UNDERSTANDING TRANSFORMATIONS OF GEOGRAPHIC INFORMATION

Nicholas Chrisman
chrisman@u.washington.edu
Department of Geography, University of Washington
Seattle WA 98195-3550

ABSTRACT

Transformations have been presented as an organizing principle of analytical cartography. To date, the theories have focused on geometric distinctions, such as point, line and area. This paper presents a new scheme for geographic transformations based on measurement frameworks as the principal distinction. Transformations between measurement frameworks can be viewed in terms of a neighborhood and a rule to process attribute information. This scheme provides a way to organize most of the operations performed by GIS software.

BACKGROUND: TRANSFORMATIONS

While the dominant school of cartography views cartography as a communication process, there has always been another group focused on transformations. The most classic transformation involves the mathematical conundrum of transferring the nearly spherical Earth onto a flat piece of paper, the process of map projection. No cartographic education is complete without a thorough understanding of projections. For centuries, a cartographer could ensure a place in posterity by inventing another solution to the map projection problem. Even Arthur Robinson, whose career was dedicated to thematic cartography and map communication, may have greater recognition for his compromise world projection through its adoption by the National Geographic Society.

The key importance of a map projection is not in the mathematical details. Projections demonstrate how measurements taken on one kind of geometric model can be transferred to another model, subject to certain constraints. For instance, in moving from the earth to a plane, it is possible to preserve either the geometric relationships of angles (conformality) or of area (equivalence), but not both. This operation on geographic data became the basis for Tolber's view of analytical cartography.

While much of Tobler's work dealt with projections, he also advanced a 'transformational view of cartography' (Tobler 1979b) that considered all operations as transformations of information content. Analytical cartography has remained a minority component of the discipline, though some continue to extend it (Nyerges 1991; Clarke 1995).

Analytical cartography developed in the era of Chomsky's transformational grammars, an attempt to systematize linguistics that had far-reaching influence throughout the social sciences. Tobler (1976) set out a systems of transformations based largely upon the geometric component of geographic information. This approach informed the three by three (point, line,area) or four by four (with the addition of volume) matrix in Clarke (1995, Figure 11.1, page 184) and Unwin, among others [Figure 1].

From \ To	**Point**	**Line**	**Area**
Point	Point -> Point	Point -> Line	Point -> Area
Line	Line -> Point	Line -> Line	Line -> Area
Area	Area -> Point	Area -> Line	Area -> Area

Figure 1: Cartographic transformations as viewed by Clarke and Unwin

In this matrix, a buffer around a road would be considered a line-to-area transformation, but so would the conversion from a contour line to a TIN. There is little in common between these operations because the relationships implied by the lines are so different. There is no denying that the geometric primitives are important, but they may not tell all the story. The geometric form of input and output are a weak guide to the actual operation that might be performed. Many of the most complex operations are lumped into the diagonal, along with operations that make very minimal changes. This matrix based on the dimensionality of the objects is clearly insufficient to explain the operations performed in a GIS.

Measurement frameworks

What is missing from the standard explanation is an explanation of the different reasons for using a line to represent geographic phenomena. While there are many reasons, they can be organized in according to the fundamental choices made in obtaining the underlying measurements. Geographic information includes a spatial component, a temporal component, and some set of attributes. As Sinton (1978) recognized, each data model imposes a different set of rules on these components. A measurement framework (Chrisman, 1995, 1997) is a conceptual scheme that establishes rules for control of other components of a phenomenon that permit the measurement of one component.

The broad groupings of measurement frameworks listed in Figure 2 provide a clearer foundation for transformations of geographic information. The different forms of control have only been recognized in specifics, not as a part of a scheme that is broadly applicable. The role of control is critical to understanding transformations. It is particularly important to note that the representation used my not be the same as the measurement framework. It is quite possible to represent a choropleth measurement in a raster data structure, or a set of pixels as vectors. In both cases, additional losses of resolution and accuracy can occur.

Figure 2: General groupings of measurement frameworks

	Attribute Controlled Frameworks
Isolated Objects	
Spatial Object	Single category distinguished from void
Isoline	Regular slices of continuous variable
Connected Objects	
Network	Spatial objects connect to each other, form topology (one category or more)
Categorical Coverage	Network formed by exhaustive classification (multiple categories, forming an exhaustive set)
	Space Controlled Frameworks
Point-based Control	
Center point	Systematic sampling in regular grid
Systematic unaligned	Random point chosen within cell
Area-based Control	
Extreme value	Maximum (or minimum) of values in cell
Total	Sum of quantities (eg. reflected light) in cell
Predominant type	Most common category in cell
Presence / absence	Binary result for single category
Percent cover	Amount of cell covered by single category
Precedence of types	Highest ranking category present in cell
	Temporal Frameworks
Snapshots	[any other measurement framework can be repeated over time]
Transactions	Discrete events are located freely in time
	Relationship Controlled Frameworks
Measurement by pair	Control by pairs of objects
Triangular Irregular Network	Control by uniform slope (gradient & aspect)
	Composite Frameworks
Choropleth	Control by categories (names of zones) then control by space

TAXONOMIES OF GIS OPERATIONS

The GIS literature has a series of alternative schemes used to present the different kinds of operations. Perhaps the most widely cited is Tomlin's (1983; 1990) *Map Algebra.* This scheme is essentially a sequence to present map operations, ranging from the simple to the complex. The simple operations work on a single map, followed by those that work locally on two maps, and so on. However, Tomlin's scheme fails to include all possibilities (and thus provide the 'algebra' promised), because it forces all measurements into a single raster representation and does not distinguish between a representation scheme and a measurement framework. Furthermore, Tomlin's terminology for the operations becomes a bit obscure for the more complex operations. Goodchild (1987) followed the flow of Tomlin's logic, adding some neglected elements, such as information attached to pairs of objects. Burrough (1992) argued for "intelligent GIS" essentially by recognizing more spatial relationships. Recently, Albrecht has described a method to develop commonalities between GIS operations using a sematic network. This approach seems to rely upon a survey of users, thus is vulnerable to the limited perspectives and training in those surveyed. It still seems worthwhile to develop a taxonomy of GIS operations based on transformations between measurement frameworks.

A Theory of Transformations

Any data model consists of a set of objects, relationships between them and a set of axioms (integrity constraints) that control the meaning of the data. Given data within a particular measurement framework, it is most direct to produce a result in the same framework. Thus, a grid of values with a 10 meter spacing can be most easily processed into another grid with ten meter spacing. To generate a different result, new sets of assumptions may be required. These assumptions are required to fill in the gaps in either space, time, or attributes in the original source.

The theory present here contends that transformations between most forms of geographic information can be performed with two sets of assumptions: one to handle space, thus creating a neighborhood, and the other to handle attributes, a rule of combination. Temporal transformations can be handled as special forms of neighborhoods. Neighborhoods can be defined rather flexibly, following the general scheme of Tomlin - moving from the purely local relationships inside one object through immediate neighbors to more complex relationships based on distance and perhaps other considerations. The rules of combination have not been considered as carefully in the GIS literature. Hopkins (1977) described some of the tools to handle map overlay based on Stevens' levels of measurement, but this scheme does not cover all cases. Rules of combination can be grouped into three broad classes based on the amount of information used in the process (Chrisman, 1997). A *dominance* rule simply selects one of the available values based on some criteria (such as taking the largest value). A *contributory* rule uses all the values, giving each an opportunity to contribute to the composite result.

Addition is the most classic contributory rule. Finally, an *interaction* rule uses not just each value, but the pairwise combinations of values.

This taxonomy of attribute rules serves to explain the differences among the approaches to area-based spatial control frameworks. Once the grid cell is imposed on the landscape, there is some kind of rule that takes all the possible attribute values and picks the value. In some cases, this is a rule like "highest value" (as on an aeronautical chart), which is a dominance rule. In other cases, an optical system adds the energy detected. Thus, the rules are a part of the original geographic measurement as well.

This approach to transformations will be introduced by an example. While a three-by-three or four-by-four matrix can be quickly comprehended, a seventeen-by-seventeen matrix (for all the frameworks listed in Figure 2) is difficult to describe or communicate. A subset of measurement frameworks used for surfaces will illustrate the approach.

The rows and columns of Figure 3 list some of the major alternatives for the representation of surfaces. The first "Points with Z" refers to "Spatial objects" where a continuous surface value is measured at an isolated point feature. The second representation is isolines, closed contours that measure the location of a given surface value. Digital Elevation Matrix (DEM) refers to a regular, spatially controlled measurement of elevations. The fourth is the Triangulated Irregular Network (TIN) whose triangles establish relationships of slope between spot heights.

Figure 3: Surface-oriented transformations

In \ Out	Points (w.Z)	Isoline	DEM	TIN
Points (w.Z)	Interpolation	Interp. & trace	Interpolation	Triangulation
Isoline	Interpolation	Interp. & trace	Interpolation	Triangulation *
DEM	Interpolation	Interp. & trace	Resampling	Triangulation *
TIN	Extraction	Tracing	Extraction	Simplify/ Refine

* denotes a triangulation operation that may produce overly dense triangles without some filtering.

The cells in this four-by-four matrix give a label for the procedure that converts information in the row dimension to the column dimension. The three-by-three matrix in the upper left (lightly grey) is filled with one form or another of interpolation. This operation provides a good example of how a transformation combines relationships and assumptions (axioms) to produce new information.

Interpolation

Interpolation involves a transformation to determine the value of a continuous attribute at some location intermediate between known points. Part of this process requires relationships – knowing which points are the appropriate neighbors. The other part involves axioms – assumptions about the behavior of the surface between measured locations. The balance between these two can vary. Some methods impose a global model, such a fitting a *trend surface* to all the points. Most methods work more locally. The top left cell in the matrix poses the classical problem: given a set of point measurements, assign values to another set of points. This requires two steps. First one must discover the set of neighboring points for each desired location, using a variety of geometric procedures. Then one must apply some rule to determine the result.

Once the neighbors are collected, the problem of assigning a value resolves itself into the rules of combination. A dominance rule will not yield a smooth surface, since it will assign the same value to a neighborhood (usually the Voronoi polygon). A contributory rule usually involves a distance weighted average of the neighbors. Various forms of interaction rules are in use as well. SYMAP had a much-copied interpolation system that weighted points so that distance and orientation to other points were considered. Each method operates by using certain relationships, plus some assumptions about the distribution of values between points. The differences between various forms of interpolation reflect various assumptions about the nature of the attributes.

The process of producing a DEM with uniformly spaced points is just a special case of interpolation for scattered points. To produce isolines, instead of requesting a value at some arbitrary point, the contour specifies the height, and the interpolation discovers the location. Functionally, this is not very different, since the procedure for a weighted average can be algebraically restructured to give a coordinate where the surface has a given value. The manual procedures for contour drawing involved linear interpolation on what amounts to a triangulation (Raisz 1948). In addition to the interpolation, the construction of isolines requires *tracing*, the process of following the contour from neighborhood to neighborhood. Usually, this procedure involves some assumptions about the smoothness of the surface, since the shape of the contour cannot be really estimated from the original point measurements. Tracing also involves relationships between adjacent contours, even those not created with the same neighborhood of points. Parallel contours imply slope gradient and aspect properties, along with other interactions caused by ridges and courselines (Mark 1986). Thus, tracing contours involves many more relationships than a simple decision about the value at a point.

If the input consists of a set of contour lines, the procedure for scattered points still applies. Interpolation will need to establish neighbors, but neighbors between adjacent contours as well as along the lines. Finding the nearest point on the two adjacent contours does not ensure a correct reading of features such as ridges or courselines. This straight line is a simplification for the line of steepest descent. Linear interpolation then proportions the value between the two contour values.

When the input values are organized in a grid structure, the matrix provides the means to access neighbors directly. To produce output for scattered points, the rules can be applied on the immediate neighbors in the grid. To trace contours, the grid values are used to estimate values in the area between them.

Producing a matrix output from a matrix input is a common requirement. Unlike the vector method where the coordinates can be transformed fairly directly, a matrix is delineated orthogonal to a given spatial reference system and with a given spacing. If a different cell size or orientation is needed, the values will have to be converted by *resampling*. For continuous variables, there is no real difference between resampling and interpolation. Sometimes, a simple dominance rule is used; each new grid cell gets the value of the nearest input grid cell. As long as the spacing is not wildly mismatched, this may produce a reasonable representation. For remotely sensed sources, the 'nearest neighbor' interpolation retains a combination of spectral values actually measured by the sensor. It does mean that each value has been shifted from the position at which it was measured by as much as 0.707 times the original pixel distance. Alternatively, it is common to use a contributory method to weight the change over distance using a various formulae, such as bilinear, cubic convolution, or higher order polynomials. Each function imposes different assumptions about the continuity of the surface.

By contrast with the nature of interpolation problems, a TIN provides its own definition of the neighborhood relationships; it also defines without ambiguity the linear interpolation over the face of the triangle. A transformation from a TIN source has much less work to perform. Once a point can be located inside the proper triangle, it is a matter of extraction. Conversion from one TIN to another is a generalization problem of refining or simplifying the representation inside a set of constraints.

Generalizing from the example of surfaces

As the explanation of surface transformation shows, a transformation can be explained in terms of a neighborhood relationship and a rule to process attributes. Temporal relationships can also be included as a form of neighborhood. This leads to a four-way taxonomy of transformations based on the degree to which the information is inherent in the data model or must be constructed through other kinds of information. This can be seen as a two-by-two matrix based on whether the neighborhood is implicit or discovered and the attribute assumptions are implicit or external.

Case 0: Transformation by extraction – When the source contains all the information required, it provides both the neighborhood relationship and the attribute assumptions to make a transformation look easy. Extraction is usually unidirectional. For example it is possible to create isolated objects from a topological vector database without much trouble, as long as the desired features are identified somewhere as attribute values.

Case 1A: Transformation based on attribute assumptions – In some cases, the transformation keeps the geometric entities intact, and works just with the attributes of those objects. Some of the steps performed on a base layer of polygons fall into this class, but the simplest form involves a raster with a uniform set of pixels. A common example is the transformation which takes a few axes of continuous spectral data and produces classes.

Case 1N: Transformation with geometric processing only – It is even rarer to use just the geometric component. Given two coverages of polygons, it is possible to convert the areas of one into attributes of the other. This is performed entirely as a geometric procedure, using the identifiers of the polygons to tabulate the areas in the correct attribute columns.

Case 2: Complete transformation – The most interesting forms of transformations are ones that combine geometric (neighborhood) constructions along with attribute rules. The interpolation problem discussed above is an archetype, because the two phases are quite distinct. Areal interpolation also falls into this class, even though it deals with areas not points, the two phases combine in much the same way. This relationship shows that this taxonomy combines those functions which are similar in their purpose, not just in the geometric form of their input.

A buffer around a road is also a complete transformation. It uses a simple neighborhood rule (all space within a certain distance of the road), and a simple dominance rule (areas near any road overrule anything else). A polygon overlay produces the geometric raw material for a suitability analysis. The next steps must take up the combination of the attributes now placed in contact. Various approaches to suitability use dominance, contributory or interaction rules, depending on the fit to the purpose. The general scheme of attribute rules that apply to spatial neighborhoods also apply to overlay processing and simple operations inside one measurement framework. This scheme incorporates Tomlin's successive broadening of neighborhood, but adds the formalization of the attribute rules. The important distinctions are not those of geometric form, but related to the basic structure of how the information was constructed.

CONCLUSION

A unifying scheme for transformations requires only two elements: a geometric neighborhood plus a rule to combine or process attributes. The rules fall into three classes (dominance, contributory, and interaction) based on the treatment of multiple attribute values. Viewed in this way, the operations of a GIS (including map overlay analysis, neighborhood operations, plus the items now treated as transformations) can all be relocated as various kinds of transformations.

ACKNOWLEDGEMENTS

Parts of this paper first appeared in *Exploring Geographic Information Systems,*© 1997 John Wiley and Sons and are used here with permission.

REFERENCES

Albrecht, J. (1995). Semantic net of universal elementary GIS functions. *Proceedings AUTO-CARTO 12,* 235-244.

Burrough, P. A. (1992). Development of intelligent geographical information systems. *International Journal of Geographical Information Systems* *6*(1): 1-11.

Chrisman, N. (1995). Beyond Stevens: A revised approach to measurement of geographic information. *Proceedings AUTO-CARTO 12,* 271-280.

Chrisman, N. (1997). *Exploring Geographic Information Systems.* John Wiley, New York.

Clarke, K. (1995). *Analytical and Computer Cartography.* Second edition. Prentice Hall, Englewood Cliffs NJ.

Goodchild, M. F. (1987). A spatial analytical perspective on geographical information systems. *International Journal of Geographical Information Systems 1*(4): 327-334.

Hopkins, L. (1977). Methods of generating land suitability maps. A comparative evaluation. *American Institute of Planners Journal 43:* 386-400.

Mark, D. (1984). Automated detection of drainage networks from digital elevation models. *Cartographica 21*(3) 168-178.

Nyerges, T. (1991). Analytical map use. *Cartography and Geographic Information Systems. 18:* 11-22.

Raisz, E. (1948). *General Cartography.* McGraw Hill, New York.

Sinton, D. (1978). The inherent structure of information as a constraint to analysis. Mapped thematic data as a case study. In *Harvard Papers on Geographic Information Systems,* vol. 6, ed. G. Dutton. Addison Wesley, Reading MA.

Tobler, W. (1976). Analytical cartography. *The American Cartographer 3*: 21-31.

Tobler, W. (1979a). Cellular geography. In *Philosophy in Geography*, eds. S. Gale and G. Olsson, p. 379-386. Reidel, Dordrecht NL.

Tobler, W. (1979b). A transformational view of cartography. *The American Cartographer 6*: 101-106.

Tomlin, C. D. (1983). Digital Cartographic Modeling Techniques in Environmental Planning. unpublished Ph.D., Yale University.

Tomlin, C. D. (1990). *Geographic Information Systems and Cartographic Modeling*. Prentice Hall, Englewood Cliffs NJ.