Lecture 11
Data Manipulation and Analysis

15-12-2017

Ref:
1. Lecture Notes by Dr. Nitin Kumar Tripathi, AIT Thailand
2. http://libraries.mit.edu/gis/teach/iap05/vectorprocessing.ppt

Data Manipulation & Analysis

- Manipulation:
  - Deals with handling spatial data for a particular purpose.
- Analysis:
  - Deals with the discovery of general principles underlying the total phenomenon
  - KNOWLEDGE discovery is finding INFORMATION hidden in DATA
  - e.g. if GDP and Literacy are related or not, etc.

Operations in Data Manipulations & Analysis

1. Reclassification and Aggregation
2. Geometric Operations
   - Rotation, Translation, Scaling,
   - Geometric Rectification
3. Centroid Determination
4. Data Structure Conversion (R2V, V2R)
5. Spatial Operations

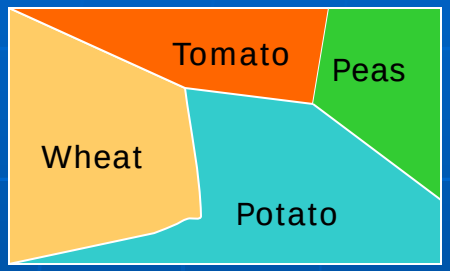
   Connectivity and Neighborhood Operations
6. Measurements
   - Distance and Directions
   - Statistical Analysis, Descriptive Statistics
   - Regression, Correlation and Cross Tabulation
7. Modeling

1 Reclassification and Aggregation

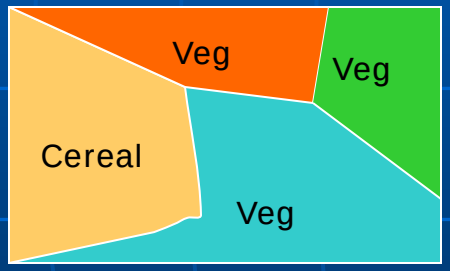
- Data may not be compatible with the user need and/or further analysis
- Data may be at different resolution
- Some time multiple attributes are there and we want to aggregate / simplify the attributes (Aggregation)
- In Raster, Aggregation is used for increasing Cell Size
  - e.g. if there are 2 features, urban and rural areas, and we want to aggregate cells, then new feature will be selected based on Majority Rule.
- In Vector, it is used for Sliver Removing
  - Slivers are small polygons (unwanted) generated after digitizing, or after R2V conversion.

Poly Aggregation

For Vector Data



Crop Type Map



Recoded Map of Cereal and Veg.

Redundant Boundaries Removed

Raster Aggregation

1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	2	1	1	1	1	1	1	1
1	1	1	2	2	1	1	1	1	1
1	1	1	2	2	2	2	2	2	2
1	1	1	1	2	1	2	2	2	2
1	1	1	1	2	2	2	2	2	2
1	1	1	1	2	2	2	2	2	2

1 = urban
2= sub-urban

1	1	1	1	1
1	2	1	1	1
1	1	2	2	2
1	1	2	2	2

After applying majority Filter on 2x2 windows

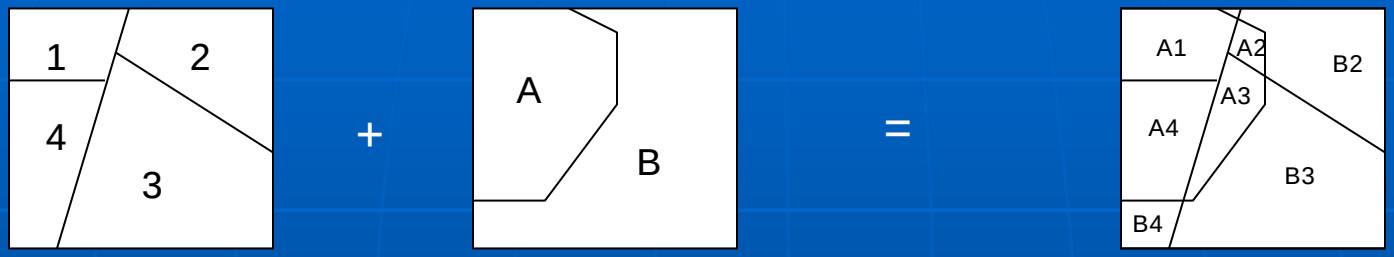
In Case of Tie?

Select 2 or 1 or declare a Mixed Area 3 (users choice)

Map Overlay and Dissolve

- Overlay and Dissolve involves Composition (Integrating) or Extracting (Disintegrating) of multiple maps in order to create new map.
- Polygon Overlay:
  - combining two or more polygon maps
  - intersection of polygons of two or more maps
  - It will create new map having many smaller polygons and attributes from both the maps (input map & overlay map)
- Mathematical Overlay:
  - is performed for the purpose of area and measurement, as well as multiple attribute modeling
  - it involves weighting of various parameters and summation of the weighted parameters to create maps such as Site Suitability maps, or capability assessment maps.

Polygon + Polygon Overlay



Landuse Map

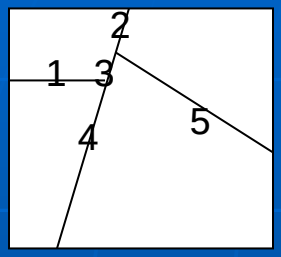
ID	LandUse
1	Agri
2	Forest
3	Residents
4	Commercial

District map

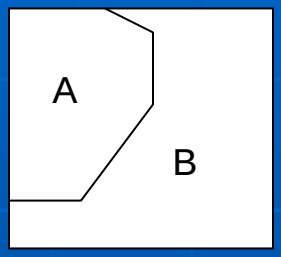
ID	District
A	Lahore
B	Qasur

ID	Landuse	District
A1	Agri	Lahore
A2	Forest	Lahore
A3	Residents	Lahore
A4	Commercial	Lahore
B1	Agri	Qasur
B2	Forest	Qasur
B3	Residents	Qasur
B4	Commercial	Qasur

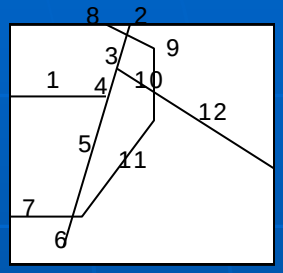
Overlay: Line + Polygon



+



=



Road Map

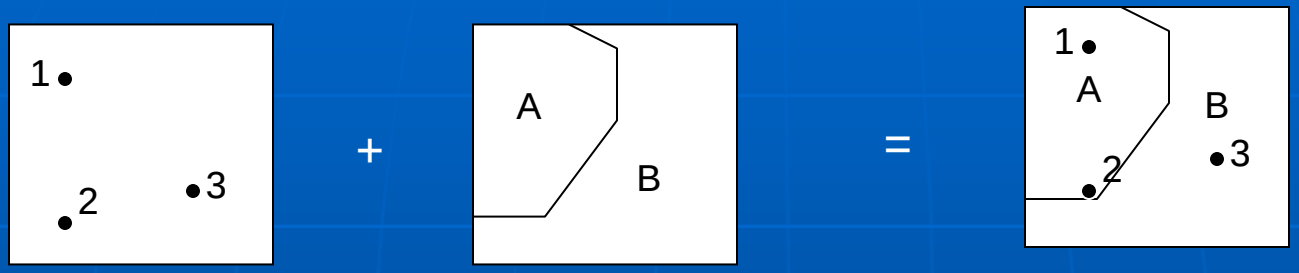
ID	Road Name
1	
2	
3	
4	
5	

District map

ID	District
A	Lahore
B	Qasur

ID	Road	District_L	District_R
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			

Overlay: Point + Polygon



Hospital Map

ID	Hospital
1	AAA
2	YYY
3	ZZZ

District map

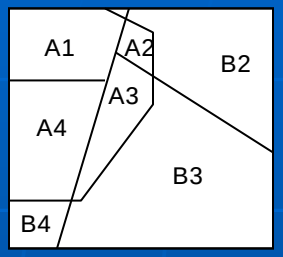
ID	District
A	Lahore
B	Qasur

ID	Hospital	District
1	AAA	A
2	YYY	A
3	ZZZ	B

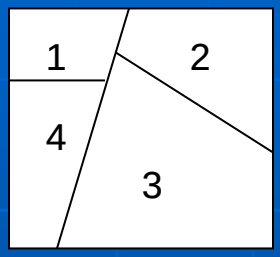
Map Overlay and Dissolve

- Map Dissolve
  - Involves inverse of polygon overlay
  - To extract a single attribute from a multiple attribute polygons
  - Similar to Aggregation
  - Helps in extracting a new map having less attributes from a map having multiple attributes

Map Dissolve

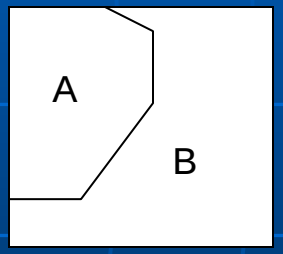


Dissolved based on Landuse



ID	Landuse
1	Agri
2	Forest
3	Residents
4	Commercial

Dissolved based on District



ID	District
A	Lahore
B	Qasur

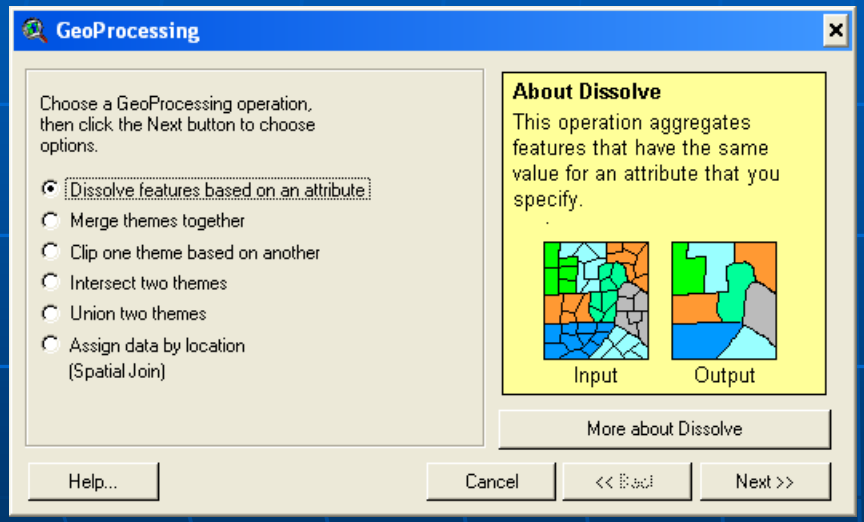
ID	Landuse	District
A1	Agri	Lahore
A2	Forest	Lahore
A3	Residents	Lahore
A4	Commercial	Lahore
B1	Agri	Qasur
B2	Forest	Qasur
B3	Residents	Qasur
B4	Commercial	Qasur

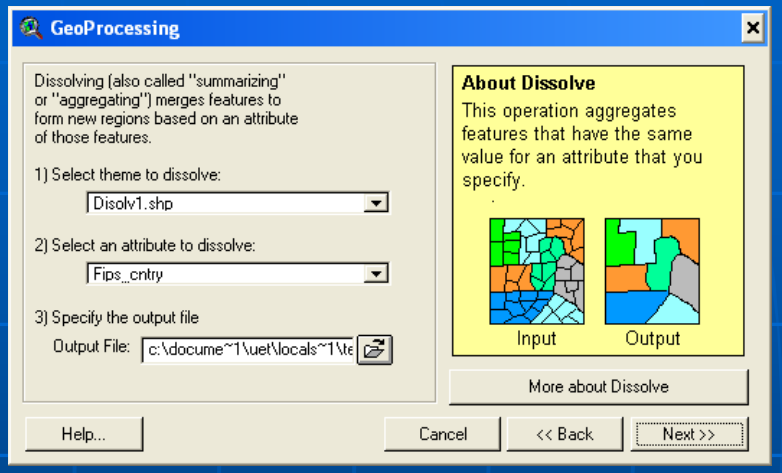
Aggregation in Arcview

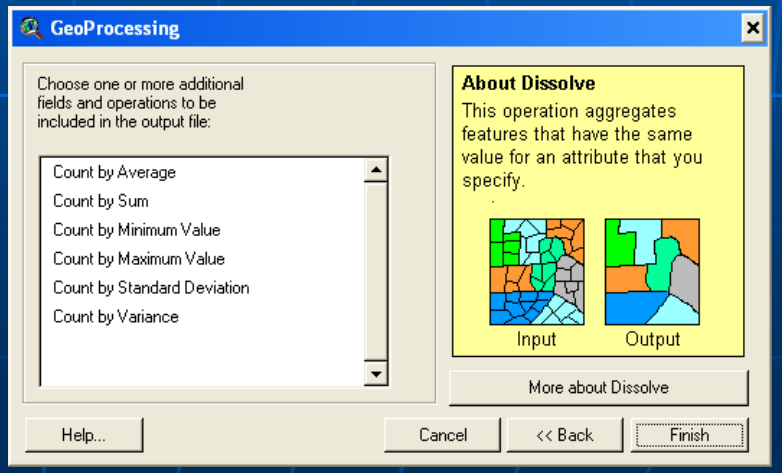
- **Dissolve**
  - This process will dissolve features in one theme based on the value of an attribute you choose. Technically this is the same as showing the attribute table for a theme, then choosing to Summarize an attribute where you elect to merge_shapes.

Aggregation in Arcview

DISSOLVE

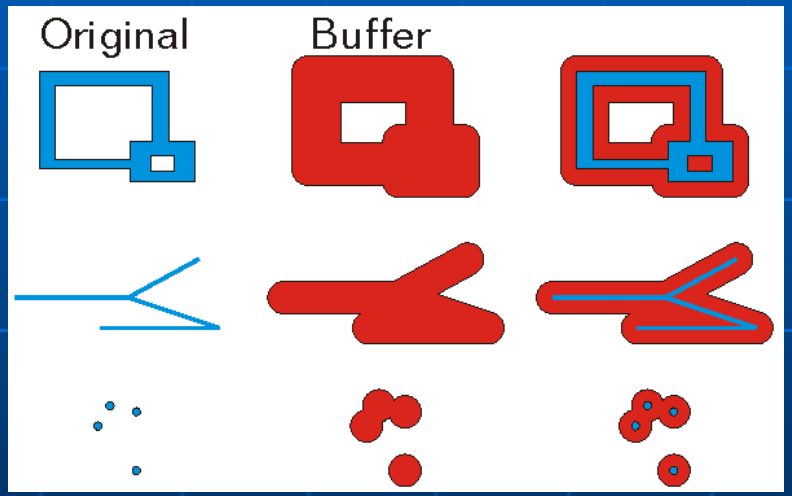






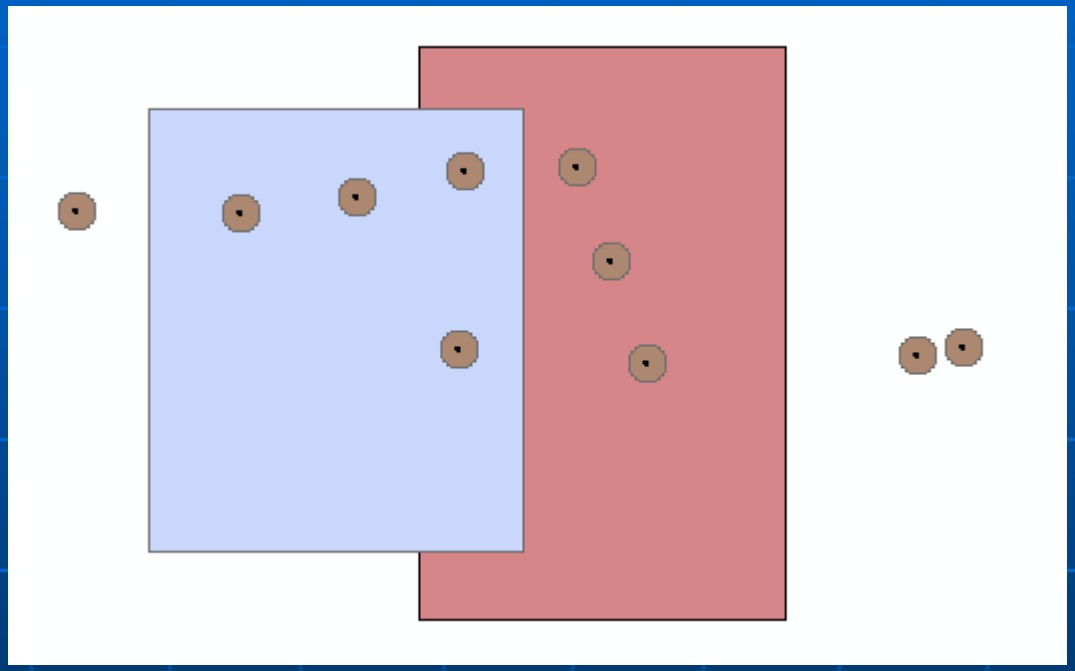
Buffer Generation

- Generation of new polygon around points, lines, or polygon features, at user specified buffer distance
- Square or circular buffer can be calculated
- If Buffer Option is disabled in Arc view, then select the Distance Units in ‘View: Properties’



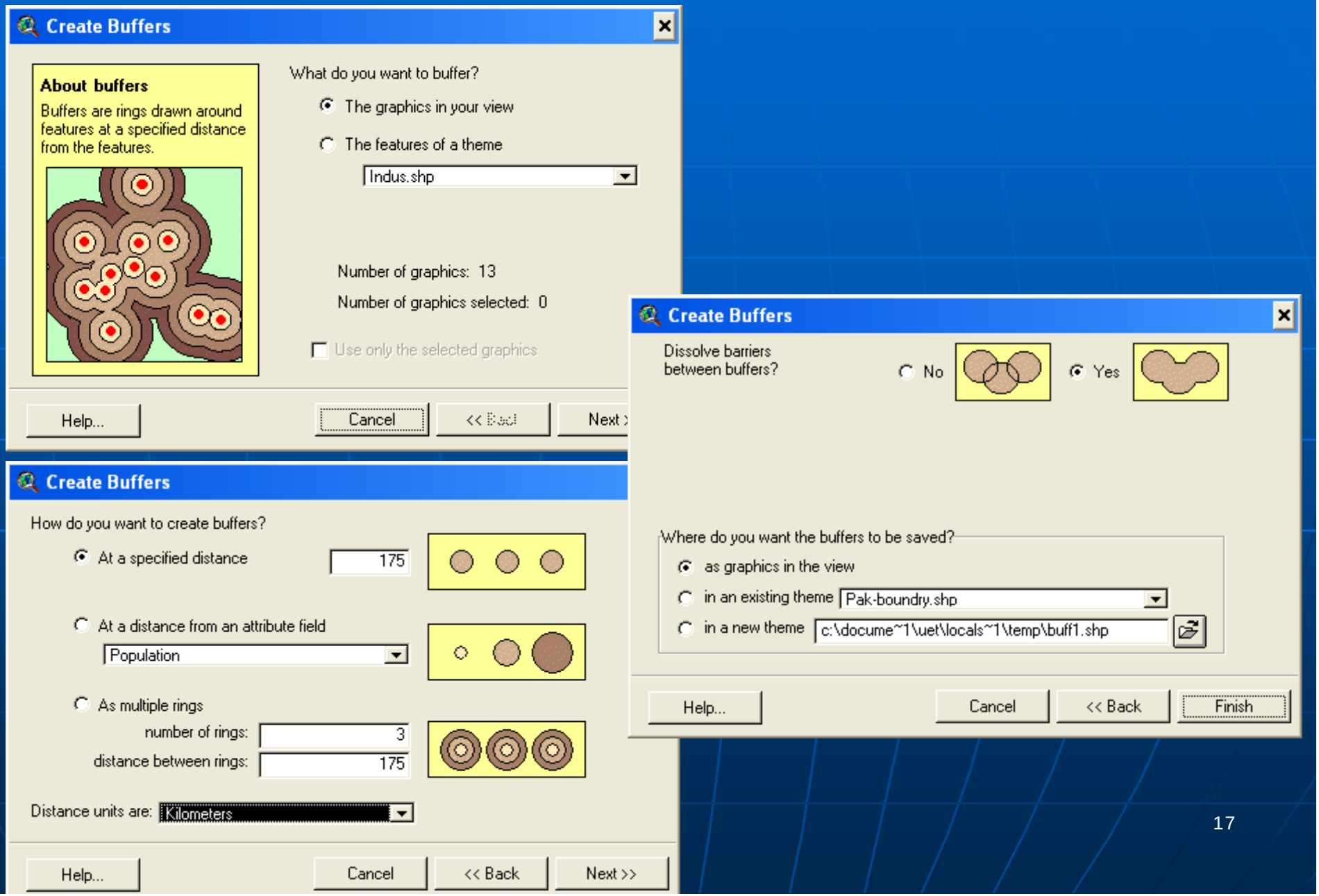
Buffer

Buffers don't share the attributes of the feature that they surround. Use spatial Joins to add the attributes.

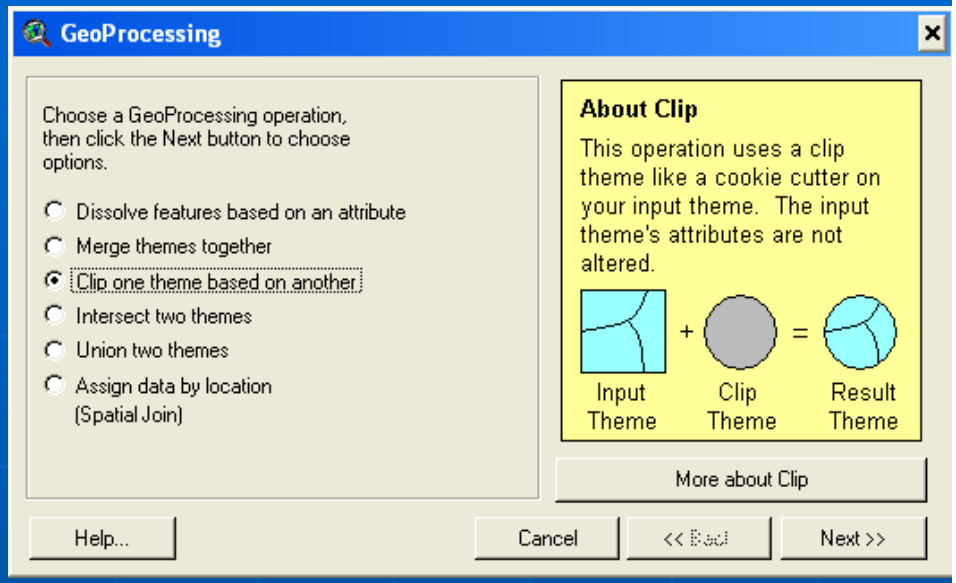


Original points (black) are surrounded by a buffer of 25 meters.

Ref: http://libraries.mit.edu/gis/teach/iap05/vectorprocessing.p


Create Buffers
About buffers
Buffers are rings drawn around features at a specified distance from the features.
What do you want to buffer?
The graphics in your view
The features of a theme
Indus.shp
Number of graphics: 13
Number of graphics selected: 0
Use only the selected graphics
Help...
Cancel
<< Back
Next >
Create Buffers
How do you want to create buffers?
At a specified distance
175
At a distance from an attribute field
Population
As multiple rings
number of rings:
3
distance between rings:
175
Distance units are:
Kilometers
Help...
Cancel
<< Back
Next >>
Create Buffers
Dissolve barriers between buffers?
No
Yes
Where do you want the buffers to be saved?
as graphics in the view
in an existing theme
Pak-boundry.shp
in a new theme
c:\docume~1\uet\locals~1\temp\buff1.shp
Help...
Cancel
<< Back
Finish

Overlay in Arcview



- **Clip**
  - "Input" theme , "overlay" theme
  - One of these themes must be a polygon theme (called the "overlay" theme)
  - It will be used to define the clipping region.
  - The clip process uses the clipping region as a cookie cutter.
  - Only those features in the other theme (called the "Input" theme) within the clipping region are stored in the new Shapefile.
  - The features of the Input theme can be either of polygons, lines, or points.
  - The features in the new Shapefile will be of the same type as the input theme's features. The feature attribute table for the new Shapefile contains the same items as the input theme's attribute table.

Clip

Input Features

Clip Features

Output Feature Class

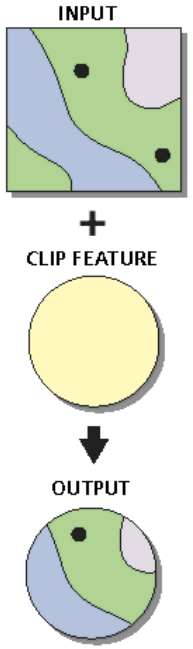
XY Tolerance (optional)

Decimal degrees

Clip

Extracts input features that overlay the clip features.

Use this tool to cut out a piece of one feature class using one or more of the features in another feature class as a cookie cutter. This is particularly useful for creating a new feature class—also referred to as study area or area of interest (AOI)—that contains a geographic subset of the features in another, larger feature class.



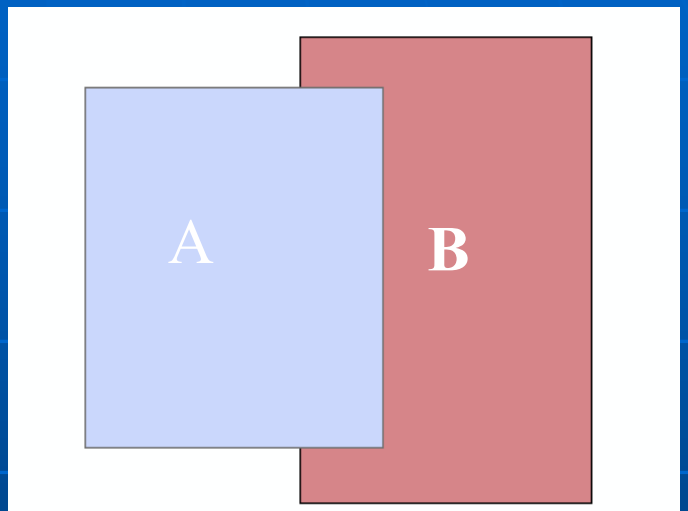
OK | Cancel | Environments... | << Hide Help | Tool Help

Type here to search

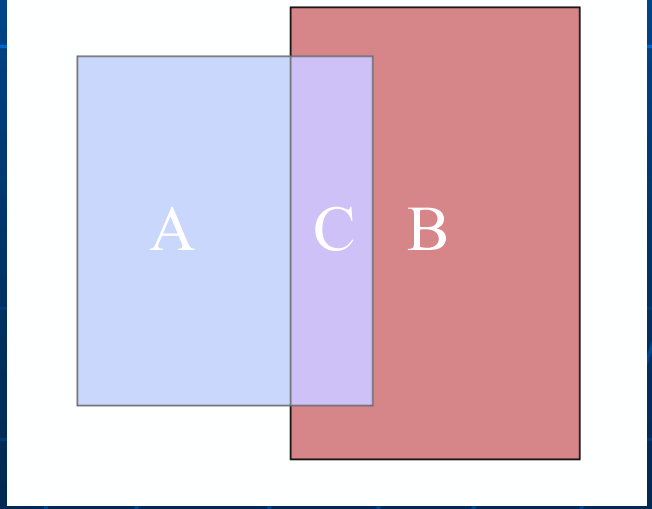
ENG 4:52 PM 22-12-17

Clip

Two polygons, A and B, Overlap. Clip A using B as a cookie cutter.

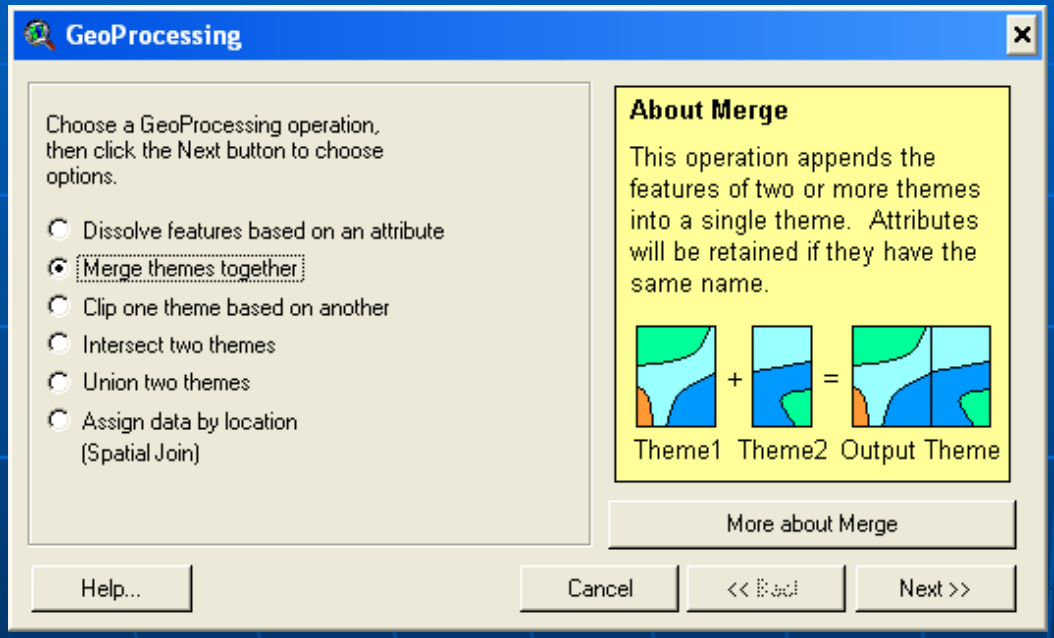


Clip operation creates a new polygon, C, which is the intersect, or overlap, of A and B. Attributes of A do not appear in C.



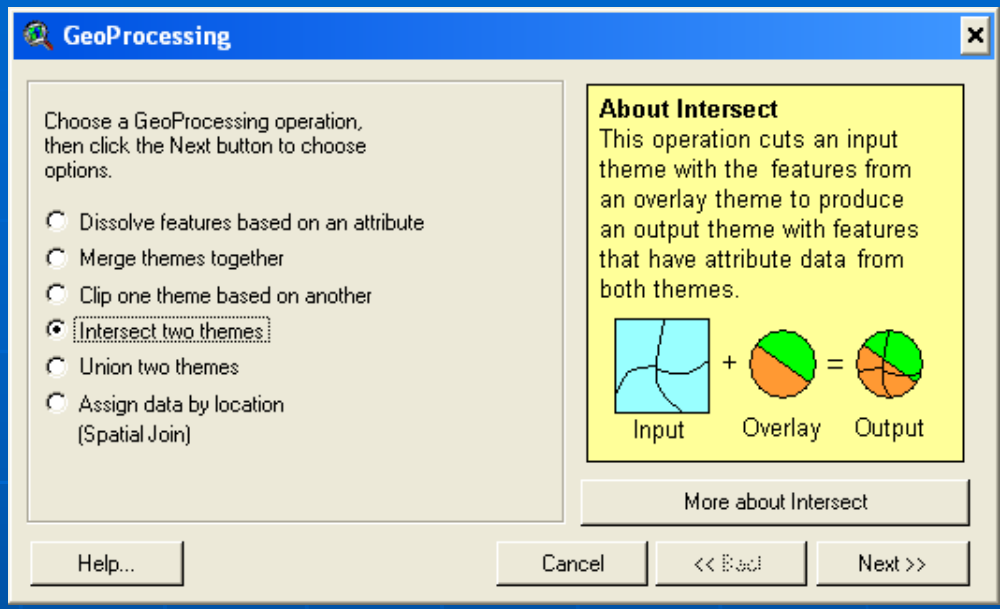
Overlay in Arcview

Merge



- The new theme will contain the fields of one of the input themes.
- If each of the other input themes have at least the same fields, then all cells in the new theme's attribute table will be populated.
- If any of the other input themes have additional fields, that data will not be included.
- If any of the other input themes are missing the fields then no data will be added to those fields for the features of that other theme.

Overlay in Arcview

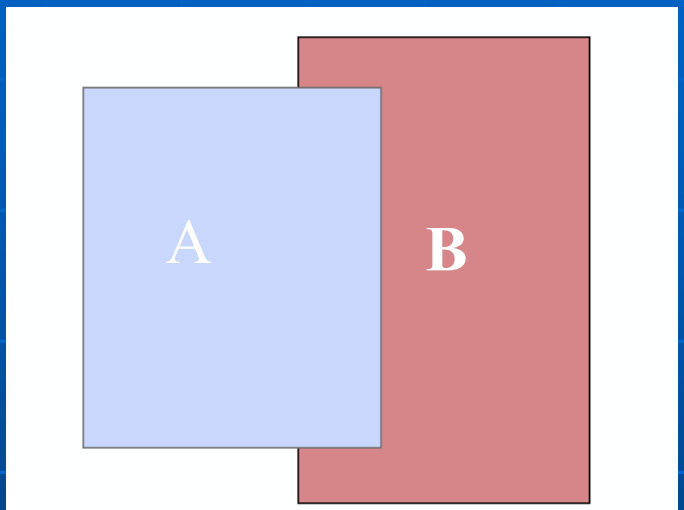


- Intersect
  - “Input" theme , "overlay" theme
  - The overlay theme must be a polygon theme.
  - Its features will split the input theme.
  - Any features in the input theme that are not overlaid by features in the overlay theme will be ignored.
  - Only the areas that are in common to the overlay and input themes will be included in the output theme.
  - The input theme’s features can be either polygons or lines.
  - The output Shapefile’s features will be of the same type as the input theme features.
  - The attribute table for the output Shapefile includes the attributes from the input and overlay theme.

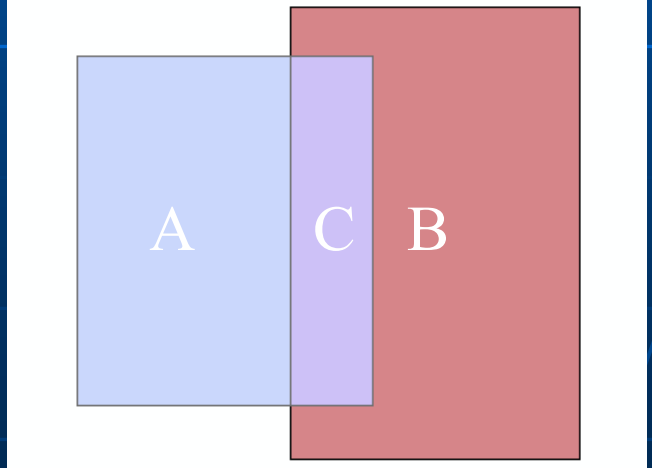
Overlay in Arcview

Intersect

Two polygons, A and B, Overlap. Find the Intersection of A using B.

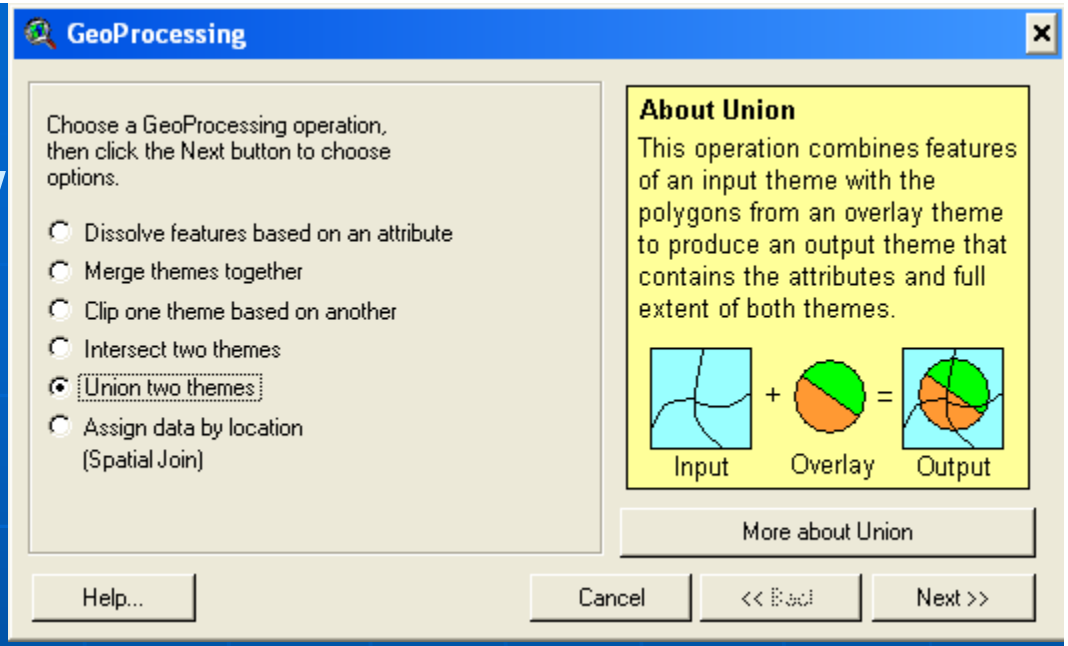


Intersect operation creates a new polygon, C, which is the intersection, or overlap, of A and B. Attributes of A and B do appear in C.



Overlay in Arcview

Union

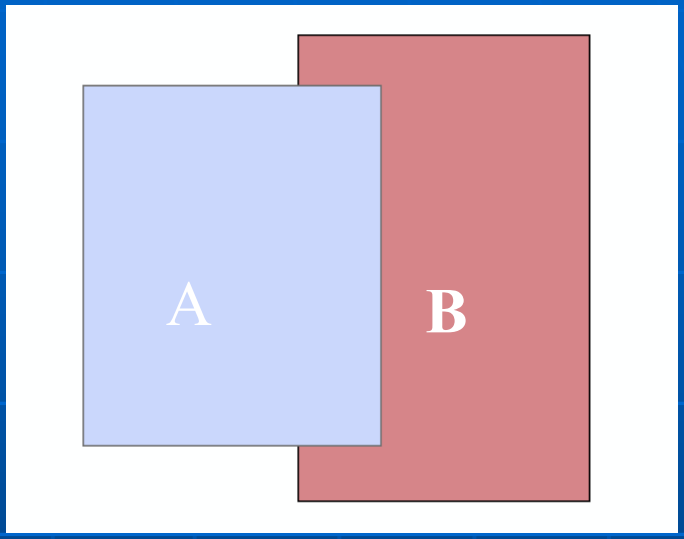


- "Input" theme , "overlay" theme
- One of these themes must be a polygon theme (called the "overlay" theme)
- The output theme contains the combined polygons and attributes of both themes.
- The polygons of the input theme are split at their intersection with polygons of the overlay theme.
- The feature attribute table for the output theme contains attributes from the input and overlay themes' attribute tables.

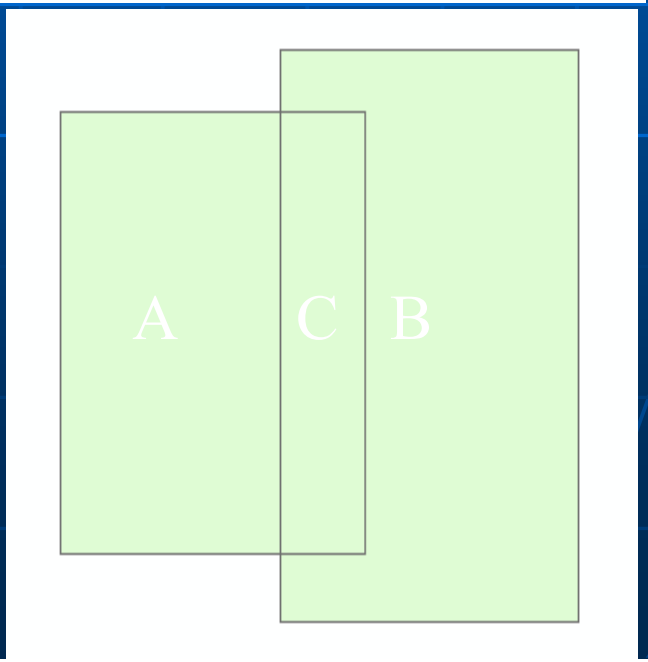
Overlay in Arcview

Union

Two polygons, A and B, Overlap. Find the Union of A & B.

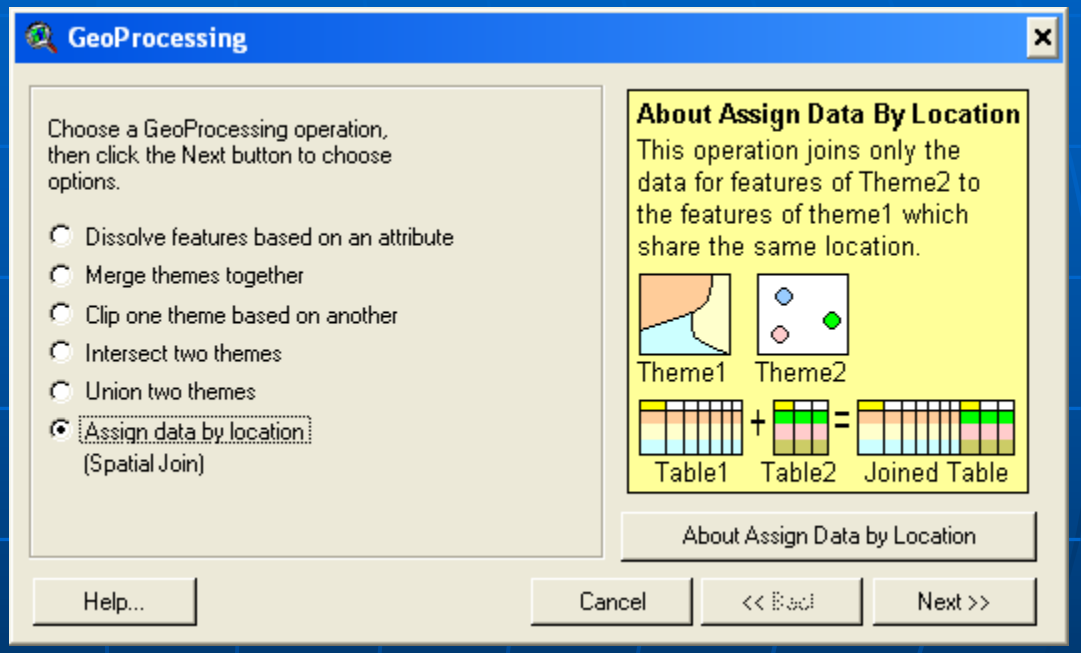


Union will create a new polygon, C, which is the intersect, or overlap, of A and B.
A, B and C all will be part of the new Union theme and will retain their attributes.



Assign by Location

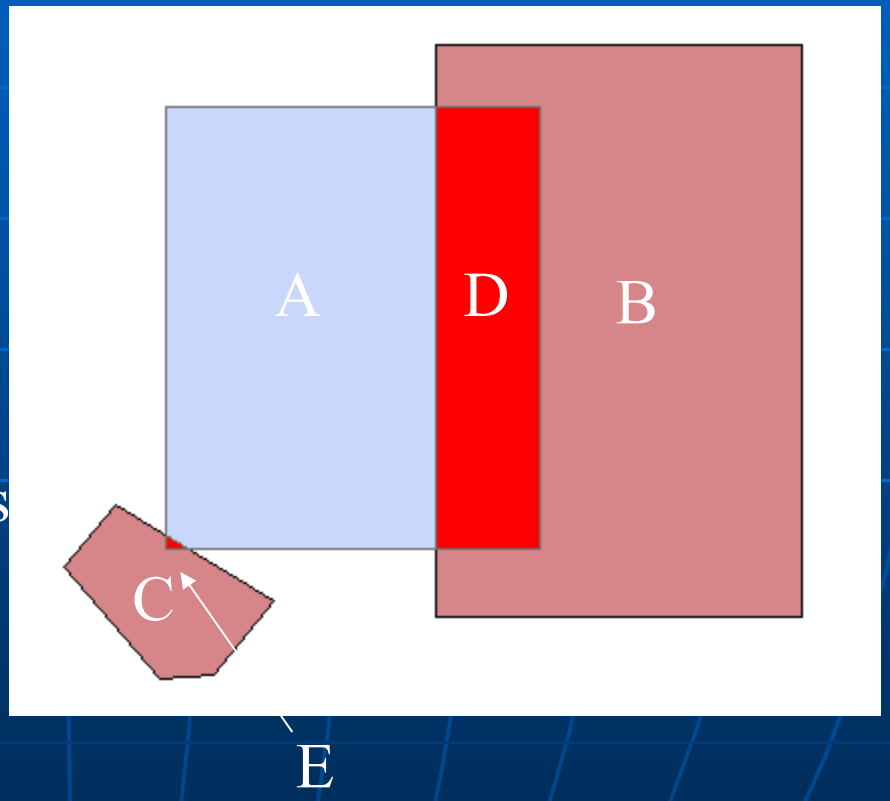
- The assign data by location operation will perform a spatial join between the two selected themes.
- A spatial join is similar to an attribute join; however it is based on the spatial relationship between the features in the two themes.
- For each feature represented in the destination table, ArcView looks to see if it has one of the following spatial relationships with any feature represented in the source table, and if it does, that feature's record from the source table is joined into the destination table:



Sliver polygons

Overlay operations often produce sliver polygons, which may or may not be meaningful.

The intersection of polygon A with a layer containing polygons C and D produce a layer with polygons D and E. E is a sliver polygon and may be considered noise.



Precision vs accuracy in overlay operations

- Sliver polygons – meaningful?
  - Decide by size, dimensions, number of arcs, but there is no hard and fast rule.

Map Abstraction

- Calculation of Centroid
- Automatic Contouring
- Proximal Mapping
- Reclassification
- Conversion to Grid

Centroid Determination

- Average location of a line or polygon
- Centre of Mass of a two-or-three dimensional objects
- Measured by averaging the location of infinitely small area elements within polygon or
- By averaging the coordinates of raster cells

Connectivity operation

- Network Analysis
  - Optimum corridor or travel Path Selection
  - Finding shortest path by time, or by distance
  - Hydrology: Discharge estimation
  - Identify the boundaries of separate water sheds, from the run-off direction

Statistical Analysis

- Descriptive Statistics
  - Mean, Medean, Mod, St Deviation, Range
- Histogram Statistics
- Extreme Values
- Correlation or Cross Tabluation

Raster Data Overlay

- Applies on cell to cell basis
- Arithmetic Operation on Raster Data
  - R=P+Q
  - R=(P*P-Q*Q)/2
  - .....
- Logical:
  - If P>30, R= 1; else R=0
    - .....

P and Q are input Raster Images
R is the resultant image