Geographical information science[a]

MICHAEL F. GOODCHILD
National Center for Geographic Information and Analysis,
University of California, Santa Barbara, CA 93106, U.S.A.

Abstract. Research papers at conferences such as EGIS and the International Symposia on Spatial Data Handling address a set of intellectual and scientific questions which go well beyond the limited technical capabilities of current technology in geographical information systems. This paper reviews the topics which might be included in a science of geographical information. Research on these fundamental issues is a better prospect for long-term survival and acceptance in the academy than the development of technical capabilities. This paper reviews the current state of research in a series of key areas and speculates on why progress has been so uneven. The final section of the paper looks to the future and to new areas of significant potential in this area of research.

“Based on keynote addresses by the author at the Fourth International Symposium on Spatial Data Handling, Zurich, July 1990 (Goodchild 1990). and EGIS 91. Brussels, April 1991 (Goodchild 1991).

1. Introduction

The geographical information system (GIS) community has come a long way in the past decade. Major research and training programmes have been established in a number of countries, new applications have been found, new products have appeared from an industry which continues to expand at a spectacular rate, dramatic improvement continues in the capabilities of platforms, and new significant data sets have become available. It is tempting to say that GIS research, and the meetings at which this research is featured, are simply a part of this much larger enthusiasm and excitement, but there ought to be more to it than that.

What, after all, is the purpose of all of this activity? Expressions such as ‘spatial data handling’ may describe what we do, but give no sense of why we do it. This was one of the themes behind Tomlinson’s keynote address at the First International Symposium on Spatial Data Handling in Zurich in 1984 (Tomlinson 1984). The title of the conference suggests that spatial data are somehow difficult to handle, but will that always be so? It suggests a level of detachment from the data themselves, as if the U.S. Geological Survey were to send out tapes labeled with the generic warning ‘handle with difficulty’. It is reminiscent of the name of the former Commission on Geographical Data Sensing and Processing of the International Geographical Union. A quick review of the titles of the papers at that or subsequent meetings should be enough to assure anyone that their authors are concerned with much more than the mere handling and processing of data—from a U.S. perspective, that the community is more than the United Parcel Service of GIS.

Geographical information systems are sometimes accused of being technology driven, a technology in search of applications. That seems to be more true of some periods of the 25-year history of GIS than of others. For example, it is difficult to suggest that Tomlinson and the developers of the Canadian Geographical Information System (CGIS) (Tomlinson et al. (1976)) were driven by the appallingly primitive hardware capabilities of 1965. On the other hand the prospect of a menu driven, full color, pull-down menu raster GIS in the 386-based personal computer on one’s desk has clearly sold many systems in the past few years. Technological development comes in distinct bursts, and so does the technology drive behind GIS. It may be the motivation behind the desire to handle spatial data, but it fails to explain many of the diverse research efforts being reported at meetings and in the literature.

There have also been phases when applications have driven GIS. CGIS itself was an application in search of a technology, and the drive was sufficiently strong to lead to the prototype of the first map scanner, and to numerous other technological developments (Tomlinson et al. 1976). McHarg had worked out the principles of the map overlay technique (McHarg 1969) long before Berry and others automated them in MAP and its derivatives (Berry 1987); school bus routing software has been around much longer than the problem’s implementation in a standard GIS. But again, much of the subject matter of GIS research lies well beyond any reasonably foreseeable application.

There have also been phases when applications have driven GIS. CGIS itself was The widespread distribution of Landsat and SPOT imagery, and the availability of digital elevation models and street files in many countries have certainly led to applications well beyond those used to justify the data's compilation. TIGER, for example, appears to be spawning its own industry of updaters, repackagers and application developers, although it exists in principle only to serve the needs of the 1990 U.S. Census.

However, although the driving seat of GIS is undoubtedly crowded, I would like to deal in this paper largely with the fourth driver located apparently irrelevantly in the back seat, the 'S' word. It seems to me that there is a pressing need to recognize and develop the role of science in GIS. This is meant in two senses. The first has to do with the extent to which GIS as a field contain a legitimate set of scientific questions, the extent to which these can be expressed and the extent to which they are generic, rather than specific to particular fields of application or particular contexts. To what extent is the GIS research community driven by intellectual curiosity about the nature of GIS technology and the questions that it raises? And if GIS can be motivated by science, then what are their subfields, what are their questions, and what is their agenda? The second sense has to do with the role of GIS as a toolbox in science generally—with GIS for science rather than the science of GIS. What do we need to do to ensure that GIS, and spatial data handling technology, play their legitimate role in supporting those sciences for which geography is a significant key, or a significant source of insight, explanation and understanding?

To do this we must first establish that spatial, or rather, geographical, data are unique, and that their problems cannot therefore be subsumed under some larger field. We must also establish that there are problems which are generic to all geographical data, or at least establish that it is possible to distinguish those that are from those that are not. For example, the accuracy of attributes on a choropleth map of crime statistics would seem to be very little informed by knowledge of attribute accuracy for geographical data generally, but to require instead a level of understanding of the specific problems of crime statistics. However, the accuracy of population estimates for an arbitrarily defined polygon may well be known from, or at least informed by, the general properties of the modified areal unit problem (Openshaw 1977).

2. What is unique about spatial data?

In many facilities management systems, the role of the GIS is to provide an alternative key to data, a method of access based on geographical location. In essence, a spatial database has dual keys, allowing records to be accessed either by attributes or by locations. However, dual keys are not unusual. The spatial key is distinct, as it allows operations to be defined which are not included in standard query languages. For example, it is possible to retrieve all point records lying within an arbitrary, user-defined polygon, an operation which is not defined in standard query languages such as SQL. In essence, the spatial key is multidimensional, but again multidimensional keys are known from other areas, and analogues of point in polygon retrieval can be defined for non-spatial dimensions.

What distinguishes spatial data is the fact that the spatial key is based on two continuous dimensions. It is possible to visit any location (x, y) in the real, geographical world, defined in principle with unlimited precision, and return a value for a variable, for example, topographic elevation z. Terrain is thus characterized by an infinite number of tuples <x, y, z>. In network applications z is defined only for locations on the network, but the number of tuples is still infinite if variation is continuous along this one-dimensional structure of links and nodes. Time series also have continuous keys, but are rarely conceived, measured or represented as continuous, and there appears to be little commonality of interest in the problems of temporal data handling. By contrast, there is ample evidence of commonality in the spatial data handling disciplines.

Many of our data models, particularly polygon networks and triangulated irregular networks (TINs), reflect an underlying view of space as continuous and the need to accommodate the user who wishes to determine z at some arbitrary and precise (x, y). One implication of this is that there exists a multiplicity of possible conceptual data models for spatial data, and that the choice between them for a given phenomenon is one of the more fundamental issues of spatial data handling.

Another distinctive feature of spatial data is what Anselin (1989) refers to as spatial dependence, the propensity for nearby locations to influence each other and to possess similar attributes. Without spatial dependence, there would be no reasonable prospect of creating even approximate views of continuous spatial variation within a discrete, finite machine. It is not uncommon for tuples which have similar values of a key to have similar values of other attributes, but the structure of spatial dependence is unusual, relying as it does on both dimensions of the (x, y) key, with similarity determined by a metric.

Finally, geographical data are distributed over the curved surface of the earth, a fact which is often forgotten in the limited study areas of many GIS projects. We have worried for centuries about how to portray the earth's surface on a flat sheet of paper, and have developed an extensive technology of map projections. However, as a result we have few methods for analyzing data on the sphere or spheroid, and know little about how to model processes on its curved surface. Moreover, we tend to have treated GIS displays as if they were virtual sheets of paper, and insisted on viewing geographical data as if they were projected to a flat surface, instead of exploiting the potential of electronic display to create views of the globe itself. We need to develop the appropriate techniques for working with the globe, and making use of solid modeling rather than conventional two-dimensional graphics, if we are to understand geographical processes at the global scale and contribute effectively to global science. We must rescue the orthographic projection from its present obscurity.

3. The content of geographical information science

Having established that geographical information has unique properties and problems, we can now review the set of generic questions which might make up a geographical information science. This can be done in a largely linear fashion, from data collection to analysis, although some themes tend to cut across this simple arrangement. However, it seems appropriate to begin this review with a disclaimer. What I present in this paper is in many ways my own view, and I would expect it to be challenged. I think my own biases will become clear in what follows. Because of the field's diversity and dynamism it is difficult, if not impossible, for any one individual to attempt a general overview. What follows is therefore almost inevitably incomplete and uneven.

Research is often identified as either pure or applied—driven by basic and innocent human curiosity or by the practical everyday needs of human society. Many GIS are a response to human needs for information management and analysis, and in that sense one might expect GIS research to be more applied than pure. However, one view of pure research is that it is research that has not yet found application; pure research is a long-term investment just as applied research is a short-term investment. From an academic perspective, pure research is often associated with higher prestige, but applied research with greater funding. I have tried to cover the full range from pure to applied, feeling that both are important to GIS. At the same time 'basic research' is the primary purpose of the U.S. National Center for Geographical Information and Analysis, and the center is very fortunate in being funded to do research the applications of which may lie years or even decades into the future.

During the design phase of the CGIS in the 1960s, it became clear that the only practical way to input the large number of maps needed would be by some form of scanning device (Tomlinson et al. 1976). At that time no scanner for map-sized documents existed, and it was necessary to invent one. A prototype drum scanner was built by IBM Canada and successfully tested, at what by modern standards would be regarded as vast expense. Other parts of the CGIS design team were busy inventing other, equally fundamental and now familiar solutions to technical GIS problems, such as the Morton order.

In the almost three decades of development of GIS that are now behind us, similar 'how to do it' research has produced a large number of algorithms, data structures, spatial indexing schemes and other technological solutions. Some of these are unique to GIS, but many have been reinvented in several related disciplines. The Morton order, for example, occurs in the literature of several spatial data handling fields under different names (Samet 1989), and descriptions of algorithms for finding Thiessen polygons are spread over a wide range of journals. At the same time there is a growing sense in GIS research that our emphasis has changed, as more and more of the underlying technical problems of GIS are solved. Attention has moved from primitive algorithms and data structures to the much more complex problems of database design, and the issues surrounding the use of GIS technology in real applications. The following sections identify some of these key issues.

3.1. Data collection and measurement

If spatial reality is continuous and subject to complex structures of spatial dependence, then how should it be compiled and measured? More generally, how do people perceive the real world of geographical variation, structure it and learn about it? Although many of these questions are part of the research agendas of remote sensing, photogrammetry, geodesy and cognitive psychology, the lines of demarcation are far from distinct. Should GIS or remote sensing concern the problems of transferring information from one technology to the other, and more importantly making good sense of it? Is it a GIS or remote sensing if ancillary geographical information is used to improve the accuracy of classification or if an image is used to update a GIS layer? Ultimately it matters little to which of the many pigeon holes we assign each topic. There are undoubtedly substantial scientific questions here, which require a depth of understanding of the nature of spatial variation, and one person's remote sensing may well be another's geographical information science.

The process of discretization, with its implied generalization, abstraction and approximation, takes place as data are collected, interpreted or compiled, and choices are made at this stage that affect the ultimate uses of the data. When those uses change, as they have been doing with the widespread use of GIS, it may be necessary or beneficial to rethink the process of data collection. For example, with digital management and delivery of census data, is it still appropriate to conduct a census on a decennial basis? Is the traditional approach to geological field mapping the most appropriate if the eventual objective is a digital three-dimensional representation of the subsurface? How will topographic mapping change now that it is cost-effective to survey new features using the Global Positioning System? Geographical data collection is often the domain of specialists in well established disciplines, so it may be many years before these kinds of questions are investigated or answered. To date the introduction of GIS seems to have had very little effect on the process of data collection.

3.2. Data capture

Enormous strides have been made in the technology for capturing digital geographical data in the past decade, and the systems now on the market are capable of a high level of intelligence in interpreting scanned map documents. The problem remains the poor quality of the documents, and the ambiguities that are caused by aspects of map design. As a result, manual digitizing remains a widely used approach, despite its high cost, tedium, and failure to show significant improvements in efficiency. Two trends may change this situation substantially in the next few years. One is the increasing avoidance of the map document as a step in the data compilation and input process. Surveying and photogrammetry are moving away from compilation using paper maps, and the more interpretive fields such as land use, vegetation or soil mapping are likely to follow suit. The digital total station is likely to be followed by the digital plane table and perhaps even the digital field geology notebook. The other is the long recognized possibility that comparatively minor changes in a map's design can make it vastly easier to scan and interpret (Shiryaev 1987).

3.3. Spatial statistics

As spatial data are always an approximation or generalization of reality, they are full of uncertainty and inaccuracy. A change of data model or scale can introduce a loss of information, as can digitizing or scanning. Processing in a finite machine also inserts its own form of uncertainty, although this is often insignificant in relation to the errors inherent in the data themselves. Many human geographical constructs are implicitly uncertain, including spatial objects ('Indian Ocean', 'Europe') and their relationships ('in', 'across'). Whether we think of uncertainly in set theoretical terms through notions of fuzziness or in statistical terms through the calculus of probabilities, the study of spatial data uncertainty, its measurement and modeling, and the analysis of its propagation through the processes of spatial data handling are undoubtedly part of geographical information science. How should one compile an accurate representation of geographical variation for input to a database? How should one represent the uncertainty or inaccuracy present in a digital representation? How can uncertainty be propagated from database to GIS products?

Geographical data bring their own special set of problems to spatial statistics. Whereas in medical imaging the problem may be to determine the true location of objects from 'dirty' pictures (Besag 1986), in geographical images there is often no clear concept of truth, as objects are often the products of interpretation or generalization. We need much better methods of measuring and describing uncertainty, particularly in the complex spatial objects common in GIS. We need better methods for dealing with the world as a set of overlapping continua, instead of forcing the world into the mould of rigidly bounded objects. Most of the answers to these questions will have to come from spatial statistics, but geographical information specialists must provide the motivation and the examples, and define the overall objectives and constraints.

Although all geographical data are uncertain to some degree, all of the current generation of GIS follow the common practice in cartography and represent geographical objects as if their positions and attributes were perfectly known; data quality may or may not be addressed in a separate statement. The consequences of uncertainty for GIS products are never estimated. Recent research has followed several different and productive lines in attempting to address the problem of data quality. One is to match precision to accuracy. In a locational sense, this means using limited precision in data representation and processing, most often through the use of a raster whose size is determined by data accuracy. Various forms of quadtree structure have also been used to fit locational precision to known levels of accuracy. There have been several recent papers on finite resolution processing in GIS (e.g. Franklin 1984, Dutton 1989) and finite resolution geometry is an active research area in mathematics.

Another productive approach has been to incorporate techniques from geostatistics, notably kriging, as the statistical basis of these techniques makes uncertainty explicit. We now have several useful models of digitizing error, and its consequences for estimated easures such as area (e.g. Chrisman and Yandell 1988, Keefer et al. 1988). Finally, there have been several successful efforts to model geographical data sets as random fields, or derivatives of random fields, and to use this approach to model uncertainty in GIS objects (e.g. Goodchild 1989). Between all of these methods, we probably now have an adequate set of models of accuracy from which to build an error-tracking GIS. However, spatial statistics are not an easy field, and many of these techniques go well beyond elementary statistics in their conceptual sophistication.

3.4. Data modeling and theories of spatial data

Data models are the logical frameworks which we use to represent geographical variation in digital databases. As each must be an approximation, the choice between alternative models constrains not only the functions available, but also the accuracy of products. Of all the developments in GIS in the past decade, perhaps the most exciting has been the proliferation of data models, and the growing literature on their relative merits. The debate over raster and vector goes back to the earliest days, but has now been joined by debates over objects, layers, the philosophy of object orientation, hierarchical models of complex objects, and the entire range of possibilities inherent in time dependence and three dimensions. Despite the interest, we still do not have a complete and rigorous framework for geographical data modeling, even in the static two-dimensional case, and without one it is difficult to see how GIS can escape the constraints imposed by specific system implementations. How much capability is being lost by forcing contemporary applications into the multilayer raster model used by many systems, or the point-line-area coverage model used by many others? This is both a pure and an applied research problem. On the one hand, we must develop a comprehensive framework for geographical data modeling, with an associated terminology, to provide the basis for standards and an ideal against which specific systems can be measured. On the other hand, an abstract framework is of little value if it does not influence practice, through implementation in the products of the vendors. Here the real issue is whether it is possible to enlarge or 'retrofit' the data model underlying an existing product, or whether any attempt to do so is doomed to cause inconsistency and incoherence.

These issues are precipitating lively discussion over the entire question of the degree to which we view, analyse, represent and model the world as discrete or continuous, as a collection of objects or a set of fields. Do we think in terms of variables with defined values everywhere in space, or of an empty space littered with possibly overlapping objects? In essence, these issues have brought the GIS debate from the comparative obscurity of internal data structures to the much more general issues of how we understand geographical variation. Everyday human experience sees a world of objects, but the science of natural processes deals more with continuous variation (Frank and Mark 1991). Thus the object oriented debate threatens to pit the New Agers against the embattled remnants of the Enlightenment, and what could be more stimulating than that?

3.5. Data structures, algorithms and processes

Many of the results of basic research which have accumulated over the past 25 years in this field of research concern internal representations of data, and the algorithms which operate on them. The quadtree (Samet 1989), band sweep algorithms for overlay (White 1977), analysis of computational complexity (Preparata and Shamos 1988) and the arc-node data structure (Peucker and Chrisman 1975) are all intellectual breakthroughs of lasting significance. Many challenging problems remain, for example in the design of efficient algorithms to minimize overposting and in other areas of cartographic design, or in developing better methods for converting between various terrain data models. Many systems now handle data through database management systems, and data structure issues have moved more and more into the realm of computer science. We seem, however, to have reached a point where all of the simpler, more generic problems have been solved, and where what remains is a set of difficult, context-specific problems. It seem clear, for example, that further advances in the conversion of terrain data models (for example, from contour to TIN) will require a much better understanding of the nature of terrain (Mark 1979), and will perhaps have to be specific to terrain type (e.g. fluvial versus glacial). There will also continue to be a need for research on efficient methods of storage and access to deal with the enormous volumes of data likely to become available in the coming decade.

3.6. Display

Geographical information systems have often been criticized for failing to give adequate attention to principles of cartographic design (Buttenfield and Mackaness 1991), or for regarding the map as a simple store of information rather than a tool for communication. If we think of the database as the truth, then a map is no more than a store, as there is often a simple correspondence between objects in the database and objects on the map. However, if the database is seen merely as an approximation of the geographical truth, then the design of output displays is critical, as it can affect the user's view of the world. Such simple things as the choice of background color, or the contrast between adjacent polygons (McGranaghan 1991) can have a significant effect.

The capabilities of electronic display go far beyond those of conventional cartography. We need research on the design of animated displays, three dimensional display, the use of icons and metaphors in user interfaces, continuous gradation of color and tone, zoom and browse, multiple media including voice and pointing devices, multiple windows which allow simultaneous access to spatial and temporal series of multivariate data. We need to use the electronic medium to think far beyond improvements to the design of choropleth maps. All of these are fundamental problems to a science of geographical information.

3.7. Analytical tools

A GIS is a tool for supporting a wide range of techniques of spatial analysis, including processes to create new classes of spatial objects, to analyse the locations and attributes of objects, and to model using multiple classes of objects and the relationships between them. It includes primitive geometric operations such as calculating the centroids of polygons, or building buffers around lines, as well as more complex operations such as determining the shortest path through a network. The functionality of leading products continues to grow, with no obvious end in sight.

Despite widespread recognition that analysis is central to the purpose of a GIS, the lack of integration of GIS and spatial analysis, and the comparative simplicity of the analytical functionality of many systems continues to be a major concern. In the early days of the statistical package SAS, there was a very rapid increase in the range of tests and techniques implemented in the system. Unfortunately, the same has not been true of GIS, and remarkably little progress has been made in incorporating the range of known techniques of spatial analysis into current products.

There are many reasons for this. One obvious reason is the heavy emphasis in the GIS marketplace on information management rather than analysis. The lucrative markets for GIS technology have comparatively unsophisticated needs, emphasizing simple queries and tabulations. Another is the relative obscurity of spatial analysis, a set of techniques developed in a variety of disciplines, without any clear system of codification or strong conceptual or theoretical framework. Even now it is difficult to identify more than a handful of texts (e.g. Haining 1990, Upton and Fingleton 1985). Although one might expect that GIS could provide the basis for a system of codification for spatial analysis, the poor level of current understanding of geographical data models is a major difficulty. Tomlin (1990) has made one of the few attempts to add some sort of structure or framework to the proliferation of GIS functions, which in the case of ARC/INFO is already around 103. We badly need a taxonomy of spatial analysis, developed perhaps from an enumerated set of data models, but going well beyond the primitive geometrical operations.

At this stage, integration of GIS and spatial analysis is proceeding slowly, in at least three different modes. Some analytical capabilities are being added directly to GIS, for example in the recent expansion of functionality in several modules for network analysis. Some progress is being made in loosely coupled analysis, where an independent analysis module relies on a GIS for its input data, and for such functions as display. However, still missing is an effective form of tight coupling, in which data could be passed between a GIS and a spatial analysis module without loss of higher structures, such as topology, object identity, metadata, or various kinds of relationships. At present this is impossible, to a large extent because of a lack of standards for data models. Instead, coupling has to occur at a lower level, and higher structures have to be rebuilt on an arbitrary basis.

Integration between GIS and spatial analysis might also take the form of a language, whose primitive elements would represent the fundamental operations of spatial analysis. The beginnings of such a language already exist in the macro languages of many of the current generation of GIS, and in various attempts to extend SQL to spatial operations. However, all of these are specific to, and heavily dependent on limited data models, and there is remarkably little similarity between them at this time. At Santa Barbara we have been attempting to define a common language from an analysis of the languages used by a variety of current GIS, but a more satisfactory solution would begin with the conceptual framework provided by a comprehensive data model. Another problem in integrating GIS and spatial analysis is that in the former discretization of space is explicit, whereas in many forms of spatial analysis it is often either implicit, or unspecified. Many forms of spatial analysis are written on continuous fields, and fail to deal with the uncertainties introduced by the inevitable process of discretization. For example, in GIS there can be no measure of slope that is independent of discretization, and similarly the length of an area object's boundary is dependent on its digital representation. However, slope and length commonly appear as unqualified parameters in spatial models. In this sense, the integration of GIS and spatial analysis is a two-way process, in which the inadequacies of both GIS and spatial analysis must be addressed.

Most of the current generation of GIS provide some sort of macro or script facility, allowing the user to define products from complex sequences of operations, but to invoke them with a single instruction. Although these often include the ability to construct customized environments and interfaces, they do not as yet provide tools which are specific to the needs of spatial analysis. One limited exception is Prime/Wild's ATB, a set of tools constructed on top of System/9 which allows the user to work with complex analyses, visualize their sequences and manage intermediate results. Tools like this will be needed increasingly if GIS are to move into an era of more sophisticated analysis and decision support, because it is not uncommon for relatively simple GIS products to involve processing tens of layers through similar numbers of primitive steps. We need to research methods for keeping track of data lineage and error propagation, backtracking to recover intermediate results, and preventing the user from combining operations in incorrect or meaningless ways (Lanter 1990). We also need research on ways of incorporating this sort of analysis into the GIS acquisition and planning process.

This emphasis on complex multistage analysis and the generation of products from a multilayered database seems very different from research on knowledge based systems, spatial reasoning and spatial query. One of the attractions of the GIS field is its breadth of applications, and the correspondingly extreme variety of environments for the design of user interfaces. In data modeling, the important question is not whether extended relational or object oriented models are better for geographical data, but what types of geographical data are best modeled by each approach. Similarly, the important research issue in the design of user interfaces is to determine the optimal environment for each of the many types of GIS application. What is best for a vehicle navigation system may be entirely different from what is best for a forest resource manager with a deeply seated fear of keyboards and VDUs, either color or monochrome.

3.8. Institutional, managerial and ethical issues

Research is just beginning to appear on the issues involved in implementing and managing GIS, especially in large institutions. This is difficult research, and generalizations are not discovered easily. However, the success of several large projects in the U.S.A., and the discussions surrounding several large acquisitions by federal agencies, have created the opportunity for a number of useful case studies. Many more are needed, particularly given the importance of such research for improving the institutional environment in the future. We need a much better understanding of the processes of adoption of GIS technology and its effects on organizations; of the value of geographical information and the benefits .of GIS; and of processes for utilizing geographical information in decision making. Theoretical frameworks for addressing many of these issues already exist in the relevant social science disciplines, and we need to make much more effective use of them in tackling the specific issues of GIS.

Despite the problems involved in adopting any new technology, GIS have been widely adopted in local government, utilities and resource management agencies. In fact, the introduction of GIS has had a major effect on the management of geographical information in society. At the same time there is increasing concern over the power of GIS for surveillance and invasion of privacy. The research community has a responsibility to monitor and study the more substantive aspects of the GIS phenomenon, including its significance to society as a whole. What will GIS mean to the balance of power in society? Will they be a technology available only to the empowered, or will they somehow serve to even the distribution of power? Thus far there have been remarkably few studies of the ethics of GIS.

4. Tests of commonality

The preceding sections have looked at various candidate areas for inclusion in a geographical information science. In each case there are clearly challenging scientific questions to be posed and researched. There is no reason to believe that the list is complete, or that there are not additional and substantive questions in other related areas. In each case the spatial context appears to be distinctive, although clearly it is more so in some than others. For example, we might debate whether the spatial context was distinctive in the area of decision theory, but the issue seems clear-cut for data modeling.

In the NCGIA research plan (NCGIA 1989), we argued that the absence of solutions to issues such as these constituted impediments to the effective applications of GIS technology. Other discussions of the GIS research agenda have come to similar conclusions, although with different emphases (Craig 1989, Maguire 1990, Masser 1990). Many are old issues, recognized long before the advent of GIS in fields such as cartography, geodesy and geography. Some may not be unique to GIS. For example, it is not immediately obvious that GIS technology diffuses in a fundamentally different fashion, or shows fundamentally different patterns of adoption from other technologies. Is the measurement of GIS benefits a unique problem, or an example of the more general problem of measuring the benefits of information technology? Of course these questions are in themselves research issues.

At the same time it is very important to identify those areas where GIS have created new and unique issues that are not common to other fields. In the early days of GIS, it was possible to argue that the technology was filling an existing gap, and making possible tasks that had been previously identified, but that were not easy to carry out manually. The use of GIS for suitability analysis, by overlaying layers (Tomlin 1990), mirrors the manual technique popularized by McHarg, although admittedly adding some interesting new capabilities. CGIS was justified on the grounds that the computer was a cost effective alternative to hand measurement of overlaid areas. But GIS make it possible to do things with data that the data's gatherers may never have envisioned. GIS technology is producing radical changes in the way geographical data are collected, handled and analyzed, and it will be many years before the impact of existing technology is felt, let alone the impacts of future developments.

Here are some of the issues that seem unique to GIS: how, to model time-dependent geographical data; how to capture, store and process three-dimensional geographical data; how to model data for geographical distributions draped over surfaces embedded in three dimensions; how to explore such data, for example, what exploratory metaphors are useful; and how to evaluate the geographical perspective on information and processes relative to more conventional perspectives?

These are important issues for GIS, and the GIS community needs a strong commitment to research if it is going to make significant progress on them. As issues that arise within the context of GIS, they are not of major concern in other disciplines. However, at the same time the GIS community can benefit enormously from interdisciplinary research. Statisticians can make a very valuable contribution to solving the error problem in GIS, and research in cognitive psychology may be helpful in designing the cognitive aspects of user interfaces in GIS.

This argument leads naturally to a proposed definition of GIS research: research on the generic issues that surround the use of GIS technology, impede its successful implementation, or emerge from an understanding of its potential capabilities. Is this 'research about GIS' or 'research with GIS'? In a sense it is both, because these are issues that are both fundamental to the technology of GIS, and also issues that must be solved before the technology can be successfully applied. If the problems of doing research with GIS are generic, then they are best tackled as part of the GIS research agenda.

However, problems that are specific to the application of GIS in a particular field clearly need to be addressed in the context of that field, and with the benefit of its expertise. Accuracy issues provide a useful example. There are aspects of the accuracy problem that span a wide range of types of geographical data, and need to be solved using generic models of uncertainty, analogous to the role played by the Gaussian distribution in the theory of measurement error. However as noted earlier, an analysis of crime data using a GIS will also raise problems of accuracy that are specific to that particular application, and need an understanding of the processes operating in criminology and in the collection of crime data if they are to be understood fully. However, mere existence of scientific questions is far from an adequate basis for a science. Is there a commonality of interest here? Can these subfields find sufficient basis for interaction that they will develop the lasting accoutrements of a science, such as journals, societies, books and philosophers? Will researchers in these subfields behave as a group of scholars? Is there a valid analogy between the systems and science of geographical information on the one hand (tools supporting researchers) and statistical packages and statistics on the other? Statistics is a highly formalized discipline, but more technologically oriented groups can be found in such areas as exploratory data analysis, statistical visualization and applied statistics. Certainly the relationship between science and tools is stormy at times, but nevertheless vital to the success of both. The ongoing debate over the value of statistical software in teaching statistics has interesting implications for the same issue in GIS.

It may be useful to look briefly at the arguments for a commonality of interest in geographical information science, first in principle and then in practice. The field is small—rhetoric about growth in the industry aside, no one would suggest that the field of GIS is a major discipline. It is distinct, with its own reasonably unique set of questions. And it is certainly challenging and innately appealing. On the negative side, it is multidisciplinary, competing with longstanding cleavages and rivalries. It lacks a core discipline, unlike the statistical analogy, where there has been a steady growth in the number and size of academic departments for the past few decades. One of the claimants to the core, geography, has traditionally been a non-technical field, and in some areas of social geography there is a strong and fundamental antipathy to technological approaches.

In practice, commonality of interest is evident in the proliferation of GIS meetings, and we are beginning to see a supply of books and journals. However, the scientific track at GIS meetings is often small. People who attend GIS meetings need a constant supply of novelty, whether in scientific research or vendor products, and will soon desert if the supply dries up.

5. Options for the future

Looking back over nearly three decades of GIS research, it is clear that the greatest progress has been made on the best defined and easiest problems, where solutions lay in advances in the technology itself. Rapid progress was made on algorithms and data structures in the 1970s and 1986s, but many of the difficult problems of data modeling, error modeling, integration of spatial analysis and institutional and managerial issues remain. Some of these may be unsolvable: for example, there may simply be no generalities to be discovered in the process of adoption of GIS by government agencies, however easy it may be to pose the research question.

Other issues have already been solved in a pure research sense, but implementation remains a major question of applied research. In accuracy, for example, a substantial set of techniques has been defined, but the problem of moving them into actual application remains. The academic research environment is set up to pursue significant areas of research, but is generally poor at providing the means of implementation. For that we need a software industry that is tightly coupled to the research community, but able to find the resources to motivate development. More importantly, we need an education system that responds rapidly to new research and is able to build new concepts quickly into its programmes. Unfortunately, the higher education sector is too often characterized by conservatism, and it may take many years for new ideas to work themselves into the curriculum.

Research in GIS is like geographical data—the more closely one looks, the more interesting issues appear. GIS research has only begun to tackle the important issues in the research agenda. We are in an enviable position, working in a field with such strong motivation and such a strong underlying industry, and with such an interesting set of problems spanning so many disciplines and fields.

I hope I have shown in this paper that the handling of spatial information with GIS technology presents a range of intellectual and scientific challenges of much greater breadth than the phrase 'spatial data handling' implies—in effect, a geographical information science. The term 'geographical' seems essential—much of what GIS research is about concerns the geographical world and our relationships with it, and the term is much richer than 'spatial'. The change in meaning of the 'S' word—from systems to science—seems to be going well, as evidenced by the success of the spatial data handling series of conferences, the move of the AutoCarto series to fully refereed papers, the new texts, subscriptions to the International Journal of Geographical Information Systems, and submissions of GIS papers to such established journals as Geographical Analysis, Computers and Geosciences, Computer Vision, Graphics and Image Processing and publications of the Regional Science Association and the IEEE.

I hope I have also shown that a strong scientific programme serves not only itself, but also the needs of industry and GIS users. GIS need a strong scientific and intellectual component if they are to be any more than a commercial phenomenon, a short-lived flash in the technological pan. It is too easy to see current GIS as a hardware and software technology in search of applications, and to see the field of GIS as defined by the functional limits of its major vendor products. We need to move from system to science, to establish GIS as the intersection between a group of disciplines with common interests, supported by a toolbox of technology, and in turn supporting the technology through its basic research. As currently perceived, GIS sometimes seem about as close to a science as FORTRAN is to algebra.

In recent years we have seen a growing cleavage in GIS between two traditions, that of spatial information on the one hand and that of spatial analysis on the other. The spatial information tradition stresses large inventory databases, and gives geography the role of an access mechanism. The spatial analysis tradition stresses rich functionality and a range of data models, and gives geography a fundamental role in analysis and modeling. The two traditions share common data structures and algorithms, and rely on the same sources of data and hardware. However, this is not enough to convince the academy of the existence of a scientific field. To claim this we need to take a broader view, and to include data modeling accuracy, cognition, reasoning, human-computer interfaces (HCI) and visualization, and to show how these are integral parts of both traditions.

Without such arguments, the GIS field will fragment, and the GIS storm will blow itself out. Associations as fundamentally disjoint as the Association of American Geographers and AM/FM will find it impossible to justify joint sponsorship of conferences. Vendors will specialize in data input workstations, spatial analysis workstations or facility management systems, with little potential for interaction or integration. This would be tragic.

How can we ensure a lasting future for both geographical information systems and science? Disciplines are like tribes, with their own totems, symbols and membership rules, languages and social networks. The GIS tribe is currently very cohesive; it is well funded, the field is exciting and much useful research is being done. However, in the longer term the field has not done well at behaving as a science, and the academy is still doubtful about whether it needs to be taken seriously. Science is hard and places heavy obligations on its practitioners. We have been too busy, and technology has been moving too quickly. Too much of our literature is in conference proceedings, which bring fast exposure but only to limited audiences, and lack sufficient quality control. Few people have had the time to write the textbooks or to identify the intellectual core, or to publish the good examples.

I believe we ensure the future of GIS by thinking about science rather than systems, and by identifying the key scientific questions of the field and realizing their intellectual breadth. Geographical information systems are a tool for geographical information science, which will in turn lead to their eventual improvement. We need to speak to the academy, both directly and through key articles and texts, on the philosophy, methodology and foundations of the field, and by placing GIS papers in strong journals. All three communities—users, vendors and researchers—have vita! And symbiotic roles to play, and we will serve all three best by playing ours in the fullest possible sense.

Acknowledgments

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

References

ANSELIN, L., 1989, What is special about spatial data? Alternative perspectives on spatial data analysis. Technical Report 89-4 (Santa Barbara, CA: National Center for Geographic Information and Analysis).

BERRY, J. K., 1987, Fundamental operations in computer-assisted map analysis. International Journal of Geographical Information Systems, 1, 119-136.

BESAG, J., 1986, On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society, B48, 259-302.

BUTTENFIELD, B. P., and MACKANESS, W. A., 1991, Visualization. In Geographical Information Systems: Principle and Applications edited by D. J. Maguire, M. F. Goodchild and D. W. Rhind (London: Longman).

CHRISMAN, N. R., and YANDELL, B., 1988, A model for the variance in area. Surveying and Mapping, 48, 241-246.

CRAIG, W. J., 1989, URISA's research agenda and the NCGIA. Journal of the Urban and Regional Information Systems Association, 1, 7-16.

DUTTON, G., 1989, Modeling locational uncertainty via hierarchical tesselation. In Accuracy of Spatial Databases, edited by M. F. Goodchild and S. Gopal (London: Taylor & Francis), pp. 125-140.

FRANK, A. U., and MARK, D. M., 1991, Language issues for GIS. In Geographical Information Systems: Principles and Applications, edited by D. J. Maguire, M. F. Goodchild and D. W. Rhind (London: Longman).

FRANKLIN, W. R., 1984, Cartographic errors symptomatic of underlying algebra problems. Proceedings, International Symposium on Spatial Data Handling, Zurich, pp. 190-208.

GOODCHILD, M. F., 1989, Modeling error in objects and fields. In Accuracy of Spatial Databases, edited by M. F. Goodchild and S. Gopal (London: Taylor & Francis), pp. 107-114.

GOODCHILD, M. F., 1990, Keynote address: spatial information science. Proceedings, Fourth International Symposium on Spatial Data Handling, Zurich, I, 13-14.

GOODCHILD, M. F., 1991, Keynote address: progress on the GIS research agenda. Proceedings, EGIS 91, Brussels, pp. 342-350.

HAINING, R., 1990, Spatial Data Analysis in the Social and Environmental Sciences (Cambridge: Cambridge University Press).

KEEPER, B. J., SMITH, J. L., and GREGOIRE, T. G., 1988, Simulating manual digitizing error with statistical models. Proceedings GIS/LIS 88 (Falls Church, VA: American Society of Photogrammetry and Remote Sensing/American Congress on Surveying and Mapping), pp. 475^83.

LANTER, D. P., 1990, Lineage in GIS: the problem and a solution. Technical Paper 90-6 (Santa Barbara, CA: National Center for Geographic Information and Analysis).
MAGUIRE, D. J., 1990, A research plan for GIS in the 1990s. The Association for Geographic Information Yearbook 1990, (London, Taylor & Francis), pp. 267-277.
MARK, D. M., 1979, Phenomenon-based structuring and digital terrain modeling. GeoProcessing, 1, 27-36.
MASSER, I., 1990, The Regional Research Laboratory initiative: an update. The Association Jor Geographic Information Yearbook 1990 (London: Taylor & Francis), pp. 259-263.
MCGRANAGHAN, M., 1991, Modeling simultaneous contrast on choropleth maps. Technical Papers, 1991, ACSM/ASPRS/Auto-Carto 10 Annual Convention, Baltimore, MD, March 25-29, 1991, Vol. 2, pp. 231-240.
MCHARG, I. L., 1969, Design with Nature (New York: Doubleday).
NCGIA (National Center for Geographic Information and Analysis, 1989, The research plan of the National Center for Geographic Information and Analysis) International Journal of Geographical Information Systems, 3, 117-136.
OPENSHAW, S., 1977, A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modeling. Transactions of the Institute of British Geographers, 1 (NS), 459^72.
PEUCKER, T. K., and CHRISMAN, N. R., 1975, Cartographic data structures. American Cartographer, 2, 55-69.
PREPARATA, F. P., and SHAMOS, M. I., 1988, Computational Geometry: An Introduction (New York: Springer-Verlag).
SAMET, H., 1989, The Design and Analysis of Spatial Data Structures (Reading, MA: Addison-Wesley).
SHIRYAEV, E. E., 1987, Computers and the Representation of Geographical Data (New York: Wiley).
TOMLIN, C. D., 1990, Geographic Information Systems and Cartographic Modeling (Englewood Cliffs, New Jersey: Prentice Hall).
TOMLINSON, R. F., 1984, Keynote address: geographical information systems—a new frontier. Proceedings, International Symposium on Spatial Data Handling, Zurich, 1, 2-3.
TOMLINSON, R. F., CALKINS, H. W., and MARBLE, D. F., 1976, Computer Handling of Geographical Data (Paris: UNESCO).
UPTON, G. J., and FINGLETON, B., 1985, Spatial Data Analysis by Example (2 volumes) (New York: Wiley).
WHITE, D., 1977, A new method of polygon overlay. Proceedings, Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Harvard.