Progress in Human Geography 15,2 (1991) pp. 194–200

Geographic information systems

Michael F. Goodchild

National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA 93106, USA

I Geographic information science

In his keynote address to the Second International Symposium on Spatial Data Handling in Seattle in 1986, Jack Dangermond, President of the Environmental Systems Research Institute, commented that: 'GIS is clearly not a new science, but rather a technology which requires a considerable scientific knowledge base for many of its data management functions'. As the GIS industry expands, and as geography programmes introduce more and more courses in GIS, the meaning of Dangermond's 'considerable scientific knowledge base' is becoming increasingly clear. Effective use of the technologies of spatial data handling – GIS, remote sensing, automated cartography, photogrammetry – requires a sound understanding of the nature of geographic information, and the issues involved in compiling, storing and using it. As the technical problems of handling spatial data have been solved over the past years, the GIS research community has moved more and more to the generic issues surrounding the technology. This is in some ways a practically motivated trend – good research is badly needed on the impediments to better and more effective use of GIS technology. On the other hand many of the generic issues of geographic information are challenging intellectual problems in their own right, with a history extending back well before the advent of digital technology. For example, recent work on digital generalization (Buttenfield 1989; McMaster 1989; Carstensen 1989) is merely revitalizing a longstanding issue in cartography – the GIS context gives it new motivation, and the greater breadth that derives from the capabilities of the new technology.

The various GIS research agendas that have appeared in print in the past year or so are all implicit attempts to define the domain of Geographic Information Science. The National Center for Geographic Information and Analysis agenda (NCGIA 1989) is based in part on the five research areas identified in the National Science Foundation's solicitation (Abler 1987): spatial analysis and spatial statistics; theories of spatial relations; artificial intelligence and expert systems; visualization; and social, institutional and economic issues. It argues that the application of GIS technology, particularly in support of scientific research, is impeded by a number of factors, ranging from problems with the technology itself to a lack of understanding of its impacts on organizations. The Urban and Regional Information Systems Association agenda (Craig 1989) overlaps substantially, but places greater emphasis on the social, institutional and economic issues of managing GIS applications. In the UK, Natural Environment Research Council (NERC) and Economic and Social Research Council (ESRC) have funded a joint research programme aimed again at generic issues in geographic data handling, and ESRC's Regional Research Laboratories are committed to a similar agenda (Shepherd *et al.*, 1989).

Goodchild (1990) lists nine research areas for Geographic Information Science, each with its own set of intellectual and scientific questions: data collection and measurement; spatial statistics; data modelling; data structures and indexes; algorithms and processes; display of spatial data; analytic tools; decision theory and risk analysis; and reasoning and cognition. Only in decision theory and risk analysis does there seem to be much doubt that the spatial context is distinct – is the theory of spatial decisions really different from any other kind of decision theory? In the other eight geography clearly provides a distinct set of issues.

Among the remaining eight topics, geography as a discipline seems to have most to say about data modelling, analytic tools and reasoning and cognition, and cartography to be most closely associated with display. Data modelling deals with the question of how the infinite complexity of the geographical world can be represented within a discrete, finite machine, and is emerging as perhaps the most significant issue in GIS. The selection of an appropriate data model requires a deep understanding of the nature of geographical phenomena and the processes which formed them. The available data models for topographic elevation, to take a simple example, include DEMs, TINs and contours. In the past the choice has often been constrained by immature technology, but GIS has now progressed to the point where data modelling can be based on knowledge of the phenomenon being modelled, and where choices clearly reflect that knowledge. Many of these issues emerged at a conference on GIS Design Models and Functionality held at the University of Leicester in March 1990 (selected papers from the meeting will appear in *Computers and Geosciences*; for early reviews see *GIS World* June/July 1990).

II Standards

As the GIS field matures, there are increasing calls for standardization, to allow greater interchange of data between systems, and to reduce confusion over terminology. Several terms are notorious for their ambiguity: 'layer', 'segment', 'data structure', 'topological', 'object-oriented'. In part the confusion arises because the field is multidisciplinary. 'Topology' is well defined in mathematics, but a GIS professional without mathematical training will more often guess an intuitive meaning than seek out a formal one.

A long-term effort sponsored by the US Geological Survey (USGS) to develop a national cartographic data standard is finally coming to fruition with the submission of the standard to the National Institute of Standards and Technology (NIST) (the draft standard was published as DCDSTF 1988). It attempts to standardize both terminology, for real geographical entities and for their digital representations, and data exchange formats. From a GIS perspective, however, the standard remains incomplete, as it deals primarily with standards for creating digital databases from maps – cartographic standards – rather than from geographical reality. The question of data modelling is dealt with only in the map context, and is thus restricted to the data models used on maps. Map data models must be compatible with the pen, that is, with a technology capable of drawing lines of constant width (Goodchild, 1988), and are designed to communicate through the process of visual perception. GIS, on the other hand, is much less restricted in its data models, and is more concerned with precise analysis than with visual perception. GIS data models remain diverse, largely incompatible and incomplete. Although

Intergraph's TIGRIS, IBM's GFIS and ESRI's ARC/INFO have all been successful in the public utility market, they require entirely different approaches to data modelling, implying three very different views of geographical reality. Thus we are a long way from any standard of data modelling. Unfortunately the fact that all three companies have significant shares of the market will likely work against any emergence of a data model standard in the near future. At this time the efforts of such organizations as USGS, FICCDC (Federal Interagency Coordinating Committee on Digital Cartography), ASTM (American Society for Testing Materials), NIST and NCGA (National Computer Graphics Association) are understandably more concerned with format standards (for a recent review of the US standards arena see *GIS World* April/May 1990).

III Accuracy

A GIS is a precise machine, working typically to seven or fourteen digits of coordinate precision. When a map showing population density by county or areas of soil class is input, it is easy to assume that these attributes apply homogeneously and precisely to the areas with which they are associated. In a common type of GIS application, three or four such layers might be overlaid and used to compute an index, perhaps of suitability for construction of some facility. Ground checks of the truth of the composite layer – that a selected point does indeed have all of the characteristics ascribed to it by the various layers – are almost always disappointing. The problem, of course, is that the source map is inaccurate – the areas of soil class are not in fact homogeneous, and are not bounded by lines of sharp transition. Boundaries on soil maps have been called 'lines that do not exist surrounding places that have nothing in common'. To the cartographer, the inaccuracy is partly unavoidable – with pen and paper it is difficult to show the continuous variation of soils in any other way – and partly irrelevant – the map is intended to give a visual impression of spatial variation, not an exact inventory. Cartographic quality and GIS accuracy are very different issues.

Ideally, we would like every product of a GIS to be accompanied by confidence limits, based on knowledge of the uncertainty present in the database and the processes of error propagation in each GIS operation. This idea lay behind the decision to focus the first research initiative of the NCGIA on Accuracy of Spatial Databases. The papers presented at the initial meeting in December 1988 have appeared as a book (Goodchild and Gopal, 1989) and cover many dimensions of the accuracy issue in GIS. Effects of aggregation, disaggregation and modifiable areal units are one area of particular concern in human geography, and GIS provides the basis for a much more comprehensive approach (Flowerdew and Green 1989; Arbia 1988; 1989).

If accuracy is a problem in spatial data handling because of the high precision of the handling system, then one rational approach would be to reduce precision to match accuracy. In effect this is what happens in a raster system. Dutton (1984; 1989; see also Goodchild and Yang, 1989) makes this point, and proposes a tesseral scheme for global data that would replace common co-ordinate referencing with a finite-resolution hierarchical key.

There has been substantial progress in recent years in understanding the propagation of uncertainty that occurs in spatial data handling. Newcomer and Szagin (1984) and Veregin (1989) have discussed error propagation during simple Boolean overlay; Lodwick (1989; Lodwick and Monson, 1988) and Heuvelink *et al* (1989) have analysed the sensitivity of weighted overlay results to uncertainty in the input layers and weights; and Arbia and Haining (1990) have analysed a general model of uncertainty in raster data. However overlay is perhaps the most simple of the primitive spatial data handling operations. Much more research is needed on the effects of buffering (dilating) uncertain objects, and of changing from one data model to another, e.g., raster/vector conversion.

If traditional map data models tend to give a misleading impression of the reliability of data, it is not surprising that many cartographic products fail to make the user aware of uncertainties. The move to standards of data quality (DCDSTF, 1988) is a significant step toward greater awareness of the reliability of spatial data, and includes concepts of lineage, consistency and completeness as well as the more statistical issues discussed here. It is particularly important as GIS users continue to apply spatial data to purposes for which they may not have been designed.

IV New applications

With a booming industry behind it, it is not surprising that GIS continues to find new areas of application, particularly in the social sciences. Beaumont (1988), Achabal and McIntyre (1987) and Goodchild (1989) discuss the status of GIS applications in marketing. Twigg (1990) looks at the problems of applying GIS technology to health data, while a forthcoming collection edited by Zubrow, Allen and Green (1990) will review applications in archaeology and anthropology.

In planning, Kim, Wiggins and Wright (1990) include several chapters on GIS applications, while French and Wiggins (1988) and French, Wiggins and Heffernon (1989) have analysed the spread of GIS technology through the profession. Couclelis (1989) discusses the use of GIS at various levels of abstraction in the planning process. She finds that the current generation of GIS assumes a container-like view of space, which is appropriate to the lower levels of physical planning, but that the more relative, interaction-based views of space necessary at higher levels of abstraction are largely missing.

V New data models

The data models embedded in the current generation of GIS are largely map-based, designed to allow information from maps to be captured easily into a digital database. The vestor model sees the world as a collection of point, line and area objects, and defines their locations and attributes, while the raster model defines what is present in every one of an array of finite elements. Both have abundant analogues in vision and graphics. But both models are seriously deficient as bases for describing real geographical variation, as distinct from the contents of maps. They are static, but the world is dynamic. They are two-dimensional. And they fail to deal with the human propensity for giving objects identities that are independent of their physical manifestation at any particular scale.

The modelling of time-dependent spatial data has been discussed at length by Langran and Chrisman (1988) and Langran (1989). From a computer science perspective, a database is populated by discrete, well defined objects that exist for defined periods of time. From a geographical perspective, objects are often interpreted and poorly defined, and may move and change form through time. Thus although databases may support simple appearance and disappearance of objects through time, the geographical case is generally much more complex and raises fundamental questions of data modelling (consider, for example, the data model implicit in Hagerstrand's time-space prism, Hagerstrand 1970).

Raper (1989) has collected a valuable volume of papers on modelling three dimensional phenomena, and a NATO-sponsored conference on the same topic was held in Montecito, CA in December 1989. The necessary software concepts for 3D modelling have been available in computer graphics for some time, and hardware is advancing very rapidly. Systems capable of drawing and rendering complex solid objects defined by up to 10^4 polygonal facets per second are already available for $10 000. But there are immense problems in bringing these capabilities into common use in the geosciences. Although there are models in, for example, oceanography and groundwater hydrology that generate results over a 3D finite element mesh, the more usual case is data-poor – geoscientists typically do not have large 3D datasets ready for visualization. Because there is no map analogue in the 3D case, there is no tradition of compiling datasets, and it will therefore be some time before effective use can be made of this technology. What seems to be needed is a 3D GIS oriented not to analysis but to compiling models of 3D structures from the assorted evidence available.

VI Spatial cognition

If GIS is a window on the world, designed to allow the user to find out about geographic reality, then issues of spatial cognition are clearly part of the agenda of geographic information science. How can we improve the design of GIS user interfaces, which are currently notoriously clumsy, to better mirror human processes of learning and reasoning about space (Gould 1989; Gould and McGranaghan 1990)? How do we design specialized user interfaces for vehicle navigation systems (Freundschuh *et al.*, 1989; Freundschuh *et al.*, 1990), or for helping the visually impaired? Is it possible to design natural language interfaces that correctly interpret locational prepositions (Haller, 1989; Haller and Mark, 1990; Mark, 1989; Mark and Frank, 1989)? Compare for example the use of 'on' in 'is the house on the lake?' and 'is the boat on the lake?'. The language of spatial relations has been the subject of the second research initiative of NCGIA, and a NATO-funded meeting in Spain in July 1990.

VII Conclusion

As we noted at the outset, the trend in GIS research is currently away from the technology itself, and towards the generic issues of geographic information that transcend the technology. Systems have matured to the point where it is now possible to see the larger picture, and to realize the importance of Dangermond's 'considerable knowledge base'. At the same time GIS research is still very dependent on the health of the GIS industry. Research results need to be implemented and that is best achieved by ensuring a healthy industry with a steady flow of innovators.

Abler (1988) describes GIS as a critically important tool for geographic research: '. . . the telescope, the microscope, the computer and the Xerox machine of regional analysis and synthesis'. Arguments about geographic information science suggest a more fundamental relevance, implying a set of scientific issues underlying geography and intimately linked to our understanding of the geographical world around us. In areas such as accuracy and user interface design it seems now clear that the development of GIS technology is leading to a critical re-evaluation of old assumptions and practices.

Abler, R.F. 1987: The National Science Foundation National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems 1 (4)*, 303–26.

— 1988: Awards, rewards and excellence: keeping geography alive and well. *Professional Geographer* 40: 135–40.

Achabal, D.A. and **McIntyre, S.H.** 1987: Guest editorial: how information technology is reshaping retailing. *Journal of Retailing 63, 321–25.*

Arbia, G. 1988: *Spatial data configuration in the statistical analysis of regional economics and related problems.* Dordrecht: Kluwer.

— 1989: Statistical effect of spatial data transformations: a proposed general framework. In Goodchild, M.F. and Gopal, S., editors, *Accuracy of spatial databases*, New York: Taylor and Francis, 249–59.

Arbia, G. and **Haining, R.P.** 1990: Error propagation through map operations. Unpublished manuscript.

Beaumont, J.R. 1988: Guest commentary: an overview of decision support systems for retail management. *Journal of Retailing* 64, 361–74.

Buttenfield, B.P. 1989: Scale-dependence and self-similarity in cartographic lines. *Cartographica* 26 (1), 79–100.

Carstensen, L.W. 1989: A fractal analysis of cartographic generalization. *Cartographica* 16 (3), 181–89.

Couclelis, H. 1989: Geographically informed planning: requirements for planning-relevant GIS. Paper presented at North American meetings, Regional Science Association, Santa Barbara, November.

Craig, W.J. 1989: URISA's research agenda and the NCGIA. *URISA Journal* 1 (1), 7–16.

Digital Cartographic Data Standards Task Force (DCDSTF) 1988: The proposed standard for digital cartographic data. *The American Cartographer* 15 (1).

Dutton, G. 1984: Geodesic modeling of planetary relief. *Cartographica* 21, 188–207.

— 1989: Modeling locational uncertainty via hierarchical tesselation. In Goodchild, M.F. and Gopal, S., editors, *Accuracy of spatial databases*, New York: Taylor and Francis, 125–40.

Flowerdew, R. and **Green, M.** 1989: Statistical methods for inference between incompatible zoning systems. In Goodchild, M.F. and Gopal, S., editors, *Accuracy of spatial databases*, New York: Taylor and Francis, 239–47.

French, S.P. and **Wiggins, L.L.** 1988: Computer adoption and use in California planning agencies: implications for education. *Journal of Planning and Education Research* 8 (2), 97–108.

French, S.P., Wiggins, L.L. and **Heffernon, K.M.** 1989: Planning agency experiences with automated mapping and geographic information systems. Paper presented at North American meetings, Regional Science Association, Santa Barbara, November.

Freundschuh, S.M., Gould, M.D. and **Mark, D.M.** 1989: Issues in vehicle navigation and information systems. *Technical Report 89–15.* Santa Barbara: National Center for Geographic Information and Analysis, University of California.

Freundschuh, S.M., Mark, D.M., Gopal, M.D. and **Couclelis, H.** 1990: Verbal directions for wayfinding: implications for geographic information and analysis systems. In *Proceedings, Fourth International Symposium on Spatial Data Handling, Zurich*, University of Zurich.

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

— 1989: Geographic information systems and market research. *Papers and Proceedings of Applied Geography Conferences* 12, 1–8.

— 1990: Keynote address: spatial information science. In *Proceedings, Fourth International Symposium on Spatial Data Handling, Zurich*, University of Zurich.

Goodchild, M.F. and **Gopal, S.** 1989: *Accuracy of spatial databases.* New York: Taylor and Francis.

Goodchild, M.F. and **Yang, S.** 1989: A hierarchical spatial data structure for global geographic information systems. *Technical Report 89–5.* Santa Barbara: National Center for Geographic Information and Analysis, University of California.

Gould, M.D. 1989: Human factors research and its value for GIS user interface design. In *Proceedings, GIS/LIS 89*. Bethesda, MD: ASPRS/ACSM, 541–50.

Gould, M.D. and **McGranaghan, M.** 1990: Metaphor in Geographic Information Systems. In *Proceedings, Fourth International Symposium on Spatial Data Handling, Zurich*, University of Zurich.

Hagerstrand, T. 1970: What about people in regional science? *Papers, Regional Science Association* 24, 7–21.

Haller, S.M. 1989: Spatial relations and locative phrase generation in a map context. *Technical Report 89–14*. Santa Barbara: National Center for Geographic Information and Analysis, University of California.

Haller, S.M. and **Mark, D.M.** 1990: Knowledge representation for understanding geographic locatives. In *Proceedings, Fourth International Symposium on Spatial Data Handling, Zurich*, University of Zurich.

Heuvelink, G.B.M., Burrough, P.A. and **Stein, A.** 1989: Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information Systems* 3 (4), 303–22.

Kim, T.J., Wiggins, L.L. and **Wright, J.R.** 1990: *Expert systems: applications to urban planning*. New York: Springer Verlag.

Langran, G. 1989: A review of temporal database research and its use in GIS applications. *International Journal of Geographical Information Systems* 3 (3), 215–32.

Langran, G. and **Chrisman, N.R.** 1988: A framework for temporal geographic information. *Cartographica* 25 (3) 1–14.

Lodwick, W.A. 1989: Developing confidence limits on errors of suitability analyses in geographical information systems. In Goodchild, M.F. and Gopal, S., editors, *Accuracy of spatial databases*, New York: Taylor and Francis, 69–78.

Lodwick, W.A. and **Monson, W.** 1988: Sensitivity analysis in geographic information systems. *Paper RP886*. Denver: Department of Mathematics, University of Colorado.

McMaster, R.B. 1989: The integration of simplification and smoothing algorithms in line generalization. *Cartographica* 26 (1), 101–21.

Mark, D.M. 1989: Cognitive image schemata for geographic information: relations to user views and GIS interfaces. In *Proceedings, GIS/LIS 89*, Bethesda, MD: ASPRS/ACSM, 551–60.

Mark, D.M. and **Frank, A.U.** 1989: Concepts of space and spatial language. In: *Proceedings, AutoCarto 9*. Bethesda, MD, ASPRS/ACSM: 538–556.

NCGIA 1989: The research plan of the National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems* 3 (2), 117–36.

Newcomer, J.A. and **Szajgin, J.** 1984: Accumulation of thematic map errors in digital overlay analysis. *The American Cartographer* 11 (1), 58–62.

Raper, J. 1989: *Three dimensional applications in GIS*. New York: Taylor and Francis.

Shepherd, J.W., Masser, I., Blakemore, M. and **Rhind, D.W.** 1989: The ESRC's Regional Research Laboratories: an alternative to the NCGIA. In *Proceedings, AutoCarto* 9, Bethesda, MD: ASPRS/ACSM.

Twigg, E. 1990: Health based geographical information systems: their potential examined in the light of existing data sources. *Social Science and Medicine* 30 (1), 143–55.

Veregin, H. 1989: Error modeling for the map overlay operation. In Goodchild, M.F. and Gopal, S., editors, *Accuracy of spatial databases*, New York: Taylor and Francis, 3–18.

Zubrow, E., Allen, K. and **Green, S.** 1990: *Interpreting space: GIS in archaeology and anthropology*. New York: Taylor and Francis, forthcoming.