UNIT 13 - THE VECTOR OR OBJECT GIS

UNIT 13 - THE VECTOR OR OBJECT GIS

Compiled with assistance from Holly Dickinson, State University of New York at Buffalo

---

For Information that Supplements the Contents of this Unit:

[Links to the following resources have been omitted.]

- Organizing Attribute Data (Foote and Huebner/Geographer's Craft) -- Flat files and spreadsheets, hierarchical files, relational files.
- Representation and Data Quality (Chrisman/U of Washington) -- Primitives for representation: attribute, space, points located in a coordinate system, lines formed from line segments; understanding Topology; alternative definitions of data quality.
- Example of Topological Relationships (Foote and Huebner/Geographer's Craft) -- Nodes, arcs, polygons, contiguity.
- Vector Data Maintenance and Analysis -- Introduction; vector data analysis (graphics); overlay operations; neighborhood operations; connectivity functions.
- Vector View of the World (Foote and Huebner/Geographer's Craft) -- Plus, graphical comparisons between raster and vector systems.

- REFERENCES
- EXAM AND DISCUSSION QUESTIONS
- NOTES

This unit begins a two part introduction to vector GIS. We have placed these units here since we feel this discussion benefits from an understanding of the previous introduction to spatial data concepts in Units 10 to 12. However, with a little revision, it is possible to move this module so that it follows Units 4 and 5 on raster GIS.

UNIT 13 - THE VECTOR OR OBJECT GIS

Compiled with assistance from Holly Dickinson, State University of New York at Buffalo

A. INTRODUCTION

Vector data model

- based on vectors (as opposed to space-occupancy raster structures)
- fundamental primitive is a point
- objects are created by connecting points with straight lines
  - some systems allow points to be connected using arcs of circles
- areas are defined by sets of lines
  - the term polygon is synonymous with area in vector databases because of the use of straight-line connections between points
- very large vector databases have been built for different purposes
  - vector tends to dominate in transportation, utility, marketing applications
  - raster and vector both used in resource management applications

B. "ARCS"

- when planar enforcement is used, area objects in one class or layer cannot overlap and must exhaust the space of a layer
- every piece of boundary line is a common boundary between two areas
- the stretch of common boundary between two junctions (nodes) has various names
  - edge is favored by graph theorists, "vertex" for the junctions
  - chain is the word officially sanctioned by the US National Standard
  - arc is used by several systems
- arcs have attributes which identify the polygons on either side
  - these are referred to as "left" and "right" by reference to the sequence in which the arc is coded

- arcs (chains/edges) are fundamental in vector GIS

Storing areas

- two ways of storing areas:

- polygon storage
  - every polygon is stored as a sequence of coordinates
  - although most boundaries are shared between two adjacent areas, all are input and coded twice, once for each adjacent polygon
  - the two different versions of each internal boundary line may not coincide
  - difficult to do certain operations, e.g. dissolve boundaries between neighboring areas and merge them
  - used in some current GISs, many automated mapping packages

- arc storage
  - every arc is stored as a sequence of coordinates
  - areas are built by linking arcs
  - only one version of each internal shared boundary is input and stored
  - used in most current vector-based GISs

C. DATABASE CREATION

- database creation involves several stages:
  - input of the spatial data
  - input of the attribute data
  - linking spatial and attribute data

- spatial data is entered via digitized points and lines, scanned and vectorized lines or directly from other digital sources
  - once the spatial data has been entered, much work is still needed before it can be used

Building topology

- once points are entered and geometric lines are created, topology must be "built"
  - this involves calculating and encoding relationships between the points, lines and areas
  - this information may be automatically coded into tables of information in the database

Editing

- during this topology generation process, problems such as overshoots, undershoots and spikes are either flagged for editing by the user or corrected automatically
  - automatic editing involves the use of a tolerance value which defines the width of a buffer zone around objects within which adjacent objects should be joined
  - tolerance value is related to the precision with which locations can be digitized

  - these edit procedures include such functions as snap, move, delete, split, join, etc.

Relationship between digitizing and editing

- digitizing and editing are complementary activities
  - poor digitizing leads to much need for editing
  - good digitizing can avoid most need for editing
  - both can be very labor-intensive

- the process used to digitize area objects can affect the need for later editing:

- in "blind" digitizing all linework is digitized once as "noodles" in any order
  - it is unlikely that the building and cleaning operations will be able to automatically sort out area objects unambiguously from the resulting jumble

- some systems require the user to identify junctions between digitized "noodles" explicitly
  - usually by touching a special button on the cursor

  - mistakes in building topology are less likely

- some systems require the user to digitize each individual arc/chain separately

  - much easier to sort out polygons - less need for editing

- some systems support the building of topology "on the fly"
  - the system searches constantly for complete area objects as digitizing proceeds
  - the user is informed by a sound or by blinking as soon as the object is detected

Edgematching

- compares and adjusts features along the edges of adjacent map sheets

- some edgematches merely move objects into alignment

- others "join" the pieces together logically - conceptually they become one object
  - the user "sees" no interruption

- an edgematched database is "seamless" - the sheet edges have disappeared as far as the user is concerned

D. ADDING ATTRIBUTES

- once the objects have been formed by building topology, attributes can be keyed in or imported from other digital databases

- once added to the database, attributes must be linked to the different objects
  - attributes can be linked by pointing to the appropriate object on the screen and coding its corresponding object ID into the attribute table

- unlike many raster GIS systems, attribute data is stored and manipulated in entirely separate ways from the locational data

E. EXAMPLE ANALYSIS USING VECTOR GIS

- compare with example analysis in Unit 4 (The Raster GIS)

Objective

- identify areas suitable for logging

- an area is suitable if it satisfies the following criteria:
  - is Jack pine (Black Spruce are not valuable)
  - is well drained (poorly drained and waterlogged terrain cannot support equipment, logging causes unacceptable environmental damage)
  - is not within 500 m of a lake or watercourse (erosion may cause deterioration of water quality)

Procedure

- database consists of three layers
  - note: polygons do not entirely fill the space in each case
    - hence, areas not included fall in polygon ID 0

- buffer hydrography out to 500 m

- merge buffer and lake

- extract Jack pine polygons (species = Jack pine)

- extract drained soil polygons (drainage = 2, therefore soil = A)

- overlay buffer, Jack pine and soil polygons

- build topology

- extract polygons not in the buffer but in others (buffer = n, Jack pine = y, drainage = y)

Result

- loggable area shown in final map

REFERENCES

Beard, M.V. and N.R. Chrisman, 1988. "Zipping: a locational approach to edgematching," The American Cartographer 15:163-72. Describes a solution to the edgematching problem.

Chrisman, N.R., 1990. "Deficiencies of sheets and tiles: building sheetless databases," International Journal of Geographical Information Systems 4:157-67. A more general discussion of building edgematched databases.

ESRI, 1990. Understanding GIS: The ARC/INFO Way, ESRI, Redlands, CA. A general introductory tutorial for ARC/INFO, a well-known contemporary GIS.

Tomlinson, R.F., H.W. Calkins and D.F. Marble, 1976. Computer Handling of Geographical Data. UNESCO Press, Paris. Excellent semi-technical description of CGIS, an early vector-based system.

EXAM AND DISCUSSION QUESTIONS

1. List and describe the processes involved in constructing a vector database by digitizing maps.

2. By using simple sketches, describe and illustrate typical problem cases which lead to difficulties in building area- object topology in a vector database, and the strategies which various GISs use to minimize editing effort.

3. Discuss the applications of GIS, in relation to the vector data model. Give examples of cases where the model

would be particularly inappropriate in comparison with raster.

4. Why did the designers of CGIS choose a vector data model, and yet use scanning as the major method of input?

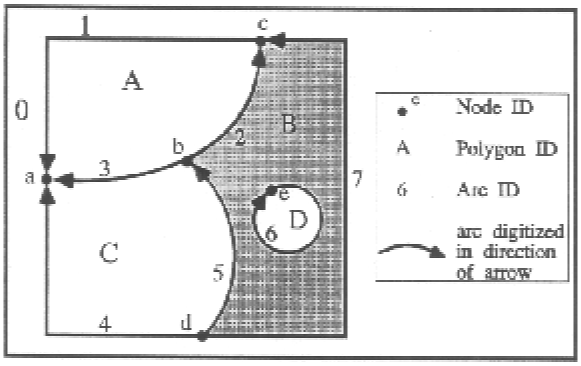
Last Updated: August 30, 1997.

UNIT 13 IMAGES

	point	line	area
scale ↓	city wells	highway political boundary streams	ag. land urban land city highway airport

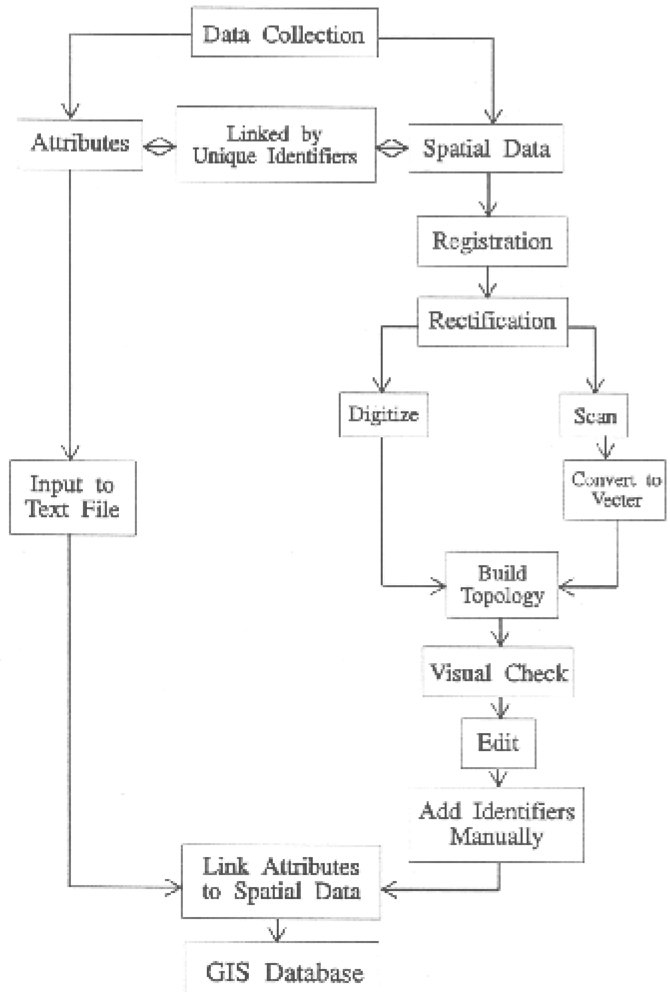
Example Attributes:

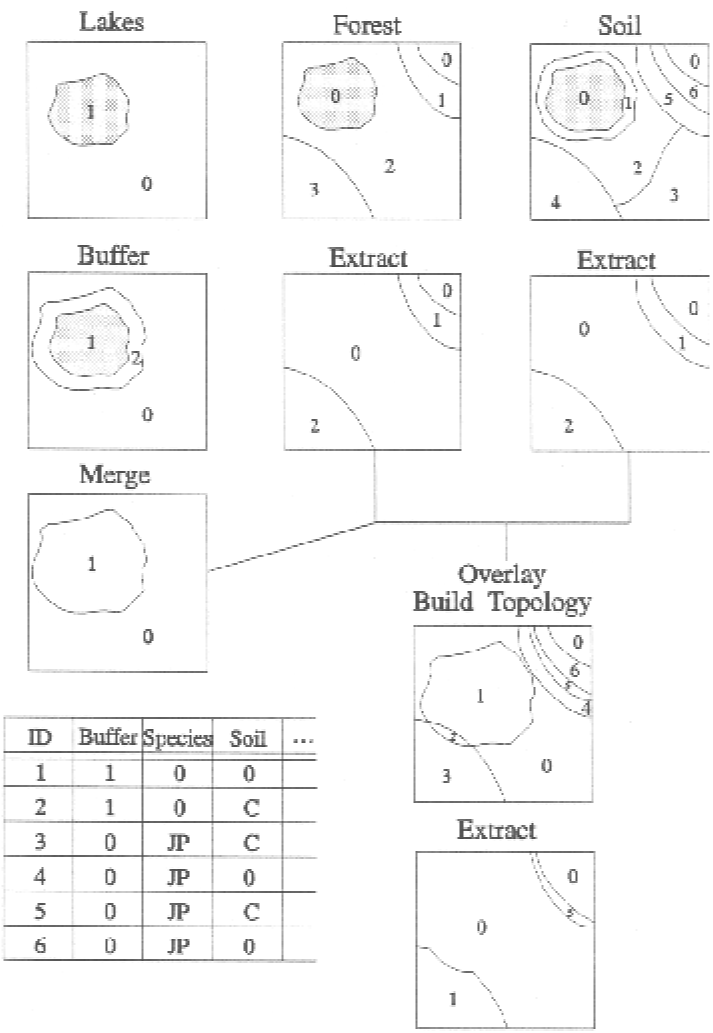
city: population, name
wells: depth
highway: number
political boundary: type
streams: name
ag. land: growth potential, acreage
urban land: urban landuse type, acreage
airport: name



Arc ID	Left Poly	Rt Poly	From Node	To Node
1	A	0	c	a
2	A	B	b	c
3	C	A	b	a
4	0	C	d	a
5	C	B	d	b
6	B	D	e	e
7	B	0	d	c

Polygon ID	No. of Arcs	List of Arcs
A	3	-1, -2, 3
B	4	2, -7, 5, 0, -6
C	3	-3, -5, 4
D	1	6


Data Collection
Attributes
Linked by Unique Identifiers
Spatial Data
Registration
Rectification
Digitize
Scan
Convert to Vector
Input to Text File
Build Topology
Visual Check
Edit
Add Identifiers Manually
Link Attributes to Spatial Data
GIS Database



ID	Buffer	Species	Soil	...
1	1	0	0	
2	1	0	C	
3	0	JP	C	
4	0	JP	0	
5	0	JP	C	
6	0	JP	0	