THE DATA STRUCTURE OF BASIS

Paul M. Wilson, Planning Consultant
Donald A. Olmstead, Principal Planner
Association of Bay Area Governments
Hotel Claremont
Berkeley, California 94705

1. Background: Small Systems, Big Applications

Small computer systems, as they have increased in power and decreased in cost, are being widely used for handling geographic data. Small systems have severe limitations, however, when the area to be covered is a large and complex region. This paper will focus on the issue of data structure - the arrangement of data in the data base and its correspondence to the geography of the region - since this element of system design is especially critical when constrained by the limitations of a small computer.

2. A Bay Area Example

An example of a minicomputer-based geoprocessing tool is the Bay Area Spatial Information System (BASIS), an automated geographic data base for the 7000-square mile San Francisco Bay region. Built over the last three years, it was designed to serve as a common data base for planning decisions in the Bay Area. Although created primarily for regional applications, it contains enough detail to be useful in some county and city level problems. It has been used in projects such as earthquake damage assessment, location of hazardous solid waste disposal sites, airport noise mitigation, and an industrial lands inventory.

BASIS is a grid cell system; space is represented by an array of one-hectare cells. Coverage of the region (including land, Bay, lakes, and parts of the ocean important for coastal planning) requires over two million of these cells. (Figure 1 shows parts of the BASIS area plotted at different scales.)

Cell size is clearly a critical design decision in a system of this nature; the tradeoffs are between level of detail (where smaller cells imply the ability to capture and maintain data with finer grain) and cost (smaller cells mean more total cells to cover a given area, which means higher costs for storage and processing). The hectare cell size used in BASIS was chosen after much discussion, and reflected consideration of the detail v. cost tradeoff as well as anticipation of probable applications.

3. Hardware Configuration

BASIS runs on a medium-scale minicomputer system. It is built around a Univac V76 processor, and utilizes a large (88mb) disk for data storage. Two peripheral devices are included specifically for geoprocessing applications: a digitizer is used to encode mapped data and an electrostatic plotter is the graphic output device.

These peripheral devices are connected to the computer through hardwired lines. This direct connection enables high speed transmission of data and eliminates the need for manual transfer of cards or tape. It also permits an interactive approach to the data entry process, so that potential errors can be flagged for immediate attention by the digitizer operators.

4. Data Management Software

Grid cell structures have been extensively used for many years to handle geographic data bases. The concept of a cell as storage location for data is easy to understand, and software for analyzing cell-based

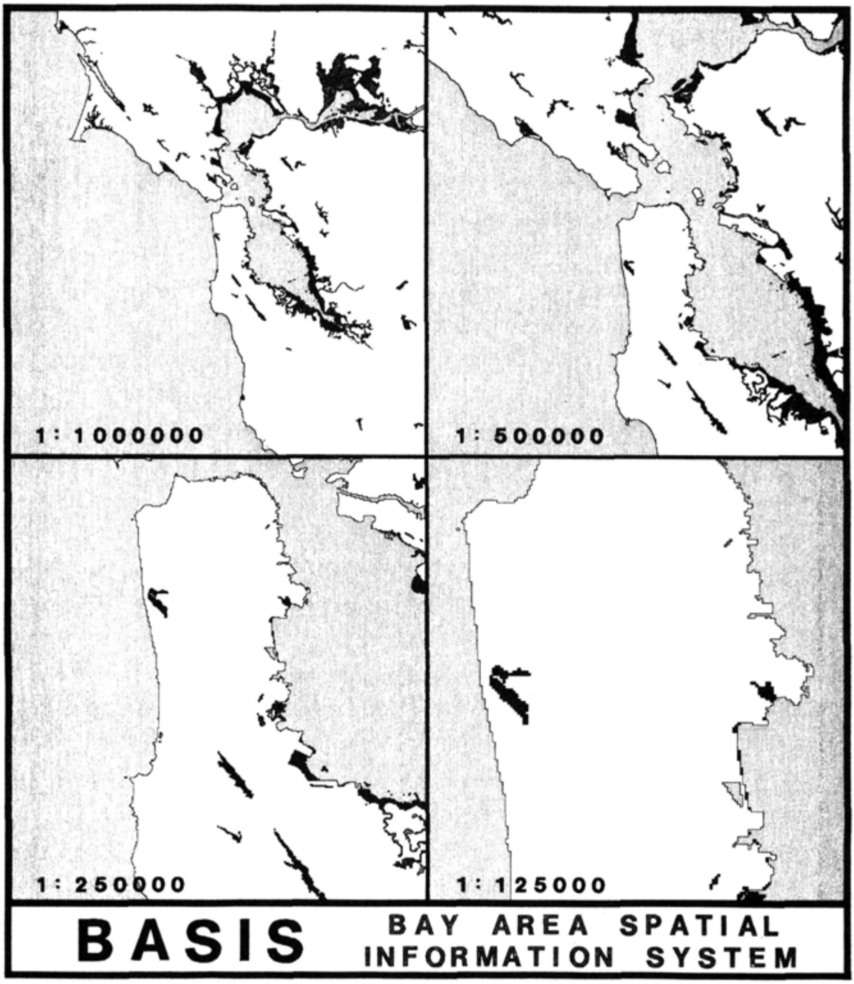

FIGURE 1 - Plot of coastal features in a portion of the area covered by BASIS, shown at four different scales (original plot has been photographically reduced so scales are no longer correct).

data is relatively simple. Problems arise with size, however: a combination of factors such as a large region, small cells, and many data types lead to large storage requirements.

Storing up to eighty separate data items for each of two million grid cells - and being able to quickly access any one of those cells - is a considerable task on any size computer system. It is made even more difficult on a small system with limitations in memory size, disk capacity, and I/O speed. Operating a large geographic data base on a relatively small computer system, then, requires considerable attention to the design of data management structure.

Design of the structure for BASIS had, as a major objective, on-line access to the entire data base. This type of access is essential for efficient updating and editing. Also, access to only part of the data base (some data types for the entire region or all data for a part of the region) would be inadequate for many applications.

Two major elements in the structure of a grid cell data management system are the location of the data for any specified ground location and the arrangement of the data base on some mass storage mechanism such as a disk. The BASIS approach to this problem of data management relies on two key concepts: direct access to data grouped by kilometer cell, and bit plane coding.

Direct access techniques

The first concept deals with the relationship of data space and "ground" space; that is, how to establish the correspondence between a given location on the earth (a defined grid cell) and the space on the disk (file name and record number) where data concerning that location on earth is stored. This correspondence is, of course, fundamental to any geoprocessing system. The linkage should, if possible, support direct access; that is, the data for a cell anywhere in the region should be accessible without having to search.

This relationship is established in BASIS by using a lookup table, which is arranged so that table position corresponds directly to ground position (e.g., the location of the data for the cell at Row 126, Column

240 in the ground coordinate system is contained the same row and column of the lookup table). Each element of the table points to the disk location (file designation and record number) of the data for the corresponding grid cell. Figure 2 diagrams this method.

To minimize the size of this lookup table, the hectare cells are grouped by square kilometer. Since each kilometer cell can be viewed as a ten by ten array of hectare cells, the location of each hectare within the kilometer space is implicit. Some entries in the lookup table will be blank; those positions which correspond to grid cells for which there is no data (cells outside the region or in the ocean) will have no pointer. This way, only areas of interest occupy storage space. By having the disk file designation in the lookup table, all references to different files can be invisible to the applications programmer. The entire data base can be viewed as one logical disk file; the table handles the correspondence between this single logical file and multiple physical files.

Bit plane coding

A second major consideration in data management structure is the question of storage efficiency. For most applications, it is desirable to have the entire data base accessible at any time; this becomes difficult when the data base is large and the amount of on-line storage is relatively small. In BASIS, where there are over two million spatial units and up to eighty data items for each cell, the problem becomes one of storing 160 million units of data on a device (disk drive) with a capacity of 44 million words.

BASIS uses a method of data compaction called bit plane coding to reduce space requirments. Instead of each separate data value occupying one full word (16 bits) of storage, it is assigned the minimum number of bits needed to retain the full range of that variable. For example, a data value that is strictly in/out, such as the delineation of a flood plane, can be coded with only one bit. Similarly, many of the data types used in this type of system require only a few bits for complete encoding. The result is more types (or planes) of data packed into the same storage space.

In theory, the structure of such a data base is not very different from the usual approach: it can be viewed as a three-dimensional array, where rows and columns represent two dimensions and different types of data are stacked vertically. (Figure 3.) The essential difference in bit plane coding is that each vertical element, instead of being a constant size in bits, has a height (i.e., number of planes) which differs according to its requirements.

The relative efficiency of this technique in minimizing storage space is dependent on the nature of the data; types with large ranges require many bits, and will therefore approach the "natural" encoding scheme where each data item occupies one word. The greatest gain is, of course, for those types of data where the coding can be accomplished with only one bit. Many types of spatial data do fit under this heading; any data with an on/off, in/out, present/not present character can be coded this way, and with good data definition other (and not so obvious) types can be fit into a few bits.

Use of the bit plane technique does require care, however. Considerable thought must be given to the definition of each data type and its range of potential values. And additional software, probably in assembly language, will be required. The additional machine operations required to pack and unpack the bit planes will increase execution time. Overall, the added work required to implement this type of coding scheme is worthwhile only if there is a severe constraint on the amount of on-line disk storage available and immediate access to all data is important.

5. Operational Environment

Creating an operational system is very different from designing an equivalent system in a research setting. Solving the technical problems - the selection of hardware, design of data base structure, development of software - does not ensure that the system will be used, or that it will receive the support necessary to maintain it. Institutional factors, such as funding and management responsibility, are also part of the overall system design.

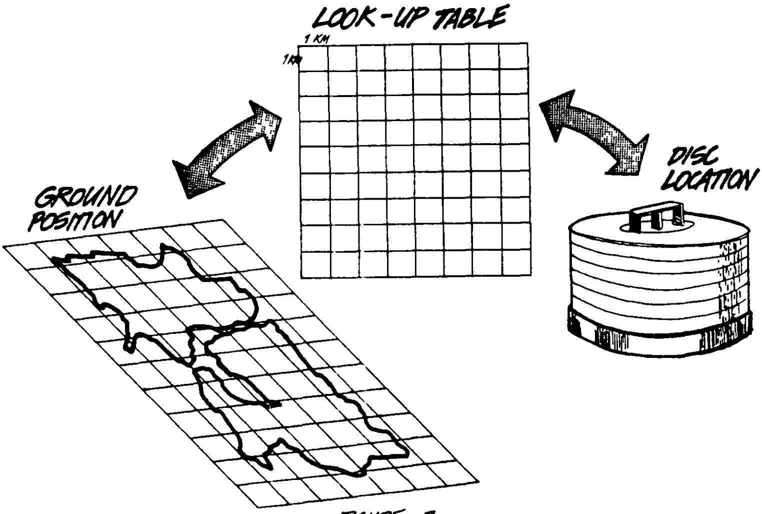

FIGURE-2

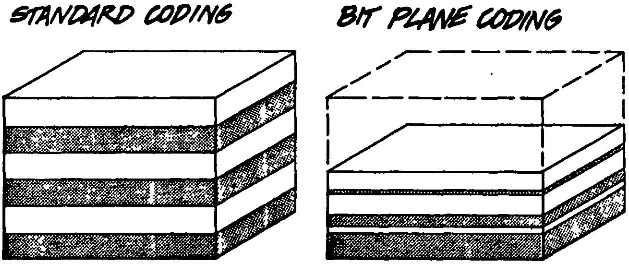

FIGURE-3

BASIS resides on a minicomputer operated by ABAG (the Association of Bay Area Governments), the regional planning agency for the Bay region. Since the computer is used for other ABAG computing functions BASIS must compete for use of system resources. This results in a very different operational environment from a system dedicated to the geoprocessing function. The data management design is based on these constraints; for example, the absence of personnel to handle media transfers (such as loading the disk from tape) is an added motivation for having the entire data base on a single disk pack.

6. Conclusions

BASIS has confirmed that a very large geographic data base system can indeed be based on a minicomputer system. This is demonstrated by the projects to which the system has been successfully applied.

The process of designing BASIS has also yielded insight into the problems of data structure in grid cell systems. The techniques of direct access and bit plane coding have been important in building the system in a way that is understandable and responsive to the user.

Finally, it is important to recognize the nontechnical aspects of system design. While this paper has concentrated on technical factors, it should not be assumed that these considerations are sufficient to build a successful system. A solid technical design is essential if the system is to be useful, but technical excellence by itself is not enough. Unless the organizational environment has been designed with equal care, the system is unlikely to be a success.

References

A discussion of bit plane coding is contained in J.L. Pfaltz's paper "Representation of Geographic Surfaces within a Computer" in John C. Davis and Michael J. McCullagh (eds.), Display and Analysis of Spatial Data (London: John Wiley and Sons, 1975).

Further discussion of BASIS applications is contained in several publications available from ABAG.