

Vector Overlay Processing - Specific Theory

Cartographic Modelling and GIS.

Cartographic modelling is a general, but well-defined methodology that is used to address diverse applications of GIS in a clear manner. It is a technique used for both vector and raster based GIS, and as the term suggests, cartographic modelling involves *models* (ie. of geospatial information) represented in *cartographic* form (ie. as Maps). (Tomlin, 1990, p xi)

Cartographic modelling is used to simultaneously analyse both the spatial and thematic characteristics of geospatial information. The thematic component of geospatial information is analysed via statistical operations on the data (for example, taking the Average and Standard Deviation of the data), where the spatial characteristics of geospatial information are described through spatial analysis techniques (which are based on coordinate data).

Introduction to Overlay in GIS.

Geographical Information Systems are used to perform a number of fundamental spatial analysis operations, such operations can use any number of analytical processes. However almost all spatial analyses utilise six fundamental operations: Topological Map Overlay, Buffer Generation, Feature Extraction, Feature Merging, and the two relational database operations - Join and Relate.

The main aim of this module is to clearly define and present the concepts and algorithms associated with topological map overlay processes, specifically in reference to vector based systems.

Topological Map Overlay.

Topological map overlay creates new features and attribute relations by overlaying the features from two input map layers. Features from each input layer are combined to create new output features. Attributes of each input feature are combined from the two input layers to describe each new output feature, thus creating new attribute relationships. (ESRI, ch 2)

Often, the manipulation of multiple data layers is required to achieve the objective of the overlay operation. This is done in a stepwise fashion - two input layers are combined to form an intermediate layer, this intermediate layer is then combined with a third layer to form another intermediate layer, and so forth until the desired resultant map layer is achieved. (Star and Estes, p 148)

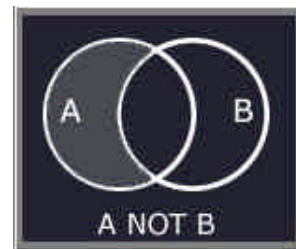
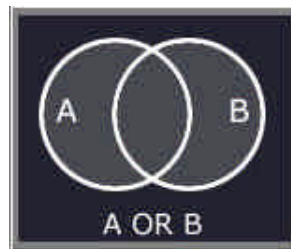
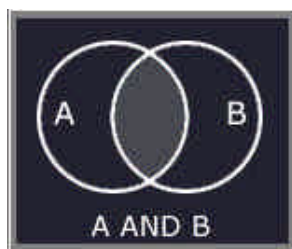
Boolean Logic in GIS.

Boolean Algebra is useful for performing operations on the attributes (which can 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. Boolean algebra uses the logical operators **AND**, **OR**, **NOT** to determine whether a particular condition is true or false.

Each attribute can be thought of as defining a *set*. Consider two sets (set A and set B);

- ?? The **AND** operator (\cap) is the *intersection* of two sets - for example those entities that belong to both set A and set B
($A \cap B$)
- ?? The **OR** operator (\cup) is the *union* of two sets - for example those entities that belong to either set A or to set B
($A \cup B$)
- ?? The **NOT** operator (\neg) is the *difference* operator identifying those entities that belong to A but not B
($A \neg B$)

These simple relations can be visualised through the use of Venn diagrams.
(Burroughs and McDonnell, 1998, pp 162, 164-166)

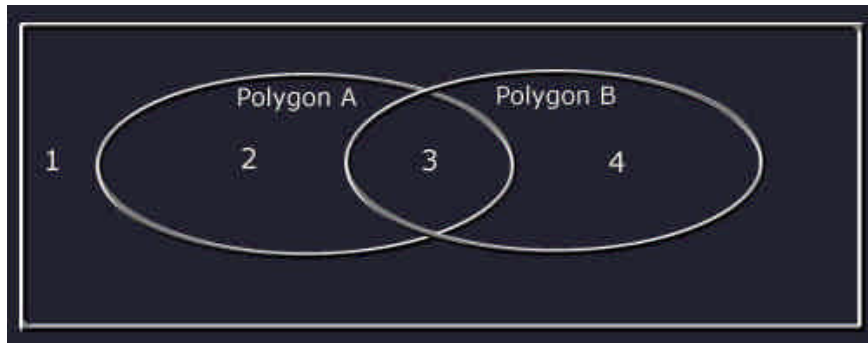


Boolean Logic in GIS - Interactive Example Summary.

Consider a simple GIS consisting of two polygons A and B, each stored on a different layer in the system. Using the logical operators **AND**, **OR**, **NOT**, there are a total of sixteen boolean processing combinations possible. Each of which, when applied to these two layers, will produce a different output layer.

The diagram below shows a representation of this output layer with:

- ?? **1** corresponding to the region that is contained in neither Polygon A nor Polygon B
- ?? **2** corresponding to the region that is entirely contained in Polygon A
- ?? **3** corresponding to the region contained in both Polygon A and Polygon B
- ?? **4** corresponding to the region that is entirely contained in Polygon B



The following table shows the sixteen Boolean Operations (and their outcomes) which can be applied to the two overlaid polygons (A and B). The Region Output results correspond to the boolean values resultant from the Boolean Operation, 0 indicating the regions which are to be excluded from the output map/layer, and 1 indicating the regions which are to be included in the output map/layer.

Outcomes of boolean operations applied to two polygons.

OPERATION	REGION OUTPUT RESULT			
KEY:	Region 1	Region 2	Region 3	Region 4
? - AND ? - OR				
\neg - NOT				
NULL	0	0	0	0
$(\neg A) ? B$	0	0	0	1
$A ? B$	0	0	1	0
B	0	0	1	1
$A ? (\neg B)$	0	1	0	0
$(A ? (\neg B)) ? ((\neg A) ? B)$	0	1	0	1
A	0	1	1	0
$A ? B$	0	1	1	1
$(\neg A) ? (\neg B)$	1	0	0	0
$(\neg A)$	1	0	0	1
$(A ? B) ? ((\neg A) ? (\neg B))$	1	0	1	0
$(\neg A) ? B$	1	0	1	1
$(\neg B)$	1	1	0	0
$(\neg A) ? (\neg B)$	1	1	0	1
$A ? (\neg B)$	1	1	1	0
\neg NULL	1	1	1	1

(G.Hunter, 1998)

Vector Based GIS Overlay.

In a vector-based system topological map overlay operations are much more complex than the raster-based case, as the topological data is stored as points, lines and/or polygons. This requires relatively complex geometrical operations to derive the intersected polygons, and the necessary creation of new nodes (points) and arcs (lines), with their combined attribute values.

(Star and Estes, pp 147-148)

In a vector-based system, topological map overlay operations allow the polygon features of one layer to be overlaid on the polygon, point, or line features of another layer. Depending on the objectives of the Overlay operation, different output features can result.

(ESRI, ch 2)

Classification of Vector Overlay Operations.

Topological vector overlay operations can be classified via two methods:

- ✍️ **T**hrough the *elements* contained in the layers to be overlaid (ie. whether the layers contain *point*, *line* or *polygon* elements), or
- ✍️ **B**y *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).

When classifying the vector overlay operation via method one, the element types each layer contains are considered. The following table identifies which overlay options exist for each possible combination of element types contained in the two input layers.

Vector Overlay Operations - Classification through Input Layer Element Types.

Input Layer Element Types	Points	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.

Vector Overlay Processing - Algorithms

Since vector-based topological map overlay operations involve overlaying the point, line, or polygon features of one layer on the polygon features of another layer, the three following processing algorithms are fundamental:

- ✍ Point-in-Polygon
- ✍ Line-in-Polygon
- ✍ Polygon-on-Polygon (ie. Polygon Overlay)

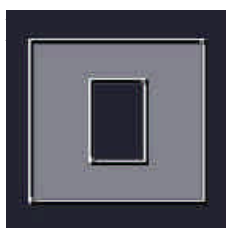
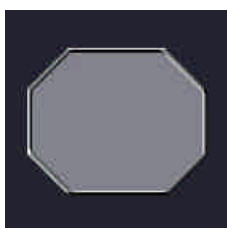
Point-in-Polygon Processing.

Point features of one input layer can be overlaid on polygon features of another input layer, Point-in-Polygon analyses identify the polygon within which each point falls. The result of a Point-in-Polygon overlay is a set of points with additional attributes (ie. those attributes of the polygon which the point lies within). (ESRI, ch 2)

The basic algorithm used to perform Point-in-Polygon analyses is detailed below:

- ?? Usually a minimum bounding rectangle for the polygon is defined within the system - by its maximum and minimum coordinates. It is easy to determine if a point (or line end) is inside or outside the rectangle's extent. If the point lies outside the minimum bounding rectangle, then it also must lie outside the polygon and the analyses is complete. However if the point falls inside the minimum bounding rectangle then the following further processing is required.
- ?? From the point, a line parallel to an axis is drawn (usually either the X or Y axis). This parallel line extends from the point (or line end) to beyond the extremities of the polygon, with its direction usually towards the highest values of this axis. Note - For this method to work, this parallel line can only extend in one direction from the chosen point or line end, it is therefore otherwise known as a "half line".
- ?? The system then counts the number of times this "half line" intersects with the polygon boundary. If the result is an even number (or zero), then this indicates that the point lies outside the polygon. If the result is an odd number, then this indicates that the point falls inside the polygon.

The Point-in-Polygon algorithm described above works very well for special cases of "island" polygons, polygons with holes, and concave polygons. Examples of these special cases are shown below:



However, problems occur if a point falls:

- ✍ Exactly on a boundary,
- ✍ On a node or a vertex, or
- ✍ When a line segment is collinear to the half-line.

In these cases, individual GIS packages use their own “decision rules” to determine whether these point types are or are not classified as being included a polygon. (G.Hunter, 1998)

An applied example of the use of Point-in-Polygon Analyses would be to determine the number of bore holes located in a particular parcel of a cadastral map. The resultant map/layer will assign the variables attached to the cadastral parcel to each of the bores located within that parcel. Once this operation is complete, the number of bore holes located in the particular parcel can be easily determined.

Line-in-Polygon Processing.

Polygon features of one input layer can be overlaid on lines (arcs) of another input layer. A line can be made up of many segments, Line-in-Polygon analyses therefore identifies which polygon (if any) contains each line or line segment. The result of a Line-in-Polygon overlay is a new layer containing lines with additional attributes (ie. those attributes of the polygon within which the line falls). Sometimes a line segment falls directly on a polygon boundary instead of within the polygon. In this special case, the additional line attributes will contain the attributes of both polygons - lying to the left and right sides of the line. (ESRI, ch 2)

As lines and polygons are both made up of line segments, Line-in-Polygon analysis requires the determination of whether any of these overlaid line segments intersect. The task of determining whether two line segments intersect consists of a simple mathematical calculation, however the complexity of this operation is increased by the number of line intersection checks that need to be made for a complete Line-in-Polygon overlay analysis. Therefore Geographical Information Systems use the following algorithm to minimise the number of calculations required.

- ?? Minimum bounding rectangles of both the line and the polygon are used to reduce the number of computations required. If no intersection occurs then a check is made to determine whether the minimum bounding rectangle of the line falls completely outside the minimum bounding rectangle of the polygon (defined by the element's minimum and maximum coordinates). If this is the case then the line definitely does not lie within the polygon, and the analyses is complete, otherwise the following further processing is required:

- ?? As the line may be made up of many line segments, each line segment has to be tested for intersection or inclusion within the polygon. If the line segment lies outside the polygon minimum bounding rectangle, then that segment also lies outside the polygon and can be disregarded, otherwise the following processing must continue:
- ?? Testing whether a line segment is totally inside a polygon or not can be difficult because polygons can have concavities or holes within them, therefore it is not enough to simply determine if both end-points of a line segment lie within the polygon. To deal with this problem, the polygon and line segment are both rotated such that the line segment lies parallel to one of the axis (X or Y).
- ?? The next step uses the “half-line” test (as described in the Point-in-Polygon Analyses theory above) along the axis parallel to the line segment to determine whether each segment end-point is in or out and note all segment intersections with the polygon. Note that half-line intersection points are not necessarily also segment intersection points.
- ?? If the results of the half-line testing show that both points are in and there were no segment intersections, therefore the whole line lies inside the polygon. Otherwise, if a point starts outside, then that first part of the line segment is outside the polygon until the first segment intersection point, the next part of the line segment is inside the polygon until the next segment intersection point, and so on.
- ?? This procedure is followed for each of the line segments making up the line currently being analysed.

(G.Hunter, 1998)

An applied example of the use of Line-in-Polygon analysis would be to determine the location of walking tracks within a particular region of crown land. The resultant map/layer will assign the variables attached to the crown land to each of the walking tracks located within the region.

Polygon-on-Polygon Processing.

This process merges overlapping polygons from two input layers to create new polygons in an output layer. The result of a Polygon-on-Polygon overlay is an output layer containing new polygons with merged attributes (ie. those attributes from each of the two overlapping polygons).
(ESRI, ch 2)

Note: As polygons are made up of line segments, Polygon-on-Polygon analysis requires the determination of whether these overlaid line segments intersect. The processing for Polygon-on-Polygon analysis is therefore essentially the same as for Line-in-Polygon analysis (as detailed in the Line-in-Polygon theory above).

Vector Overlay Processing - Examples

Vector Overlay processing uses the previously mentioned boolean logical operations **AND**, **OR**, and **NOT**: ($A \cap B$, $A \cup B$, $\neg A$), which are used in a GIS setting to determine the INTERSECTION or UNION of two layers, and the COMPLEMENT of one layer in the system.

Vector Overlay processing also uses the boolean relations **INCLUSION** and **EXCLUSION**: ($A \subset B$, $A \not\subset B$), which are used in a GIS setting to determine which part of a layer is either **CONTAINED** or **NOT CONTAINED** within another layer in the system.

For example, consider two layers in a GIS (A,B):

$A \cap B$ determines the part of layer A which is **CONTAINED** in layer B.

$A \setminus B$ determines the part of layer A which is **NOT CONTAINED** in layer B.

Numerous Vector overlay operations exist, all of which are derivative of the following four basic operations; **UNION**, **INTERSECT**, **INCLUDE** and **EXCLUDE**.

UNION

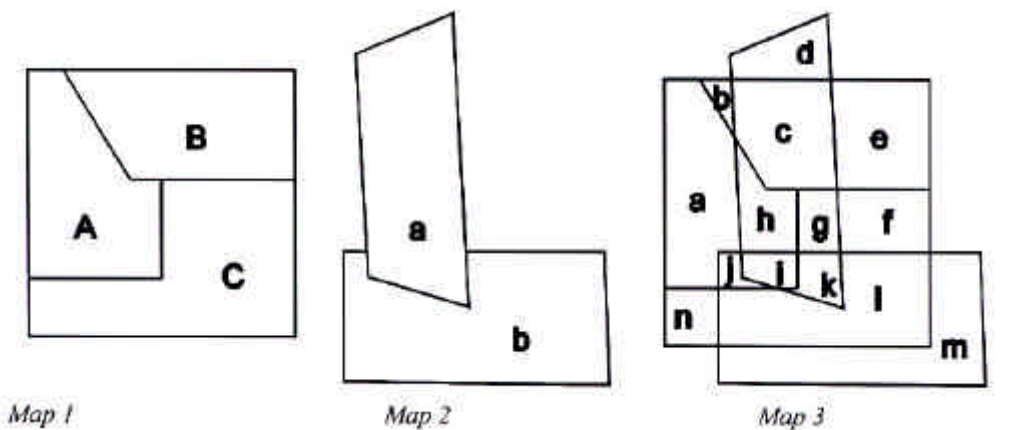
Analogous to the boolean logical operator **OR**, where all elements from both input layers will be present in the output layer. This is illustrated in the following example:

To create a new layer (MAP 3) containing areas of agricultural land use (MAP 1) or environmentally protected areas (MAP 2), perform the operation:

MAP 1 **UNION** MAP 2.

This can be written in boolean algebraic form as:

MAP 1 \cup MAP 2 = MAP 3



INTERSECTION

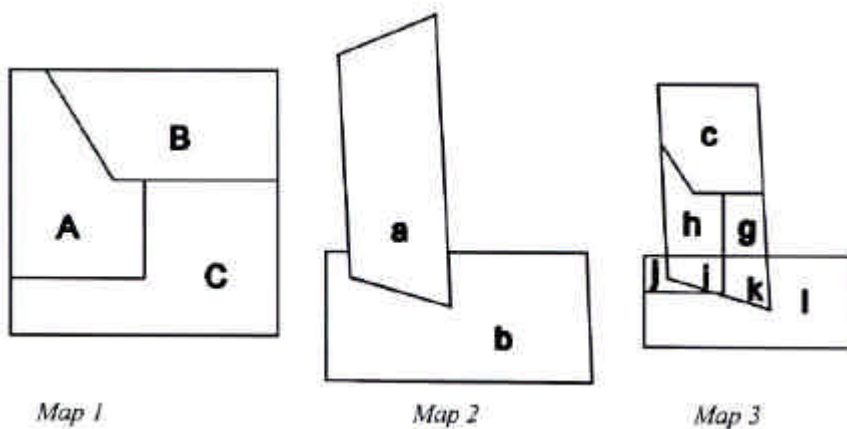
Analogous to the boolean logical operator **AND**, where only the elements from both input layers that coincide in the same geographic extension will be part of the output layer. This is illustrated in the following example:

To create a new layer (MAP 3) containing areas of agricultural land use (MAP 1) that have an average annual temperature of over 10 degrees (MAP 2), perform the operation:

MAP 1 **INTERSECTION** MAP 2.

This can be written in boolean algebraic form as:

MAP 1 **?** MAP 2 = MAP 3



INCLUDED (CONTAINED)

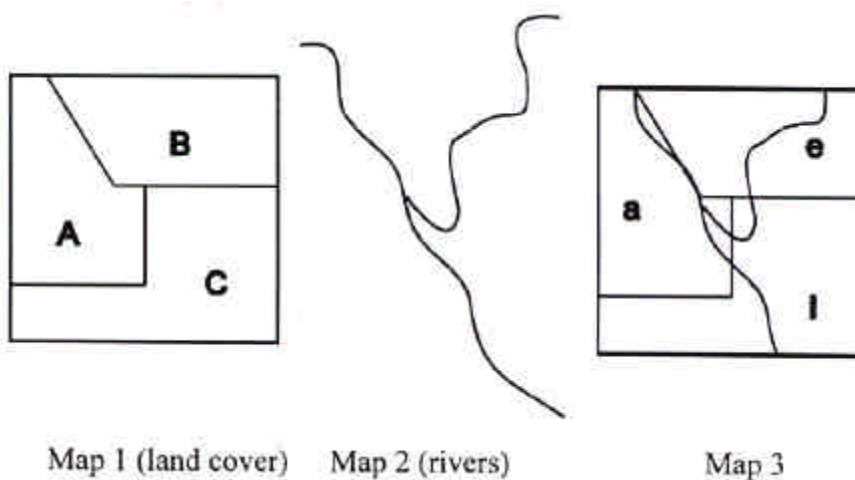
The output layer will contain all elements present in the first input layer plus all elements present in the same geographic extension of the second input layer. This is illustrated in the following example:

To create a new layer (MAP 3) containing the hydrological network (MAP 2) added to the land cover map of Region X (MAP 1), perform the operation:

MAP 1 **INCLUDE** MAP 2.

This can be written in boolean algebraic form as:

$$(\text{MAP 2} \cup \text{MAP 1}) \cap \text{MAP 1} = \text{MAP 3}$$



EXCLUDED (NOT CONTAINED)

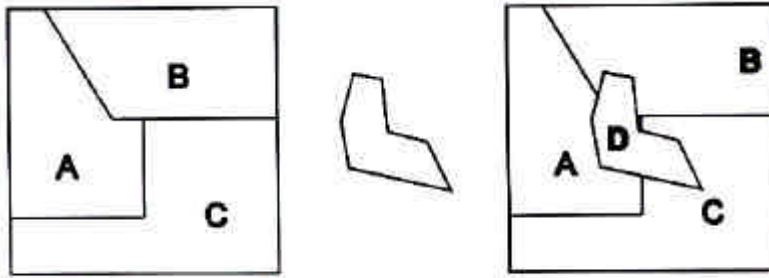
The output layer will contain only the elements of the first input layer that do not match within the geographical extent of the second input layer. This is illustrated in the following example:

To create a new layer (MAP 3) containing agricultural areas (MAP 1) which are located far from the noisy roads (Map 2), perform the operation:

MAP 1 **EXCLUDE** MAP 2.

This can be written in boolean algebraic form as:

$$\text{MAP 1} \setminus \text{MAP 2} = \text{MAP 3}$$



Map 1 (agricultural land) Map 2 (areas near the road) Map 3

As can be seen in the above examples, Polygon-on-Polygon overlay can be a useful tool for modelling purposes.
(F.Escobar, 1998)

References

- Burroughs, P.P. & McDonnel, R.A. 1998, *Principles of GIS*, Oxford University Press, pp. 162 - 166.
- Escobar, F. 1998, Vector Overlay Processes, Sample Theory, The University of Melbourne.
- ESRI (Environmental Systems Research Institute) 1990, *PC Overlay Users Guide*, ch.2.
- Hunter, G.J. 1998, Boolean Operations, 451-620 Lecture Notes, The University of Melbourne.
- Sikorski, R. 1969, *Boolean Algebra*, Springer-Verlag, Berlin, pp.3-13.
- Star, J. & Estes, J. 1990, *Geographic Information Systems An Introduction*, Prentice-Hall, New Jersey, pp. 147-148.
- Tomlin, D. 1990, *Geographic Information Systems and Cartographic Modeling*, Prentice-Hall Inc., New Jersey, pp. xi-xiv, 14-22, 96-99, 119-122.