GEOG 204

LECTURE 4
Spatial Analysis

1

Mid-term Exam

- Tuesday 17th Oct
  - During the Lecture Time slot
  - Multiple choice questions and Short Answer Questions
  - Materials covered in the lectures (Sept to Oct 10th included)
  - 45 Minutes in Duration

2

Mid-term Prep

Multiple Choice

1. In a vector data model, how many coordinate pairs does a single straight line have
   a. 1   b. **2**   c. 3   d. 4   e. none of the above

2. Data downloaded from a website are known as vector data.
   a. True   b. **False**

3

Mid-term Prep

Multiple Choice

3. Which of these terms is used to describe a surface that approximates the shape of the earth
   a. Grid   b. Round   c. Geoid   d. Globe   e. none of the above

4

Mid-term Prep

Short Answer Questions

1. What is the difference between a GIS data model and GIS file format?
   - *Here you would be expected to explain your understanding a data model contrasting that with a file format.*
      - *Data model – how geographic objects are represented for computational purposes*
      - *File format – how data arranged and compressed in a computer file*

2. In the context of representing geographic phenomena, what the field-based view of the world?
   - *The field-based view represents geographic space as a continuous distribution of phenomena. This representation is done with the aid of the raster-data model where a grid of cells is used to represent every point in the geographic space.*
   - *The field-based view conceptualises space as a continuous surface modeled by a mathematical function such that one is able to get a value at any point on this surface. This continuous surface is represented with the raster-data model where a grid of cells is used to represent the values from such a mathematical function.*

Mid-term Prep

Short Answer Questions

3. List three properties distorted of geographic objects by map projections

4. Describe a primary method of collecting vector data

Spatial Analysis

- In a GIS,
  - Data are usually grouped into layers (or themes).
  - The analysis functions of a GIS use both spatial and non-spatial data
  - Analysis functions are used
    - for maintenance of the data
    - Computing new information from the data
- Spatial Analysis can be viewed as:
  - Operations that use spatial data to derive new spatial information.
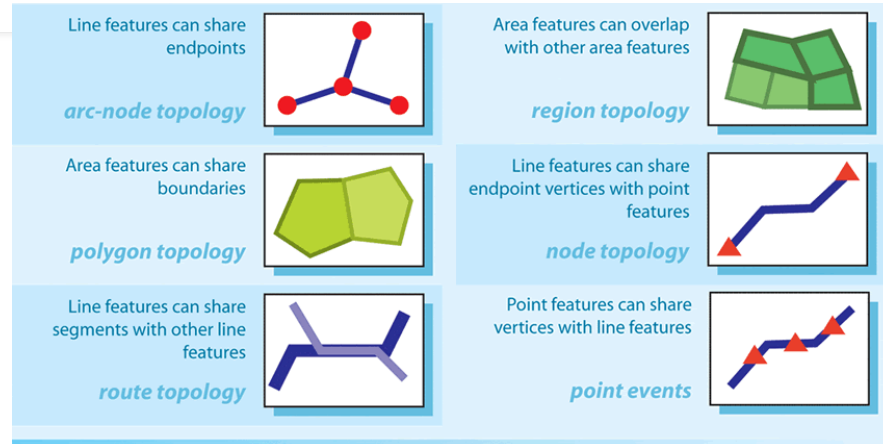 - The most distinguishing purpose of a GIS

7

7

Topology

- Topology: The spatial relationships between adjacent or neighboring features.
  - Based on 3 basic elements, Adjacency, Containment and Connectivity
  - Allows analyses like as contiguity, containment, connectivity, overlap…
    - For example, it is used in routing through linear networks. Routes cannot be established through the network if the line features do not share nodes/vertices.
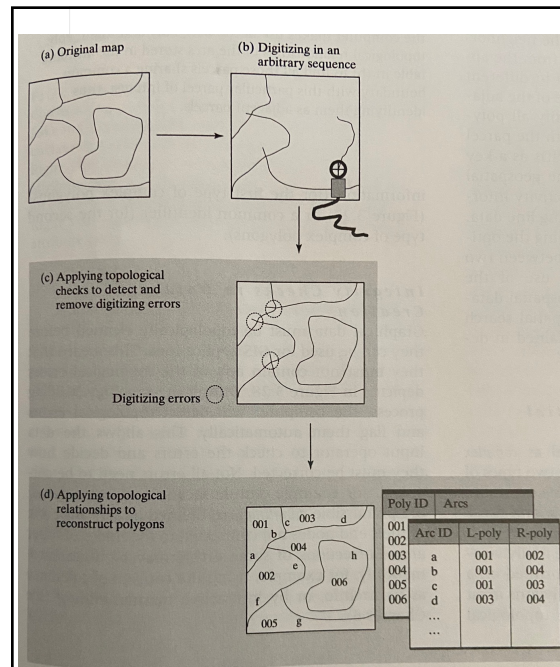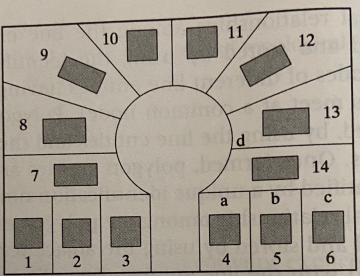
8

8

Topology

Line features can share endpoints *arc-node topology*		Area features can overlap with other area features *region topology*	
Area features can share boundaries *polygon topology*		Line features can share endpoint vertices with point features *node topology*	
Line features can share segments with other line features *route topology*		Point features can share vertices with line features *point events*	

Source: ESRI

9

9

Topological Relationships: Areas of Usage

- Data input and representation
  - Identify and store adjacency, connectivity & containment information
    - field data collection, digitizing
    - E.g. Left/Right of polygon or line

(a) Original map

(b) Digitizing in an arbitrary sequence

(c) Applying topological checks to detect and remove digitizing errors

Digitizing errors

(d) Applying topological relationships to reconstruct polygons

Poly ID	Arcs
001	
002	
003	
004	
005	
006	

Arc ID	L-poly	R-poly
a	001	002
b	001	004
c	001	003
d	003	004
...		
...		

Lo and Yeung (2006) Concepts and Techniques of Geographic Information Systems, Prentice Hall

10

10

Topological Relationships: Areas of Usage

- Spatial Search by Topological Relationships
  - Find features by adjacency, containment and connectivity
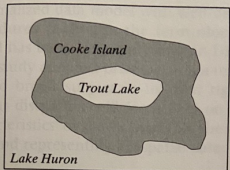
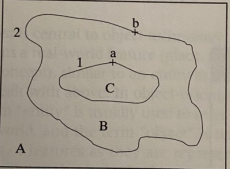Lo and Yeung (2006) Concepts and Techniques of Geographic Information Systems, Prentice Hall

11

11

Topological Relationships: Areas of Usage

- Construction of Complex Spatial Relationships
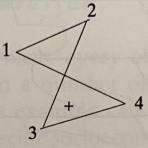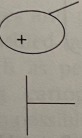 - E.g. Holes on polygons
  - Multipart polygons
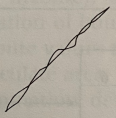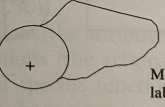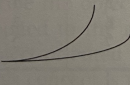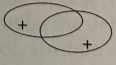
Lo and Yeung (2006) Concepts and Techniques of Geographic Information Systems, Prentice Hall

12

12

Common network topological errors	Common polygon topological errors
Floating node	Weird polygon
Dangling arcs (not all dangling arcs are errors)	Sliver polygon
Floating arc	Missing label
Overlapping arc	Overlapping polygons

Topological Relationships: Areas of Usage

- Integrity checks and balances
  - Essentially the removal topological errors
    - E.g. No dangling arcs or overlapping polygons

Lo and Yeung (2006) Concepts and Techniques of Geographic Information Systems, Prentice Hall

13

13

Data Representation and Analysis

- Vector and Raster data models are ideally suited for solving different problems
  - Raster Model: phenomena that exhibit spatial variability
    - Temperature, topography
    - Environmental aspects at a regional or national scale
      - Habitat, biodiversity analysis

14

14

Data Representation and Analysis

- Vector and Raster data models are ideally suited for solving different problems
  - Vector Model: phenomena necessitates the need for location and individuality of spatial objects
    - Land title management, transportation management and planning, facilities management
    - Environmental aspects at a local scale (as opposed to a regional or national scale)

15

15

Spatial Analysis

- Analysis occurs using either:
  - **Spatial location** e.g. wetlands near a pipeline; houses within 5km
    - Distinguishes GIS from a non-spatial analysis
  - **Data attributes** e.g. wetlands with black spruce; houses in a given price range
    - Distinguishes GIS from 'non-GIS' mapping software
  - **Comparative spatial statistics**
    - e.g. the correlation of variables across space
    - e.g. wetlands clustered at local scale, uniform at regional scale
    - Statistical Spatial Data Analysis with ArcGIS, QGIS,….

16

16

Spatial Analysis
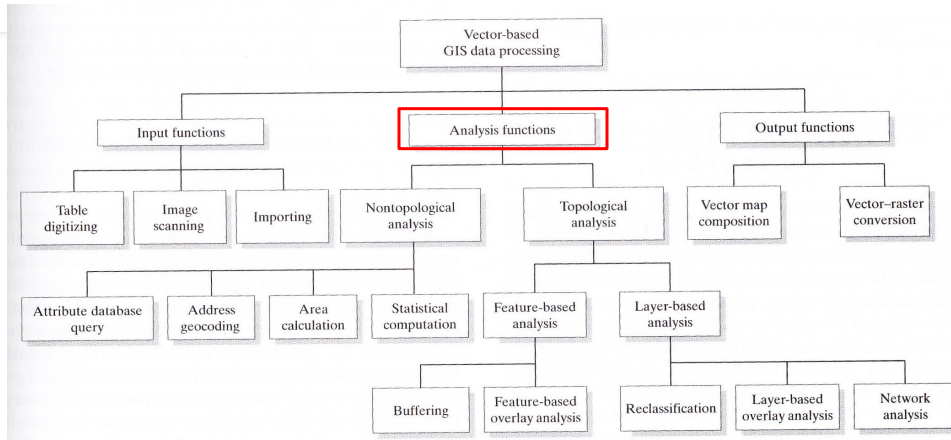
Source: C.P. Lo, A.K.W. Yeung (2007) Concepts and Techniques of Geographic Information Systems

17

17

Spatial Analysis
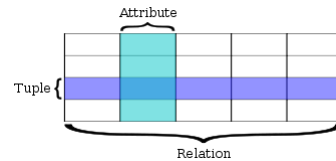
- Analysis functions
    - Non-topological functions
        - Attribute values of spatial data

    - Topological functions
        - Individual data objects
            - Feature-based
                - individual data objects
        - Layer-based
            - all objects in a layer

18

18

Nontopological Analysis

- SQL: <u>Structured Query Language</u>
- Goal: select only the features with (or without) particular values
- Relational table

19

Nontopological Analysis
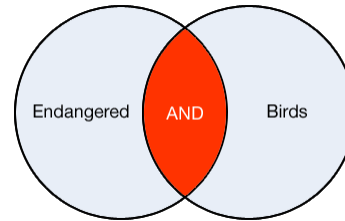
- Attribute Database Query Operations
  - **SELECT**: Extract data items in specified rows of a table

    SELECT <attribute_name> FROM <table> WHERE <condition_statement>

  - **JOIN (RELATIONAL JOIN):** Merges two tables based on the values in the columns of the tables

    [a select statement] *<table_1> JOIN <table_2> ON table_1.attribute= table_2.attribute* [a condition statement]
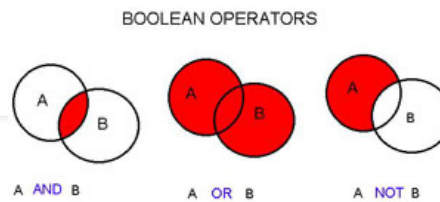
20

Nontopological Analysis

- **Boolean Logic** (after George Boole, a 19th century mathematician)
  - "Create an expression reducible to a true or false condition".
- **SQL can use:**
  - **Boolean Operators**
    - **And** (narrows the selection)
    - **Or** (expands the selection)
    - **Not** (excludes/reduces results)
    - **XOR** (A or B, but not both)
    - **Like (Similarity)**
      - **% is wildcard**

  - Relational Operators (=, >, <, >=, <=)
  - Arithmetic Operators (+, - *, /)

21

Boolean Operators (to combine conditions)

BOOLEAN OPERATORS

A AND B   A OR B   A NOT B

Selection by attribute

type = pine AND age > 100   ... selects all old growth pine

type = pine OR age > 100   selects all pine and any type older than 100

These may require the use of brackets to avoid ambiguity in complex queries e.g.

type = pine OR type = fir AND age > 100   selects any pine plus old growth fir

type = (pine OR type = fir) AND age > 100   selects old growth (pine and fir)

22

Nontopological Analysis

- Statistical Computation
  - Extension of attribute data query
  - Conventional statistical measures such as:
    - mean, maximum, minimum, range, standard deviation, frequency
    - Note: These are aspatial (non-spatial) statistics
    - Note: Can't do statistics on text field

23

23

Nontopological Analysis

- Spatial Calculation
  - Spatial calculation of position, area, perimeter, distance, proximity
    - Algorithms are well known
  - An intermediate step to obtain new attribute data to support specific spatial analyses
  - Be sure that data are preprocessed
    - projection, coordinate transformation, measurement unit conversion

24

24

Statistics and the Field Calculator

25

Selection in ArcGIS

- Manual selection
- By Location
- By Attributes

- Show attributes from
  - All layers
  - Selected layer
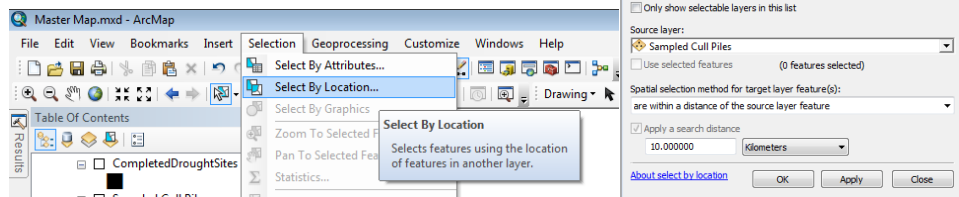 - Visible layers
  - Top-most layer

Once you have selected something, the selection will be all ArcGIS sees by default.
CAUTION when you export, analyze, compare

Select by Location

- Selects features from one data set based on their spatial relationship with another layer e.g.
  - Are within a distance of
  - Are within
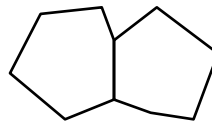 - Are completely within
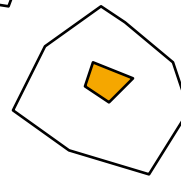 - Share a line segment with
  - Intersect

27

Topological Functions

- Topology: spatial relationship between entities

  - Adjacency (polygons)

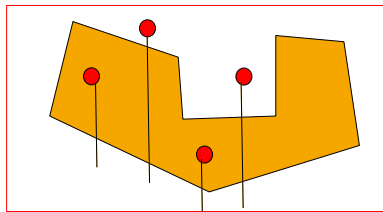 - Containment (e.g. points in polygons)

  - Connectivity (lines)

28

Topological Functions

- Feature-Based Analysis: Find relationships between features of one layer and those of another
  - Common functions: Point in polygon, adjacency, connectivity, buffering
    - Point in polygon matching: whether a given point feature falls inside a polygon

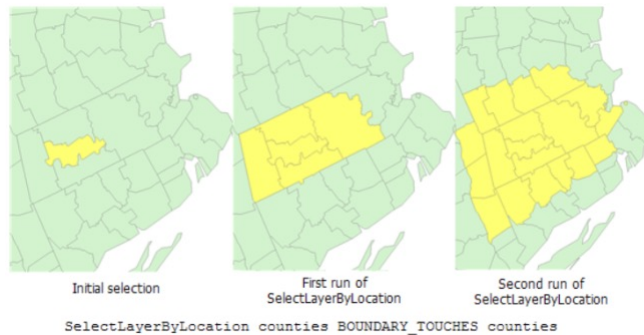**Plumb-line algorithm**

Odd number of intersections: **In**
Even number of intersections: **Out**

29

29

Touch = adjacency

Initial selection | First run of SelectLayerByLocation | Second run of SelectLayerByLocation

SelectLayerByLocation counties BOUNDARY_TOUCHES counties

Are connected with

Initial selection | First run of SelectLayerByLocation | Second run of SelectLayerByLocation
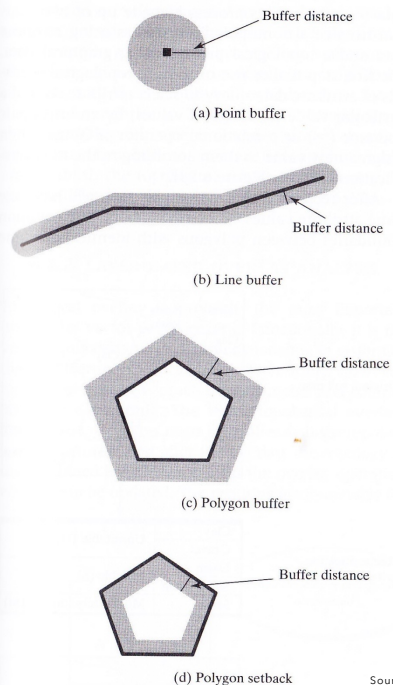
30

Topological Functions

- Buffering
  - A buffer is a zone covering a specified distance around a spatial feature
    - Buffer zones are polygons.
    - Buffer distance is usually user defined
    - Boundaries of overlapping buffer zones are automatically dissolved to give a single coherent buffer polygon

31

31

(a) Point buffer

(b) Line buffer

(c) Polygon buffer

(d) Polygon setback

Useful when evaluating charactering of an area surrounding a feature.

- properties/population with walking distance of a bus stop
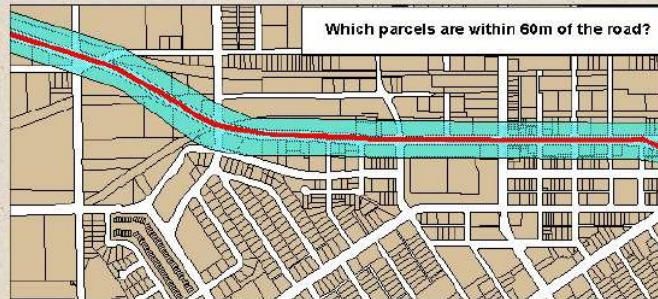- delineation of restricted zones around sensitive sites

32

Source: C.P. Lo, A.K.W. Yeung (2007) Concepts and Techniques of Geographic Information Systems

32

Applications of BUFFER

Which parcels are within 60m of the road?

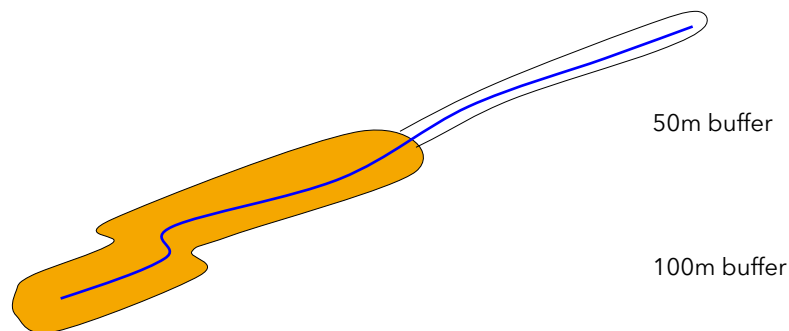For example: consideration of widening of existing roadway, to identify parcels affected by underground utility line additions, etc.

33

Topological Functions

- *Variable buffer distances* (assigned by attribute)
  - Different buffer size depending on stream attribute
  - e.g. stream protection status
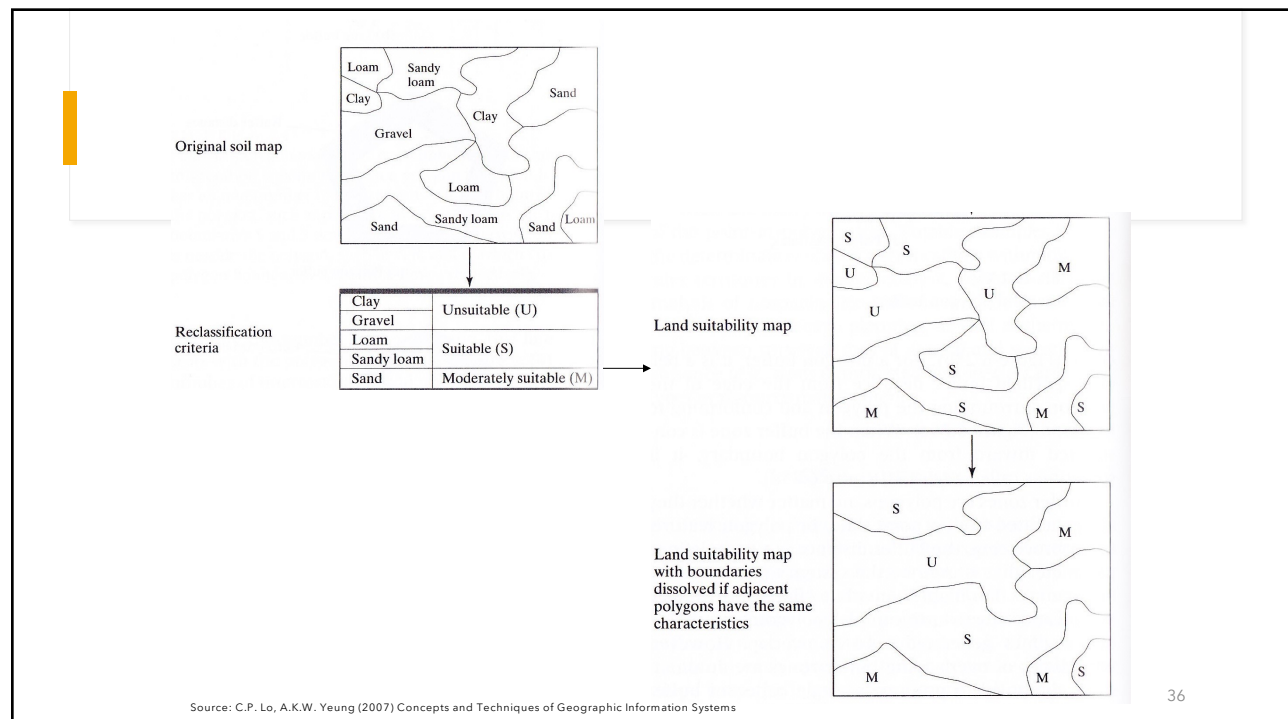

50m buffer

100m buffer

34

Topological Functions

- Reclassification
  - Database simplification process
    - reduces (modifies) the categories in attribute data
  - Two step process
    - Nontopological: select attribute e.g. a range of values and assign them a new class
    - Topological: dissolve according to new classification scheme

35

35

Original soil map

Reclassification criteria

Clay	Unsuitable (U)
Gravel	
Loam	Suitable (S)
Sandy loam	
Sand	Moderately suitable (M)

Land suitability map

Land suitability map with boundaries dissolved if adjacent polygons have the same characteristics
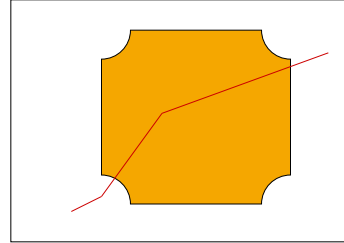
36

36

The Concept of GIS Overlay

- Overlay addresses the relationship of the overlap between spatial features.

- Overlay combines the spatial *and* attribute data of two input themes.

Three input feature types, overlay cover is always polygon:

1) point-in-polygon, ⟶ points are output
2) line-in-polygon, ⟶ lines are output
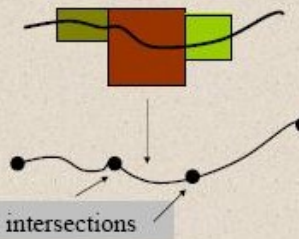3) polygon-in-polygon ⟶ polygons are output

37

Line Feature Types and Map Overlay
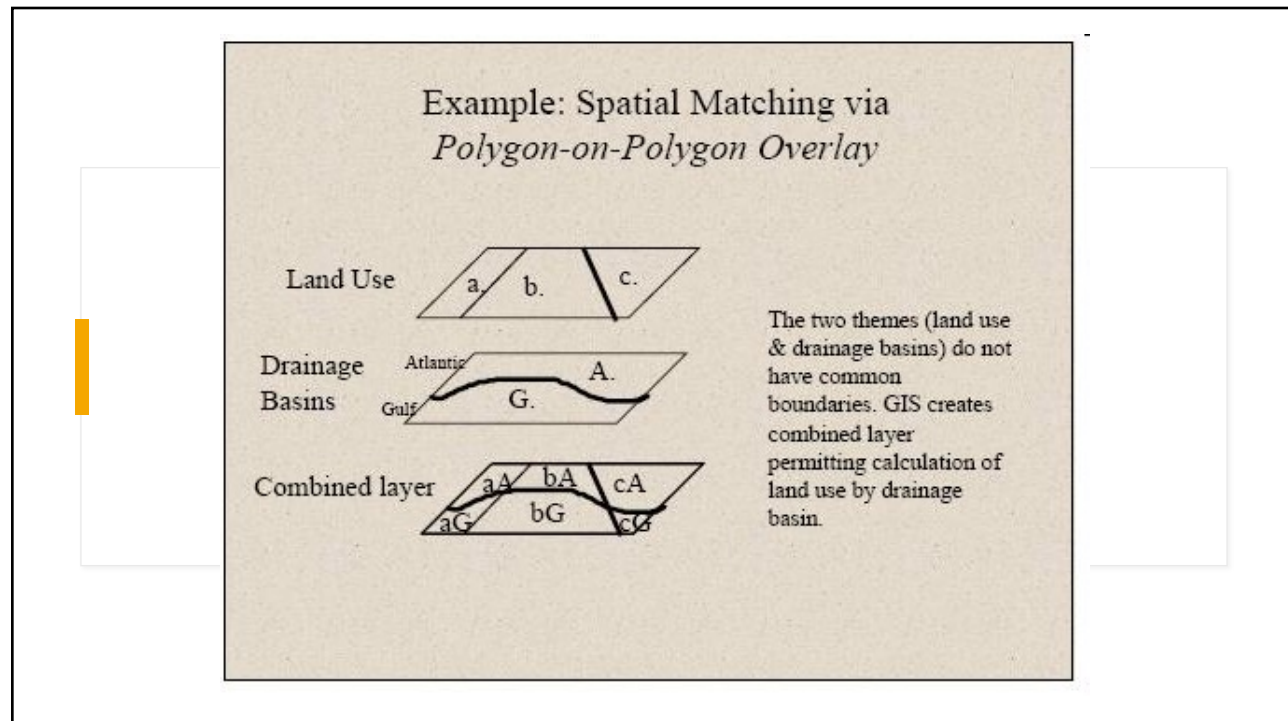
- Line - in - Polygon

  Input Theme   Overlay Theme

  Two new intersections

  – Attach attributes of polygon to line that traverses through polygon
  – Example - What types of soils does a proposed road cross?
  – Line is actually broken into different line segments at polygon boundaries. Each line segment takes on attributes for the polygon it overlays.
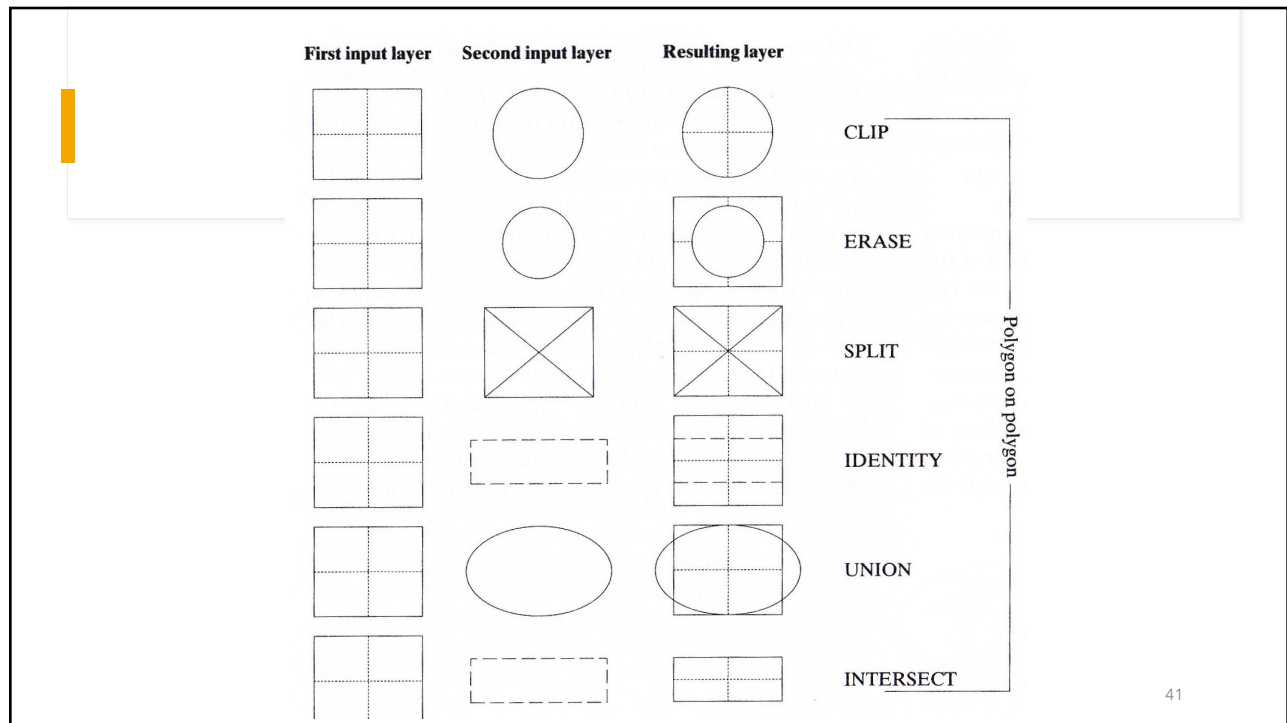
38

Example: Spatial Matching via *Polygon-on-Polygon Overlay*

The two themes (land use & drainage basins) do not have common boundaries. GIS creates combined layer permitting calculation of land use by drainage basin.

39

Topological Functions

- Topological Overlay Operators
  - UNION: Overlays polygons and keeps all areas in both layers. Never clips data
  - INTERSECT: Keeps portions of the first input layer that fall within the second input layer
  - IDENTITY: Overlays polygons and keeps all input features. The input features that overlap identity features will get the attributes of the identity features.
  - CLIP: Cuts out the first input layer using the second input layer as a cookie cutter
  - ERASE: Erases part of the first input layer using the second input layer
  - SPLIT: Divide the polygons in the first input layer into a number of smaller polygons based on the second input layer

40

40

First input layer	Second input layer	Resulting layer		
			CLIP	
			ERASE	
			SPLIT	Polygon on polygon
			IDENTITY	
			UNION	
			INTERSECT	

41

41