TRIBULATIONS OF AUTOMATED CARTOGRAPHY AND HOW MATHEMATICS HELPS

MARVIN WHITE, *Statistical Research Division, Bureau of the Census, Washington, D.C.*

INTRODUCTION

Errors in digital maps seem to have survival instinct. Costs routinely exceed estimates and budgets. Sharing data among computer-assisted cartographers is so problematic it is usually avoided. In short, difficulties abound! The application of mathematics, particularly comprehensive mathematical theories, alleviates these difficulties.

The plan for this paper is to consider in turn several common situations in computer-assisted cartography where difficulties are often encountered, and each time discuss how a mathematical approach overcomes or avoids the problem. Such problems include complicated calculations, expensive procedures, and plain inadequacies. Following the examples, a syllabus for the most important mathematical topics is presented. In each case some particular part of mathematics already provides a solution but that knowledge of the subject is not sufficiently widespread. Solutions are too lengthy to be given in this paper but can be found in the references.

DIFFICULTIES IN COMPUTER-ASSISTED CARTOGRAPHY

A recurring mathematical theme in the following sections might better be regarded as philosophical. It is fundamental and must not be dismissed as academic. It is the *ontological* question: What exists? Confined to a map the question becomes 'what are the elementary objects in this map?'

Failure to answer this question is the source of vexing difficulties. As simple and easy as answering this question may seem, inability to answer it is quite common. The problem is not a matter of resolving details or avoiding minor errors, but rather is related to defining gross categories of objects in the context of a theory of maps. Enter mathematics. A mathematical theory of maps provides those categories and much more, as we shall see in the following examples.

A second recurring concept in the sequel is *neighborhood.* The common usage of 'neighborhood' corresponds well to its mathematical usage. A neighborhood can be small or large but neighboring people and neighboring points are always near to each other in some sense. The mathematical concept is more definite and more abstract but keeping the ordinary meaning in mind won't lead you too far astray. Neighborhoods are fundamental in *topology*, the branch of mathematics that deals with geometrical properties that remain invariant under deformation. Topology is fundamental in the mathematical theory of maps.

Boundary Traces

Costs of maintenance. A very common method for storing geographic region location is to trace the boundary of each region. The regions may be political jurisdictions, areas of similar land cover, and so on. Usually the most difficult problem is maintaining and correcting the data. Large programs have been abandoned due to the high cost and error proneness of updating boundary traces.

The difficulty with updates and corrections is caused by a very fundamental deficiency in the mathematical approach taken. The deficiency illustrated in Figure 1a is that elementary geometrical objects, the points and lines, have not been identified (as they are in Figure 1b). The ontological question lies unanswered. In fact, there is only one boundary line separating A and B, and only one point at which A, B, and C have a common boundary. However, the independent boundary traces weave around each other and no point is common to the boundary traces for A, B, and C, which causes difficulty in updating and gives rise to slivers and gaps (discussed below).

Independent traces no matter how carefully done will fail to coincide, so that making a change in the boundary of A will always lead to difficulties in making the corresponding change in an adjacent region, say B. The remedy is to actually record single lines and points once and refer to the single record many times. This means digitizing the common boundary separating A and B once and somehow recording the connection with regions A and B and the connection with the B-C and A-C boundaries at their single point of intersection. That is, answer the ontological question first then record data only about objects that actually exist

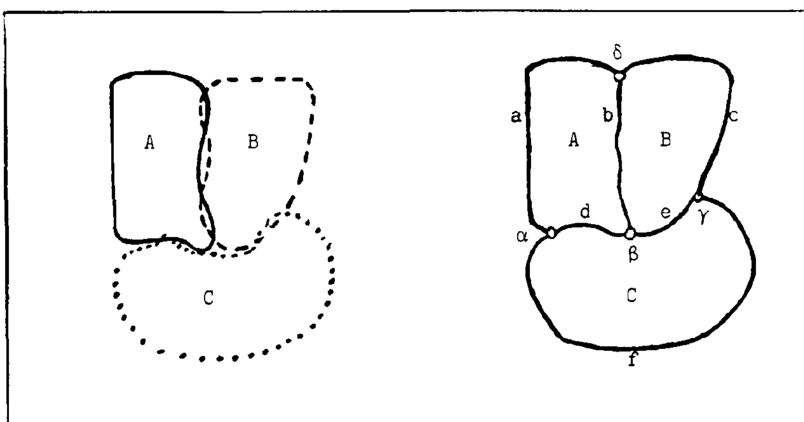

FIGURE 1. a *Boundary traces for regions A, B and C.* b The same map with elementary objects identified.

and are distinct. Figure 1b illustrates the solution. Every point is identified and labelled with a lower case Greek letter, each line with a lower case roman letter, and each region with an upper case roman letter (as before).The labels are not necessary, but are shown to specify what elements have been identified.

Slivers and Gaps. A related difficulty is the occurrence of slivers and gaps in plots of boundary traces. These are the small spaces in Figure 1a that are eliminated in 1b. To avoid these slivers and gaps in plots, some people merely use a thicker pen, which of course only masks the underlying error. This problem is another aspect of the same deficiency that causes updating to be so expensive, failure to answer the ontological question (What are the ingredients of this map?).

Many people fall into the trap of thinking that precise measurement of coordinates will identify points exactly. But measurements are always subject to error and consequently cannot provide unique identification. What precise measurements can do is make the problem appear less severe, but this is an illusion. It is simpler and better not to fall into the trap at all but to identify the elements and subsequently take one measurement. Coordinates and their proper role are discussed in more detail below.

Islands. A different but still common problem with boundary traces arises because not all regions are bounded by polygons. For example, some have interior islands and lakes, and some features are simply linear, like rivers. The problem is representing such features.

Papers have been published on schemes for introducing imaginary lines to connect components of a boundary and schemes for retracing linear features so that they appear to be bounding empty regions. Figure 2a illustrates such schemes. They become quite complicated as more islands, lakes, and isolated tree structures are added.

Here knowledge of topology, particularly homology theory, would avoid the problem altogether. In a sense, the whole problem is invented; it does not naturally arise. Lakes and islands are simply recognized as holes in the surface,

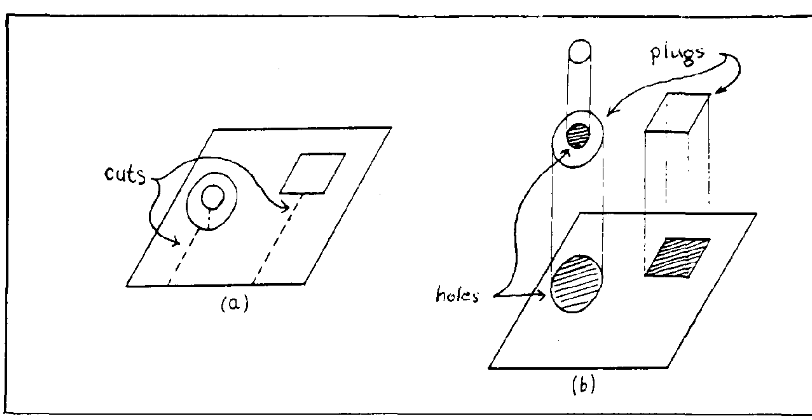

FIGURE 2. a *Artificial cuts are an unnecessary complication.* b *The solution comes from homology theory.*

and topology teaches us what facts are important to record and which structures are essentially different and which are abstractly the same (this theory alters our answer to the ontological question – holes and plugs to fill them now exist as elements of the map). No imaginary lines are ever needed. Figure 2b shows the topological representation, holes filled with plugs that may themselves have holes.

Overlays and Separates

The cartographic profession has taken the production and registration of separations to a fine art. Many automated systems have attempted to mimic the fine art but suffer deficiencies quite similar to those inherent in polygon tracing. The cost of updates is very high, and slivers and gaps appear and resist attempts to eliminate them.

The cause is the same as before, failure, indeed inability within the scheme, to answer the ontological question correctly. The solution is also the same, take the trouble to answer the question: identify the elements that the map comprises. Overlays present an example of unwittingly answering the ontological questions incorrectly. Map sheets are fundamental objects in an overlay system but they are not in any way related to the message conveyed by the map; overlays and sheets are strictly a matter of convenience in handling paper maps and have no place in an automated system.

The use of overlays in an automated system illustrates that simply mimicking the motions of a manual process does not always succeed. Human geometrical intuition is at work in the manual process but is absent in computers. The remedy, regarded from another viewpoint, is to supply to the computer what humans grasp intuitively in the form of mathematical theories.

For overlays this solution may require more effort than for polygon boundary traces, because the source maps may be separations. In fact, it is sometimes better to encode each overlay separately and use the computer-assisted system to help identify the elementary features, but it is not better to maintain the separate overlays, because the above-mentioned problems ensue. Inconsistencies between the features on one overlay and those on others begin to appear and then accumulate. The weight of accumulated inconsistencies has been the cause for abandoning expensive computer-assisted mapping systems.

Polygon overlay. The so-called 'polygon overlay' calculations also present difficulties. Gaps and slivers are generated by nearly coincident lines. This problem was discussed at length above. Another difficulty is just the amount of computing that seems to be required. The problem is to limit the search for coincident or intersecting features to a small number. Here topology again offers a solution. This time the most fundamental topological concept, ***neighborhood,*** is the key.

In polygon overlay calculations, the search for intersecting features can be done neighborhood by neighborhood or by considering, in turn, each feature from an overlay as an update, and limit the search to a neighborhood surrounding the feature. Of course you must know in a mathematical (topological) sense just what neighborhoods are, and how to record information about neighbors. All this is thoroughly explicated in combinatorial topology.

2-D		3-D	
Primal	Dual	Primal	Dual
0-Cell	2-Cell	0-Cell	3-Cell
1-Cell	1-Cell	1-Cell	2-Cell
Boundary	Co-Boundary	Boundary	Co-Boundary

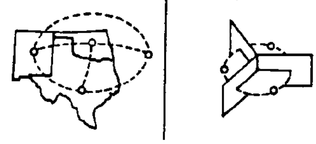

FIGURE 3. *Poincaré duality.*

Error detection. Discovering errors in maps of any kind is no simple task. In computerized maps the problem is worse than with ordinary maps and is most difficult in overlays. To be sure, error detection is difficult for boundary traces too, but the source of problems is the same as for overlays. The common practice is to examine plots (on paper or CRT) for errors, make corrections (actually changes that one hopes are corrections) and continue until no visible errors remain, no time remains, or no money remains, whichever is first. This method is especially ill-suited to discovering geographic naming errors, that is, errors in delineating various boundaries and labelling with geographic names and codes, such as city names, census tract numbers, or ZIP codes.

Neighborhoods are again the key to using the computer to discover and localize errors. Naturally, the ability to use neighborhood calculations has as a prerequisite knowledge of topology and its applications to system design.

Dual Incidence Matrix Encoding (DIME)

DIME is popular for encoding street maps but is equally applicable to plat maps, county maps, or any map. DIME was devised on the basis of topology and graph theory, and was purposely designed to aid in detecting and controlling errors. It avoids many of the problems discussed above, like gaps and slivers.

Orientation and duality. People do encounter difficulties using DIME, the most perplexing relating to orientation. Surfaces have a property called orientability, which determines whether one can consistently designate 'left' and 'right' sides to any closed path on the surface. Those who are not familiar with elementary topology are surprised to learn that there are non-orientable surfaces like the Moebius strip, which have only one side. Clerical coding is notorious for swapping left and right and some computer programs have systematically produced orientation errors. The only way to avoid these problems is to understand orientability.

The fact that DIME takes advantage of a geometrical symmetry called duality, as its name implies, leads to a difficulty for those not familiar with duality. In DIME a piece of linear feature (boundary, river, road, etc.) is recorded as a single record also containing information about the immediate neighborhood of the feature in a symmetrical fashion. Figure 3 explains this symmetry in topological

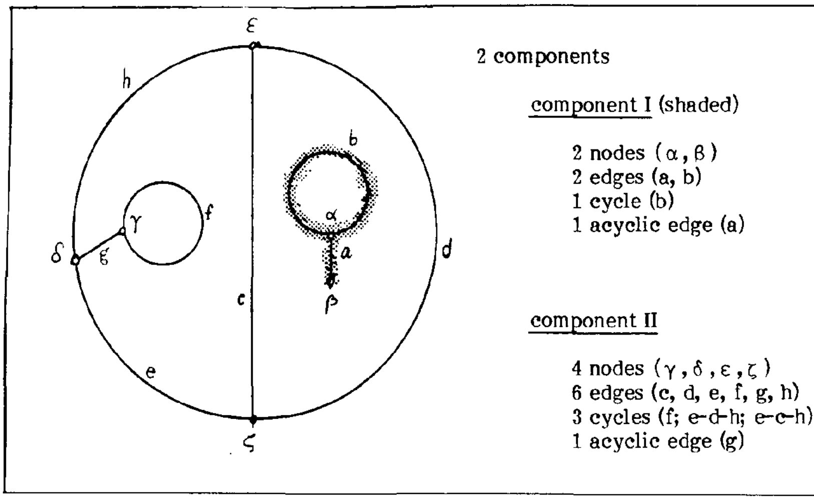

FIGURE 4. *Kirchoff's analysis.*

terms. Duality depends on the dimensionality of the space; duality in 2-D differs from duality in 3-D but the general idea is the same. The solid line and dashed line diagrams are duals. In 2-D 0-cells (points) are dual with 2-cells (areas): each area in the primal is replaced by a single point in the dual and vice versa. Because of this symmetry, a number of procedures apply equally to 0-cells and 2-cells even though they seem quite different. This is a source of confusion for the uninitiated but a source of efficiency for the cognoscenti.

Chaining. The users of DIME files thus have some unfamiliar mathematics forced upon them, and that alone is a difficulty. It is even more difficult to use the plethora of information captured in a DIME file. To make a choropleth map many people make a copy of the file but with orientation reversed, sort the original and copy together (calling the aggregate a NICKLE file) and chain DIME segments around each region to produce a boundary trace file. Chaining presents the most difficulties, especially where there are islands and lakes. A thorough understanding of the Jordan curve theorem (see any topology text) would permit one to dispense with the entire NICKLE-sort-chain process. In case one wants to use a plotting package that demands a boundary trace for input and so is obliged to generate boundary traces, knowledge of Kirchoff's theorem (see Biggs, Lloyd and Wilson) greatly simplifies the chaining algorithm. Figure 4 shows the results of the Kirchoff algorithm. The purpose of the Kirchoff analysis is to isolate distinct parts (*components*) and within components to classify each element of the graph. Kirchoff invented the method for analyzing electrical networks, but we use it for reducing complex cartographic networks to simple sub-elements.

Any uses of a DIME file are better understood in the context of the algebra of combinatorial topology. The uses include plotting maps, editing, service area definition, transportation modeling, and geocoding. The algebra is based on the *boundary* and *coboundary* operators and retrieves neighborhoods of any kind from secondary storage. It is a foundation for every application. A powerful system called 2-D based on this algebra has been implemented on a microcomputer and is in use producing a computer-aided emergency vehicle dispatch database. Using this algebra, 2-D returns police and fire jurisdiction, and nearest major intersection for an incident location specified by address or intersection. This is an example that obviously demands neighborhood searches, but others are often not so obvious.

Triangulation

Partitioning the mapped region into triangles is the standard starting point in topology. It is used in cartography for Digital Terrain Modeling (DTM) and contour mapping. In these applications the elevation at each vertex is recorded and the triangulation is effectively a mesh cast over the surface. The main application in DTM is cut-and-fill calculations for earth-moving projects. Algorithms for interpolating elevations and computing volumes are quite efficient using a triangulation. There are no difficulties at this level but there is of course considerable supporting mathematics.

Conservative extension. Difficulties do arise at another level. If one tries to incorporate the elevation data of the triangulation into another cartographic file or tries to annotate natural features on the triangulation, it becomes evident that not everything in nature is a triangle.

Although a triangulation is the starting point for topology where it is called a simplicial complex, the theory quickly advances to the more general cellular complex, in which the cells are not merely triangles (simplexes) but may be any shape. Likewise in automated cartography triangulations can be generalized into cellular complexes. The advantages are that the cells can be made to conform to natural features and that the theory is conservatively extendible.

The first advantage is easy to understand. The cells can be constructed by adjoining neighboring triangles thus producing a more natural cell and retaining the elevation data of the triangulation. If it is desired, the triangulation itself can be retained as interior structure in each cell describing the cells shape in 3-dimensional space.

The second advantage of cellular complexes is due to the fact that new cells can be built by joining adjacent cells. In contrast, joining adjacent triangles does not return a triangle (Figure 5). The theory and programs for a triangulation are not conservatively extendible to aggregates but for a cellular complex they are. As an example, consider a detailed city map with ZIP code boundaries shown. If the details were all represented in triangles and the programs were made to deal only with triangulations, abstracting a ZIP map would require entirely new programs. For a cellular complex no new programs are needed, because ZIP areas are also 2-cells. This abstract-seeming advantage is a very practical matter in multipurpose systems and in developing new applications on a existing system.

Raster Images, Grids, Hexcells, and Quadtrees

A completely different scheme of storage is to store a picture image or some coarsened version of a picture. A television image is a raster image composed of

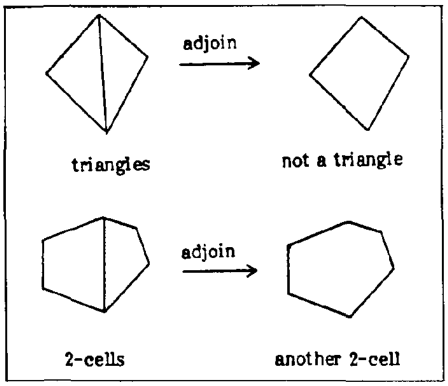

FIGURE 5.

rows of dots, each dot (also called a pixel) having a single hue and brightness. Grid cells are usually coarser than pixels in a raster image but serve the same function. Hexcells, like grid cells tile the plane but in a hexagonal network rather than a square network. Copious storage is needed for all of these methods, and thus techniques to reduce storage occupied are important.

Quadtrees are a particularly efficient method of saving storage for a gridded image.

Picture processing has become a well-known field because of its application to satellite imagery. Contrast enhancement, multi-band spectral analysis, and image editing have all been applied quite usefully in automated cartography. These techniques each demand certain mathematical skills, which confer substantial advantage. The result is a very versatile camera system with some powerful analytic capabilities.

Vectorizing. But treating maps pictorially falls far short of the potential for computer-assisted cartography. For example, analyzing data by political jurisdiction cannot be done on a picture-map; the map content must be understood for the analysis. This is where difficulties arise.

The ontological question in this case is even harder to answer than for overlays and boundary traces, because photo-interpretation is required. Automating this process has never succeeded completely. Some systems 'vectorize' the image, but there always remains a large editing job to identify points that should be the same but were interpreted as separate and vice versa. Solving this problem, if it is ever solved, will require considerable mathematical expertise.

Coordinates

To many people, using coordinates and automating cartography are nearly synonymous, and to these people automated cartography is impossible without coordinates. Furthermore, higher precision coordinates are regarded as a sine qua non of higher quality automated cartography.

This is a fallacy and a pernicious one. The damage is twofold. First, system builders try to use coordinates to identify points thinking that they are somehow cleverly getting both coordinates and identifiers for the price of one. Second, exceedingly expensive programs are initiated to gather coordinates as a preliminary to building and using the entire database, preventing useful programs from being started and forcing some to be halted before ever reaching the useful stage.

Identifiers. Coordinates, being calculated from measurements, are subject to measurement error. This makes using coordinates as identifiers somewhat fuzzy, and in unfortunate cases even high order digits are susceptible to changes. Worse, tracing the history of changes becomes impossible, since points cannot be moved; rather, they vanish entirely and a new point materializes at another location. Trying to get identifiers to do two jobs, their own and carrying data, is a mistake, because ultimately the jobs conflict and there is no space saving anyway. Identifiers must be constant and absolutely distinguishable, but data, for example coordinates, must be unconstrained. Data values are variable but identifiers must be fixed and unique. This is the ontological question revisited.

Expensive prerequisites. The second problem, undue expense at the outset, is especially unfortunate because it is so easy to avoid. One need only know enough about topology to understand that a space with coordinates (which must be a metric space) is in fact a topological space, to avoid the initial coordinate gathering expenses. Crude measurements along with an accurate record of the topological facts, which must be gathered in any case, will suffice. Later, when the system has proven its worth, it can be enhanced with more precise measurements. For a specific example consider a plat map in a computer-assisted cartography system. The important facts shown by a plat map are: who adjoins whom; which adjoiners share a common boundary line, and which share only a corner – in short, topological data. Survey notes show detailed measurements but these are only partially copied over to a plat map. Similarly an automated cartography system can provide a digital plat map, and leave precise coordinates for later concentrating first on correct topology.

To be sure, it is important to digitize and record coordinates. It is also important to provide access to geographic data via coordinates. In addition, more precise coordinates permit more precise and better-looking maps. The important point is that the role of coordinates in cartography, although quite important, is not what is commonly believed. Geometric topology is the subject that clarifies the facts.

Systems of Programs

Automated cartography is realized through computer programs, of course. Systems of programs are formal objects and can be studied using the theory of formalized theories. Nowhere is the formality of computer systems more evident than in relational databases.

There have been attempts to represent maps in relational DBMSs. The relations explicitly stated in the schema in most of those attempts bear little resemblance to the actual geometric relations. One author went so far as to assert that there are so many relations that it is nearly impossible to state them.

Nothing could be further from the truth. The topological theory of maps has only two: boundary and coboundary. Add to these the metrical descriptions of

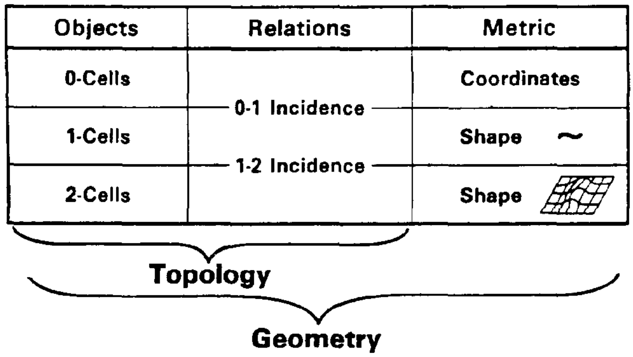

FIGURE 6. *The geometry of a map requires only three objects, two relations and metrical descriptions of the objects.*

points, lines, and regions, namely location, shape, and surface shape, respectively. These provide a complete representation of the geometry of a map. Five relations isn't so many. And the formalized theory states explicitly what they are. Figure 6 lists the elements and relations of geometry.

Failure of systems analysis. We hear in lecture after lecture and sermon after sermon that the first task in designing a system is to list all the uses of the system. This approach never succeeds in producing a sound system for a subject as demanding as cartography. The first task is to discover or invent a comprehensive mathematical theory that applies, and this requires a great deal more than listing uses. Only within a theory can the ontological question be answered and the list of uses stated in a coherent way.

SYLLABUS

The student of the mathematics of automated cartography will find many texts and monographs on the topics. A few are listed in the references below. Analytic geometry and in some cases differential geometry are already required for manual cartography. These are needed for using coordinates and computing projections. Picture processing, for example interpreting satellite imagery, requires knowledge of elementary statistics, information theory, and signal processing. These are special topics needed only in a few automated cartography systems. This syllabus includes only topics needed by every system developer.

Topology. Automation requires a deeper understanding of the nature of space than one gains in high school geometry courses. In particular, topology is the subject that reveals the elementary but not obvious nature of space and is essential in answering the ontological question for maps. The most important subjects are:

Homeomorphism (continuous deformation)
Complexes
  Cells (the elementary objects)
  Incidence (the elementary relations – boundary and coboundary)

Manifolds (smooth surfaces)
Orientability
Duality
Homology

Graph theory. The 1-dimensional skeleton of a 2-dimensional complex is a graph. Many analytical algorithms needed in automated cartography are graph theoretical. The important topics are:

Graphs
Paths and circuits (including Kirchoff's theorem)
Connectivity
Planarity (connection to topology)
Duality (a related connection to topology)

Model theory. Computer systems are formal. Model theory is part of mathematical logic and treats formalized theories and their interpretation. This subject is not well-known even among mathematicians – it is a specialty. Nevertheless, the elementary results of model theory are helpful in building computer systems. The topics are:

First order theories
Free and bound variables
Quantification
Models

CONCLUSION

We have considered only a few common and well-known cases where applied mathematics alleviates severe difficulties in automated cartography. Many more such examples exist. Quite often very elementary mathematics solves the problem; we saw several cases that only needed a proper answer to the ontological question. Finding a comprehensive theory that applies is the crucial and difficult step. In the context of such a theory the problems vanish or become quite tractable.

For cartography, geometry and especially topology are the theories that organize the subject and illuminate the structure of maps. Graph theory is also quite useful. I hope you are encouraged to learn about the topics in the syllabus. It will make you much more effective as a cartographer or programmer in our field.

BIBLIOGRAPHY

ALEXANDROFF, P. 1961. *Elementary concepts in topology.* Dover, New York.
This little book was written as a companion to Hilbert and Cohen-Vossen Geometry and the Imagination, and gives a clear concise statement of the important ideas in combinatorial topology-simplexes, cells, dimension, manifolds, homology, etc.

BARNES, D.W. and MACK, J.M. 1975. *An algebraic introduction to mathematical logic,* Springer-Verlag, New York.
Formalized theories and models are presented.

BIGGS N., LLOYD E. and WILSON R. 1976. *Graph theory 1736–1936*, Clarendon Press, Oxford.
Original papers in graph theory and related topology with interpretations and historical remarks by B, L. & W. Kirchoff's paper is included.

CHANG, C.C. and KEISLER, H.J. 1973. *Model theory*, North Holland, Amsterdam.

CORBETT, J.P. 1979. *Topological principles in cartography*, Bureau of the Census, Washington, D.C.
The application of topology of 2-dimensional space to cartography. This monograph addresses many of the examples directly.

DEO, N. 1974. *Graph theory with applications to engineering and computer science*, Prentice-Hall.
Provides many useful algorithms and insight into graph theory from an application viewpoint.

LEFTSCHETZ, S. 1979. *Applications of algebraic topology*, Springer-Verlag, New York.
Develops relevant topics in topology, graph theory and duality and applies them to circuit theory, surfaces and planarity. Leftschetz provides especially clear explanations of the important topics.

MOISE, E. 1977. *Geometric topology in 2 and 3 dimensions*, Springer-Verlag, New York.
A text intended for graduate students in mathematics. Moise takes care to alert the student to pitfalls and difficulties.

VEBLEN, O. 1931. *Analysis situs*, American Mathematical Society, New York.
An early treatise in topology. Veblen develops the incidence matrix method of topological computation, much like methods used for DIME.

WHITE, M.S. 1983. Technical requirements and standards for a multipurpose geographic data system. *The American Cartographer*.
An application of the topological theory of maps.