INT. J. GEOGRAPHICAL INFORMATION SCIENCE, 1999, VOL. 13, NO. 7, 617–637

Research Article

A transformational approach to GIS operations

NICHOLAS CHRISMAN
Department of Geography, University of Washington, Seattle WA 98195-3550
USA
e-mail: chrisman@u.washington.edu

Abstract. GIS research has a continuing thread devoted to classifying the operations performed by analytical software. Prior efforts to classify GIS operations have limitations and inconsistencies, often arising from an attempt to establish overly direct links between geographic procedures and arithmetical operations. The transformational view of cartography provides a more solid basis for classifying GIS operations. This paper presents a new scheme for geographical transformations based on measurement frameworks as the principal distinction. Transformations between measurement frameworks can be summarized in terms of a spatial neighbourhood and a rule to process attribute information. This scheme organizes most analytical GIS operations by their geometric and attribute assumptions.

1. Demystifying the conversion of data into information

A geographical information system (GIS) converts raw data into more useful information (an anti-entropic feat) through a maze of analytical procedures that are poorly explained and often misunderstood. Most accounts place a strong emphasis on the analytical toolkit (embodied in software) as the agent of the conversion. The GIS is presented as a marvellous machine that can distil valuable information from huge stockpiles of ordinary data. Such a claim should not be propagated without documenting the mechanisms that enhance the content of a data source. While software can compute any number of new relationships, only a few of these hold potential value. Users of these software tools should be better informed of the applicability of various software tools to their particular situation so that the information serves the intended purpose.

Some authors define GIS around a specific range of tools to integrate diverse forms of geographic information (Cowen 1988, Maguire 1991). Yet, behind the grand claims, the procedures of integration vanish into a maze of special cases and an ever-expanding list of software functions. Students and neophyte users are at a loss to navigate the maze we have built. Yet, despite the lack of articulate description, practical applications of GIS routinely manage to use the tools to create new information not inherent in the available sources. The goal of this paper is to present a taxonomy for GIS operations that is sufficiently comprehensive and flexible while reducing the maze to a clear structure. This paper will begin with a review of the

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online ©1999 Taylor & Francis Ltd
http://www.tandf.co.uk/JNLS/gis.htm
http://www.taylorandfrancis.com/JNLS/gis.htm

previous approaches to organize the diversity of GIS operations before it presents its new approach.

2. GIS operations: the literature

Research on GIS converges from a number of origins. For example, there are clear examples of the use of map overlay in landscape planning dating to the beginning of the twentieth century (Steinitz *et al.* 1976). The geographical literature on integration of sources goes back through the nineteenth century (Harvey 1997a), at least in intent if not in the particular use of map-based operations. The self-conscious consideration of technique does not reach back so far, and it is in this recent period that authors have tried to classify GIS operations.

2.1. *GIS literature*

The GIS literature has a series of alternative schemes used to present the different kinds of operations. In fitting with the practical origins of GIS, the early descriptions of operations had no particular theoretical model. They were lists of known functions, organized often by data flow (with the dominant approach leading from input through processing to output displays). The IGU Commission on Geographical Data Sensing and Processing produced a number of such lists in the 1970s, but the paper produced by Dangermond (1983) has appeared in a number of versions. After considering various forms of input and editing, the list of 'manipulation techniques' is organized into ten major headings: data retrieval, map generalization, map abstraction, map sheet manipulation, buffer generation, polygon overlay and dissolve, measurement, grid cell analysis, digital terrain analysis and output techniques. Some of these are extremely specific and others are expansively inclusive. This list served its function to introduce the currently known capabilities, but it does not give a framework to understand functions that might not be currently available. The list produced by the Technology Exchange Working Group of the Federal Interagency Coordinating Committee For Digital Cartography (Guptill 1988) is even more of a checklist of current capabilities. A later book chapter on 'GIS Functionality' (Maguire and Dangermond 1991) shows only limited improvement. There is a detailed consideration of structuring data input, a short exhortation to avoid any kind of transformation or restructuring of these sources, then two categories for query and analysis. Complex GIS operations are not given any particular structure. To a large extent, this kind of unordered list of functions remains the state of software descriptions from the commercial sector.

Perhaps the most widely cited scheme for spatial data handling operations is Tomlin's (1983, 1990) *Map Algebra*. This scheme presents map operations in a sequence from the simple to the complex. The idea of the algebra is that the distinct operations can be combined to produce more complex results. The use of the term algebra emphasizes the symbolic manipulation of numbers—representing attributes. (Tomlin calls the more complex combinations 'models', though he does not confuse them with data structures.) The simple operations work on a single map, followed by those that work locally on two maps, and so on. This scheme starts from the syntactic properties of unary and binary operators, extending these to consider the spatial scope of an operation: local, focal, and extended neighbourhoods. Each operation is seen as a primitive element of an arithmetic that can be combined in various orders to suit more complex purposes. For this to work, each operation must produce the same structure it had to start with. This assumption contradicts the observation that cumulative GIS operations can change the raw data content into more useful information. This scheme simplifies what the tool does to pushing numbers around. For example, if there is a grid of population, the sum of all the cells is the total population of the region. If a 'Focal Sum' operation is employed, each cell will record the total population within some search radius (perhaps how far one could be expected to walk to a bus stop). However, the resultant grid will no longer have the property that the sum will equal the total regional population, due to double counting. Tomlin's algebra deals with the syntactic constraints on the operation, but leaves the interpretation of meaning to the user.

Tomlin's terminology for operations becomes a bit obscure for the more complex operations. Viewsheds and other more complex functions are far from the simple manipulation of a fixed set of immediate neighbours, yet they are classed with 'focal' operations. The concept of 'extended neighbourhoods' is not enough to deal with iterative application and more complex algorithms. Joseph Berry's reformulation of Tomlin's scheme names four 'classes' of operations: 'reclassify map categories, overlay maps on a point-by-point or region-wide basis, measure simple or weighted distance and connectivity, characterize cartographic neighbourhoods' (Berry 1987, p. 123). This list is driven by the number of input values (one map, two maps, etc.), and seems too closely tied to a given package of software. Hadzilacos (1996, pp. 242–243) also offers four classes of operations, based on similar, but more generic logic: derive computable attribute, compute spatial, reclassification and overlaying. His analysis of these groups is not based on distance as Tomlin, but still largely syntactic.

Goodchild (1987) followed the syntactic direction of Tomlin, adding some neglected elements from the viewpoint of a spatial analyst. In particular, spatial analysis (such as location-allocation) involves relationships between primitive objects. Goodchild produced an arrangement based on three dichotomies: one class/two class (the syntactic concern about number of parameters), attributes/ geometry (the source of the data), and single object/object pair, resulting in six 'classes of spatial analysis':

1. Operations requiring access only to the attributes of one class of objects. In this case, the model reduces to a simple table and it is likely the analysis can be handled by a statistical package.
2. Operations requiring access to both attributes and locational information for a single class of objects; examples include calculating simple spatial descriptive statistics such as location of mean centre and dispersion.
3. Operations which create object-pairs from one or more classes of objects.
4. Operations which analyse attributes of object-pairs; examples include autocorrelation indices and nearest neighbour analysis.
5. Operations requiring access to attributes and locational information for more than one class of objects or object-pairs. Spatial interaction modelling requires access to origin and destination objects and attributes of associated object-pairs.
6. Operations which create a new class of objects from an existing class, including generation of Thiessen polygons from points or buffer polygons around line segments. (Goodchild 1987, p. 332)

Despite a discussion of surfaces in the paper, this list seems focused on handling discrete objects in location-allocation problems, not polygon overlay, neighbourhood filters, viewsheds, and all the other operations that form the core of Tomlin's toolkit.

The focus on pairs of objects is important; ten years later very few software packages provide tools to handle them. But object pairs are also too limited to handle the one-to-many relationships that occur with neighbourhood operations. In summary, Goodchild's list simply unfolds the established divisions between location and attribute without providing any particular help in understanding what GIS software does.

Burrough (1992) included a scheme to organize GIS operations inside his plea for 'intelligent' systems which could treat uncertainty more effectively. He specified three kinds of functions which can be applied to determine a new value for a point: values of all attributes of the point itself, values derived from contiguity (neighbourhoods), and values derived from temporal association. This, like the exposition in both editions of his textbook (Burrough 1986, Burrough and McDonnell 1998), is based on Tomlin's map algebra. The issue of inexactness is left for nine 'classes' of operations (paraphrased from (Burrough 1992, pp. 3–6)):

1. New attribute from exact-valued attributes of exact objects
2. As Class 1, allowing non-exact attributes but of exact objects
3. New attribute based on attributes in a neighbourhood (discrete objects)
4. As Class 3, for continuous surfaces
5. New attribute assigned to neighbouring points (e.g. buffers)
6. New vector objects created
7. Geometrical measures of objects
8. Summaries and reports generated
9. Data management (rectification, projection, scale change, join, snap, etc.)

This list hides the clarity of Burrough's attempt to add intelligence to analysis, as it does not provide for any symmetry between vector discrete objects, continuous surfaces and grid cells. The linkages between the three kinds of functions and the nine classes of operations, while promising, remain undeveloped. Considering the call for greater 'intelligence', the nine classes end up sounding like a list of functions in a commercial package, with some slight additions for inexactness.

A group of four GIS researchers (Giordano *et al.* 1994) offered an evaluation of prior taxonomies of GIS operations and their own proposal. They evaluated a set of lists including Dangermond's (1983) and others that are primarily unstructured lists. They evaluated these taxonomies in terms of structure, scope, detail, and internal consistency. They seem to approve of hierarchical structure for GIS operations. While this might make for easy pedagogy, like most of the other taxonomies listed above, they underestimate the transformational basis of important GIS functions. It is fairly easy to show that a 'from-to' graph connecting all possible data structures is not reducible to a tree. Their 'conceptual model' places operations into three general groups input, analysis, and output. Though they make great rhetorical importance of analytical functions, they also seem to suggest that analytical functions do not change the form of the data: 'These functions require data of a specific structure, and return data with the same structure' (Giordano *et al.* 1994, p. 49). Thus, their analysis stage adopts Tomlin's syntactically defined algebra with complex operations built from combinations of simpler steps. They do consider how successive operations influence the 'uncertainty' associated with a feature in time, space and attribute, but the paper remains fairly firmly linked to the Tomlin approach, and it finds that the resulting taxonomy applies more directly to raster software than it does to vector implementations. The only explicit mention of transformations in their scheme is in the category 'restructuring' that occurs both in input and output groupings. Restructuring includes the data structure transformations presented by Clarke (1995) in table 2 (see below). The analytical functions listed (logical operations, arithmetic operations, overlay operations, geometric property operations, geometric transformation operations, geometric derivation operations) do include operations that change the nature of the input (such as interpolation and overlay), though the use of the term transformation refers to projections and co-ordinate manipulations that do not alter the data representation. Their taxonomy was motivated by a need to describe software functions across vendors, but it does not provide much assistance understanding how GIS integrates information from different sources. The paradigm of input-processing-output, which dates back to the earliest days of computer cartography (Tobler 1959) masks the transformations that occur in each step.

In a more inductive approach, Albrecht (1995) has described a method to develop commonalities between GIS operations using a semantic network. This research relies upon a survey of users (in this case, students of GIS) who are asked to organize commands into groups. While such research may provide useful evaluation of teaching methods, it cannot provide much assistance with operations to which the students have not been exposed.

2.2. *Transformations in cartography*

While the dominant school views cartography as a communication process, there has always been another thread of research focused on transformations that has more direct bearing on classifying GIS operations. The classic transformation involves the mathematical conundrum of transferring the nearly spherical Earth onto a flat piece of paper, the process of map projection.

The key importance of a map projection is not in the mathematical detail. Projections demonstrate how measurements taken from one kind of geometric model can be transferred to another model, subject to certain constraints (Maling 1973, Snyder 1987). 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. These properties can be demonstrated by mathematical analysis (differential equations). From the starting point of map projections, Waldo Tobler broadened into a 'transformational view of cartography' (Tobler 1979) that considered all operations as transformations of information content.

Tobler developed his analytical cartography in the era of Chomsky's transformational grammars, an attempt to redirect linguistics. Chomsky (1968) sought a universal grammar hidden behind the confusing details of actual languages. The concept of 'deep structure' shifted focus from the syntax of language to its underlying meaning. The work of Chomsky also motivated interest in the artificial grammars of computing, providing some basic theory to what had been practical issues. More widely, the concept of transformations swept the intellectual scene, as in Leonard Bernstein's Norton Lectures on the structure of music (Bernstein 1976). Tobler associated his work with Bernstein's general intent, while stating that his transformational view cartography was not particularly 'Chomskian' (Tobler 1979, p. 101). Tobler did present his work as a part of changing the paradigm of cartography, moving transformations from the issues of projections and generalization to the whole of cartography. The article mentions dozens of geographical transformations with no particular organization, other than mathematical relationships of inverses, closure, differentiability, and so on. There is one passing reference to the geometric primitives (point, line, area) and a three-by-three table of possible transformations (Tobler 1979,

p. 104). Although Tobler did not detail its contents, this matrix appears in Clarke (1995, figure 11.1 page 184) and Unwin (1981), among others (table 1).

In this matrix, a buffer around a road would be considered a line-to-area transformation, but so would a conversion from contour lines to a triangular irregular network (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 do not encompass the variation in transformations. The geometric form of input and output offers only a weak guide to the 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. Basically, the representational primitives (point, line, area) do not encapsulate the 'deep structure' of cartography completely.

Clarke codified many of the procedures that Tobler mentioned into another matrix based on transformations between 'data structures' (Clarke 1990, figure 10.09 p. 199, Clarke 1995, figure 12.8 p. 243). He presents a group of four generic data structures, and labels the possible transformations between them (table 2).

The diagonal includes two kinds of operations. Scaling refers to changes associated with cartographic scale (meaning generalization, filters and other similar operations), while 'dimensional' refers to the operations presented in table 1 that move between different dimensions of objects. The off-diagonal elements of this matrix are not presented in great detail; the term 'structural' refers to changes in data structure similar to the conversions between raster and vector or reverse. The discussion of the cells in this four by four matrix occupies three pages in Clarke's book, just enough to whet the appetite. Clarke's matrix presents a critical aspect of GIS operations beyond syntactical issues, but the explanation of the elements of this matrix must be extended and presented in greater detail.

Table 1. Cartographic transformations based on dimension of object.

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

Table 2. Data structure transformations (after Clarke).

	Entity by entity	Topological	TIN	Grid
Entity by entity	Scaling dimensional	Structural	Structural	Vector to raster
Topological	Structural	Scaling dimensional	Structural	Vector to raster
TIN	Structural	Structural	Scaling dimensional	Structural
Grid	Raster to vector	Raster to vector	Structural	Scaling dimensional

2.3. *Conclusions from the literature*

There a number of useful aspects of the various taxonomies listed above. Each one responds to a deep-seated urge to do science by naming things. There is a risk in advancing another taxonomy in that it might be treated as just another set of names. The GIS literature has focused mostly on the mechanics of the procedures (how they obtain their parameters, what kind of mathematical operations are performed), and has tried to extend the analogue with algebra beyond its applicability. These taxonomies shrink back from the semantics, preferring to settle for the syntax, leaving the meaning for the users. While a GIS can indeed perform many operations in sequence, the process of creating new information requires transformations that produce different content. In place of the four criteria advanced by the Giordano *et al.* paper (structure, scope, detail and internal consistency), I offer a taxonomy that tries to deal with the issue of information content by extending the framework of transformational cartography. A transformational approach can focus on the differences in the information, rather than looking at differences in the software.

3. Measurement frameworks

What is missing from both the cartography and the GIS literature is an explanation of the different reasons for using geometric primitives to represent geographical phenomena. Not all lines are created equal; not all attributes can be combined at will. For example, some overlays produce useful analytical results (such as change detection from land use snapshots at different dates), while others (such as census tract population characteristics overlaid with toxic release plumes) are much easier to misinterpret. While each attaches attributes to polygons, the measurement rules differ so much that the usefulness of the tool is called into question. As a further example, an iso-declination line on an aeromagnetic survey, a road in a highway maintenance file, and a county boundary on a choropleth map might all be represented as lines. However, the meaning of some further operation (such as a buffer around these lines) would be totally different. Each line could have an attribute value of 2.7, but the numerical value tells you nothing about the operations that might be appropriate. Burrough's call for greater intelligence in GIS processing might be better served by recognizing the diverse meanings attached to geographic representations, not trying to force-fit them into some deductively derived schema. The division between measurement and representation has been recognized by a number of authors, such as Unwin (1981), but it is hard to clarify the distinction when software is organized so routinely as if they were the same.

The heart of this paper's proposal arises from a consideration of the fundamental choices made in obtaining measurements. It is quite commonplace to observe that geographical information includes a spatial component, a temporal component, and some set of attributes. Each of the taxonomies mentioned above repeat this common division. Sinton (1978) moved from recognizing the fundamental components of geographical information to recognizing that the seemingly programming-oriented differences between vector and raster data structures derive from different attitudes towards measurement. He discerned three possible roles for the spatial, temporal and attribute components. In order to measure one component, one of the others had to be fixed while the third serves as control. In this context, control denotes a mechanism of restraint on the variation of that component. For example, a tide gauge must be fixed in one location and control the rate of temporal sampling so that the height of the water (the attribute) can be measured. Other forms of measurement might control for a specific height of tide, then record the exact time (or close a floodgate). Temporal databases often require two forms of control, so that no component is fixed (Langran 1991). The role of control is one example of the differences in ontologies; Smith and Mark (1998) have suggested various forms of 'fiat' by which objects are made crisp in order to make them tractable. A measurement framework (Chrisman 1995, 1997) is a conceptual scheme that establishes rules for control of other components of a phenomenon to permit the measurement of one component.

Each of the 'measurement frameworks' listed in table 3 collects different forms of geographical measurement which cannot be directly combined with the other frameworks unless certain assumptions are made. The choice of control implies decisions about resolution and accuracy; some portion of the phenomenon is lost in order to emphasize or preserve others. These trade-offs are an important part of spatial data handling and the core of the transformational approach described by Tobler and Clarke.

It is particularly important to note that the technique used for representation may not be the same as the measurement framework. Such a composite situation imposes the limitations of each method in succession. 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. Some measurement frameworks associate directly with certain representation structures, while others require substantially more effort. These transformations imposed for data representation are what Giordano and others called 'restructuring' in their input grouping and Clarke termed 'structural' transformations. It is important to recognize that these are not simple technical steps, but may require careful consideration of the assumptions imposed. While the Giordano paper places the emphasis on the 'uncertainty' accumulated, the issue is more the transformations of meaning.

Table 3 groups a series of measurement frameworks according to a scheme developed from Sinton's original concept. The basic concern was to contrast spatial control (usually in the form of regular tessellations of pixels or cells) with attribute control. This extended list begins with a set of temporal frameworks not normally given much attention in the spatially-centred world of GIS. Sinton's scheme has played a key role in efforts to extend GIS to the spatio-temporal domain (Langran 1991).

In Sinton's paper, as in Burrough's work (Burrough 1986, 1996, p. 4, Burrough and McDonnell 1998, p. 2), the choropleth map is usually treated as a simple instance of a vector approach. In its cartographic data structure, a simple vector **representation** may suffice, but a choropleth map involves more steps in its creation. In the case of the categorical coverage, the boundary lines are measured as carefully as possible to serve as boundaries of the particular category being captured. For a choropleth, the boundary lines are not placed on the map to measure the attribute, but to locate the zones. The zones are used as a template to aggregate some attribute phenomena; that is why a common geometry can be used for so many attributes. So, choropleth must be recognized as a composite. First there was an attribute control to identify certain zones. The lines measure the boundaries of those categories to any resolution required. On the base of these zones, an attribute rule must be applied to measure the value within that zone. Sometimes the rule is addition, sometimes it is an average or a density. These distinctions are at least as important as the 'level of measurement' for the attribute discussed by Chrisman (1995, 1998).

Attribute rules apply in most of the area control situations, both choropleth and raster. The rules listed under 'area-based control' are the terms used in practical

Table 3. Expanded list of measurement frameworks.

	Temporal frameworks
Snapshots	Temporal repetition of any other measurement framework
Transactions	Discrete events located freely in time
Period	Function (mean, max, min, etc.) applied to temporal series
	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
Categorical coverage	Network induced by exhaustive classification
	Space controlled frameworks
Point-based control	
Centre 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 (e.g. 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
	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
Space-time series	Control by time-period and by zone (usually sum)

applications of gridded databases for landscape planning. The 'predominant type' rule differs from the categorical coverage in that the grid cell specifies the spatial unit on which the rule applies. By contrast, the categorical coverage as a formal model does not presuppose any particular resolution (Frank *et al.* 1997). In practical terms, of course, a categorical coverage is rarely totally pure. There is usually an explicit 'minimum mapping unit' or an implicit sense of scale. Still, the spatial unit of a categorical coverage is not as rigid as the pixel in a spatial control framework. These two rules are the classic opposition of the raster-vector debate; many other rules might be invented—there is no real limit.

Some measurement frameworks do not fit into Sinton's simple taxonomy. The Triangulated Irregular Network (TIN) is often considered as a special structure for surfaces. There is substantial confusion about TINs, largely due to the sequence of operations imposed by many software packages. A TIN representation produced by a software package has only the information content that the software can glean from the source (basically, it is the result of a transformation). As a measurement framework, TIN does consist of point measurements, just like simple isolated points with a Z height, but the TIN also includes a set of relationships between points – a set of adjacencies in the form of triangles. Ideally, these triangles have been chosen to represent areas of consistent slope (gradient and aspect) (Peucker and Chrisman 1975, Males and Gates 1978, Mark 1979). The edges of triangles should follow ridges, stream courses, and breaks in slope. This is a much more complex measurement operation than simply controlling the height as a contour does. The payoff comes in the relationships captured. A TIN structure created by planar Delaunay relationships cannot be guaranteed to capture the slope relationships.

The list of measurement frameworks also makes a distinction between isolated objects (what Clarke calls 'entity by entity', see table 2) and connected objects. Despite the continued treatment of vector databases and the increased attention to object-oriented methods, these two groups are often confused or treated as simply more or less intelligent data structures. Some of the differences are based on more fundamental measurement decisions. Isolated objects might be points, lines, or areas. Each object is recognized as distinct from the surrounding void. By contrast, connected objects are defined by their relationship to each other. A given road must lead to a junction with other roads, not only as a data editing concern, but as a matter of its geographical meaning. A land use polygon is created by locating the boundary with some adjacent use. These differences of measurement impose different rules on the use of the information.

Even more importantly, attribute values depend on the rules used to measure them. As suggested above, population figures are inherently associated with addition (being counts), while population densities require weighted averages. An overlay of a 'presence/absence' wetland coverage with a predominant type land cover raster database could produce the impression of wetland loss simply due to the difference in the rule applied to generate the two sources. Incompatibilities between sources, derived from differences in measurement rules, often cause a GIS analysis to produce meaningless results.

Much of the diversity of measurement frameworks has not been apparent because GIS software has such limited choices for representations (data structures). Programmers have made representation structures the limiting factor, not the assumptions behind the information. Users have been at the mercy of the limited language provided. 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 (Codd 1981). The concept of measurement framework provides a more direct approach to the axioms relevant for GIS than the previous literature.

4. Understanding transformations

With the more solid foundation of measurement frameworks, a more useful taxonomy of GIS operations can be built. The approach developed here begins with the transformational approach developed by Tobler and Clarke described above. Before presenting the specific taxonomy, it is important to consider the theoretical grounding for a transformational approach.

4.1. *Theoretical grounding*

Tobler's transformations are developed as an application of mathematical analysis. His interest in a transformation focuses on properties such as inverses, commutability, and so on. These interests parallel the algebraic interests of most of the GIS literature, yet transformations suggest more than syntax.

In the studies of science and technology, the term 'translation' has been applied to the process by which ideas and practice from one group become useful to others

(Latour 1988, Star and Greisemer 1989, Fujimura 1992). This word highlights the connection to meaning (semantics) as well as to form or syntax. It also signals that information may be lost or rearranged. Clarke (1997, p. 95) associates translation problems with the successive flaws in communication through a number of persons as in the children's game of 'telephone'. Galison (1997) goes even further in using the linguistic metaphor, and emphasizes the incompleteness of many interactions across boundaries in the sciences. Instead of complete languages, he refers to pidgins and activity in 'trading zones'. In GIS, movement of information from one form to another induces massive changes of entities and relationships, perhaps more dramatic than those hinted by any linguistic metaphor (transformational grammars or translations). While human languages might serve as vehicles to express the same meanings, geographical measurement frameworks choose to emphasize or suppress different elements of the world (Harvey 1997b, Harvey and Chrisman 1998).

The term applied in this paper—transformation—is chosen to evoke the heritage of the transformational cartography that provided much inspiration for GIS, even though the linguistic evocation is more deterministic than intended. In Tobler's (1969, 1979) original formulation, transformations were most interesting if they had an inverse, meaning that the information was not lost in the process. Few of the practical GIS transformations have symmetrical inverses. Most sequences of operations produce significantly different results if performed in a different order (Heuvelink and Pebesma 1999). The process of translation is partial, and the differences in assumptions limit the ability to return to the prior representation. Despite the lack of mathematical symmetry, partial translations seem to be much more common in interdisciplinary relationships than perfect inverses and lossless encodings. After all, the underlying direction is counter-entropic; GIS operations (if they work) create new information by making explicit the relationships implicit in the source material. Others (such as Giordano *et al.* 1994, p. 48)) have addressed the same problem, but not with the same focus on how transformations can be made to work.

4.2. *Components of transformations*

Given data within a particular measurement framework, it is most direct to produce a result in the same framework. A grid of values can be most easily processed into another grid with completely identical spatial reference, resolution, and all the other assumptions. To generate a different result even in something as simple as the cell spacing, rotation or grid origin, many new assumptions may be required. These assumptions are required to fill in the gaps in either space, time, or attributes in the original source. It is the contention of this paper that measurement frameworks serve as a guide to the assumptions involved.

Transformations between most forms of geographical information can be performed with two components: one to handle space, thus creating a neighbourhood, and the other to handle attributes, a rule of combination. Temporal transformations can be handled as special forms of neighbourhoods or special attribute rules. Neighbourhoods can be defined rather flexibly, following the general scheme of Tomlin—moving from the purely local relationships inside one object through immediate neighbours to more complex relationships based on distance and 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' (1946) levels of measurement, but his 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 common contributory rule. Finally, an *interaction* rule uses not just each value, but the pairwise combinations of values. This grouping of rules does not match the syntactic distinctions between unary and binary operators, but it emphasizes the relationships between input and output.

This taxonomy of attribute rules serves to explain the differences among the approaches to area-based spatial control frameworks (table 3). 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 the minimum flying height by degree square on an aeronautical chart), which is a dominance rule. In other cases, a remote sensing detector accumulates the energy incident over a slice of time and space (a contributory rule). Similarly, in a choropleth context, the difference between 'raw values' such as population counts and 'derived' ratios such as population come from the underlying rules of those attributes (Chrisman 1995, 1998). These attribute rules are often part of tacit knowledge associated with the discipline responsible for the data source. Unlike the geometric component of GIS, the attribute rules cannot be determined without understanding the whole measurement operation.

4.3. *An example: surface transformations*

This approach to transformations is introduced by an example. While a three-by-three or four-by-four matrix can be quickly comprehended, a nineteen-by-nineteen matrix (for all the frameworks listed in table 3) is difficult to describe or communicate. A subset of measurement frameworks used for surfaces will illustrate the approach.

The rows and columns of table 4 list some of the major alternatives for the representation of surfaces, a surface-specific selection from the four 'data structures' used by Clarke (1995, figure 12.8) in table 2. 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. Both 'Points with z' and Isolines are forms of Clarke's 'entity by entity' approach because each object is isolated from the others. Contours may have implicit topological relationships of nestedness, but since the lines do not cross, there are no nodes in the graph, and neighbourhoods are not easily constructed. Given a segment on a contour loop, it requires substantial geometric processing to locate the nearest points on the contours above and below. Digital Elevation Matrix

Table 4. Surface-oriented transformations.

In/out	Points with z	Isoline	DEM	TIN
Points with z	Interpolation	Interpolation & trace	Interpolation	Triangulation
Isoline	Interpolation	Interpolation & trace (extraction)	Interpolation	Triangulation*
DEM	Interpolation	Interpolation & trace	Resampling (extraction)	Triangulation*
TIN	Extraction	Tracing	Extraction	Simplify/refine

* produces overly dense triangulations without a filtering or simplification step.

(DEM) refers to a regular, spatially controlled measurement of elevations (a grid). The fourth is the Triangulated Irregular Network (TIN) whose triangles establish relationships of slope between spot heights. As described above, TINs in their pure form have lines chosen to represent ridges and courses, not just triangles connecting known points.

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. In two cases (for isolines and DEMs) it is possible to transform into the same structure without interpolation, simply by extraction. The conditions for these two cases are highly restrictive: for contours it would require selecting a larger contour interval as an exact multiple of the original (vertical dimension); for DEMs it would require a grid spacing at an exact multiple of the original (horizontal dimension). Both of these are special cases; anything else requires interpolation. The interpolation operation provides a realistic example of how a transformation combines relationships and assumptions (axioms) to produce new information.

4.5. *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 neighbours. The other part involves axioms—assumptions about the behaviour of the surface between measured locations. The balance between these two can vary. Some methods impose a global model, such as 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 neighbouring points for each desired location, using a variety of geometric procedures, and then apply some rule to determine the result.

Once the neighbours 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 neighbourhood (usually the Voronoi polygon). A contributory rule usually involves a distance weighted average of the neighbours. 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 (Shepard 1968). Splines use a different local model, based on the physics of a spring. 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 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 co-ordinate 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 neighbourhood to neighbourhood. 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 neighbourhood 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 can still be used. Interpolation will need to establish neighbours, in this case neighbours 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. Because contours are often densely sampled, some interpolation methods will not reach out to the neighbouring contours. Thus some software packages only permit conversion from isolines to TIN, and thence to other structures.

When the input values are organized in a grid structure, the matrix provides the means to access neighbours directly without a geometric search. To produce output for scattered points, the rules can be applied on the immediate neighbours 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 co-ordinates 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 neighbour' 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 $\sqrt{2}/2$ times the original pixel distance. Alternatively, it is common to use a contributory method to weight the change over distance using various algorithms, such as bilinear, cubic convolution, or higher order polynomials. Kriging uses a more statistical estimation technique, but ends up with similar mechanisms. 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 neighbourhood 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.

5. Generalizing from the example of surfaces

Summarizing the description of surface transformations, a transformation can be decomposed into a neighbourhood relationship and a rule to process attributes. Temporal relationships can also be included as a form of neighbourhood. 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 inferred through other relationships. This can be seen as a two-by-two matrix based on whether the neighbourhood is implicit or discovered and the attribute rules are implicit or external (table 5).

Case 0: Transformation by extraction—When the source contains all the information required, it provides both the neighbourhood relationship and the attribute assumptions to make a transformation look easy. Extraction is typically 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. In the reverse direction, it usually does not work. Isolated objects cannot become a connected network without substantial geometric calculation to create the connectivity and contiguity relationships. When stored as isolated objects, the topological information will be lost. In general, an extraction usually reduces the information content.

Case 1A: Transformation based on attribute rules—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 takes continuous spectral data in a raster and produces classes based solely on spectral values, not surrounding texture. This classification process takes cellular measurements of reflectance in the cell and produces a code referring to a category. The rule applied can be either supervised (based on closest match to selected prototypes) or unsupervised (based on a clustering algorithm and nearest neighbours in spectral space). Attribute processing is not always in the direction of reducing information content, as in classification. In some cases, a set of categories can be evaluated along a continuous axis, using some external source to upgrade the categories into continuous numbers, for example.

Case 1N: Transformation with geometric processing only—Though most GIS processes deal with the attributes, some 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 through polygon overlay, using the identifiers of the polygons to tabulate the areas in the correct attribute columns. An example of this transformation was applied to demonstrate that the tax assessments of swamp and waste by parcel did not match the wetlands zoning in Westport, Wisconsin (Sullivan *et al.* 1985, Chrisman 1997, pp. 225–229). In this case, and in most others, the geometric relationships create a new source of information only latent in the originals.

Case 2: Complete transformation—The most interesting forms of transformations are ones that combine geometric (neighbourhood) constructions along with attribute rules. The surface interpolation problem from scattered points discussed above

Table 5. Cases of transformation.

	Attribute rules	
Neighbourhood construction	Implicit	External
Implicit	0	1A
Discovered	1N	2

is an archetype, because the two phases are quite distinct. First, the geometric neighbourhood locates the points to be applied in a certain context, then an attribute rule (either linear or more sensitive to spatial configurations) generates the new attribute for the unknown point.

This taxonomy of GIS operations incorporates Tomlin's emphasis on neighbourhood, but adds the formalization of the attribute rules. It distinguishes between operations that reduce information content and those that increase it. If information content is to be increased, there must be additional assumptions required to justify the increase. Sometimes this is latent in the geometry or in the attributes, but in other cases it must come from external sources. The important distinctions are not those of geometric form (point, line, area), but related to the basic structure of how the information was formulated. A transformation uses the content of an existing data base to restate the information in another form, a kind of translation between dialects or even language families. Current GIS literature has focused on the geometric primitives at the expense of the attribute rules implicit in different measurement frameworks; it has focused on syntax, not semantics.

5.1. *Re-presenting surface transformations*

The four combinations of neighbourhoods and attribute rules occur in handling surfaces. Some of the cases are described as interpolation, although they require quite different processing. The transformation matrix in table 4 can be reinterpreted in table 6.

In this revised matrix some of the interpolation procedures are recognized to be different from the others. Only the operations that start with isolated objects (points or isolines) require complex processing to generate neighbourhoods, so these are the only ones to require full Case 2 transformations. When these isolated sources are transformed to non-TIN form, an attribute rule is required as well. Attribute rules are rarely dominance rules. A strictly linear interpolation is a form of a contributory rule, but most interpolation procedures apply distance weighting or more complex directional interaction rules.

The use of strictly geometric criteria for contour sources can produce stair-step surfaces in areas of low relief (Gousie and Franklin 1998). Mark (1986) described geomorphically specific rules as a solution, though it requires knowing the relevant landform processes that apply. Schneider (1998) develops a more complete set of criteria for 'geomorphologically sound' rules. When producing triangles, the attribute rule is given by the assumption of a flat triangle, so these situations fall into case 1N. The choice of points from contours can be rather complex (Christensen 1987).

Table 6. Surface-oriented transformation.

In/out	Points with z	Isoline	DEM	TIN
Points with z	2: N-found A-imported	2: N-found A-imported	2: N-found A-imported	1N: N-found A-flat Δ
Isoline	2: N-found A-imported	2: N-found A-imported	2: N-found A-imported	1N: N-found A-flat Δ
DEM	1A: N-implicit A-imported	1A: N-implicit A-imported	1A: N-implicit A-imported	0: N-implicit A- flat Δ
TIN	0: N-pt in Δ A-flat in Δ	1A: tracing rule (linear?)	0: N-pt in Δ A-flat in Δ	2: Simplify Refine

The DEM source makes it much more direct to extract the immediate neighbours of a given point; the neighbourhood rule for a DEM is inherent to the structure. These operations are usually lumped together with other interpolation procedures because they share similar attributes rules, but they are Case 1A transformations. Interpreted literally, the transformation from a DEM to a TIN is a mechanical extraction (Case 0), however, this usually produces overly dense triangulation, requiring some form of filtering (Heller 1990, Falcidieno and Spagnuolo 1991, Little and Shi 1998). Also, generating isolines from a TIN falls into 1A because most users are unwilling to accept the straight-line contours inherent in a TIN model. Some form of smoothing rule is usually imported to make the contours conform to expectations. The only extractions (Case 0) involve triangulations in some way. It is quite direct to estimate the elevation for any point in a TIN, as long as each triangle is considered to be flat. This operation can simply be iterated at regular intervals to generate a DEM. The DEM to TIN operation, as explained above, generates overly dense triangles. Most algorithms to generate triangles from contours have a similar flaw, since they use the contour line as one side of each triangle. More complex processing (of a Case 2 variety) is required to produce more reasonable triangulations (Bello-García *et al.* 1992).

Ironically, perhaps the most complex transformation involves producing a TIN from a TIN. There is a long line of articles about hierarchical triangulation (Peucker and Chrisman 1975, Mark 1979, Heller 1990), but the process definitely falls into Case 2, requiring substantial geometric and attribute manipulation (Junger and Snoeyink 1998). The process of refining a TIN—adding resolution—is not possible without some external source of measurements. Otherwise the new triangles will be co-planar with some neighbours.

6. Beyond surfaces

The matrix of surface operations samples a tiny portion of GIS operations, drawn to represent the data structure conversions that Clarke and others already recognized as transformations. Yet, the generic nature of the explanation applies to a much broader range of functions. The ‘analysis functions’ should not be treated as primitive steps in an algebra where order makes no difference, but as transformations that alter the structure at each turn. Tobler’s search for invariances and inverse transformations diverted attention away from the asymmetric relationships where information is created or destroyed.

Of course, some operations are indeed extractions (Case 0). However, the result of an extraction is usually weakened so that it can not be plugged back into the more complex structure. Extractions are not the place to look for the special value of GIS in integrating sources from different frameworks.

Case 2 complete transformations fill this role. A buffer around a road is a good example of a complete transformation. It uses a simple neighbourhood rule (all space within a certain distance of the road), and a simple dominance rule (areas near any road overrule anything else). The buffer edge should be seen as a kind of isoline, since it records a fixed distance from a particular feature. There is no need for the ‘special’ category of ‘geometric operations’ included in most of the earlier taxonomies; buffers are just a very simple pair of neighbourhood and attribute rules. Often the most critical operation in producing the buffer is choosing the categories from which to search. This step is usually done separately with an attribute tool from Class 1A (termed reclassification or grouping).

A polygon overlay, in its full application context, also involves a complete transformation. Vector-based processors labelled 'polygon overlay' produce the geometric raw material (a preliminary 1N), but that is not the end of the story. The real analysis requires some attribute combination from the joined tables. Though this might be separated in time, it forms a part of a complex transformation. Before it is complete, the intermediate stages can be described as incomplete transformations. In the raster situation, the geometric stage (1N) has been pre-computed in placing all the data into a common grid structure. This is why the raster overlay operators seem to apply just to the attributes (1A). The division between 'logical operations', 'arithmetic operations' and so on simply classifies the kinds of attribute rules. Whether in vector or raster representation, many polygon overlay operations produce meaningless results. The fault does not lie in the tool itself, but in the incompatibility of the measurement frameworks used as input. While this paper does not demonstrate how each operation should be applied, it does provide a guide to how to characterize the meaning of operations.

The real lesson from a transformational approach reorients thinking from the particular steps that a software package might use. The underlying transformations might be cut up in various stages, with intermediary results obscuring the database. If you look through to the final products, the increase in information value is most likely to come from complete transformations.

7. Conclusion

A unifying scheme for transformations requires only two elements: a geometric neighbourhood 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, neighbourhood operations, plus the items now treated as data structure conversions) can all be relocated as various kinds of transformations.

The list of measurement frameworks is open-ended. The relationship-controlled frameworks demonstrate that the old raster-vector poles do not bisect all possibilities. There is substantial diversity within each major grouping, causing a need for care in choosing transformation algorithms. How does a GIS integrate from different sources? It uses geometric neighbourhoods and attribute combination to adjust for differences in measurement frameworks. A GIS tries to make the incommensurable compatible enough to be merged.

Integrating information from different sources is the strength of GIS, as well as its major weakness. This paper provides some guidance to the process of incomplete translations performed by GIS analysis. Transformations between different measurement frameworks are not just preparatory steps, but the very heart of GIS analysis. Each measurement framework makes choices of which elements to emphasize, and which to simplify. These choices set limits to what can be extracted and what information must be rebuilt or provided externally.

Acknowledgments

Parts of this paper first appeared in *Exploring Geographic Information Systems*, ©1997 John Wiley and Sons and are used here with permission. Francis Harvey has contributed greatly to successive drafts of this work. The second and third draft of this paper benefitted from the commentaries of the anonymous reviewers. Revision was supported in part by National Science Foundation Grant 98-10075.

References

Albrecht, J., 1995, Semantic net of universal elementary GIS functions. In *Proceedings, AUTO-CARTO 12* (Bethesda: American Congress on Surveying and Mapping), pp. 235–244.

Bello-Garcìa, A., Gonzàlez-Nicieza, C., Ordieres-Merè, J. B., and Menèndez-Dìaz, A., 1992, A Contour Line Based Triangulating Algorithm. In *Proceedings of the 5th International Symposium on Spatial Data Handling* (Columbus: International Geographical Union), **2**, pp. 411–423.

Bernstein, L., 1976, *The Unanswered Question: Six Talks at Harvard* (Cambridge MA: Harvard University Press).

Berry, J. K., 1987, Fundamental operations in computer-assisted map analysis. *International Journal of Geographical Information Systems*, **1**, 119–136.

Burrough, P. A., 1986, Principles of Geographical Information Systems for Land Resource Assessment (Oxford: Clarendon Press).

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

Burrough, P. A., 1996, Natural objects with indeterminate boundaries, In *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank (London: Taylor & Francis), pp. 3–28.

Burrough, P. A., and McDonnell, R. A., 1998, *Principles of Geographical Information Systems* (Oxford: Oxford University Press), second edition.

Chomsky, N., 1968, *Language and Mind* (New York: Harcourt, Brace and World).

Chrisman, N. R., 1995, Beyond Stevens: a revised approach to measurement for geographic information. In *Proceedings AUTO-CARTO 12* (Bethseda: American Congress on Surveying and Mapping), pp. 271–280.

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

Chrisman, N. R., 1998, Rethinking levels of measurement for cartography. *Cartography and Geographic Information Systems*, **25**, 231–242.

Christensen, A. H. J., 1987, Fitting a triangulation to contour lines. In *Proceedings AUTO-CARTO 8* (Bethseda: American Congress on Surveying and Mapping), pp. 57–67.

Clarke, K. C., 1990, *Analytical and Computer Cartography* (Englewood Cliffs NJ: Prentice Hall).

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

Clarke, K. C., 1997, *Getting Started with Geographic Information Systems* (Upper Saddle River, NJ: Prentice Hall).

Codd, E. F., 1981, Data models in database management. *SIGMOD Record*, **11**, 112–114.

Cowen, D. J., 1988, GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing*, **54**, 1551–1555.

Dangermond, J., 1983, A classification of software components commonly used in geographic information systems, In *Design and Implementation of Computer-Based Geographic Information Systems*, edited by D. J. Peuquet and J. O'Callaghan (Amherst NY: IGU Commission on Geographical Data Sensing and Processing), pp. 70–91.

Falcidieno, B., and Spagnuolo, M., 1991, A new method for the characterization of topographic surfaces. *International Journal of Geographical Information Systems*, **5**, 397–412.

Frank, A. U., Volta, G. S., and McGranahan, M., 1997, Formalization of families of categorical coverages. *International Journal of Geographical Information Science*, **11**, 215–231.

Fujimura, J. H., 1992, Crafting science: standardized packages, boundary objects, and 'translation', In *Science as Practice and Culture*, edited by A. Pickering (Chicago: University of Chicago Press), pp. 168–211.

Galison, P., 1997, *Image and Logic* (Chicago: Chicago University Press).

Giordano, A., Veregin, H., Borak, E., and Lanter, D., 1994, A conceptual model of GIS-based Spatial Analysis. *Cartographica*, **31**, 44–51.

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

Gousie, M. B., and Franklin, W. R., 1998, Converting elevation contours to a grid. In *Proceedings Spatial Data Handling 98* (Vancouver: Simon Fraser University), pp. 647–656.

Guptill, S. C., 1988, A process for evaluating geographic information systems. *US Geological Survey Open File Report 88-105* (Reston VA: Technology Exchange Working Group, Federal Interagency Co-ordinating Committee on Digital Cartography), pp. 1–57.

Hadzilacos, T., 1996, On layer-based systems for undetermined boundaries, In *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank (London: Taylor & Francis), pp. 237–255.

Harvey, F., 1997a, From geographic wholism to geographic information system. *Professional Geographer*, **49**, 77–85.

Harvey, F., 1997b, Agreeing to disagree: the social construction of geographic information technology. In *Proceedings GIS/LIS '97* (Bethseda: American Society for Photogrammetry and Remote Sensing/American Congress on Surveying and Mapping), **1**, pp. 808–815.

Harvey, F., and Chrisman, N. R., 1998, Boundary objects and the social construction of GIS technology. *Environment and Planning A*, **30**, 1683–1694.

Heller, M., 1990, Triangulation algorithms for adaptive terrain modeling. In *Proceedings of the 4th International Symposium on Spatial Data Handling* (Zurich: International Geographical Union), **1**, pp. 163–174.

Heuvelink, G. B. M., and Pebesma, E. J., 1998, Spatial aggregation and soil process modelling. *Geoderma*, **89**, 47–65.

Hopkins, L. D., 1977, Methods of generating land suitability maps: A Comparative Evaluation. *American Institute of Planners Journal*, **43**, 386–400.

Junger, B., and Snoeyink, J., 1998, Importance measures for TIN simplification by parallel decimation. In *Proceedings Spatial Data Handling 98* (Vancouver: Simon Fraser University), pp. 637–646.

Langran, G. E., 1991, *Time in Geographic Information Systems* (London: Taylor & Francis).

Latour, B., 1988, *The Pasteurization of France* (Cambridge MA: Harvard University Press).

Little, J. J., and Shi, P., 1998, Structural lines, TINs, and DEMs. In *Proceedings Spatial Data Handling 98* (Vancouver: Simon Fraser University), pp. 627–636.

Maguire, D. J., 1991, An overview and definition of GIS, In *Geographical Information Systems: Overview, Principles and Applications*, **vol 1**, edited by D. J. Maguire, M. F. Goodchild and D. W. Rhind (Harlow: Longmans), pp. 9–20.

Maguire, D. J., and Dangermond, J., 1991, The functionality of GIS, In *Geographical Information Systems: Principles and Applications*, **1**, edited by D. J. Maguire, M. F. Goodchild and D. Rhind (Harlow: Longmans), pp. 319–335.

Males, R. M., and Gates, W. E., 1978, ADAPT—A spatial data structure for use with planning and design models, In *Harvard Papers on Geographic Information Systems*, **3**, edited by G. Dutton (Reading MA: Addison Wesley).

Maling, D. H., 1973, Coordinate Systems and Map Projections (London: Philip).

Mark, D. M., 1979, Phenomenon-based data structuring and digital terrain modeling. *GeoProcessing*, **1**, 27-36.

Mark, D. M., 1986, Knowledge-based approaches for contour-to-grid interpolation on desert pediments and similar surfaces of low relief. In *Proceedings of the Second International Symposium on Spatial Data Handling* (Columbus: International Geographical Union), pp. 225–234.

Peucker, T. K., and Chrisman, N. R., 1975, Cartographic data structures. *The American Cartographer*, **2**, 55–69.

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

Schneider, B., 1998, Geomorphologically sound reconstruction of digital terrain surfaces from contours. In *Proceedings Spatial Data Handling 98* (Vancouver: Simon Fraser University), pp. 657–667.

Shepard, D., 1968, A two-dimensional interpolation function for irregularly spaced data. In *Proceedings, Twenty-third National Conference, Association for Computing Machinery* (New York: ACM Press), pp. 517–524.

Sinton, D. F., 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*, **6**, edited by G. Dutton (Reading MA: Addison Wesley), pp. 1–17.

Smith, B., and Mark, D. M., 1998, Ontology and geographic kinds. In *Spatial Data Handling 98* (Vancouver: Simon Fraser University), pp. 308–320.

Snyder, J. P., 1987, Map projections—a working manual. *Professional Paper 1395* (Reston: US Geological Survey).

Star, S. L., and Greisemer, J. R., 1989, Institutional ecology, 'translations', and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology. *Social Studies of Science*, **19**, 387–420.

Steinitz, C., Parker, P., and Jordan, L., 1976, Hand-drawn overlays: their history and prospective uses. *Landscape Architecture*, **66**, 444–455.

Stevens, S. S., 1946, On the theory of scales of measurement. *Science*, **103**, 677–680.

Sullivan, J. G., Chrisman, N. R., and Niemann, B. J., Jr., 1985, Wastelands versus Wetlands in Westport Township, Wisconsin. In *Proceedings URISA* (Washington: Urban and Regional Information System Association), **1**, pp. 73–85.

Tobler, W. R., 1959, Automation and cartography. *Geographical Review*, **49**, 526–534.

Tobler, W., 1969, Geographical filters and their inverses. *Geographical Analysis*, **1**, 234–253.

Tobler, W., 1979, 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* (Englewood Cliffs NJ: Prentice Hall).

Unwin, D., 1981, *Introductory Spatial Analysis* (London: Methuen).