GEOGRAPHIC INFORMATION SYSTEMS: TOWARD A GEO-RELATIONAL STRUCTURE

Kenneth J. Dueker
Portland State University
School of Urban and Public Affairs
P.O. Box 751
Portland, OR 97207

ABSTRACT

The paper distinguishes features of computer-aided mapping and geographic information systems and identifies an emerging link. Common to both systems is the need to link graphic data and attribute data. This geo-relational structure requires a one-to-one relationship of the graphic record with the record of the table of attributes. Combining distance-referenced line data with a topological data structure is the key to a powerful geographical information system.

INFORMATION

The similarities and differences between a system for computer-aided mapping (CAM)[1] and a geographic information system (GIS) are difficult to articulate, but important. The following definitions are offered as a means of initiating this presentation of distinguishing features:

1. A GIS is defined as a specialized information system in which locational identifiers are attached to data for spatial analysis and/or mapping. Thus a GIS provides spatial information to users for decision making in management, planning, and research. Importantly a GIS allows the spatial collation of separately collected data. Cartographic modeling or overlaying is employed for analysis across layers of data to express the spatial relationship among the variables.

2. On the other hand, computer-aided mapping is a more limited display of layers of data with the ability to select layers, window, scale, and display, but without the ability for analysis across layers.

These definitions ignore a feature that both kinds of systems are stretching to meet -- a data base management system that relates objects on the map to records in data files containing attributes of those map objects. This portends that the distinctions identified in the definitions will diminish. This will be illustrated by the subsequent development of a geo-relational structured GIS.

The purpose of the paper is to distinguish better between computer-aided mapping and a GIS and to identify an emerging link. Although the distinctions are becoming less in actuality as systems evolve and mature, it is important to understand and to be able to describe clearly the difference. Belaboring the distinctions draws attention to the required properties and power of the GIS, while exploring ways by which

[1]Not to be confused with Computer-Aided Manufacturing (CAM) systems.

computer-aided mapping systems can be augmented for greater intelligence.[2]

The Urban and Regional Information Systems Association (URISA) supports this on-going process of promulgating consistent terminology and in clarifying concepts. A draft of the paper was prepared for discussion at the 1984 URISA conference in a special session on geoprocessing and the present version has benefited from feedback received.

COMPUTER-AIDED MAPPING

Computer-aided drafting systems when applied to mapping problems are hereinafter referred to as computer-aided mapping (CAM). The map is the product and the information to draw the map is not used for computation or analysis. Although, more current models of CAM have analytical capabilities, early versions merely overplotted layers of data with no analytical capability to relate data across layers. Streets, railroads, streams, boundaries, water mains, etc. were digitized as separate layers and the relationship among layers was only discernable visually. This inability to relate data across layers was due to a lack of data structure. The strings of data that made up a layer were in effect cartographic spaghetti. The strings were not structure to correspond to logical entities, i.e., points, lines, or areas. Windowing was about the only logical operation that could be performed. Within a layer no further selection by attribute could be performed. The roads, sewer lines, or streams were not segmented logically by width or diameter, or capacity. Even if they had been, selection by another attribute, say slope, would not have been possible.

More current CAM systems utilize within a layer a distance-based system of recording location on a line, such as a milepost, stream mile, or station, to select graphically from a layer <u>and</u> to provide linkage to attribute data about segments between stations or of stations. This linkage between graphic data and attribute data is extremely important. Before developing this importance, the emergence of GIS technology is described.

GIS APPROACHES: TOPOLOGICAL STRUCTURE

Early approaches to dealing with the limitations of CAM for analysis of spatial data as well as mapping data recognized the need to structure graphic data as point, line, and area, or grid cell elements.

Thematic mapping programs were built around a one-time specification of the graphic structure of points or areas, which would be related to values or value ranges for a number of attributes. In this way, a number of maps depicting a variety of data for the same graphic structure could be mapped efficiently.

Relating attributes or data across layers was more difficult. Grid data

---

[2]Dangermond and Freedman (1984), p. 4, draw a distinction between automated geographic information systems and computer-aided design-type system. The latter type are based on relationships <u>visible</u> to the user, resulting in "implicit" understandings <u>about</u> how data interrelates, whereas the topological relationships of GIS result in "explicit" data relating.

structures were adopted for this reason. The imposition of a regular lattice on layers of data allowed comparison of or computations in corresponding grid cells.

For vector formatted data, point-in-polygon, line-in-polygon, or polygon overlay were used to relate layers of data. Encoding data layers for polygon processing poses quality control problems, especially in natural resource applications. Polygon digitizing necessitates line segments of a network of areas to be digitized twice, once for each polygon it bounds. This process yields small gaps and overlaps as the same line yields two sets of coordinates when digitized. Overlays of polygons so derived results in many spurious polygons.

To avoid the spurious polygon problem, many systems adopted a chain encoding scheme wherein a line between junctions or nodes was digitized once and polygons formed by the computer, either using digitized center identifiers or by use of right and left polygon identifiers. This topological data structure was useful in editing to ensure the logical consistency of the network of areas.

Segmenting graphic records to correspond to chains connecting nodes in a topological data structure is useful for relating data across layers. It is a limited approach though in relating graphic data of line features to records in data files used by agencies responsible for operating and managing a resource or facility type characterized by a single graphic layer, say sewers or streets. Consequently, GIS technology also sought ways to relate graphic records of linear features to corresponding attribute records.

GIS EXTENSIONS: GEO-RELATIONAL STRUCTURE

The application of relational data base concepts to GIS yields a geo-relational structure. This geo-relational structure requires a one-to-one relationship of the graphic record with the record of the table of attributes. This means segmenting chains between nodes of a topological GIS data structure. Segmenting is necessary to produce graphic elements that are homogeneous with respect to attributes that may be used for selection.

Whereas a topological structure only requires a record pertaining to a section of road between intersections, a geo-relational structure requires segments that are homogeneous with respect to roadway width, depth of pavement, traffic volume, conditions of surface, etc. As these attributes change in value, segmentation would be necessary and a corresponding record in the table of attributes would be added. This geo-relational data structure allows selection and analysis by both geographic and attribute criteria. Roads or sewer lines of specific slope, age and/or condition can be selected.

In much the same way as drop-line algorithms can be used to delete boundaries that are redundant in a topological structure, segment-aggregation algorithms are needed to collapse networks into more generalized forms, when an attribute is common in value, dropped or ignored.

However, it is difficult to forsee all possible segmentations imposed by combinations of attributes. Consequently, a combined topological and distance-based system (station, milepost, stream mile, etc.) has the greatest potential for implementation of a geo-relational structured

GIS. Thus, segmentation of topological entities is not needed. The distance-based system (stations, road miles, stream miles) constitutes means by which attribute data can be used to select graphic elements.

With distance referenced line data it is possible to both select lines within arbitrary polygons and to select attribute records for distance units for lines contained in the polygon that possess specific attribute levels. Thus we could select roads of a given width, functional classification, and volume-to-capacity ration range within specified counties and between specified elevations. Or all two-mile sections of rivers downstream of specified discharge points having specified low flow rates could be selected. Property owners within 200 yards of the stream with stream withdrawal permits could be selected.

Distance referenced line data with associated attribute data that is coordinate based and topologically structured constitutes a basis for a powerful knowledge-based system. The addition of topologically structured and coordinate-based networks of areas, and coordinated point data add greater intelligence to the digital map data for inclusion within a geographic information system.

CONCLUSION

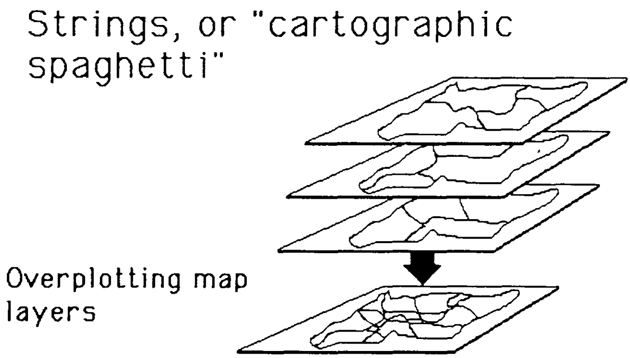
An intentially brief articulation of GIS concepts have been presented. An extreme distinction between CAM and GIS is stressed to draw attention to the importance of data structure for computer-assisted analysis. These distinctions are summarized and illustrated in the attached figures.

A distinction was then drawn within the GIS field that identifies a stand-alone topological data structure used in natural resource application. This topological GIS approach was compared to a geo-relational GIS that has great potential for facilities management applications.

This paper results from observations by the author that the concepts contained herein are understood by those working in the field, but poorly articulated. This is an attempt to clarify and explain the concepts and by so doing, hopefully encouraging standard terminology for communicating to potential users of GIS.

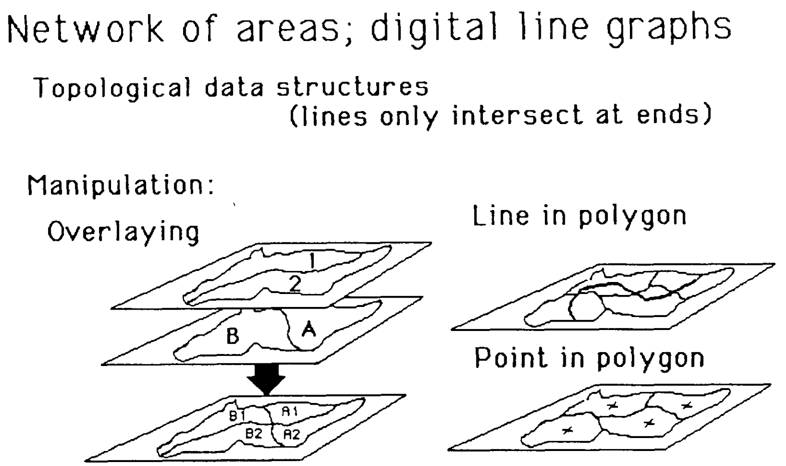
REFERENCES

Dangermond, J. and C. Freedman 1984, "Findings Regarding A Conceptual Model of a Municipal Data Base and Implications for Software Design," Environmental Systems Research Institute, Redlands, CA. (unpublished).



CAM (Computer-Aided Mapping)

Figure 1

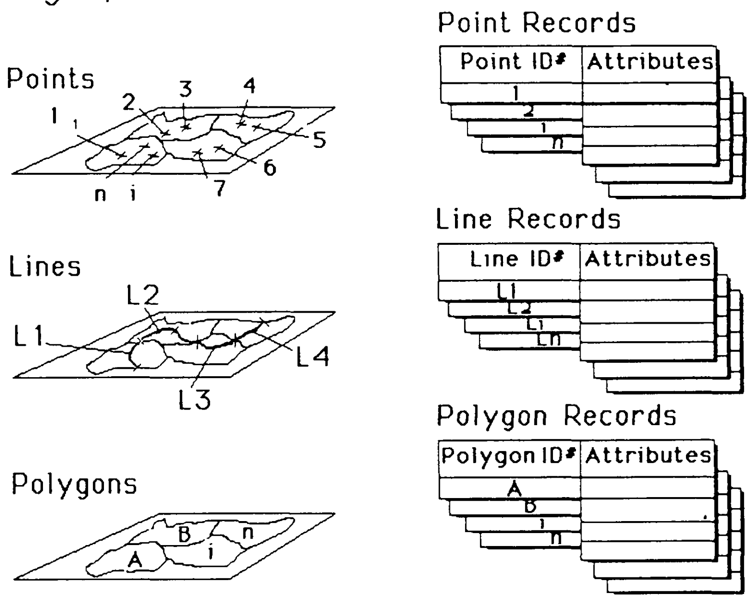


GIS (Geographic Information System)

Topological Structure

Figure 2

Correspondence of graphic records to non-graphic records



GIS (Geographic Information System)

Geo-Relational Structures

Figure 3

Station	Row Width	Pavement Width	Pavement Condition	Traffic Count	Culvert Diameter
0+00	100	26	3		
10+33					24
50+00			2		
63+21					48
100+00		24			
121+11					24
148+50				500	

Figure 4