

The Use of Topology on Geologic Maps

By Ronald R. Wahl

U.S. Geological Survey
P.O. Box 25046, MS980, Denver Federal Center
Denver, CO 80225
Telephone: (303) 236-1320
Fax: (303) 236-5349
e-mail: rwahl@usgs.gov

ABSTRACT

Topology, as a term used in the context of a digital representation of spatial data in two dimensions, pertains to the relationships among feature geometry objects such as points, lines and polygons. Topology helps define as well the functional relationships among geologic map objects (feature classes or sub-themes, and themes or feature data sets), the feature geometries of which provide the framework for these representations. Geologic maps contain abundant data about the locations of different kinds of rock materials, as well as relational information about the contacts between contrasting units and the nature of these contacts (for example, depositional contacts, intrusive contacts, faults, etc.). Various tools have been developed to aid in creating geologic map databases that contain the complex information and relationships expressed in geologic maps, but not all tools have been equally successful.

ESRI's ArcGIS software represents one of the latest developments in this process, and it differs from earlier tools in that a topological framework is optional. Topology in ArcGIS is expressed in the form of rules that govern the allowable relationships among geometric entities and objects, which if carefully followed, lead to digital database structures that reduce the likelihood of errors and inaccuracies and that facilitate automated analysis. The main point of this paper is that the application of topological structure in ArcGIS is beneficial to converting geologic maps into digital databases that are both rigorous and efficient.

I recommend the following kinds of topology for making geologic maps with ArcGIS:

1. A boundary line must not cross over itself or lie on top of itself.
2. Polygons must not overlap one another, must not have gaps between one another, and map unit polygon edges must be covered by boundaries (contacts and faults).
3. Points that are measurements on a fault or other structure represented by a line must be covered by the line with which they are associated.

INTRODUCTION

Dave Viljoen (Geological Survey of Canada) in 1997 presented a paper about geologic map topology at the first Digital Mapping Techniques workshop. He said, "Topology is the [expression of] relationships that can exist among the geometries of spatial features...(and) is a major step towards minimizing the number of points, lines, and polygons to digitally capture and manage a geologic map." Data accuracy and validation are enhanced by the proper use of topological rules. These concepts are still valid.

In addition to the making of geologic maps, topology enhances analysis of the features on a geologic map in the following general cases:

- Polygons that outline a part of an image
- Point to polygon (is a point on the inside, outside, or on the boundary of the polygon?)
- Polygon adjacency
- Buffer distance from a line (or arcs of a type). An example of this is the following query from a geologic map: "Show as a polygon the region where every point on the constructed polygon boundary is ten miles from the San Andreas fault system."

HISTORY OF TOPOLOGY IN GEOLOGIC MAPS

Long before geographic information systems (GIS), geologic mappers and cartographers were less concerned with questions about topology than they are today. For example, mappers colored early geologic maps by hand mostly with watercolors. These maps had topologic problems. Lines bounding the areas on the maps representing rock outcrops did not necessarily close. Application of color with a brush meant that the color sometimes went outside these lines. The only topological question that was asked was "Are any adjacent polygons of the same color?"

In the 1950s, the basic geologic map-making process was the same, but by then, the materials used to make geologic maps had changed. Geologic maps were made using many scribe coats and peel coats each of which

contained information primarily separated out by color. Mappers and cartographers proceeded with great care to insure that all of the layers of mylar that were used were accurately registered with every other layer. Minor misregistrations meant that the old maps weren't accurate under close inspection. Contacts and faults features overshot or undershot other lines on the final geologic map, and this was particularly true for lines in other map themes like hydrography. Rock outcrop color fills went across contact and fault lines.

In the late 1970s and early 1980s, an Israeli company, SCITEX, brought equipment to the United States that originally was used to make patterns for wallpaper and cloth. I attended the 11th Annual Conference on Computer Graphics and Interactive Techniques, at Anaheim, CA in April 1984, where SCITEX showed equipment that was able to scan and then print maps. These scans could be digitized and subsequently edited to make map separates that regular printing presses could use. The USGS soon adopted this technology to make maps, including geologic maps. The scanned map was digitized within the SCITEX system. This process forced the map makers to ensure line closures to make closed areas (or polygons) for rock outcrops and to ensure that contacts and faults aligned precisely with rock outcrop boundaries.

Selner and others (1986) wrote a software package named GSMAP for the IBM PC environment to digitize and then print geologic maps and diagrams on Hewlett-Packard Graphics Language pen plotters. The USGS (author unknown) developed a translator to change GSMAP files to SCITEX format. Now, USGS geoscientists could digitize their own maps and deliver their digital maps to the SCITEX. However, the SCITEX installations were so complex they required many technicians to operate it and it was a "black box" for many geologists. In general, geologists still had no knowledge about the topology requirements for their maps.

In the mid-to-late 1980s, ARC/INFO from ESRI was slowly being introduced to USGS geologists. ESRI, convinced of the usefulness of the SCITEX system, wrote a series of commands to exchange files with the SCITEX. In the process, USGS geologists discovered intersecting lines with dangles (both over and under shoots), pseudo-nodes, and other new problems identified by the topology of an ARC/INFO coverage. Geologic maps improved from the cartographic standpoint because of the enforced topology. These improvements included:

- no over or undershoots when lines intersect, leading to
- no unclosed polygons so that polygon fills are exact, and
- line sharing is exact among themes.

While ARC/INFO was the old name for ESRI's

premier product, the name ArcInfo Workstation, included with Arc GIS, will be used below.

For ArcView 3.x, theme means one shape file set and the associated attributes. Theme as used in ArcGIS is a synonym for a feature data set. The basic idea for both ArcView and ArcGIS is that lines used for contacts or faults are not exactly the same lines that are stored for polygon edges. ArcInfo Workstation coverages can have more than one attribution (or theme) for a line or polygon so that the actual storage of lines is shared among the various map themes.

ArcGIS introduces a new view of topology. Successful implementations of this topology can lead to geologic map databases that are more internally consistent and better prepared for analysis and comparison.

TOPOLOGY

The word "topology" will be used in the following discussion to mean a topologic paradigm. For ArcGIS, a topology is: "...a set of integrity rules and software tools that defines the behavior of spatially integrated geographic features and feature classes" (ESRI, 2002a). While old ARC/INFO and the Geographic Resources and Support System (GRASS <http://grass.itc.it>) each had a topologic paradigm, ArcView, through version 3.3 did not.

Arc-Node Topology

The topology model used in the Spatial Data Transfer System (SDTS) is essentially the topology model used in ArcInfo Workstation. SDTS topology including the topological Vector Profile (TVP) is extensively described in USGS, 1991. In ArcInfo Workstation, the following simplified definitions are compared to those from SDTS TVP.

ArcInfo Workstation Vector or GT-Complete Chains:

- Have a line with a node at each end
- Have zero or more vertices
- Have direction—"From Node" and "To Node"
- Can form polygons. If polygons are constructed, each line is attributed with an entry that shows the polygons to the left and right of the line as the line is traversed in the direction of the "To Node" from the "From Node". A line with a sense of direction is properly called a vector.

ArcInfo Workstation Polygons or GT-Polygons:

- Are a closed figure made of one or more arcs or GT-complete chains
- Support islands (polygons completely internal to a larger polygon).

Label Points:

- ArcInfo Workstation can have label points in a polygon coverage, but otherwise it has no point topology.

features (OGC, 1999a) and has specified geometries for these features in OGC, 1999b. ESRI, in implementing ArcGIS, has built upon the geometries in OGC, 1999b.

I first present a comparison of the geometries of ArcInfo Workstation to ArcGIS, and between the OGC and ArcGIS geometries (Table 1).

Open GIS Consortium Geometry and Topology

The Open GIS Consortium (OGC), of which ESRI is one of many members, has agreed upon standards for the collection of digital data describing spatial objects called

Topology—OGC and ArcGIS

OGC geometries

Topology is implicit in the OGC geometry descrip-

Table 1. Comparison of geometries for OGC, ArcInfo Workstation, and ArcGIS.

ArcInfo Workstation, and ArcGIS:

ArcInfo Workstation	ArcGIS	Notes
Coverage feature classes	Geodatabase feature classes	Each geodatabase can contain one or more feature data sets. Each feature data class can contain one or more feature classes
Tic	Point*	*The spatial extent and coordinate system are defined as part of the dataset’s spatial reference. The tics could optionally be managed as a point feature class.
Bnd	*	
Arc	Polyline	**With topological relationships to other line and
Node	Point**	
polygon feature classes.		
Point	Point	***A label location with a topological association to its polygon.
Polygon	Polygon	
Polygon label	Point***	
Route System	Line with measures	****The best strategy is to use generic feature classes whenever possible to ensure support for open, multipurpose, standards-based data. However, the geodatabase is extensible and it is possible to add custom feature classes.
Region	Polygon	
Annotation	Annotation	
NA	Dimension	
NA	Custom features****	

From: ESRI, 2002a

OGC geometries and ArcGIS:

OGC Geometries (OGC, 1999b)	ArcGIS Geometries (ESRI, 2002a)
Point	Point
Line, LineString	Polyline
LinearRing	Polygon Boundary
Polygon (one or more linear rings)	Polygons with islands

tions but no basic topology model exists. Topologic relationships are determined from the Dimensionally Extended Nine-Intersection Model (DE9-IM) shown in Table 2.

The matrix in Table 2 is difficult to implement directly in a GIS, so the OGC has specified the following methods (or program functions) based on DE9-IM model for use in computer programming languages. They are:

- Equals
- Disjoint
- Intersects
- Touches
- Crosses
- Within
- Contains
- Overlaps
- Relate

ArcGIS—Topology in a Geodatabase

ESRI has embraced OGC geometries and has developed topological rules and software to implement them as a part of the geodatabase structure. In a geodatabase, spatial relationships among feature classes in a feature dataset are defined by topology, and use specific rules for each basic geometry type. A map maker can choose to create topology for feature classes in a feature dataset. The primary spatial relationships that one can model using topology are adjacency, coincidence, and connectivity.

There are three types of topology available in the geodatabase: geodatabase topology (over 20 topology rules), map topology, and geometric network topology. Map topology is a dynamic topology derived from the creation of polygons from one or more polyline feature classes in a feature data set. Network topology uses a feature data set with one or more feature data classes that

have intersecting lines, and this topology allow the tracing of a designated path through these lines. An example of network topology would be a bus route through a city road network. I will concentrate only on geodatabase topology in this paper. For a complete discussion of the topology rules under ArcGIS, see ESRI, 2002a and ESRI, 2002b, and ESRI’s on-line documentation.

Line Sharing

When geodatabase topology is appropriate, geodatabases create a place for the implemented topological rules in a separate feature class in a feature data set. When the chosen topology rules are in force, an individual feature in a feature class can be chosen separately even though line work is shared among a number of feature classes. This is quite different than the ArcInfo Workstation model. In ArcInfo Workstation, only one set of lines exist in a coverage no matter how many attributes each line may carry. Even though lines can apparently be broken into pieces when a route is built from a line coverage (using dynamic segmentation), lines are inviolate in such a case and extra attribution is added in additional tables to contain the route construct.

Polygons in a Geodatabase

Unlike ArcInfo Workstation, only closed linear rings form polygons under the OGC definitions (see above and OGC, 1999b). This means that if two polygons are adjacent, they do not share adjacent edges, which in turn means that the polygons can overlap (one polygon is drawn over an adjacent polygon) or have gaps between the two nearly adjacent edges. This gap looks like a polygon between two other polygons, but this “gap polygon” is not directly known as a polygon by the geodatabase or ArcGIS. An island (one polygon wholly contained in an-

Table 2. The OGC Dimensionally Extended Nine-Intersection Model.

	Interior	Boundary	Exterior
Interior	$dim(I(a) \cap I(b))$	$dim(I(a) \cap B(b))$	$dim(I(a) \cap E(b))$
Boundary	$dim(B(a) \cap I(b))$	$dim(B(a) \cap B(b))$	$dim(B(a) \cap E(b))$
Exterior	$dim(E(a) \cap I(b))$	$dim(E(a) \cap B(b))$	$dim(E(a) \cap E(b))$

From OGC (1999b) Table 2.1 —The DE-9IM

Where:

- a* is a geometry, and *b* is another geometry
- I* is the interior of a geometry
- B* is the boundary of a geometry
- E* is the exterior of a geometry
- dim* is the dimensionality of the resulting geometry
- \cap is the symbol for intersection.

other polygon requires, in theory, two linear rings for the containing polygon. The outer polygon tracks clockwise and the inner polygon tracks counter-clockwise. A third linear ring tracking clockwise is then necessary for the inner polygon. Then, if polygons have edges that need to carry attribution (as in geologic mapping), another feature class (rock outcrop boundaries for example) should cover the polygon edges exactly. Topological rules to accomplish these tasks are among those provided in ArcGIS, (ESRI, 2002a, and ESRI, 2003). These rules are summarized on a PDF poster (ESRI, 2002b).

OTHER GIS SYSTEMS WITH OGC TOPOLOGY

A number of other GIS systems and RDBMs are equipped to implement geometry and (perhaps) topology for vector features. I list some of these below.

GIS Packages

- Vivid Solutions—JUMP (Java Unified Mapping Platform) <http://www.vividsolutions.com/products.asp?catg=spaapp&code=jump>.
- Vivid Solutions—JTS library (Java Topology Suite) <http://www.vividsolutions.com/products.asp?catg=spaapp&code=jts>.
- PostGIS (Based on PostgreSQL and JTS) <http://postgis.refractor.net> and <http://www.postgresql.org>.
- GRASS5.7—Based on PostGIS.

RDBMS Packages with Geometry

The Relational Database Management Systems (RDBMS) below can store geometry of the OGC type, but do not always have a topology software suite for use with geometry in the RDBMS.

- Oracle Spatial can be modified with the Laser-Scan Radian topology software <http://www.oracle.com/index.html> and http://www.laser-scan.com/technologies/radius/radius_topology.
- Informix (DB2)—Data Blades can store geometry <http://www-306.ibm.com/software/data/informix>.
- MySQL—Stores geometry <http://www.mysql.com>.
- PostgreSQL—Stores geometry and is an Object-RDBMS.

CONCLUSIONS

Topology greatly aids the validation of spatial data, which leads to increased accuracy and rigor among feature classes. It also means easier geodatabase maintenance

tasks, and that the geodatabase is physically smaller on a disk. The speed and capability with which analyses can be made is enormously increased.

I want to write a word about splined curves and other interpolations to give smooth curves. Mathematically, splines revolve around the process of taking derivatives of a function. This is an inherently “noisy” process. By noisy, I mean that the result of adding together derivatives of a function is an inaccurate process, the higher the derivative order the less accurate the result. Further, if the same vertices are interpolated from opposite sides, the direction of curvature from these opposite sides will be in the opposite direction. The net result is that vectors made from vertices tracked from opposite directions along the same path will not lie on top of one another. As a consequence, spline-generated curves for adjacent polygon edges cannot be subject to topological rules.

Recommendations

I recommend that the geologic community study implementation of topology in ArcGIS for use with geologic mapping and recommend minimal standard topology rules. I further recommend that this study consider in addition: network topology, and relationship classes.

Recommended Topology Rules

I have found the following topology rules to be useful and suggest that these rules form a basis for standard usage of topology rules:

1. Polygons
 - a. Adjacent polygons must not partially overlap (or overlap) each other.
 - b. Adjacent polygons must not have gaps between them.
2. Arcs (Edges)
 - a. An line must not self-overlap or cross over itself—like a freeway off-ramp.
 - b. Polygon boundaries must be covered by contacts and faults.
3. Points

Structure measures associated with a contact or fault must be covered by the appropriate line, modified from ESRI, 2002b.

REFERENCES

Note: Some of the ESRI documents referenced below are best found on a computer with ArcGIS installed. This insures that the reader will have the version that coincides with the version to which the reader has access. The top of the installation is the directory (folder) named “arcgis”.

- ESRI, 2003, ArcGIS: Working with geodatabase topology: An ESRI White Paper, Environmental Systems Research Institute, Inc., Redlands, CA, 23p., accessed at <http://support.esri.com/index.cfm?fa=knowledgebase.whitepapers.viewPaper&PID=43&MetaID=524>.
- ESRI, 2002a, ArcGIS 8.3 geodatabase topology rules: Environmental Systems Research Institute, Inc., Redlands, CA, accessed at http://arcgis/arcexe83/Documentation/topology_rules_poster.pdf.
- ESRI, 2002b, What is ArcGIS?: What_is_ArcGIS.pdf, Document file with ArcGIS, Environmental Systems Research Institute, Inc., Redlands, CA, 80p., accessed at <http://arcgis/arcexe83/Documentation,>.
- Open GIS Consortium (OGC), 1999a, The OpenGIS Abstract Specification Topic 5: Features, Version 4: OpenGIS Project Document Number 99-105r2.doc, accessed at <http://www.opengis.org/docs/99-105r2.pdf>.
- Open GIS Consortium, 1999b, OpenGIS Simple Features Specification For SQL, Revision 1.1: OpenGIS Project Document 99-049, accessed at <http://www.opengis.org/docs/99-049.pdf>.
- Selner, G.I., Taylor, R.B., and Johnson, B.R., 1986, GSDRAW and GSMAP version 3.0; prototype programs for the IBM PC or compatible microcomputers to assist compilation and publication of geologic maps and illustrations: U.S. Geological Survey Open-File Report 86-447-A, 53 p.
- USGS, 1991, Logical Specifications in SDTS_standard_oct91, accessed at http://mcmweb.er.usgs.gov/sdts/SDTS_standard_oct91/part1.html#p23.
- Viljoen, David, 1997, Topological and thematic layering of geological map information: Improving efficiency of digital data capture and management, in: D.R. Soller, ed., Digital Mapping Techniques '97—Workshop Proceedings: U.S. Geological Survey Open-File Report 97-269, p. 15–22.