Introduction

THE EDITORS

In this final Technical Issues section, attention turns to the algorithms that allow GIS to transform and link data. Section 2(a) dealt with trends in the computing industry, in computer architectures, programming environments, and communications networks. All of these are likely to influence GIS profoundly over the coming decade. The next section looked at spatial databases, where the general solutions offered by the database industry intersect with the special needs of spatial databases for high performance, multi-dimensional search, and complex structures. The third section dealt with selected technical aspects of spatial data collection. Having addressed architectures, programming environments and data, this final set of chapters examines technical issues in the processing of data in GIS, and focuses on areas where algorithms are particularly challenging and complex, and the subject of ongoing research.

Several disciplines are contributing actively to research in advanced spatial data processing algorithms. The field of computational geometry addresses specifically the computer processing of geometric objects, and has widespread applications in GIS. For example, a fundamental algorithm in computational geometry detects whether two line segments intersect, and if so, where; it finds applications in such GIS processes as polygon overlay. GIS adds a particular context to computational geometry – one of its most distinguishing features is that position can never be exact, and must always be subject to a prescribed tolerance. GIS is an important application field for 'finite precision geometry', the field that attempts to formalise and theorise about such fuzzy spatial problems.

Another distinguishing feature of many GIS applications is the importance of performance, and the ability of algorithms to scale over a very wide range of problem sizes. The field of algorithmic complexity deals formally with performance, and the effects of problem size, and is widely applied in research on advanced GIS algorithms (see De Floriani and Magillo, Chapter 38, for example).

Image processing is yet another cognate field, with many applications to geographical data and GIS. As we saw in the last section, GIS colours these applications with particular characteristics. The surface of the Earth is not flat, unlike a photographic plate, and algorithms are needed: to register images to it accurately; to register images of the same feature to each other; to 'rubber sheet' images for geometric correction; or to change from one projection to another (see Dowman, Chapter 31). Another distinguishing characteristic of image processing when applied to geographical data is the widespread absence of 'truth' (see also Barnsley, Chapter 32) – when imaging the human body, for example, it is possible to regard the label 'liver' as 'true' of a certain part of the image; but suitably precise geographical analogies are much harder to come by – the label 'lake' is confused by many problems of definition (when is a lake a swamp, or a reservoir? See Mark, Chapter 7).

The ability to link data is often cited as the distinguishing feature of GIS. Location on the Earth's surface forms a convenient common key between otherwise disparate datasets and forms of information, allowing data to be linked across the boundaries of disciplines, departments, and agencies. When events occur in the same place, or near to each other, it is easy to believe that they also influence each other, and that both need to be taken into account in making decisions. Contemporary thinking on environmental management urges us to think of all things on the Earth's surface as connected and inter-related – in Tobler's 'first law of geography', 'all things are related, but nearby things are more related than distant things' (Tobler 1970).

Several different forms of linkage can be identified in GIS. Consider two datasets, A and B, covering the same geographical area, and imported to a GIS from different sources. At one extreme, the information provided by B is entirely distinct ('independent' or 'orthogonal' in a statistical sense), and both A and B are necessary in some application. The application will likely require that A and B be overlaid, and the ability to do so will depend on the formats of the two datasets – most GIS will require that they be both raster, or both vector; if both raster, that the rasters be congruent. In another case, the common key between A and B may be a feature identifier, rather than a geographical location. Such cases occur when B provides tabular information to be added to the geographical features in A as additional attributes. Yet another case occurs when B is a source of selective updates for the information in A – for example, B might contain more accurate coordinates for the features shared between it and A. When no common feature identifier is available, the features in B must be matched geometrically to those in A, a process that has been termed 'conflation', and is itself the subject of intensive current research.

Some forms of data transformation and linkage in GIS are straightforward, and do not justify particular attention. Others are made sufficiently special by their geographical context to have emerged as strong subfields for research and development within GIS. In this section the editors have selected several of these, while recognising that the set is by no means exhaustive, and may not survive the test of time – five or ten years from now research and development attention may have turned to quite different problems.

The section contains seven chapters. In the first, Lubos Mitas and Helena Mitasova review the state of the art in spatial interpolation, a vital component in the GIS arsenal because it provides estimates of the value of a variable z at locations (x,y) where it has not been measured. Spatial interpolation is essential in resampling, when data must be shifted from one raster to another; in transformation between representations, such as from a grid to contours; or in dealing with the problems caused by incompatible reporting zones. Mitas and Mitasova review the methods currently available, discuss the bases on which they can be evaluated, and review the applications of the methods in GIS.

Data linkage across GIS layers provides the theme for Ronald Eastman's contribution. Multi-criteria evaluation is concerned with the allocation of land to suit a specific objective on the basis of a variety of attributes that the selected areas should possess. This implies an apparently straightforward GIS-based overlay exercise, yet this process is complicated on the one hand by differences in data structure (raster versus vector) and, on the other, by ambiguities in the ways in which criteria should be standardised and aggregated into a single summary coverage. Eastman reviews these problems and suggests the use of fuzzy measures as a means of reconciling and developing current practice. In this way, the harshness of using Boolean operators to identify intersection and union of data layers is replaced with an approach which also provides improved standardisation of criteria and better evaluation of decision risk (the likelihood that the decision made will be incorrect).

One aspect that distinguishes GIS from other forms of spatial data processing, notably computer-assisted design (CAD), is its emphasis on representing fields, or variables having a single value at every location on the Earth's surface. Examples of fields are elevation, mean rainfall, or soil type. Because a field is continuous by definition, it must somehow be rendered discrete in order to be represented in the finite space of a digital store. Methods for discretising fields have been discussed at many points in this volume. Among them, two achieve their objective by dividing a plane surface into regular or irregular tiles, forming a 'tessellation'. The mathematics and statistics of tessellations, and their representation and processing in GIS, are important topics for research. The third chapter in this section, by Barry Boots, reviews the state of the art. Some of this research goes well beyond the current state of GIS implementation, particularly in the area of weighted tessellations, but it is easy to see how such methods could be implemented and applied.

GIS has a long history of successful implementation of digital models of the Earth's terrain, collectively known as digital terrain models (DTMs). They include triangulated irregular networks (TINs); and the commonest form of DTM, the digital elevation model (DEM), a rectangular array of spot elevations. DTMs are available in one form or another for much of the Earth's terrestrial surface and for the ocean floor, although the sampling density and conditions of availability vary enormously. One major factor driving the development of DTM technology is its importance in military applications, particularly missile guidance systems ('cruise' missiles navigate largely by recognising the geometric form of the terrain under them).

Many sciences, hydrology and geomorphology in particular, have an interest in the form of the Earth's surface, and its influence on the environment. The availability of DTMs, and the facilities in GIS for processing them, have led to an explosion of research on DTM analysis techniques. GIS is now a very significant tool in these fields, and DTMs are also useful in such practical applications as transmission tower location. The fourth and fifth chapters in this section discuss recent DTM research from two perspectives. Lawrence Band reviews the importance of DTMs and related datasets in hydrography and the analysis of landforms; while Leila De Floriani and Paola Magillo discuss representations, and associated algorithms, for the transformation of DTM data into useful information on intervisibility.

The last two chapters in the section move somewhat away from this intensive discussion of theory and algorithms. In Chapter 39, Jorge Nelson Neves and Antonio Câmara look at the role of GIS in the expanding field of virtual reality, or Virtual Environments (VEs). VEs are clearly an important area of application for GIS, particularly if the environment being simulated is in any sense related to the real, geographical world; and even totally artificial environments must be constrained by certain characteristics of the real world if they are to be convincing. VEs require many standard GIS techniques, as well as more generic techniques of image processing and visualisation; and also require much higher performance than traditional GIS applications.

In the final chapter, Michael Goodchild and Paul Longley discuss the issues encountered in using GIS as a data linkage technology. These range from the technical issues of accuracy, compatible data formats and data models, and rules for conflation, to the capabilities of contemporary communication networks for supporting data search and sharing, to institutional issues that are beyond the scope of this section.

Reference

Tobler W R 1970 A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46: 234–40