Vector data analysis

Rama Chandra Prasad
Lab for Spatial Informatics, IIIT Hyderabad

December, 11th 2015

Vector Data Models/Structures

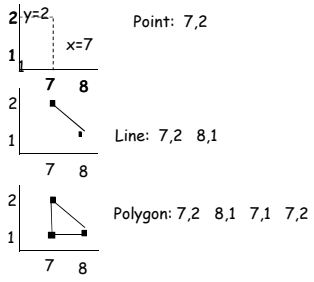
- One model for representing geographic space
- Spatial locations are explicit
- Relationships between entities/objects are implicit
- The vector data model provides for the precise positioning of features in space.
- Based on analytical geometry, a vector model builds a complex representation from primitive objects for the dimensions: **points, lines and areas**

In Summary- **Vector described as a quantity with a starting coordinate and an associated displacement and direction.**

Concept of Vector

- The basic elements of vector data model are **points;**

- point (node): 0-dimension
  - single x,y coordinate pair
  - zero area
  - tree, oil well, label location
- line (arc): 1-dimension
  - two (or more) connected x,y coordinates
  - road, stream
- polygon : 2-dimensions
  - four or more ordered and connected x,y coordinates
  - first and last x,y pairs are the same
  - encloses an area
  - census tracts, county, lake

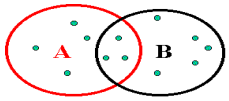


Representation of data

- Tied to scale and data type
  - City as point or polygon? Do you need to see the area of a city?
  - Road centerlines vs entire road system (usually to the curb)
- Vector objects in a GIS have defined boundaries, unlike nature (transition in forests)

Overlay operations in a GIS

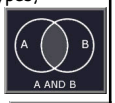
- Origins in Landscape Planning
  - Literally overlaying maps on a light table and searching for overlapping areas
  - GIS started out at GSD in the 1960s
- Set theory – polygons represent sets, overlay represents intersects and unions



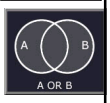
- Computational Geometry

Boolean Logic in GIS

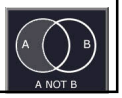
- Boolean Algebra is useful for performing operations on the attributes (which may be positional or descriptive) attached to geographic entities in a GIS.
- Boolean Logic is especially useful in computing (or modelling) new attributes in topological overlay processing for both vector and raster based systems, as they can be applied to all data types, be they Boolean, Ratio, Interval, Ordinal, or Nominal.



- AND operator ($\cap$) is the intersection of two sets - those entities that belong to both set A and B ($A \cap B$)

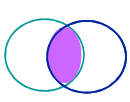


- OR operator ($\cup$) is the union of two sets - those entities that belong to either set A or to set B ($A \cup B$)

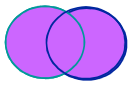


- NOT operator ($\neg$) is the difference operator - identifying those entities that belong to A but not to B ($A \neg B$)

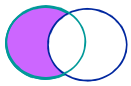
Boolean Operators



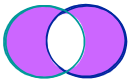
A AND B



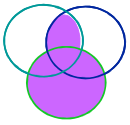
A OR B



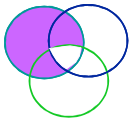
A NOT B



A XOR B



(A AND B) OR C



A AND (B OR C)

Classification of Vector Overlay Operations

Topological vector overlay operations can be classified via two methods:

1. Through the elements contained in the layers to be overlaid
2. By operation type (for example; the user wants to generate a layer comprising of the Union, Intersection, or some other boolean operation of the two input layers).

For Element based method

Element types	Point	Lines	Polygons
Points	Points Coincide	Point in Line	Point in Polygon
Lines	Point in Line	Line intersection	Line in Polygon
Polygons	Point in Polygon	Line in Polygon	Polygon Overlay

Complex databases such as GIS classify vector overlay operations via method two, where the particular overlay operation a user wishes to perform defines which element types may be contained in the two input layers.

Simplest form of overlay

Point in polygon procedure

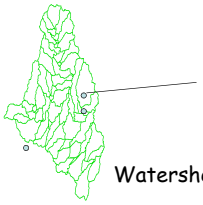
Line intersects 1 edge of polygon

Odd number of Intersections = inside, Even = outside

- Count how many intersections of the ray, originating at point A, pass through edges of the polygon

Point in polygon

- In its simplest form, the point in polygon operation determines whether a given point lies inside or outside a given polygon.
- In more elaborate forms there may be many polygons, and many points, and the tasks is to assign points to polygons. If the polygons overlaps, it is possible that a given point lies in one, many, or no polygons, depending on its location.



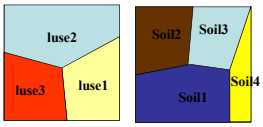
Rainfall gauging stations

Watershed boundaries

Polygon overlay

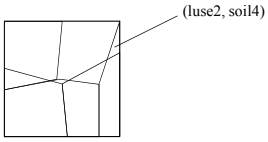
- It is similar to point in polygon transformation in the sense that two sets of objects are involved, but in this case both area polygons.
- The complexity of computing a polygon overlay was one of the greatest barriers to the development of vector GIS.
- Polygon overlay has different meanings from the field and discrete object perspectives.
- Raster overlay is simpler, but it produces a fundamentally different kind of result. When two raster layers are overlaid, the attributes of each cell are combined according to a set of rules.
- In vector overlay there is no rule for combination, and instead the result of overlay contains all of the input information, rearranged and combined.

Polygon overlay

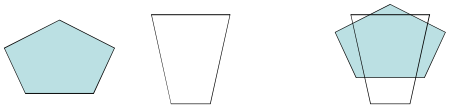


Landuse

Soil



Field case: all polygons with landuse and soil attributes



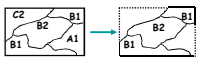
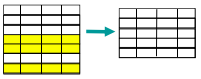
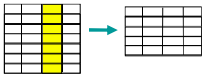
Discrete object case: 7 polygons after overlay

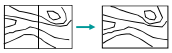
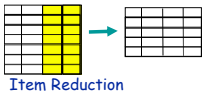
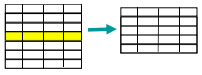
Geographic Query/Reporting

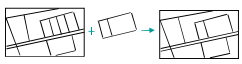
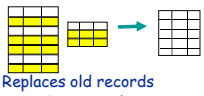
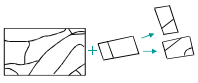
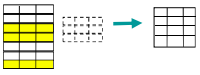
Use of database query and tabular report generation tools to perform special geographic queries and generate hard copy reports, displays, or maps.

Examples:

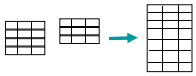
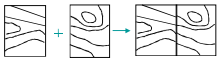
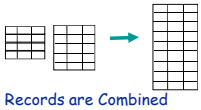
- Display and produce tabular summary of all water quality permit sites exceeding discharge standards
- Identify and display all water lines of a specified age and diameter

Logical Operators	Spatial Results	Tabular Results
RESLECT Selection by logical expressions		 Subset of existing features (same features, but fewer of them)	Subset of records (same items & same values, but fewer records)
DISSOLVE (one item) Selection by item		 Polygons with same values merge (fewer polygons)	Item Reduction (less items, fewer records, new User-IDs & adjusted areas)

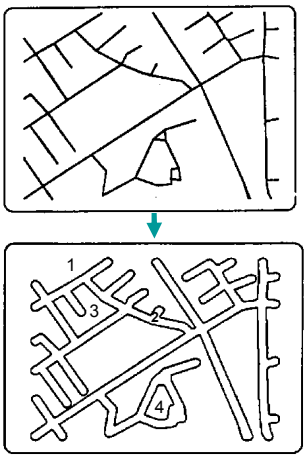
Logical Operators	Spatial Results	Tabular Results
DISSOLVE (all items) Selection by #ALL option	 Polygons with same values merge (fewer polygons)	 Item Reduction (same items, fewer records, areas adjusted & feature IDs can be maintained or offset.)
ELIMINATE Selection by logical expressions	 Polygons merged to neighbor by removing longest shared border (fewer polygons)	 Record Reduction (same items, fewer records, (areas adjusted & User-IDs of eliminated polygons are deleted)

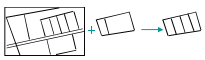
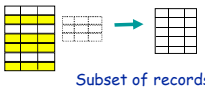
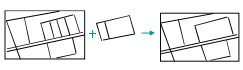
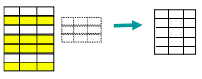
Boundary Operations	Spatial Results	Tabular Results
UPDATE POLY NET	 Replace existing polygons ~ external polygon boundary is used to 'cut & paste" new features ~ updating uses topological	 Replaces old records ~ attributes in [update_cover] must match [in_cover] ~ User-IDs are renumbered (unique values)
SPLIT POLY LINE POINT NET LINK	 Subset of existing features ~ retains features inside internal polygon boundaries ~ splitting uses topological overlay	 Subset of records ~ same attributes as [in_cover] but fewer records, areas adjusted & User-IDs are renumbered (unique values)

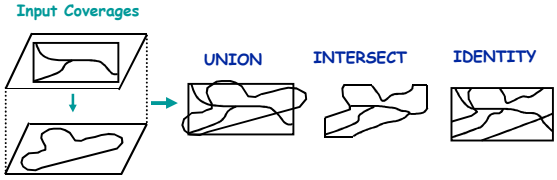
16

Boundary Operations	Spatial Results	Tabular Results
MAPJOIN POLY NET		 Merge Adjacent polygons ~polygons from many coverages are combined, including their borders Records are Combined ~ attributes in all coverages being combined must match ~ User-IDs can be offset or remain the same (Duplicate values possible)
APPEND POLY LINE POINT NET LINK		 Merge similar features adjacent or coincident ~ features from several coverages are combined ~topology must be reestablished with CLEAN or BUILD Records are Combined ~ attributes in all coverages being combined must match ~ User-IDs can be offset or remain the same (Duplicate values possible)

Buffer

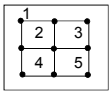


Boundary Operations	Spatial Results	Tabular Results
CLIP POLY LINE POINT NET LINK		 Subset of existing features ~ retains features inside the external polygon boundary ~ clipping uses topological overlay Subset of records ~ same attributes as [in_cover] but fewer records, areas adjusted & User-IDs are renumbered (unique values)
ERASECOV POLY LINE POINT NET LINK		 Subset of existing features ~ erases features inside the external polygon boundaries ~ erasing uses topological overlay Subset of records ~ same attributes as [in_cover] but fewer records, areas adjusted & User-IDs are renumbered (unique values)

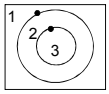


Union

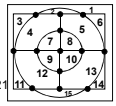
Input Coverage



Union Coverage



Output Coverage



21

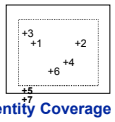
INPUT COVERA	
#	Attribut
1	
2	A
3	B
4	C
5	D

UNION COVERA	
#	Attribut
1	
2	102
3	103

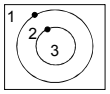
OUTPUT COVERA	INPUT COVERA		UNION COVERA	
#	#	Attribut	#	Attribut
1	1		1	
2	1		2	102
3	2	A	1	
4	2	A	2	102
5	3	B	2	102
6	3	B	1	
7	2	A	3	103
8	3	B	3	103
9	4	C	3	103
10	5	D	3	103
11	4	C	1	
12	4	C	2	102
13	5	D	2	102
14	5	D	1	
15	1		2	102

Identity - Point Option

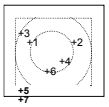
Input Coverage



Identity Coverage



Output Coverage



22

INPUT COVERAG	
#	Attribut
1	A
2	B
3	C
4	D
5	E
6	F
7	G

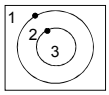
IDENTITY COVERA	
#	Attribut
1	
2	102
3	103

OUTPUT COVERA	INPUT COVERA		IDENTITY COVERA	
#	#	Attribut	#	Attribut
1	1	A	2	102
2	2	B	2	102
3	2	C	1	
4	3	D	3	103
5	3	E	1	
6	4	F	3	103
7	4	G	1	

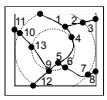
Identity - Line Option

Input Coverage

Identity Coverage



Output Coverage



23

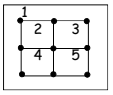
INPUT COVERA	
#	Attribu
1	A
2	B
3	A
4	C
5	A
6	D
7	A

UNION COVERA	
#	Attribu
1	
2	102
3	103

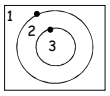
OUTPUT COVERA	INPUT COVERA		UNION COVERA	
#	#	Attribu	#	Attribu
1	1	A	2	102
2	2	B	2	102
3	2	B	1	
4	3	A	2	102
5	3	A	3	103
6	4	C	3	103
7	4	C	2	102
8	4	C	1	
9	5	A	3	103
10	6	D	2	102
11	6	D	1	
12	7	A	2	102
13	6	D	3	103

Identity - Poly Option

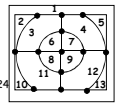
Input Coverage



Identity Coverage



Output Coverage



24

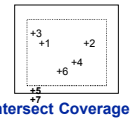
INPUT COVERAG	
#	Attribute
1	
2	A
3	B
4	C
5	D

IDENTITY COVERA	
#	Attribute
1	
2	102
3	103

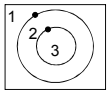
OUTPUT COVERA	INPUT COVERA		IDENTITY COVERA	
#	#	Attribute	#	Attribute
1	1		1	
2	2	A	1	
3	2	A	2	102
4	3	B	2	102
5	3	B	1	
6	2	A	3	103
7	3	B	3	103
8	4	C	3	103
9	5	D	3	103
10	4	C	1	
11	4	C	2	102
12	5	D	2	102
13	5	D	1	

Intersect - Point Option

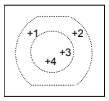
Input Coverage



Intersect Coverage



Output Coverage



INPUT COVERAG	
#	Attribute
1	A
2	B
3	C
4	D
5	E
6	F
7	G

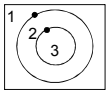
INTERSEC COVERAG	
#	Attribut
1	
2	102
3	103

OUTPUT COVERAG	INPUT COVERAG		INTERSECT COVERAG	
#	#	Attribute	#	Attribute
1	1	A	2	102
2	2	B	2	102
3	4	D	3	103
4	6	F	3	103

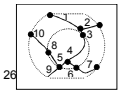
Intersect - Line Option

Input Coverage

Intersect Coverage



Output Coverage



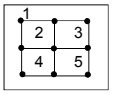
INPUT COVERAG	
#	Attribut
1	A
2	B
3	A
4	C
5	A
6	D
7	A

INTERSEC COVERAG	
#	Attribut
1	
2	102
3	103

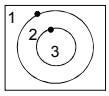
OUTPUT COVERAG	INPUT COVERAG		INTERSECT COVERAG	
#	#	Attribut	#	Attribut
1	1	A	2	102
2	2	B	2	102
3	3	A	2	102
4	3	A	3	103
5	5	A	3	103
6	4	C	3	103
7	4	C	2	102
8	6	D	3	103
9	7	A	2	102
10	6	D	2	102

Intersect - Poly Option

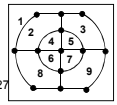
Input Coverage



Intersect Coverage



Output Coverage



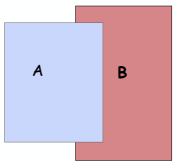
INPUT COVERAG	
#	Attribute
1	
2	A
3	B
4	C
5	D

INTERSEC COVERAG	
#	Attribute
1	
2	102
3	103

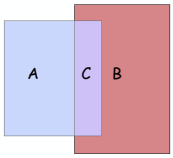
OUTPUT COVERAG	INPUT COVERAG		INTERSECT COVERAG	
#	#	Attribute	#	Attribute
1	1		1	
2	2	A	2	102
3	3	B	2	102
4	2	A	3	103
5	3	B	3	103
6	4	C	3	103
7	5	D	3	103
8	4	C	2	102
9	5	D	2	102

Clip

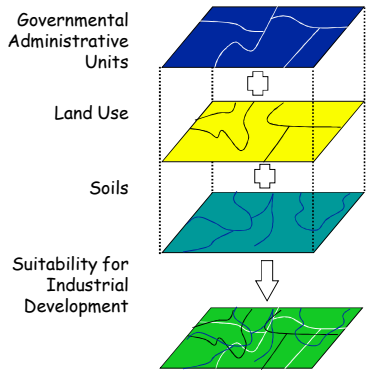
Two polygons, A nd B, Overlap. Clip A using B as a cookie cutter.



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.



Suitability Analysis Using Map Overlays



Analysis of these interrelated maps

Produces

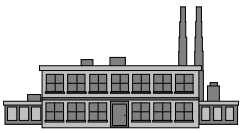
Application (Result)

Polygon Overlay

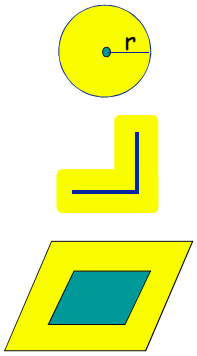
The vertical overlay of multiple map layers to derive a resulting map layer based on some logical combination of attributes from the original layers.

Examples:

- Erosion analysis based on overlay of soil, slope, and land cover layer;
- Determine parcels that fall within flood prone areas
- Site suitability for industrial development



Buffering



Buffering a Point

example: All area within one mile of the city.

Buffering a Line

example: All areas within 1000 meters of a road.

Buffering an Area

example: All areas within 500 meters of a wetlands area.

Proximity/Buffer Analysis

Generation and selection of map data within specified distance around a point, line, or polygon feature

Examples:

- Property search to find all parcels within a specified distance of a target parcel
- Environmental impact analysis requiring delineation of a buffer zone around a stream

Network Analysis

Analytical technique to evaluate flow or paths through a defined linear network.

Examples:

- Analysis of optimal routes through road network
- Analysis of flow in a water distribution system

Address Matching and Incident Mapping

Generation of maps showing the point location of features or phenomena on a suitable base map. Incident locations may be defined by x, y coordinate, street address, or other locational identifier.

Examples:

- Mapping of water sampling points
- Police incident mapping