever. The paper addresses the long-term significance of GIS from three perspectives. First, the various application fields are reviewed and used to define four views of GIS. Second, the paper looks at some new and continuing debates in GIS and their significance for cartography. The third section of the paper discusses the search for an intellectual core for GIS, and the role of the US National Science Foundation's National Center for Geographic Information and Analysis.

This paper was the keynote address at the Eighth Australian Cartographic Conference, Darwin, April 1990.

Michael F. Goodchild
National Center for
Geographic Information and Analysis
University of California
Santa Barbara 93106
U.S.A.

Introduction

Although no one believes it can last for ever, the GIS phenomenon in North America currently shows no sign of weakening. Sales are up again this year, and new vendors are entering the market. Job prospects for GIS analysts remain exceptionally good. New consulting companies are being formed, and existing ones are booming. Dataquest, a California-based market-research firm, estimates a world-wide GIS market of $500 million by 1991 (Computer Graphics World, March 1989; Fortune, April 1989). An article in Government Technology (March/April 1989) reports that Utility Graphics Consultants of Englewood, CO, surveyed over 1200 cities and counties in North America with populations over 100,000 and found that 77% would be installing geographic information systems within the next three years (there are over 70,000 units of local government in the US alone).

What is all this about, and how will the world of geographic information be different when the GIS boom has faded? How is GIS related to the more established disciplines of cartography, remote sensing, photogrammetry, geodesy and computer science? And how are we to separate the substance from the hype? New technologies are always oversold, and GIS is no exception—what is the lasting impact of this latest exercise in salesmanship likely to be?

I would like to address these issues from three different perspectives. The first will try to make some sense of the confusing and often conflicting claims for GIS and its role in different areas of application. The second will review current and in some cases unsolved issues of architecture and function. The third section of the paper will look at the prospects for a common intellectual and conceptual core, of which GIS is a temporary, technology-driven manifestation, and the role being played by the US National Science Foundation's National Center for Geographic Information and Analysis.

WHAT IS GIS?

The very rapid rates of growth which have occurred in the GIS industry in the past five years create the impression that this is a new phenomenon. In fact, the roots of GIS go back to the mid-1960s, when the term was coined independently and virtually simultaneously in two different fields. Marble and others at Northwestern University used Geographic Information System to describe a computer-based system for management and modeling of transportation networks, in connection with the large-scale transportation studies then in vogue. Almost at the same time Tomlinson and others in the Canadian federal government were developing the Canada Geographic Information System (CGIS) as a practical solution to the problems of managing and analyzing the data being collected by the Canada Land Inventory.

Although the same term was used in both cases, there is no evidence that the designers of CGIS believed that their system could be used to model traffic flows, or that a system for modeling traffic would be useful for resource management. Even today, GIS is not so much a homogeneous community as a collection of largely independent application fields loosely held together by a common set of software and hardware solutions. The development of a sense of commonality is one of the more significant elements in the evolution of GIS over the past two decades. Nevertheless, GIS still lacks a clear intellectual core, and although a consensus is developing, there are still major differences of approach between the various groups which make up the GIS community. In essence, this is a 'bottom-up' field, defined by its component parts rather than by acknowledged commonalities. It lacks almost all of the usual trappings of an intellectual core—textbooks, disciplines, academic departments, journals, and learned societies—although this situation is changing rapidly.

Definitions of GIS

It is perhaps not surprising that there is no single, widely accepted definition of GIS. 'Geographic' implies only that the database is georeferenced, or that it includes some means of accessing data by geographic location. To be called an 'Information System', a software product must integrate a variety of functions, and must allow the user to access a database without detailed knowledge of its format. A GIS can be defined by its functionality ('a system for input, storage, analysis and output of geographically referenced information'), its contents ('a system containing geographically referenced information'), or by its purpose ('a system for support of spatial decision making'). However, all three are broad enough to include a vast array of software products.

The next sections provide a brief overview of the application fields of GIS. From a sociological perspective, these are some of the more significant subcultures which make up the GIS consortium. The fields are not exclusive or exhaustive. This attempt at a taxonomy is merely intended to underscore the diversity of GIS at this stage of its development. For general texts on GIS see Star and Estes (1990), Burrough (1986), Tomlin (1990) and Aronoff (1989).

Resource Management

Maps are very efficient stores of information, but it is surprisingly difficult to obtain certain types of numeric information from them. One traditional method of measuring area, for example, requires the counting of dots on a transparent overlay, an extremely tedious and labor-intensive operation. Nevertheless, such measurements are an essential part of many aspects of resource management, for example in computing timber yields in forestry. Three operations are particularly significant in explaining the early interest in GIS applications in resource management: measurement of area, superimposition and analysis of maps of different themes, and the generation of buffer zones of specified width around map features.

Infrastructure Management

Organizations which maintain complex infrastructure need the ability to track and manage installations geographically. For example, it is important to a utility company to know the locations of requests for service in order to schedule its service vehicles efficiently. 'One-call' services which provide information on underground infrastructure rely almost entirely on geographical access to records, in order to identify any existing facilities within range of a proposed construction project. Major customers for infrastructure management using GIS include utilities, departments of transportation, railroad companies, and city and county engineering departments. The acronym AM/FM (automated mapping/facilities management) is often used for infrastructure applications of GIS.

Land Information Systems

Land information systems (LIS) maintain data on individual ownership parcels and associated attributes relevant to assessment and taxation. The term 'multipurpose cadaster' (MPC) refers to a parcel-level database used to support activities in a number of areas besides taxation, such as infrastructure management. Since an LIS is often constructed from raw survey information, there is a strong linkage between this field and the discipline of surveying, just as resource management applications are often strongly linked with remote sensing.

Vehicle Routing and Scheduling

In resources and infrastructure applications, it is widely acknowledged that the establishment of the necessary database can often absorb the majority of the system's budget. Other applications of GIS have grown around existing digital databases, in part because the costs of data input can be largely avoided. The existence of the US Bureau of the Census's TIGER and DIME files has led to a number of GIS applications, particularly in the routing and scheduling of vehicles. Vehicle navigation aids, for example, offer the capability to display to the driver a continuously updated route map. Vehicle routing systems generate instructions for following an optimum path through a street network. For a discussion of vehicle navigation aids see McGranaghan, Mark and Gould (1987).

Marketing and Retailing

Location is of paramount importance in determining the success or failure of retail establishments, so it is not surprising that geographical factors play a significant role in retail analysis. Marketing and retailing are comparatively recent application areas for GIS. Key GIS functions in this field include: geocoding, the ability to generate coordinate locations from street addresses; point in polygon operations to identify the reporting zone (e.g., ZIP code and census tract) containing a customer's location; and polygon overlay to transfer estimates of population counts between two sets of incompatible reporting zones. For a more extensive discussion of GIS applications in this area see Goodchild (1989).

Four views of GIS

There appear to be four threads underlying these applications of GIS, or four identifiable traditions:

GIS as automated mapping. GIS in this context means little more than the ability to carry out simple manipulation functions on map data before generation of the final (map) product. The new technology provides a cost-effective solution to an existing problem, rather than something new.

GIS as map analysis. CGIS was proposed and built in the mid 1960s as a tool for overlaying maps and measuring areas, both functions which are difficult and costly using traditional techniques. Systems such as the Map Analysis Package (MAP) have their roots in McHarg's (1969) use of manual overlay for landscape planning (Tomlin 1990).

GIS as inventory. Much of the commercial growth in GIS in the past few years has been due to its applications in utilities, local government and infrastructure management. In many of these applications the purpose of GIS is to provide geographical access to existing databases.

Spatial analysis and spatial decision support. Despite the power of the technology, digital mapping is often only marginally cost-effective. But benefits rise dramatically if other uses can be found for the digital data. Cowen (1988) has argued that a GIS is a spatial decision support system, existing to provide answers to queries, and to analyze data in ways which help users make decisions. Its value is to be determined by the quality of its input to the decision-making process, or the degree to which it allows the user to avoid bad or costly mistakes.

ISSUES IN GIS ARCHITECTURE

Rasters and vectors, layers and objects

Cartography clearly lies close to the core of GIS, and we have already seen one view of a GIS as a system for automating the map-making process. Maps are most often the medium for data input, either by scanning or digitizing, and are also often the medium for hard copy output. Yet GIS technology clearly raises a number of profound issues for cartography. Does the digital medium offer a host of new opportunities for conventional cartography, or is it its greatest threat? Does the 'paperless office' have its equivalent in the mapless map library? Will car navigation systems remove the need for road maps?

As an information system, a GIS presents a particular perspective to the user, who sees the real world of geographical variation filtered through the digital medium. The form of the user interface controls the view which the user sees of the contents of the database, and the data models used to build that database control the user's view of the source documents from which it was derived.

There are two major ways of representing a source document in digital form—raster and vector—and these became the basis of two different traditions at a very early stage in the development of GIS. Even today very few GIS can accommodate both structures in their databases, or process requests which refer to both. For example few GIS can take a road represented by vectors and a topographic surface represented by a raster of spot heights, and compute, store and display a topographic profile along the road. Although the choice between raster and vector should be determined by their relative abilities to represent real geographical features, instead it is determined too often by the availability of particular systems.

Recently the raster/vector issue has become part of a much larger debate over data modeling in GIS. The real world of geographical variation is infinitely complex and often uncertain, but must be represented digitally in a discrete, deterministic manner. Sometimes it is possible to define discrete features or objects in more or less rigorous fashion, but more often the digital representation is an abstraction of reality. The raster/vector debate conceals a number of more subtle and fundamental issues.

In essence, GIS databases present two very different views of reality. In one, geographical variation is represented by a set of layers, each of which records the pattern of one variable over the study area. If there are n layers, then n separate items of information are available for each and every point in the area. The variation in any one layer may be represented in numerous ways, including a raster of point samples, a set of non-overlapping polygons, an irregular set of point samples, or a TIN.

In the second approach, we think of the world as a space populated by objects of various kinds—points, lines and areas. Objects have attributes which serve to distinguish them from each other. Any point in the space may be empty, or occupied by one or more objects. GISs which take the layer view of the world often allow the user to populate a space with objects, but then insist that they be forced into the layer model.

For example, in digitizing a soil map one first creates a set of line objects littering the plane (the object view) and then 'builds topology' by using them to partition the plane into non-overlapping, space-exhausting regions. The resulting layer assigns every point in the space to exactly one region.

Object-oriented databases

It is particularly exciting that the layer/object debate is occurring in GIS at the same time that computer science is debating the significance of object-oriented (OO) databases and programming. The contexts are very different, but it is clear that OO has become a major stimulus to GIS (Egenhofer and Frank 1988a,b; Kjerne and Dueker 1988; Armstrong and Densham 1989). OO is a multi-dimensional philosophy rather than a single framework, so its significance to GIS is not at all clear at this time. The following sections briefly review some of the key ideas of OO and comment on their significance for GIS.

Object identity. In the OO approach an object is assumed to be well-defined, and to have an identity which can survive various types of operations. Many of the objects dealt with in GIS are not particularly well-defined, although the concept of 'feature' in cartographic databases is relatively well understood. Just as raster and vector models are each appropriate for different thematic layers, so layer and object views may be more appropriate for different areas of GIS application.

Encapsulation. From an OO perspective, certain operations on objects become properties of the objects themselves, rather than separate procedures executed on the objects. Armstrong and Densham (1989) argue that this makes good sense at the level of the user interface, as the user seeks to manipulate objects. The exercise of rewriting procedures into this form may also lead to significant insights. On the other hand there may be methods of spatial analysis which are not easy to write in encapsulated form.

Inheritance. In the case of survey data (Kjerne and Dueker 1988) it is important that links exist between the coordinates which form parcel boundaries and the survey data which produced them. These links can be used to propagate inaccuracies, and to keep track of the lineage of data. Concepts of inheritance are not strong in current GIS, but will have to be in the next generation as legal issues become more and more important (Epstein 1988).

Examples of the layer/object debate

Some examples might help to illustrate the layer/object issue. When the US Bureau of the Census introduced the DIME system in the 1960s it was seen as a method of representing street networks with one specific advantage: the structure allowed internal checks of data integrity. More specifically, the street segments forming a block could be checked geometrically (they all connect to form a single loop) or logically (all must point to the same block identification). A more elegant version of the same concept is present in the new TIGER files produced for the 1990 census. Every point is assigned to exactly one block or 2-cell; 2-cells are bounded by streets or 1-cells; and 1-cells are bounded by nodes or 0-cells. In essence, the TIGER model is an example of the layer view of the world, which was chosen because of the integrity checks which are possible. However street networks do not behave as a single layer; two streets can cross in the plane but not intersect, and the point of crossing is occupied by two different street objects. Thus the object view seems more compatible with reality.

A second and somewhat related example concerns transportation networks. The US rail network, for example, can be represented at a scale of 1:100,000 by approximately 300,000 line objects. Attributes of the network, such as stations, bridges or tunnels, are located by rail companies in terms of mileage along a route. However the current generation of GIS requires a line object to be homogeneous for purposes of display. Thus in order to display tunnels it is necessary to split line objects at the beginning and end of each tunnel feature. The effect is a massive and unmanageable increase in the number of line objects.

In essence the rail network example requires a third view of the world which has more in common with the layer view than the object view. The network is a one-dimensional space, with its own system of addressing, embedded in a two-dimensional space. A number of layers are defined over the network's one-dimensional space, using a variety of methods of modeling—point events (e.g. stations), segments with uniform values (e.g. track quality), or segments with linear variation (e.g. track elevation).

Map data models

The layer/object debate is becoming as important to the current GIS industry as the earlier raster/vector debate, and carries a fundamental message. Whereas the raster/vector debate was over how to represent the contents of a map in a database, the layer/object debate is over how to represent the multivariate complexity of reality. The map is no longer the source of geographical data, but a step along the way to creating an accurate view of the real world. That view is filtered or constrained by the source document as well as by the database itself.

Maps were devised as very efficient stores of geographical data, and very efficient ways of communicating information about geographical variation to the user. However these objectives were optimized over centuries within the constraints of a very limiting technology (Goodchild 1988). Maps represent the world in a static manner on a sheet of paper using lines of constant width, text, and uniformly textured fill; it is comparatively difficult to vary line width or texture continuously. By contrast, a digital display has the ability to vary color and texture continuously, rotate and change scale, and animate. Pointing devices can be used to bring up more text as required. The options available for displaying topographic variation, for example, are much richer than on topographic maps, which traditionally rely on the contour model. Contours are very effective at allowing the user to perceive the nature of topographic surfaces, but in a digital database they become comparatively inefficient representations compared to alternatives like the DEM or TIN. The accuracy with which a contour model can return the elevation at a point is variable, depending on how close the point is to the nearest contour.

The debate over GIS data models is changing the role of the map in GIS. Maps are no longer seen as the truth which digital databases must capture, but as steps in the process of capturing real geographical variation. It may be that the filtering which maps impose by their use of particular data models acts as an important part of the generalization and interpretation process. On the other hand the same filtering may severely restrict the user's ability to access the real world. This is particularly true in dealing with the issue of data accuracy, and the need to establish levels of confidence for GIS products (Goodchild and Gopal 1989).

Views of the future

Three points seem to emerge from this discussion of GIS data modeling, and to offer some expectations about the future:

Comprehensive data models. To compete effectively, the next generation of GIS must deal with more than one data model. The user must be able to choose that model which best captures his or her view of the world, and must not be restricted by GIS design choices. Rasters and vectors, layers and objects have all proven valuable in representing various types of geographical variation.

Platform diversity. On the other hand, the number of data models, the sizes of future databases and the complexity of many applications all suggest that the future lies in diversity. It will become increasingly impossible for GIS vendors to offer one product for all purposes, running on a variety of platforms under different operating systems. The vendor's resources will be entirely absorbed by the task of maintaining such a system, leaving nothing left for development. Instead, vendors will specialize. Some will serve the needs of large databases, perhaps with limited functionality, running on large systems. Others will gravitate towards high functionality analysis systems, running on small databases on workstation platforms. Data compilation and data input are other areas of potential specialization. But this will require a clearly defined common data standard, which should be comprehensive enough to include all applicable data models. Needless to say such a standard does not currently exist.

3D and time. The earlier data model discussion made no mention of either the vertical dimension or time. To some extent both can be incorporated within current data models, but a more comprehensive approach is clearly needed. Raper (1989) has recently published an extensive review of 3D data modeling. It remains to be seen whether 3D GIS is sufficiently distinct from other fields such as CAD and medical imaging, and whether fields such as geology, hydrology, geophysics, oceanography and atmospheric physics have sufficient commonality of interest.

A SCIENCE OF GEOGRAPHIC INFORMATION

The debate about GIS data modeling seems to underscore the central message of the GIS phenomenon—that different applications of geographical data share a number of common problems and issues. GIS is a coming together of an amazing array of diverse interests, and the fact that each can find some usefulness in a common technology, despite its current level of development, is important in itself.

Three decades ago remote sensing hardly existed, and disciplines like photogrammetry, cartography, geodesy and spatial analysis had their own unique technologies. With a common digital technology, it seems that many of the old disciplinary boundaries are also breaking down. What is emerging in its place may be the beginnings of a science of geographical information, with the ability to focus on generic issues rather than on the idiosyncracies of individual technologies.

The suggestion that science be organized around different classes of information makes sense in an age in which information is increasingly the limiting factor in human activity. The extent to which spatial information, as distinct from other types, is of value is open to debate, but it is clear that an enormous variety of human activities are organized spatially, and that spatial access to data is particularly important. Numerical and textual information are much easier to deal with in the linear structure of conventional computing, and spatial information was largely ignored in the first three decades of the digital revolution. Problems in spatial statistics are also less tractable than the more conventional kind, despite their importance in understanding uncertainty in spatial data, and there has never been a theory of spatial information to parallel those in other non-spatial disciplines (e.g. Coombs 1964).

The NCGIA

This growing sense of a commonality of issues lay behind the US National Science Foundation's decision in 1988 to establish a National Center for Geographic Information and Analysis (NCGIA). The Center consists of a consortium of three institutions: the University of California, Santa Barbara as the lead institution, together with the State University of New York at Buffalo and the University of Maine. David S. Simonett and Michael F. Goodchild are the Center's co-Directors, located at Santa Barbara; Terence R. Smith, Ross D. MacKinnon and Andrew U. Frank are the Associate Directors at Santa Barbara, Buffalo and Maine respectively. General oversight is provided by a 17-member Board of Directors chaired by John E. Estes.

Abler has described the process of creating the Center, which began with a proposal to the National Science Foundation (NSF) from Jerome Dobson of Oak Ridge National Laboratory in 1984 (Abler 1987). Funding for the Center from NSF is initially at a level of $1.1 million per year for 5 years.

The primary purpose of the Center was laid down in the solicitation document issued by NSF in late 1987: to conduct 'basic research on geographic analysis utilizing geographic information systems'. This puts the emphasis clearly on applications, and particularly scientific and policy-oriented applications, rather than on technical development. On the one hand, the statement of purpose implies that the technology offers significant potential for a wide range of geographically based analyses. On the other, it suggests that although GIS has been widely adopted as a technology in numerous fields, its applications to date have often been relatively unsophisticated. Openshaw (1987) has described GIS as '20th century technology being used for 19th century purposes', implying dissatisfaction with the somewhat rudimentary nature of many applications. To quote from the solicitation document, the goals of the Center are to:

—'advance the theory, methods and techniques of geographic analysis based on GIS in the many disciplines involved in GIS research;

—augment the nation's supply of experts in GIS and geographic analysis in participating disciplines;

—promote the diffusion of analysis based on GIS throughout the scientific community; and

—provide a central clearinghouse and conduit for disseminating information regarding research, teaching and applications.'

The first and third goals both give the Center a role in promoting GIS specifically as an enabling technology for science. The GIS community is currently dominated by applications in management, inventory and policy formation, and the potential scientific applications of GIS have attracted comparatively little interest to date from the scientific community. Scientific applications stress the power of GIS to place information in a spatial context, to suggest relationships based on spatial proximity, and to explore the role of distance as a causal factor. For example, GIS can assist in superimposing spatially organized data (maps) from different sources, and one can envision an 'exploratory spatial analysis' (ESA) tool analogous to the Exploratory Data Analysis (EDA) tools now common in statistics (Tukey 1977). GIS technology greatly increases our ability to view data from different perspectives and under different forms of manipulation and summary. In fact, it seems that the power of spatial organization to suggest causes, explanations and relationships is significantly superior to other forms of data organization, such as the table or even the graph. Recently ESA tools have begun to offer the capability to view data through multiple windows—tabular, graphical and cartographic—with simultaneous updating, so that movement of a cursor within the cartographic window results in continuous (and appropriate) updating of the tabular window, for example.

Applications of GIS in the social sciences have lagged behind those in the earth and natural sciences for many reasons. Funding for technical tools is less readily available; there is a general suspicion of technical approaches in many areas of social science; data may be less reliable and harder to come by; and spatial analysis has only recently become common in disciplines such as history and anthropology. The Center's role as a promoter of scientific applications of GIS applies especially to the social sciences.

Besides ESA, scientific applications require comparatively sophisticated capabilities for modeling and analysis, which are lacking in many GIS activities. The Center therefore stresses not only the incorporation of spatial modeling and analysis techniques with GIS, but also the formal, theoretical basis for GIS and the development of an intellectual and conceptual core to the field.

A useful analogy is between GIS as a scientific tool and other widely distributed software tools for science such as the statistical packages (e.g. SAS, SPSS, and BMD). These have emerged in a similar time frame to GIS, but almost entirely as products for the scientific market, whereas the scientific applications of GIS have had relatively little role in driving and directing its development. Although there is some overlap, one can argue that GIS is a supporting tool for spatial analysis in the same way that SAS, for example, is a supporting tool for statistical analysis. Given the usefulness of spatial organization and the spatial context for data, one might contend that the long term potential for GIS in this context is as large as it is for SAS.

Based on these goals, the Center's programs fall into three general areas: research, education and outreach.

Research Plan

The Center's research agenda is based on the proposition that GIS technology has enormous potential in a variety of applications, particularly science, but that numerous impediments currently exist which constrain the full realization of that potential. The Research Plan (NCGIA 1989) described these impediments in detail under five general headings (in large part these match the five suggested areas of research contained in the NSF solicitation) so they will be merely summarized here:

—Spatial analysis and spatial statistics. Impediments exist in the lack of implementation of spatial analytic methods within GIS, and also in the lack of explicit treatment of data quality (accuracy and uncertainty) in current systems.

—Spatial relationships and database structures. The power of a GIS is constrained by the methods used to represent spatial data within its database. Current systems use a limited range of data models and structures, often derived from cartographic representations, and are based on inadequate understanding of the nature of spatial relationships.

—Artificial intelligence (AI) and expert systems. Many analysis and modeling requirements of GIS are poorly structured and could benefit from AI technologies, as could the complex processes of data input and output.

—Visualization of spatial data. The electronic display offers enormous potential for improved methods of visualization, but current GIS technology largely fails to exploit the capabilities of the new medium.

—Social, economic and institutional issues. GIS technology raises numerous issues of a managerial, organizational or legal nature, and its adoption is currently impeded by the difficulties of accurately assessing its costs and benefits, and by adequate understanding of its impact on organizations.

In formulating a specific agenda of research, the Center looked for a mechanism which would allow it to operate as a three-institution, multidisciplinary, multi-investigator consortium but at the same time focus attention on well-defined topics within the broader research plan. The Center also realized that the plan defined a range of research topics which would be far too large for it to handle, and that the mechanism would therefore have to allow it to encourage and stimulate as much research outside the Center as possible, and provide effective two-way communication.

Research Initiatives

The research plan is implemented through a series of Research Initiatives, based on a model which has worked well at the University of California Santa Barbara's Institute for Theoretical Physics. An initiative lasts for between one and two years, with four or five running at any one time. It begins with a Specialist Meeting, which brings together from 25 to 50 people from three constituencies: organizations with experience of the problem and its effects, researchers with interests in solving the problem, and representatives from the vendor community who can implement the solutions. The purpose of the specialist meeting is to lay out the specific research agenda, including tasks which can be accomplished by the Center or affiliated groups and individuals in the timeframe of the initiative, but recognizing that the meeting will have a role in stimulating research by other groups and individuals as well. The subsequent research at the Center is undertaken by faculty, research assistants, and visiting researchers, and the results are reported at a suitable occasion such as a national or international conference. In several cases it has proven useful to establish a mailing network associated with an initiative, to distribute current information about research activities during the research period.

The current program of initiatives is as follows, with the names of initiative leaders and dates of specialist meetings:

1. Accuracy of Spatial Databases (Michael F. Goodchild, Santa Barbara; December, 1988). GIS are high precision systems which process data as if they were perfectly accurate: in reality, spatial data are often subject to surprisingly high levels of uncertainty and inaccuracy, which current GIS designs largely ignore.

   Over 50 people attended the specialist meeting for the first research initiative, and a book has been published from the proceedings (Goodchild and Gopal 1989). Specific research activities include a bibliography and taxonomy of spatial data errors, to raise user awareness of the problem; fundamental work on the formulation of models of error; methods for incorporating error information within spatial databases; analysis of the propagation of error through GIS processes; and development of finite resolution data structures.

2. Languages of Spatial Relations (David M. Mark, Buffalo; Andrew U. Frank, Maine; January, 1989). GIS can be seen broadly as a technology for helping people work with spatial data, and more specifically as a tool for learning and reasoning about space and spatial relationships. As such,

the technology will be most useful when its data representations and operations emulate the learning and reasoning processes of users, yet current data models and structures fall far short of this ideal. Initiative 2 is conducting research which will ultimately help to improve the way spatial data is represented digitally, and the design of GIS user interfaces.

The research agenda includes the following topics: wayfinding, driving directions and processes of spatial knowledge acquisition; analysis of the structure, cross-cultural and cross-linguistic variations of driving directions, with potential applications to vehicle navigation aid systems; cross-linguistic analysis of locative expressions, and studies of linguistic variation in natural language terms for spatial relations; user interface design, including research on multi-media interfaces, metaphors for conveying and perceiving spatial information, and the visualization of spatial relations; and formalization of spatial relationships, the algebra of spatial relations, and formal reasoning.

3. Multiple Representations (Barbara P. Buttenfield, Buffalo; February 1989). The representation of a geographical feature on a map depends on the map's scale, and the same feature will likely be represented in markedly different ways at various scales. Within a spatial database it is attractive to imagine that a feature might be given a single representation, which would be generalized or simplified for display at different scales. Because this has proven difficult to do, it is common for databases to contain several representations of the same features. Moreover, GIS databases currently provide no explicit and fully satisfactory means of logically relating the various representations of a feature. Hierarchical data structures offer potential and are one of the research topics of this initiative.

   Other topics include definition of the rules required to automate the generalization process; systems for describing the ways features change with scale; and data structures which formalize the logical relationships of multiple representations. The Center is developing a multi-agency, multi-scale database to be distributed as a standard for research work in this area.

4. Use and Value of Geographic Information (Harlan Onsrud, Maine; Hugh Calkins, Buffalo; May 1989). This is the first initiative to address the social, economic, and institutional issues raised by the adoption of GIS technology. Three research themes emerged at the specialist meeting: the need for a taxonomy of geographic information and its uses; development of objective methods for measuring the value of geographic information; and empirical studies of the diffusion of GIS technology. The taxonomy must address questions such as: what types of geographic information exist, and how do they relate to the variety of data models of spatial databases; are certain types of geographic information more or less suitable for handling in spatial databases; what role does geographic information play in human activity, who uses it, and for what purposes?

   Objective measurement of the value of geographic information is an essential component of any serious attempt to evaluate GIS benefits. Finally, research on the diffusion of GIS technology will determine the factors which control the rate of diffusion, and how they can be modified.

5. Design and Implementation of Large Spatial Databases (Terence R. Smith, Santa Barbara; Andrew U. Frank, Maine; July 1989). Two meetings were held on this topic in July 1989: a symposium with formal position papers, and a smaller workshop discussion to lay out the initiative's research agenda. The initiative will examine the technical problems which arise in handling the large spatial databases now being constructed, such as the U.S. Geological Survey's digital cartographic database.

6. Spatial Decision Support Systems (Paul J. Densham, Buffalo; Michael F. Goodchild, Santa Barbara; March 1990). The specialist meeting for this initiative was held in March 1990, and looked at the issues surrounding the development of decision support systems based on GIS technology (Densham and Goodchild 1989).

Future initiatives

Future plans call for further initiatives which will extend the range of research foci, but at the same time build on the results of previous work. Initiatives 7 through 13 have been identified:

7. Visualization of the quality of spatial information.
8. Expert systems for cartographic design.
9. Institutions sharing spatial information.
10. Temporal relations in GIS.
11. Space-time statistical models in GIS.
12. Remote sensing and GIS.
13. User interface design.

This list of initiatives represents a grouping and prioritization of the much larger set of impediments identified in the research plan. It is likely that the order will change, and that the list will be extended, in response to future changes in priorities.

Several specific issues are likely to affect future planning for research initiatives. First, although the methodology and subject matter for much of Initiatives 2, 4 and 9 belong in the social sciences, as well as many of the applications of Initiative 1, no initiative is directed explicitly at developing social science applications of GIS. To rectify this, the Center is planning a conference specifically on that topic in late 1990.

Second, GIS technology has enormous potential in modeling the global environment and the interaction between human and physical systems at global scales. Yet no initiative is explicitly concerned with global scale GIS, although there is relevant research in Initiatives 1 and 5. Data input and conversion are a topic of great significance to the GIS community but not explicitly treated in the current agenda. It may be necessary to find ways of increasing the number of initiatives which can be supported at any one time, perhaps by a mechanism which would allow initiatives to be led at sites outside the Center.

Education

The very rapid development of the GIS field in the past few years has led to an acute shortage of adequately trained staff at all levels, particularly in those areas which require a moderate level of technical skills combined with an understanding of GIS application areas. The major Center effort in education in year one has been the Core Curriculum project, which was designed as a way of quickly increasing the availability of teaching materials in GIS, so that courses could be introduced in new institutions and new disciplines.

The curriculum has been developed as a one-year sequence, as three quarter courses or two semester courses, with a total of 75 one-hour lectures. In the three quarter arrangement the courses are:

1. Introduction to GIS (lectures 1–25)
2. Technical Issues in GIS (lectures 26–50)
3. Application Issues in GIS (lectures 51–75)

The materials consist of lecture notes (typically 6–8 pages for a one-hour lecture), supporting materials (handouts, slides and overheads), lab exercises for courses 1 and 2, exam questions and discussion topics. Course 1 consists of some 250 pages of material, a selection of slides and 6 diskettes containing data and text for the lab exercises.

Between August and December 1988 the proposed outline was developed and circulated for comment, and presented at several conferences. Some 35 experts from the GIS community were then invited to contribute the texts of specific lectures. These were edited and augmented by the Center, and preliminary versions of the courses were completed by July 1989 (courses 1 and 2) and October 1989 (course 3). The project was discussed extensively at the GIS in Higher Education Conference at Ohio State University in June of 1989. During the 1989/90 academic year some 74 institutions have agreed to evaluate and test the materials by incorporating them into their offerings. A large proportion are departments of Geography, but the test sites also include marine science, geology, anthropology, engineering and other disciplines. Instructors will be providing feedback on each lecture and lab, and comments will also be obtained directly from students. These will be used to develop a second, final version of the curriculum which will be ready for distribution in the summer of 1990.

The Core Curriculum project has generated significant interest not only in the academic community but also in industry and in government agencies. At the same time it addresses only part of the overall education and training problem in GIS. Specifically, other areas which the Center would like to pursue in the future include short courses for Core Curriculum instructors; short courses in specialized topics (two such courses have been developed in year one as one-day workshops and presented at several conferences); a Case course to consist of several well-documented case studies designed on the model used in many Schools of Business; and training courses for users of specific systems, emphasizing analytic and modeling capabilities.

THE FUTURE FOR GIS

The high growth rates experienced recently in the GIS industry are exciting, but they inevitably lead to concern for the future: how long can the GIS phenomenon last, and has GIS been oversold? To some extent, growth has been sustained because new application fields have appeared, or new disciplines have become interested in GIS tools, but this cannot continue forever.

GIS is a loose collection of interests held together by common hardware and software solutions, whose long-term survival depends on the emergence of an intellectual core and the symbols and institutions normally associated with a discipline. It is difficult to think of precedents in the form of disciplines which have been founded on tools, although one might argue that computer science emerged when computing found an intellectual foundation, and developed a program of basic, fundamental research. The 20-year horizon seems to offer two alternative scenarios for GIS: a technology which failed to deliver on its promises, or a technology which blossomed into a geographical information science.

Many believe that GIS is driven by technology, rather than by any clearly understood set of objectives. Its drive is undisputable, and shows little sign of diminishing. It raises a host of issues, many of which have been recognized for a long time in more established fields like cartography, but gives them new impetus.

When the GIS phenomenon finally runs its course, and the acronym disappears from view, I hope we will be left with a heightened awareness of the common issues which underlie all of the disciplines which collect, compile and analyze geographical information, and a much better knowledge of how to deal with them.

ACKNOWLEDGMENT

The National Center for Geographic Information and Analysis is supported by the National Science Foundation, Grant SES 88-10917.

REFERENCES

Abler, R.F. (1987) The National Science Foundation National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems* 1:303–26.

Armstrong, M.P. and Densham, P.J. (1989) Object oriented locational analysis. *Proceedings, GIS/LIS '89*. ASPRS/ACSM, Bethesda, MD, pp. 717–726.

Aronoff, S. (1989) *Geographic Information Systems: A Management Perspective*. WDL Publications, Ottawa.

Burrough, P.A. (1986) *Principles of Geographical Information Systems for Land Resources Assessment*. Clarendon, Oxford.

Coombs, C.H. (1964) *A Theory of Data*. Wiley, New York.

Cowen, D.J. (1988) GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing* 54:1551–4.

Densham, P.J. and Goodchild, M.F. (1989) Spatial decision support systems: a research agenda. *Proceedings, GIS/LIS '89*. ASPRS/ACSM, Bethesda, MD.

Egenhofer, M.J. and Frank, A.U. (1988a) Object-oriented databases: database requirements for GIS. *Proceedings, International Geographic Information Systems Conference: The Research Agenda*. NASA, Washington, II:189–211.

Egenhofer, M.J. and Frank, A.U. (1988b) Designing object-oriented query languages for GIS: human interface aspects. *Proceedings, Third International Symposium on Spatial Data Handling*, pp. 79–86.

Epstein, E.F. (1988) Litigation over information: the use and misuse of maps. *Proceedings, Inter-*

national Geographic Information Systems Conference. NASA, Washington, I:177–84.

Goodchild, M.F. (1988) Stepping over the line: technological constraints and the new cartography. *The American Cartographer* 15:311–20.

Goodchild, M.F. (1989) Geographic information systems and market research. *Papers and Proceedings of Applied Geography Conferences* 12:1–8.

Goodchild, M.F. and Gopal, S. editors (1989) *Accuracy of Spatial Databases*. Taylor and Francis, London.

Kjerne, D. and Dueker, K.J. (1988) Modeling cadastral spatial relationships using Smalltalk-80. *Proceedings, GIS/LIS '88*. ASPRS/ACSM, Falls Church, VA, pp. 373–85.

McGranaghan, M., Mark, D.M. and Gould, M.D. (1987) Automated provision of navigation assistance to drivers. *The American Cartographer* 14:121–38.

McHarg, I.L. (1969) *Design with Nature*. Doubleday, New York.

NCGIA (1989) The research plan of the National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems* 3:117–36.

Openshaw, S. (1987) Guest editorial: an automated geographical analysis system. *Environment and Planning* A 19:431–6.

Raper, J., editor (1989) *Three Dimensional Applications in GIS*. Taylor and Francis, London.

Star, J.L. and Estes, J.E. (1990) *Geographic Information Systems: An Introduction*. Prentice Hall, Englewood Cliffs, NJ.

Tomlin, C.D. (1990) *Geographic Information Systems and Cartographic Modeling*. Prentice Hall, Englewood Cliffs, NJ.

Tukey, J.W. (1977) *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.