Topology and Topological Rules
Geometric properties that are maintained in spatial databases

The definition of topology

Topology is a term used around GIS that is sometimes confused with the term “topography”. The meanings of the two terms are quite unique. Topography refers to the study and depiction of physical features in the landscape. Topology used to describe the relationships between objects.

Merriam-Webster's Online Dictionary
Main Entry: topology
Pronunciation: t&-'pä-l&-jE, tä-
Function: noun
Inflected Form(s): plural -gies
Etymology: International Scientific Vocabulary
Date: 1850
1 : topographic study of a particular place; specifically : the history of a region as indicated by its topography
2 a (1) : a branch of mathematics concerned with those properties of geometric configurations (as point sets) which are unaltered by elastic deformations (as a stretching or a twisting) that are homeomorphisms (2) : the set of all open subsets of a topological space b : CONFIGURATION <topology of a molecule> <topology of a magnetic field>
- topologicist /-jist/ noun

As currently used in GIS, the term has been applied to several uses:

• Theory or model of features in space
• Mechanism allowing features to share geometry
• Set of editing tools for integrated features
• Set of validation rules
• Mechanism for navigating between features using their topological relationships

Questions for Consideration

A good way to illustrate the behaviors and utility of topology is by examining its application in answering spatial questions:

• Given a model of network space, how do we develop a data structure to support way finding?
• How is topology used during the development of spatial databases?
In order to address these questions, we must reiterate the context of the question. Network space is nearly always modeled using the object model of spatial information. Data models based on this model of spatial information involve the coordinatization of space. In turn these data models are implemented as specific vector data structures. The result is that points, lines, and polygons are stored in a variety of formats.

A specific type of vector data model that was utilized by ESRI to represent linear features and polygons is the arc-node model. The model manifests itself in GIS data as a specialized data structure that supports analysis functions. The fundamental elements of the structure are the arc or line and its beginning and ending points called nodes. In this structure nodes do not exist as points independent of arcs and all arcs begin and end with nodes.

---

**Problem Set**

Given a simple network describe the method of getting from an origin to a destination. The activity of getting from an origin to a destination is referred to as way-finding. A network is an interconnected set of arcs representing possible paths for the movement of resources from one location to another. There are many examples of networks such as pipelines, power lines, sewers, and roads.

Consider the graphic below, which is map of a portion of a road network. On this network are indicated a starting point (origin) and an ending point (destination). The process of reading the map could be used, by humans, to derive several different routes from the origin to the destination. For example, “travel east on Confusion Street, then north on Codd …”.

For a digital computer, however, this simple cognitive process is not so easily modeled. For example, how does the computer application read the database to know that Confusion Street and Codd are even connected?

The answer lies in structuring data to include information about the relationships between the roads (or arcs) in the database. In this arc-node approach the vector data are structured to have attributes related to the identification of all arcs and nodes in the network. Consider the portion of the map at right. On this map the blue numbers are node numbers, the red numbers...
are arc numbers and the green numbers are the numbers of the polygons formed by the arcs. The red arrows indicate the direction of the arc, based on the ordered sequence of coordinate pairs. By storing the so-called “from-node” and “to-node” for each arc the software application can analyze relationships between the arc features – for example that arc 7 and arc 8 share node 17. In this way the software application can analyze a series of trials to get from the origin to the destination by moving out along arcs connected by nodes. In this way, the process of way finding that involves map cognition in humans can be abstracted in a computer because the connectivity of the arc network is stored with the features.

With respect to the arc-node model, a definition of topology emerges that is a bit more specific than the dictionary definition. For these purposes, topology is the study of relationships between objects and includes those properties that remain constant under deformation. As originally modeled by ESRI, these relationships include

- Connectivity – the attachment of the objects
- Contiguity – the identification of adjacent polygons by recording the left and right polygon of each arc
- Area Definition – a closed area defined by a boundary

**Uses of topology in GIS**

Topology utility in GIS can be grouped into two broad areas. The first of these is to support spatial analyses. These can include using connectivity for network analysis; area definition to determine containment; and contiguity for neighborhood analysis.

The other use area is the support of database development. Many times knowledge of the topological conditions in a data set can be used to discover structural problems with the feature database – for example polygons that are not closed or are overlapping. It can also be used to automate feature creation and ensure feature integration.

**Other Topologies**

Just as the structure of the geodatabase changed the specific storage of feature geometries in feature coordinate files, it offered the ability to model behaviors in objects by using additional topologies. These additional topologies are established by specifying rules that determine the behavior of features with respect to other features. The figure below describes some of the types of topology supported by the geodatabase. Note that in addition to arc-node topology, which deals with the geometry of feature components in a feature, there are topologies that are topologies that affect the geometry of features based on the location of other features, sometimes in other feature classes.
The figure below displays a set of feature classes and the associated ranks and rules that apply to those participating in a topology. The behaviors that the rules enforce can affect the construction of the feature database. Many times, such rules are established to maintain physical integrity of the feature data.

The figure on the left represents a simple geodatabase. The diagram above displays topology rules that could be applied to this data.
Topology in the Geodatabase

According to its developers, in ArcGIS topology is maintained \emph{on top of the geodatabase}. This is to say that the topologies supported by ArcGIS are built on the data structure of the geodatabase. Topology appears as a database object in the geodatabase. It defines a formal mathematical model used to integrate geometry form one of more feature classes. The database object also manages topology elements are relationships between features.

Below is a view of a geodatabase called “Montgomery_full” containing a feature data set called “Landbase” that contains a topology object called “Landbase_Topology”.

The feature data set called “Landbase” contains a set of feature classes that share the same spatial reference. From this view it is not possible to tell, necessarily, which feature classes in the feature data set are participating in the named topology.
**Parameters that define topology**

Topologies are used to store three important sets of parameters and some internal feature layers. The first of these parameter sets is the *cluster tolerance*. It specifies the distance at which all vertices or boundaries are considered coincident. There is a single tolerance established for the topology. At a particular phase of database construction vertices falling within this tolerance are snapped together – this is called *validation*.

Ranks are the second set of parameters established for a topology. There are defined at the feature class level and determine how much features in that class can move relative to features. Lower ranked features snap to higher ranked features during validation. This implies that the positional accuracy was higher in some features than others. The movement of two features of the same rank is geometrically averaged as described in the figure below.

*When you validate a topology, the ranks of the feature classes in the topology control how features are snapped together. Lower-ranking features snap to higher-ranking features. Equally ranked features snap to the geometric average of their position.*
The third group of parameters is the topological rules themselves. The rules defined for a topology control the allowable relationships of features within a feature class and between features in different feature classes. During database development, users can add or remove rules from a set. The rule defines a condition in the topology, such as “land parcels must not overlap” or “census blocks must be contained in block groups”. The features are checked against the rules to identify features that are in violation.

During validation the process of feature cracking and clustering occurs. Vertices are inserted at intersection of features during cracking. Clustering involves the snapping together of vertices within the cluster tolerance. After validation the feature class will report that the feature class has been validated and may report errors.