G492 GIS Applications
Vector Data Models

I. Introduction

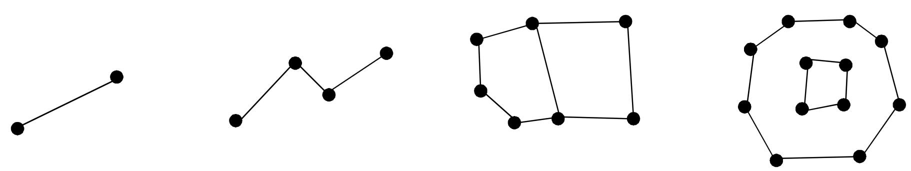
A. GIS Vector Data Model
- 1. points and x,y geocoordinates are used to constructure spatial features

B. Point-Based Spatial Features
- 1. points (uh, dah!)
- 2. lines
  - a. straight line segments
  - b. polylines
    - (1) multiple linked straight line segments
- 3. "arcs" - any curved or straight line segment that connnects two points

II. Geometric Objects

A. points, lines, polygons (enclosed areas)

B. polyline features
- 1. nodes
  - a. start node (beginning of line)
  - b. end node (end of line)
- 2. vertex
  - a. points not at the start or end of a line, that for the basis for polyline segments
  - b. vertices - series of points that form the shape of the line
- 3. smooth curves vs. segmented polylines
- 4. Line intersections, Line Joins



straight line (arc)   polyline   contiguous polygons   doughnut and island polygons

C. Polygon - enclosed areas in which the start node and end node are located at the same x,y coordinate
- 1. continguous polygons - mutually adjoining polygons with common vertices and/or nodes
- 2. isolated polygons
  - a. islands - free standing polygons enclosed by another polygon
  - b. doughnut-shaped polygons

D. Scale as a Controlling Factor of Geometric Map Feature
- 1. e.g. on a 1:24,000 topographic map a first order stream = a polyline; a wide river downstream (e.g. Willamette) = a polygon
  - a. At 1:1,000,000 the Willamette River = a polyline
- 2. e.g. the building we are in at 1:2000 = a polygon, the building we are as viewed from outer space = a point.

III. Map Feature Attributes and Topology

A. Attributes - Characteristic of a geographic feature described by numbers or characters, typically stored in tabular format, and linked to the feature.

1. E.g. attributes of a water well, represented by a point, might includedepth, pump type, location and gallons-per-minute.

B. Topology - an expression of the spatial relationships between map features

1. Map Features in a Topological Model
    a. composed of points and directed lines
        (1) directed lines - lines drawn in a particular direction and order from point to point
        (2) arcs = directed lines

2. Fundamental Components of Topological Relationships Between Map Features

    a. Connectivity - arcs connect to each other by nodes
    b. Area Definition - an area is defined by a series of connected arcs
    c. Contiguity - arcs are drawn in a given direction, from node to node, with left and right polygons of defined attributes

3. Methodology

    a. Nodes are assigned numbers (1 = first node, 2 = second node, etc.)
        (1) each node has an x,y georeference coordinate
        (2) Resulting arcs are assigned numbers, with left and right polygon attributes
    b. Vertex - points along an arc or polyline, that are not at the start or finish points (or nodes)
    c. Left-Right Polygon Lists
        (1) Each arc is identified and the attributes of the left and right polygon identified

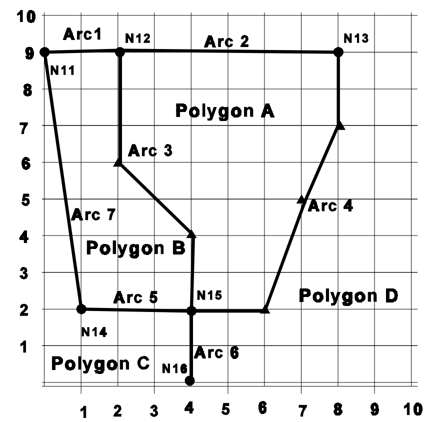
C. Non-Topological Vector Data

1. Problem with geometric features and topology - more data intensive for computer devices
    a. simple geometric features will load and run faster than topologically based features

2. Shape Files in ArcView - non-topologically based feature vector file
    a. points - i.d. by simple x,y coordinate
    b. lines - a series of points
    c. polygons - series of lines
    d. mutual polygons can have duplicates arcs overlapping one another
    e. File Format
        (1) *.shp - geometry of shape file stored here
        (2) *.dbf - attributes of shape file
        (3) *.shx - index file that links the *.shp file to the *.dbf file

In Class Exercise - Geometric Elements and Topology

The Figure at the right is a polygon map theme with polygons A, B, C, and D. The polygons are constructed from arcs 1 through 7. The arcs are composed of Nodes N 11 through N 14. The topology of the map is built upon graphical analysis of the georeference coordinates of the nodes and the arcs/polygons that they build.



The table below shows a typical topological framework for the spatial relations. The abbreviations are as follows:

Fnode The node at the beginning or start of an arc, "From Node"
Tnode The node at the end of an arc, "To Node"
Arc# The internal number assigned to identify the arc
Lpoly Attributes of the Left Polygon while "driving" from the Fnode to Tnode, along the arc.
Rpoly Attributes of the Right Polygon while "driving from the Fnode to Tnode, along the arc.

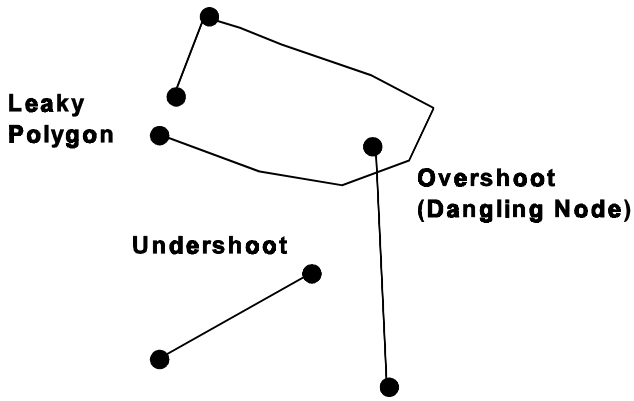
Exercise to complete. Based on the answer model for the first row below, complete the topological tables for the map to the right.

Arc Node List			Arc Coordinate List		Arc Polygon List		
Arc#	Fnode	Tnode	Arc#	x,y Coordinates	Arc#	Lpoly	Rpoly
1	11	12	1	(0,9) (2,9)	1	Polygon D	Polygon B
2	___	_____	2	_______________	2	_________	__________
3	___	_____	3	_______________	3	_________	__________
4	___	_____	4	_______________	4	_________	__________
5	___	_____	5	_______________	5	_________	__________
6	___	_____	6	_______________	6	_________	__________
7	___	_____	7	_______________	7	_________	__________

D. Topological Errors

1. topological errors arise when nodes and arcs are not properly "snapped" to one another or aligned
2. Error Types
   a. dangling nodes - nodes dangle in space without being snapped to another node
   b. undershoots - nodes are short of being snapped
   c. overshoots - nodes are long on being snapped
   d. leaky polygons - polygons are not closed, nodes are not properly snapped

See diagram below for examples



IV. Map Scale, Spatial Resolution, and Spatial Data Accuracy

A. Map scale is an indicator of map accuracy
   1. The smaller the scale, in general, the lesser the accuracy, and vice versa
      a. e.g. map accuracy at 1:100,000 scale is much less than 1:24,000 scale

B. Locations Accuracy and Topological Accuracy in GIS
   1. Location Accuracy - measures the error in the absolute position of a map point or feature relative to real world, georeference coordinates.
   2. Topological Accuracy - a measure of the error in topology and attribute features of map features

C. USGS Map Standards for Accuracy
   1. USGS maps are tested and standardized so that there is no more than 10% of total position points can be more than 0.02 inches (0.5 mm) out of position at the prevailing map scale.

In Class Exercise

At a scale of 1:65,000, 0.02 inches on the map represents how much distance on the ground in meters? Show all of your work.

Given a scale of 1:24,000, 30 m error on the ground would represent how many millimeters of error on the map? Show all of your work.

Given a scale of 1:24,000 and a spatial feature resolution of 10 m, how many inches of resolution does this represent in map units? Show all of your work.

V. Vector Data Input: How to capture and produce digital, vector-based GIS maps

A. Existing Data Sets (internet download)

1. Federal Geographic Data Committee www.fgdc.gov
2. National Spatial Data Infrastructure (NSDI)
3. U.S. Geological Survey
    a. USGS Vector File Formats: DLG - Digital Line Graphs
    b. DRG - digital raster graphics (scanned maps)
    c. DEM - digital elevation models (raster-based elevation grids)
4. US Census Bureau
5. States, Counties, local municipalities
    a. e.g. Polk County GIS office

B. Metadata - information about GIS data (standardized text files, download with data)
1. identification, title, dates, accuracy
2. Georeference information, coordinates, projections
3. entity and attribute information
4. References

C. Common Vector Files Available for Download from Internet
1. Zip files *.zip - compressed groups of files
2. Shape files (a set of *.shp, *.shx, and *.dbf)
3. DLG, DRM, DEM
4. *.e00 ArcInfo export files (must be imported into ArcView using the "Import72" utility of arcview)
5. SDTS files - spatial data transfer file format
    a. promoted as a standard transfer format by the USGS
6. *.DXF files - vector line files exported from AutoCAD

D. Creating New Data

1. GPS Data
2. Digitizing Paper Maps / Digital Conversion

    a. Digitize known control points with known georeference coordinates on map
        (1) once control points are established, all other positions on the map will be calibrated and known
    b. Heads Up Digitizing
        (1) On-Screen Digitizing on Air Photos or Remote Sensing Imagery
        (2)
    c. Digitizing and Root Mean Square Error
        (1) RMS = goodness of fit between the control points and their actual georeferenced coordinates
            (a) poor calibration of control points will lead to error in the digitizing process
        (2) RMS - measures the deviation between the actual location of the control points on the original map versus the digitized location

(a) RMS listed for each control point with an average for all points

RMS for a tic = sqrt [$(actX - estX)^2 + (act Y-estY)^2$]

Average RMS for all Points = sqrt [(sum of squares of deviation in X and Y) / (no. of control points)]

(i.e. for average, add up all of the $(actX - estX)^2$ and $(act Y-estY)^2$ caluclations and divide by the total number of control points, take the square root)

where
actX = actual X coordinate location of point
actY = actual Y coordinate location of point
estX = estimated X coordinate location of point
estY = estimated Y coordinate location of point

In-Class Exercise

Calculate the RMS for each of the control tics below, and the average RMS. Coordinates are in UTM meters. Show all of your work.

Control Pt.	actX	estX	actY	estY
1	481023.334	481029.71	4966231.786	4966234.25
2	481592.256	481596.89	4966834.765	4966854.32
3	481018.448	481044.76	4966245.354	4966251.87
4	481402.309	481499.72	4966845.274	4966839.71
