

Cartographic Modeling

C. Dana Tomlin, University of Pennsylvania

Cartographic Modeling is a general methodology for the analysis and synthesis of geo-spatial data. It has been incorporated into a number of raster-oriented geographic information systems, and it can be used to address a variety of applications in a unified manner. This is done by decomposing both data and data-processing tasks into elemental units that can then be clearly and flexibly recomposed. The result is an algebra-like language in which the variables are maps and the functions are map-transforming operations. The nature of this “map algebra” can be expressed in terms of three fundamental components: a body of data, a set of data-processing capabilities, and a mechanism to control that data processing. To the extent that map algebra can be regarded as a language, these components can respectively be characterized as nouns, verbs, and expressions.

The Nouns

The primary unit of data employed in cartographic modeling is the “layer.” This can be envisioned as a single-factor map. Like any map, it is a bounded plane area depicting a geographic region such that every location within that area represents a corresponding location within the region. Formally, a “location” is that portion of the cartographic plane that is uniquely identified by a pair of planar coordinates. In the case of raster-encoded data, it is a grid cell or pixel. As a single-factor map, each layer is one on which every location is characterized in terms of exactly one of a related set of site conditions. Thus, one layer might depict every location’s soil type or its proximity to the nearest supermarket, while others might depict variations in characteristics such as population density or groundwater contamination. When multiple layers are used to represent a common region, all must be geometrically compatible with one another in terms of their spatial extent, orientation, and cartographic projection.

The set of all locations on a layer that share a common site condition is referred to as a “zone,” and each of a layer’s zones is represented by a numerical “value.” This is an integer or real number that identifies a zonal condition in terms that may be either qualitative or quantitative in nature. In the case of qualitative conditions (such as soil types or land uses), these values will be nominal and may be arbitrarily assigned. In the case of quantitative conditions (such as rankings, dates, distances, or directions), values may relate to ordinal, interval, ratio, or even cyclical scales of measurement. A special “NULL” value is also used to represent the absence of any recorded site condition.

In more general settings, the term “layer” is sometimes used in reference to a multiple-factor map. A layer of soil types, for example, might be created such that each type is characterized not only by name but also in terms of its acidity, permeability, bearing capacity, and so on. From the perspective of cartographic modeling, each of those separate characteristics would constitute a separate layer.

The Verbs

If layers are the nouns, then the verbs of this cartographic modeling language are layer-transforming “operations.” Each of these operations generates output in a form (the map layer) that can then be accepted as input to any other operation. Since multiple operations can be combined in this manner, no one of those operations needs to be particularly complicated. Just as a small number of primitive algebraic functions (such as addition, subtraction, or multiplication) can be combined into an endless variety of mathematical equations, so can a concise vocabulary of elementary map algebraic operations be combined into an open-ended array of “cartographic models.”

Each map algebraic operation is defined in terms of its effect on a single, typical location. This worm’s eye perspective gives rise to four major types of operation that are respectively referred to as “local,” “zonal,” “focal,” and “incremental.”

Local Operations

A local operation is one that computes a new value for every location as a function of its value(s) on one or more existing layers. *LocalVariety*, for example, is an operation that indicates the number of dissimilar values at each location, while *LocalCombination* associates a unique new value with each existing-value combination. *LocalRating* assigns a designated value (or value drawn from the same location on a designated layer) to all locations having a specified set of one or more existing values. And operations *LocalSum*, *LocalDifference*, *LocalProduct*, *LocalRatio*, *LocalRoot*, *LocalMajority*, *LocalMinority*, *LocalMaximum*, *LocalMinimum*, *LocalMean*, *LocalSine*, *LocalCosine*, *LocalTangent*, *LocalArcCosine*, *LocalArcSine*, *LocalArcTangent* compute familiar algebraic, trigonometric, or statistical functions of each location’s existing values.

Zonal Operations

A zonal operation is one that computes a new value for every location as a function of whatever values from one existing layer are associated with that location's zone on another existing layer. Similar to *LocalVariety*, for example, *ZonalVariety* indicates the number of dissimilar values within each zone. *ZonalCombination*, *ZonalRating*, *ZonalSum*, *ZonalProduct*, *ZonalMajority*, *ZonalMinority*, *ZonalMaximum*, *ZonalMinimum*, and *ZonalMean* also apply functions that are comparable to local counterparts but do so on a zone-by-zone basis. Several zonal operations, however, have no local counterpart. *ZonalRanking*, for example, indicates the ordinal position of each location's value among all those within its zone, while *ZonalPercentage* indicates the percentage of each location's zone that shares that location's value, and *ZonalPercentile*, indicates the percentage of each location's zone that has a lower value.

Focal Operations

A focal operation is one that computes a new value for every location as a function of the existing values, distances, and/or directions of neighboring locations. *FocalVariety*, *FocalCombination*, *FocalRating*, *FocalSum*, *FocalProduct*, *FocalMajority*, *FocalMinority*, *FocalMaximum*, *FocalMinimum*, *FocalMean*, *FocalRanking*, *FocalPercentage*, and *FocalPercentile* are all similar to local and zonal counterparts except that they are applied to neighboring values. *FocalProximity* is another focal operation that computes each location's distance to the nearest of a specified set of neighboring locations, while *FocalBearing* and *FocalNeighbor* respectively indicate the direction and the value of that nearest neighbor. And *FocalInsularity* uniquely identifies islands or clusters of like-valued locations. In each of these cases, the set locations from which existing values are drawn constitutes a "neighborhood" of which there are three major types: "lateral," "radial," and "fluvial."

Lateral neighborhoods are rectangular. Each extends above, below, left, and right of a central or "focal" location at distances that may either be specified as constants or drawn from focal location values on designated layers. Minimum limits on these distances may also be specified in order to create neighborhoods with holes. Furthermore, weights can be associated with neighborhood positions in order to modify the input values associated with neighboring locations.

Radial neighborhoods are generally circular. Each extends outward from a focal location to a distance and in directions that, again, may either be specified by constant values or drawn from the values of focal locations on a designated layer. As with lateral neighborhoods, minimum distances can be specified in order to create neighborhoods with holes, and distance or directional weighting factors can be specified as well. Radial neighborhoods can also be subjected to visibility constraints, such that a neighboring location is regarded as part of a neighborhood only if an unobstructed line-of-sight connects it to the neighborhood's focal location.

Fluvial neighborhoods are likewise defined by distance from focal locations. Here, however, that distance is not measured in terms of meters, miles, or other units of physical separation. Rather, it is measured in units such as minutes, dollars, or gallons of fuel that accumulate as a consequence of motion. Given a particular type of motion, these units may well accumulate at rates that vary according to motion-impeding site conditions. Consider, for example, a neighborhood encompassing all locations within five minutes of walking time from its focal location. Impedance may be affected not only by the medium through which this motion is being simulated but also by characteristics of the resulting motion itself: velocity, direction, duration, acceleration, momentum, changes in direction, and so on.

Incremental Operations

A more specialized form of focal processing is employed by the final group of cartographic modeling operations. An incremental operation is one that computes a new value for every location in order to characterize the size or the shape of that location's unique portion of a one-, two-, or three-dimensional cartographic form. *IncrementalLinkage*, for example, is an operation that characterizes each location according to the particular manner in which it connects to adjacent locations of similar value as part of a linear network, while *IncrementalLength* computes the total length of such connections. *IncrementalPartition* regards each location as part of either the interior or the edge of a two-dimensional zone. For the latter, it then indicates that shape of the zonal edge at that location, while *IncrementalFrontage* and *IncrementalArea* respectively measure each location's contribution to zonal perimeter and area. The measurements produced by *IncrementalLength*, *IncrementalFrontage*, and *IncrementalArea* can also be applied to non-planar surfaces by equating one of each location's values with elevation in a third dimension perpendicular to the cartographic plane. This also gives rise to operations *IncrementalVolume*, *IncrementalGradient*, *IncrementalAspect*, and *IncrementalDrainage*, which respectively measure the sub-surface volume, the steepness, the downhill direction, and the downstream direction at each location on such a non-planar surface.

The Expressions

Given the data and data-processing constructs associated with cartographic modeling, the manner in which that processing is controlled may still vary from one computing environment to another. The original pseudo-language of map algebra attempts to relate to as many of these environments as possible by employing a highly general form of verbal notation. While some software developers have adopted this pseudo-language directly, most have instead elected to employ their own variations.

Like their conventional algebraic counterparts, map algebraic expressions are imperative statements in declarative form that specify operations and processing options as well as input and output variables. Below, for example, is the first statement in a cartographic model that starts with an existing layer called *INPUT*.

BACKGROUND = *LocalRating of INPUT*
with *NULL* for ... and 0 for *NULL*

If *INPUT* is a layer on which each of a group of island-like zones is identified by its own unique value on a background of locations set to *NULL*, then this *LocalRating* statement will result in a new layer on which all of those island-like zones are set to *NULL*, while non-island locations are set to zero. Subjecting the resulting *BACKGROUND* layer to an operation given as

DISTANCE = *FocalProximity of BACKGROUND*

will then generate a *DISTANCE* layer on which every island location's value indicates its proximity to the nearest non-island location. By regarding those proximity values as third-dimensional elevations, an operation given as

DIRECTION = *IncrementalAspect of DISTANCE*

can then be used to create a *DIRECTION* layer on which each location's value indicates the compass bearing of the nearest non-island location. (This is something that could actually have been done in a more efficient, though less illustrative, manner by applying a *FocalBearing* operation directly to the *BACKGROUND* layer.) Next, a compound statement given as

ORIENTATION = *LocalRating of DIRECTION*
with (*LocalDifference of DIRECTION minus 180*)
for 180 ...

calls for the generation of an *ORIENTATION* layer by subtracting 180 (degrees) from all *DIRECTION* values of 180 or greater and thereby equating diametrically-opposing directions. Finally, the resulting directional values are averaged on a zone-by-zone basis in response to an operation specified as

OUTPUT = *ZonalMean of ORIENTATION*
within *INPUT*

A cartographic model like this might be applied just once to a particular *INPUT* layer, or it might be stored as a new map algebraic operation (perhaps called *ZonalOrientation*) that could then be applied to whatever input layer is specified when that operation is invoked.

Ultimately, it is the ability to develop such user-generated capabilities from an accessible set of primitive components that accounts for the power and the promise of cartographic modeling.

Further Readings

Tomlin, C.D. (1983). *Digital cartographic modeling techniques in environmental planning*. Doctoral dissertation. Division of Forestry and Environmental Studies, Yale University Graduate School.

Tomlin, C.D. (1990). *Geographic information systems and cartographic modeling*. Englewood Cliffs, New Jersey: Prentice-Hall.

Tomlin, C.D. (1991). Cartographic Modeling. In M. Goodchild, D. Maguire, D. Rhind, (Eds.), *Geographical information systems: principles and applications*. Harlow, Essex, United Kingdom: Longman Group Ltd.