LESSON 9
PRINCIPLES OF TOPOLOGY AND THE SPATIAL DATA TRANSFER STANDARD

Introduction

Creating the data base is one of the most expensive and time consuming aspects in GIS development. It also creates the most problems in that it becomes the source of subsequent manipulation and analysis. If not done properly then data may not be accessible to the user or may not meet the needs of the user community in terms of both accuracy and completeness.

The data are varied and complex. Unlike most database programs used in business applications, GIS data consist of both graphic and non-graphic attribute data. The data base therefore needs to describe a myriad of attributes such as digital description of map objects, logical geographic relationship between features, and attributes that describe the characteristics of the mapped features. The content of the data base depends on the user requirements.

There is no agreement as to how data base should be spelled: two or one word. In these lessons, the use of data base (2 words) indicates the form of the data. It is more generic. On the other hand, database will refer to the program where the data are stored and manipulated. Typical database programs are dBase, Rbase, Lotus Approach, etc.

Graphic and non-graphic data

Graphic data is the map, almost always stored in a digital format (in fact I do not know where it is maintained in an analog paper form as the sole source of information). The descriptions of these map features include coordinates, rules, and symbols used to define the cartographic objects. The non-graphic data contains the characteristics, qualities, or relationships of the features and their locations. The non-graphic data is stored in a conventional database format. The third item of importance is the linkage between these two different data. This linkage is critical if queries are to be made either from the graphic to the non-graphic data, or vice versa. For example, we might want to know all parcels that are zoned residential and depict them on a map. Therefore, we can search the attribute data for zoning description and flag those that are residential. Then those tagged parcels can be drawn using the graphic data.

The non-graphic data can be divided into attributes, geographically referenced data, geographic indexes, and spatial relationships. Attributes describe the graphic elements either with words or numbers. For example, a parcel may have attributes that include the owner’s name, address, and assessed value. Pipe segments within a GIS may include attribute data describing the diameter, material, installation date, and maintenance record. Geographically referenced data describe physical phenomena, man-made objects, and events that occur at a particular location. An example would be location of particular crimes that occur in a locality. These just show where the crime occurred but are not attributes to a graphic entity, such as a parcel. Geographic indexes help locate map features and data based on their geographic identifiers. As an example, we might want to locate all parcels that fall within an address range, assuming address is an identifier carried within the data base. Finally, spatial relationships deal with topological relationships such as proximity, adjacency and connectivity. For example, we may use a road network to aid us in optimizing traffic flow. In such a situation we would want to know which roads are connected to each other, whether they are one-way streets or not, if there are traffic rules to follow such as no left turn, etc.

Six different types of elements depict the graphic data: points, lines, areas, grid cells, pixels, and symbols. Because it is important that the data be portable to different platforms and software systems, the federal government has adopted Spatial Data Transfer Standards (SDTS), which are an ANSI (American National Standards Institute) standard referred to as ANSI NCITS 320-1998. A factsheet of this standard, which is produced later in these lesson notes can be found at: http://mcmcweb.er.usgs.gov/sdts/whatsdts.html. This sheet also contains a linkage to the actual standard and links to the SDTS home page http://mcmcweb.er.usgs.gov/sdts/index.html.

According to these standards, a point is a zero-dimensional object specifying the geometric location using a set of coordinates. A special kind of point is a node, which is defined as a topological junction or end point. This concept, topology, will be discussed shortly. A line is a one-dimensional object with a line segment being a directed line between two points. Some more important special lines defined by SDTS are string, arc, and chain. The string is a series of line segments. An arc is a line forming a curve as defined by some mathematical relationship. A chain is a directed sequence of non-intersecting line segments or arcs with nodes at each end. At least one software package, ARC/INFO by Environmental Systems Research Institute (ESRI), uses the term arcs instead of chain in their topological definitions.

A two-dimensional object is defined as an area. Individual areas are represented by polygons. A pixel, as we recall, is a two-dimensional picture element that is the smallest indivisible element of an image whereas a grid cell is another two-dimensional object representing a single element of a continuous surface. An example of a grid cell is a USGS (United States Geological Survey) digital elevation model. Symbols are graphic elements or pictures representing feature locations on a map. Finally, an annotation is a text or label plotted on the map used for identification of a feature.

Location is determined by establishing some type of coordinate system. In a global sense, latitude and longitude are used to located positions on a glove or spherical earth. A Cartesian coordinate system, though, is often used to locate a point on a plane surface, such as a map. There are many different types of Cartesian systems but they all share one common feature and that is that points are defined by x and y coordinates from an origin (figure 1). The main advantage of this kind of system is the ease in computation since the curvature of the earth does not have to be accounted for. This figure also shows that point A can be defined in terms of polar coordinates: Azimuth (Az) and Distance. The two most popular Cartesian Coordinates systems are State Plane Coordinates and Universal Transverse Mercator (UTM) coordinate systems.

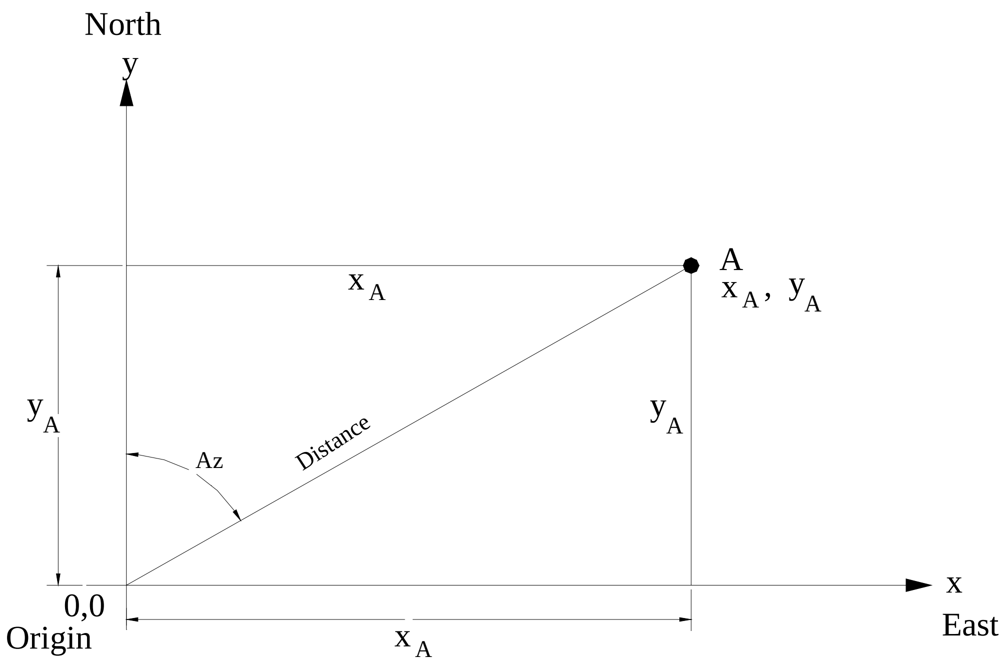


Figure 1. Cartesian coordinate system.

State plane coordinates project the position of a feature on the earth onto a surface that can be developed into a plane. It should be pointed out that in this transformation, data gets distorted. But, these systems can be designed to minimize this distortion. The State Plane Coordinate System (SPCS) , maintains the shape, or angles, in the projection. This means that an angle measured on the earth will be the same as on the projections surface. Distances and areas, though, are distorted. A projection system that maintains the angles is called a conformal projection. It must also be pointed out that this is only a general property and that angles do experience some distortion because this is inevitable. But this treatment is well beyond what we are trying to discuss here. The amount of angular distortion is much less that that of area and distance. Distance distortion is minimized in the SPCS by limiting the size of the projection to about 158 miles. The amount of distortion in distance is usually less than 1:10,000. Because of the limitation in the width, large states must have more than one zone.

As it has been stated, a conformal projection is one with a general property of controlling angle or shape distortion. In other words, the shape of an area, as an example, is more faithfully rendered in a conformal projection than other types of systems. An equidistant projection is one that preserves distances over the map while an equivalent, or equal area, projection preserves area.

Two main projection systems are used in the SPCS: Lambert Conformal Projection and Transverse Mercator Projection. The Lambert projections uses a cone as the projection surface with the apex located on the rotational axis of the earth. It is used in states that are primarily east-west in extent, such as Tennessee. The Transverse Mercator system used a cylinder as the projection surface. The center of the cylinder is perpendicular to the rotational axis of the earth. This system is used for states that are primarily north-south in extent. In both systems, the projection surface can be cut and then flattened into a plane without further distortion. When using a SPCS in a particular state it is critical to understand what units define the system. Many are defined in metric units although both US Survey Foot and International Foot are also frequently employed[1]. This lack of unity just adds confusion to the use of state plane coordinates.

The UTM projection system is similar to the Transverse Mercator projection in the SPCS except that it covers most of the earth. It divides the earth into 6-degree zones based on longitude. For the North and South poles a stereographic projection is used.

For more information on map projections in general and the SPCS and UTM coordinate systems in particular, visit the following web site: http://www.colorado.edu/geography/gcraft/notes/mapproj/mapproj_f.html.

[1] In the United States, the use of the metric system was legalized in 1866 by President Andrew Johnson even though it was used earlier by Ferdinand R. Hassler of the Coast Survey (now Coast and Geodetic Survey, NOS, NOAA) who brought an iron copy of the French meter to the U.S. One important aspect of the 1866 metric bills was the specification of the English equivalent. The foot/meter ratio was defined as 3937/1200 exact. After the International Prototype Meter was established, two copies were sent to the U.S. These were declared the fundamental standards in 1893 in what has been called the Mendenhall Order, after the Superintendent of the Weights and Measures. The same foot/meter relationship was maintained.

Because of inconsistencies in the yard and meter both in the United States and abroad, an agreement between the U.S. and the United Kingdom, in 1959, further refined this relationship. From that agreement, the yard was defined as being 0.9144 meter exact, or a foot being equal to 0.3048 meter exact. This had the effect of shortening the foot by two parts in a million. Thus, this refinement defines the international foot. Because of the problems in surveying and mapping, the old standard (3937/1200) was maintained for surveying purposes and this is called the U.S. Survey foot.

While the International Prototype Meter was an accurate standard when it was developed, there existed a need for higher accuracies in science. Therefore, in 1960 the length of the meter was redefined to be equal to 1,650,763.73 wavelengths in a vacuum of orange-red light of the krypton 86 atom. This became known as the National Prototype Meter. This modernized system was called Le Systeme International d'Unites (International System of Units) and is commonly referred to as the SI system. On December 23, 1975, President Ford signed the Metric Conversion Act of 1975 calling for gradual conversion to the metric system. In 1983 the 17th CGPM (General Conference on Weights and Measures) redefined the meter to be the distance that light travels in 1/299 792 458 of a second.

It was already discussed in an earlier lesson that the graphic data can be represented either as raster or vector format. Both systems can be displayed together on most GIS packages. If remotely sensed or digital photography are used with the GIS, the data are represented by raster data. Most often, cadastral information is shown in vector form, unless the cadastral maps where scanned in the input process.

Database management

The management of GIS data is important because it has unique characteristics not found in conventional database management systems (DBMS). In fact, the non-graphic data often utilized conventional DBMS software. The big difference is that a GIS also contains a high volume of graphic data that needs to be efficiently designed for quick find-and-retrieve procedures. Different vendors use several approaches.

A hierarchical data model stores data based on a parent-child, one-to-many relationship. An excellent example of a hierarchical system is the public land survey system . Here one township (parent) contains 36 sections (children), each section (parent) contains four quarter-sections (children), etc. This is like a family tree. The biggest problem is that the relationships need to be clearly defined, which may be an issue. One problems it that longitudinal searches require the computer system to traverse up and down the branches found in this structure.

A network data model can take on different forms. One easily understood approach is to create a tree-like structure like that found in the hierarchical model. At this level, the network model has the same one-to-many relationship. But additionally, the network model can have a many-to-many relationship. This makes for more efficient searches within the data base without traversing up and down branches.

Parcel No.	Owner	Address	Zoning	Assessed Value	....
1	K. Smith	1433 Main	R1	89,000	....
2	J. Brown	744 Elm	R1	93,000	....
3	B. Clinton	1500 Maple	C1	126,900	....
4	D. Jones	752 Elm	R2	79,500	....

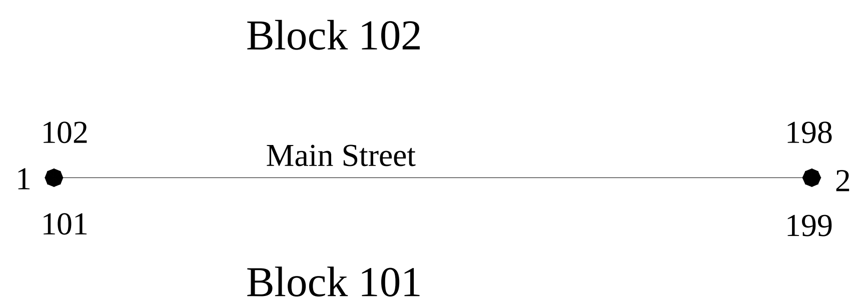
Table 1. Example of relational data structure.

A very popular model is the relational data model. This structure is based on a matrix format consisting of rows and columns. Each row is called a record and each column a field or item. An example relational model is shown in table 1. This structure allows for great flexibility in that a wide range of queries can be handled. There is no need to predefine selection keys thus allowing almost any kind of combination of values to be used as selection criteria. It is also very easy to link one item within one database with other items in another database using a common identifier.

Topology

There is a branch of mathematics that deals with relationships between objects. There are two types of objects: points, which are called nodes, and lines referred to as edges (also called links or chains). The basic relation between them is called incidence. This mathematical field is called topology. It is used in a GIS to record and manipulate the logical relationships of the map features and geographic information. The three relationships important for us include contiguity, adjacency, and area definition

Figure 2 shows the topological elements that are found within geographic data. The end points and intersection points are the nodes or 0-cell features. Between node 1 and 2 all of the spatial relationships are the same. The links between the nodes are 1-cell entities. Here it is depicted as a single line but it can consist of a series of line segments. In topology, a link must have a direction or a from-node and a to-node. This allows us to define the relationship between the adjacent polygons. In the figure, the link is running from node 1 to node 2. Thus, we can see that the polygon on the right is Block 101 and the polygon on the left is Block 102.



- From Node 1 to Node 2
- Street Name - Main Street
- Left Address Range - 102-198
- Right Address Range - 101-199
- Left Block - 102
- Right Block - 101

Figure 2. Topological elements.

Figure 3 shows a topological structure of a series of polygons, such as land-use areas. The numbers 1-8 represent the nodes. The lower case letters, a-f, represent change points. These are points on the directed line sequence from one node to another that provide shape to the polygon. The upper case letters, A-E, indicate the polygon. Table 2 shows the directed link segments which associates the link to the from and to nodes. It also identifies which polygon is to the right and left as we travel from the from_node to the to_node. Note that "0" is used to indicate all of the area outside of the site the GIS encompasses. Similar tables can be constructed for nodes and polygons. For example, a node table may look like that shown in table 3. This shows how many links intersect at the node (called valency) and identifies those links. Valency is important because it gives us a measure of the topological correctness of the data. For example, a valency of one or two is not normally found in the data. If we were to evaluate the topology, it means that this might be an error. An island, though, would have a valency of one since the beginning and end node would be the same. In this case, the topological structure does not mean that there is an error. Another example would be a street centerline network where the node identifies a cul-de-sac. A valency of two is also not frequently found, but it could indicate correct data. For example, a freeway may have nodes at entrance and exit ramps. But between these nodes, there may have been new pavement over a small section of the road. The highway department may want to identify these new sections as different from the original sections.

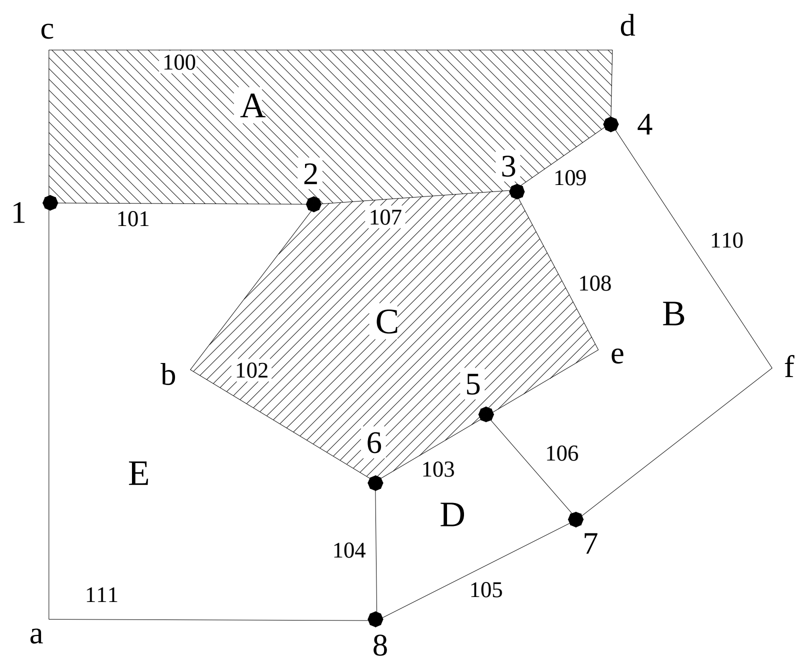


Figure 3. Topological structure.

Finally, it should also be pointed out that topological structure is not the same as Euclidean geometry, another branch of mathematics. Figure 4 shows three topologically similar geometries. They are topologically correct since they have the same relationships between the data elements. Clearly, from a Euclidean geometry point of view, they are not similar. Therefore, to properly represent the phenomenon in their correct space, coordinates need to be provided for each node and change point along a chain.

Link	From_Node	To_Node	Left_Polygon	Right_Polygon
100	`1	4	0	A
101	1	2	A	E
102	6	2	E	C
103	5	6	D	C
104	8	6	E	D
105	7	8	0	D
106	5	7	B	D
107	3	2	C	A
108	3	5	B	C
109	4	3	B	A
110	4	7	0	B
111	1	8	E	0

Table 2. Link file.

Node	Links
1	100, 101, 111
2	101, 102, 107
3	107, 108, 109
4	100, 109, 110
5	103, 106, 108
6	102, 103, 104
7	105, 106, 110
8	104, 105, 111

Table 3. Nodes data file.

How does topology help us? Lets look at a routing example. We want to find the shortest path from A to B, given that the streets are all one-way (figure 5). The optimal path would be to go to the top two blocks and then to the left 5 blocks. But, because we have to account for one-way streets, the best path is to go right 2 blocks, to the top 3 blocks, to the left 9 blocks, to the bottom 1 block, then to the right 2 blocks. The direction of the streets would be found by looking at their topological structure.

Data dictionary

An essential ingredient of any database is a data dictionary that describes the feature for all users. The problem with phenomenon is that different people in different parts of the country may call them different things. For example, a soft drink might be called pop, soda, or even Coke. In terms of spatial data, what is a road? To the highway engineer, a road has several different connotations. Does the road shown on a map mean that it is paved? What about two-track roads? Thus, the data dictionary must carefully describe the feature so that all users know exactly what it means. This information should be included within the metadata. This will be discussed in the next lesson.

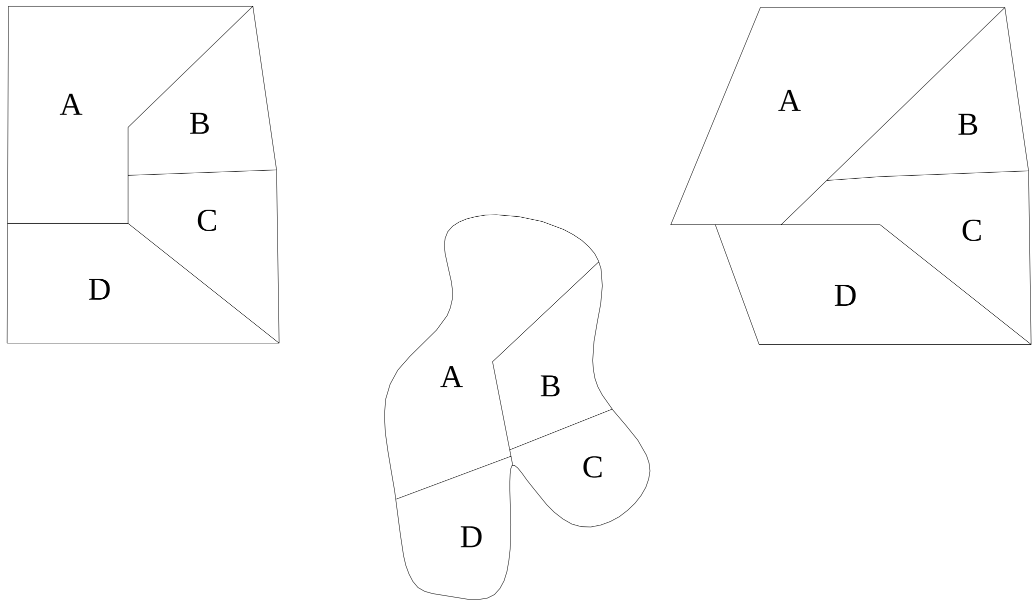


Figure 4. Topologically similar areas.

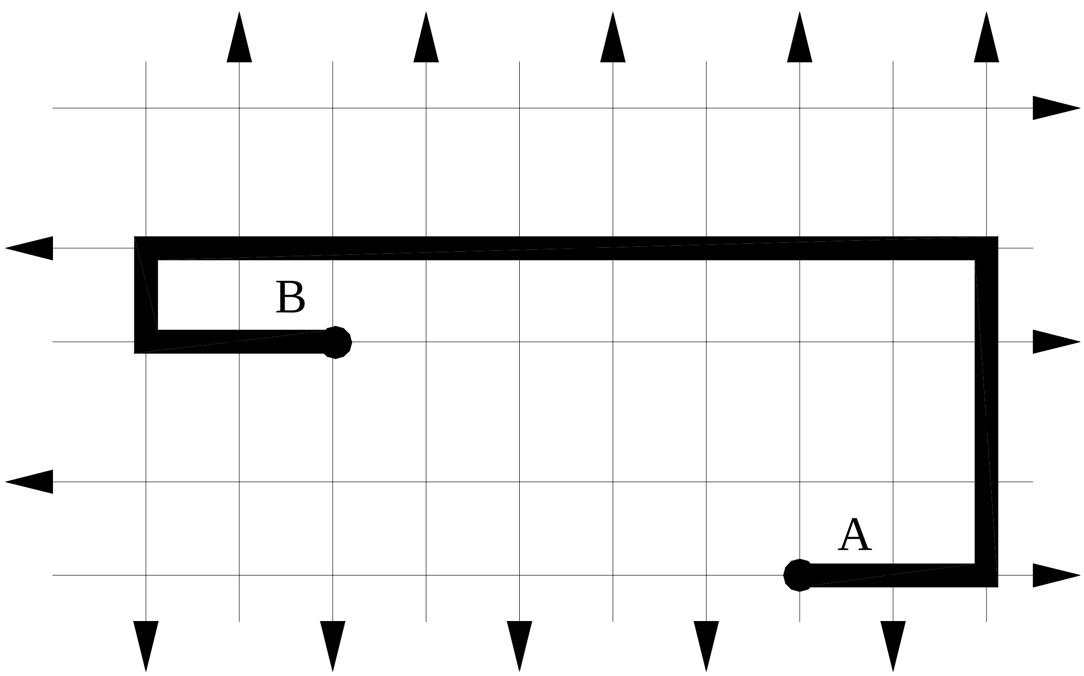


Figure 5. Topology example with vehicle routing.

Spatial Data Transfer Standard

With so many different computers and computer systems being used in geographic information systems, compatibility becomes a real issue if data are to be shared across different systems. The problem is how to facilitate distributing spatial data to different users who are using dissimilar systems? From many of our experiences in converting documents created in one word processor and trying to read it in another, we often find anomalies. GIS data is unique in that there are linkages between the graphic and non-graphic data that need to be preserved during this conversion.

The following information, slightly modified from the USGS, describes the Spatial Data Transfer Standard[2].

USGS

What is SDTS?

The Spatial Data Transfer Standard, or SDTS, is a ***robust*** way of transferring earth-referenced spatial data between dissimilar computer systems with the potential for no information loss. It is a transfer standard that ***embraces*** the philosophy of self-contained transfers, i.e. spatial data, attribute, georeferencing, data quality report, data dictionary, and other supporting metadata all included in the transfer.

Purpose of SDTS

The purpose of the SDTS is to promote and facilitate the transfer of digital spatial data between dissimilar computer systems, while preserving information meaning and minimizing the need for information external to the transfer. Implementation of SDTS is of significant interest to users and producers of digital spatial data because of the potential for increased access to and sharing of spatial data, the reduction of information loss in data exchange, the elimination of the duplication of data acquisition, and the increase in the quality and integrity of spatial data. SDTS is neutral, modular, growth-oriented, extensible, and flexible--all characteristics of an "open systems" standard.

The SDTS provides a solution to the problem of spatial data transfer from the conceptual level to the details of physical file encoding. Transfer of spatial data involves modeling spatial data concepts, data structures, and logical and physical file structures. To be useful, the data to be transferred must also be meaningful in terms of data content and data quality. SDTS addresses all of these aspects for both vector and raster data structures.

---

[2] Document accessed from http://mcmcweb.er.usgs.gov/sdts/whatsdts.html.

Components of SDTS

The SDTS specification is organized into the base specification (Parts 1-3) and multiple profiles (Parts 4-6). Parts 1-3 are related, but relatively independent, each dealing with its own piece of the spatial data transfer problem. Parts 4-6 each define specific rules and formats for applying SDTS for the exchange of particular types of data in SDTS.

The six parts of SDTS:

Part 1 - Logical Specifications

Part 2 - Spatial Features

Part 3 - ISO 8211 Encoding

Part 4 - Topological Vector Profile

Part 5 - Raster Profile

Part 6 - Point Profile

History of SDTS

The Federal Information Processing Standards (FIPS) Program was established in the 1960s to standardize federal usage of computers. FIPS are government standards for federal agencies and organizations. The administrator of the FIPS Program is the National Institute of Standards and Technology (NIST). In the mid-1970s, computers began popping up throughout many federal geographic and cartographic agencies. As the application of computers in geography and cartography grew within the federal government, the need for earth science data standards became apparent.

In 1980, the U.S. Geological Survey (USGS) was designated the lead agency in developing earth science data standards for the federal government. The USGS worked with academic, industrial, and federal, state, and local government users of computer mapping and GIS to develop a standard for transfer and exchange of spatial data. In 1992, after twelve years of developing, reviewing, revising, and testing, the resulting standard--SDTS, was approved as Federal Information Processing Standard (FIPS) Publication 173, known as **FIPSPUB 173-1, 1994**. The FIPS version has been superceded by current version, known as **ANSI NCITS 320-1998** and was ratified by the American National Standards Institute (ANSI) June 9, 1998.

Compliance with SDTS is now mandatory for federal agencies. SDTS is available for use by state and local governments, the private sector, and research and academic organizations.

To receive SDTS information in hardcopy, send your mailing address and a request for the SDTS information you would like to:

SDTS Task Force
U.S. Geological Survey
1400 Independence Road
Rolla, MO 65401

Or call 573-308-3561, FAX 573-308-3652

As we can see from the information from this flyer, profiles are used to help create the SDTS. The following information, again slightly modified from the USGS information sheet, helps define what a profile is and what profiles are currently available[3].

USGS

Implementing SDTS through Profiles

The Spatial Data Transfer Standard (SDTS) is implemented through the use of profiles. Since the SDTS is designed to support all types of spatial data, implementing all of the standard's options at one time would be a monumental task. Profiles provide the best method for successful implementation of SDTS. Profiles balance two objectives of SDTS, first to allow both encoding and decoding to be feasible, and second to ensure that all meaningful information is transferred.

Profiles balance two objectives of SDTS, first to allow both encoding and decoding to be feasible, and second to ensure that all meaningful information is transferred.

A profile is intended to provide specific rules for applying SDTS base specifications (Part 1, 2, and 3) to a particular type of spatial data. A profile can be considered a subset of the SDTS specification that defines:

- Restrictions and requirements for use of specific spatial object types
- Restrictions and requirements for use of SDTS modules, including rules for choosing among options present in the base specifications
- Module naming and file naming conventions
- Use of ISO 8211 encoding specifications, including allowable options

As the SDTS concerns itself with all spatial data types and their variations, a profile is created to handle a particular model of spatial data. Profiles are designed with enough flexibility to account for variability in user data models, thus avoiding l arge proliferation of profiles for different data models that are similar in structure. Several user data formats and data products can share a single profile. For example, the Topological Vector Profile will handle USGS DLG, Census TIGER, ESRI Coverage, and any other topologically structured vector data set.

All SDTS transfers must be encoded under a profile. Users of the SDTS are encouraged to make every effort to conform to an existing profile. The existence of many competing

[3] See web page: http://mcmcweb.er.usgs.gov/sdts/profile.html.

or conflicting profiles to SDTS would require a large number of translators and this defeats the goal of having a common transfer standard to spatial data.

Specific SDTS Profiles

Topological Vector Profile (TVP). The TVP is endorsed by FGDC as FGDC-STD 002.4. It is the first completed and NIST approved profile and is included as Part 4 of the SDTS. The TVP was developed to support geographic vector data with geometry and topology. It includes the SDTS defined spatial objects representing vector data with full topology that comprise a two-dimensional manifold. Data sets may contain point, line, polygon, and composite features. The TVP accommodates USGS vector products such as the Digital Line Graph (DLG) data. The Bureau of the Census will also be using the TVP to distribute its Topologically Integrated Geographic Encoding and Referencing (TIGER) data files.

Raster Profile and Extensions (RPE). The RPE, formerly known as SRPE, is Part 5 of the SDTS. FGDC endorsed the RPE as FGDC-STD-002.5-1999. The RPE was developed to support two-dimensional spatial data sets in which features or images are represented in raster or gridded form. This profile can accommodate image data, digital terrain models, gridded GIS la yers, and other gridded data. This profile does not permit vector objects, raster objects higher than two-dimensional, or irregular grids. The RPE will accommodate USGS raster products such as Digital Elevation Models (DEM), and Digital Orthophoto Quadrangles (DOQ) data files. The Extensions support optional use of (1) the ISO Basic Image Interchange Format (BIFF) for images, (2) the JPEG File Interchange Format (JFIF) for compressed images, and (3) Georeferenced Tagged Information File Format (GeoTIFF).

Transportation Network Profile (TNP). The TNP has been prepared in draft form by the Volpe National Transportation Systems Center for the US Department of Transportation (USDOT) Bureau of Transportation Statistics (BTS). The TNP contains specifications for an SDTS profile for use with geographic vector data with network topology. Data sets are represented by vector objects which comprise a network (sometimes non-planar) or planar graph. Excluded are raster data and geometry-only vector data. Once available, the TNP will accommodate the USDOT BTS transportation network data files.

Point Profile. The Point Profile has also been referred to as the Geodetic Profile, the High Precision Point Profile, and SDTS Part 6. A draft version of the Point Profile was prepared in mid 1996 by the National Oceanic and Atmospheric Administration - National Geophysical Data Center (NOAA-NGDC) and the USGS. This profile is designed to support a major release of geodetic control point data from NOAA's National Geodetic Survey (NGS), as well as point-only data from other agencies. The adoption of this profile as an FGDC standard was sponsored by the Federal Geodetic Control Subcommittee (FGCS) of FGDC through the Standards Working Group of FGDC. FGDC approval for the Point Profile, known as FGDC-STD-002.6, was approved in 1998.

Computer Aided Design and Drafting Pr ofile (CADD). The Computer Aided Design and Drafting Profile (CADD) contains specifications for an SDTS profile for use with vector-based geographic data as represented in CADD software. The purpose of this profile is to facilitate the transl ation of this data between CADD packages without loss of data, and support the translation of this data between CADD and mainstr eam GIS packages. CADD software makes up a large portion of the Geographic Information Systems (GIS) marketplace. CADD software allows for several types of elements, in particular, the use of three-dimensional elements and complex curves that are not comm only used by GIS. This profile allows the representation of two- and three-dimensional geographic vector data from CADD package s to be transferred via the SDTS standard. This profile supports two-dimensional vector data and three-dimensional vector data, where the third dimension is the "height" of the object. These data may or may not have topology. This profile does not suppor t raster data or two-dimensional transfers already represented by another profile. The CADD profile, approved by the FGDC in Mar ch 2000, is known as FGDC-STD-002.7-2000.

Additional profiles are being considered for non-topological vector data, and DIGEST-VPF relational vector data. A number of **papers** were presented during the SDTS Workshop held in Rolla, Missouri in 1997 that may be helpful if you are interested in creating a new profile. As other profiles become available they will be posted on this page.

Steps for the Development and Approval of an SDTS Profile

1. Have interest in putting data into SDTS.
2. Try to match your data model to existing profiles.
3. Determine that no current profile will do.
4. Contact FGDC or SDTS authority with idea for new profile.
5. Announce intentions. Seek groups with similar needs.
6. Write new profile document using most similar profile as guide.
7. Contact NIST, FGDC SWG, ANSI, or other standards group to begin formal review and approval process , if desired.
8. Develop documents describing mappings of specific products into this profile.
9. Develop sample data sets using profile.
10. Seek vendor support for translator development.
11. Work with NIST on conformance testing.
12. Begin producing or exchanging data using profile.

- To view **Parts 1 - 5** of SDTS
- To view the draft **Transportation Network Profile** document
- To see an HTML version of the draft **Point Profile** document
- For a detailed overview of **SDTS Profiles**
- For more informative articles on Profiles visit our **SDTS Information FTP Site**

There are a number of good articles on the Spatial Data Transfer Standard. Two provide a general review of the standard and are available in pdf format. The first is a short article called the "Senior Management Overview". Read this paper for this lesson. The paper can be accessed at: http://www.geo.ulg.ac.be/interne/Fichiers_Interet_General/SIG/SDTSgen.pdf. The second is a longer publication called "The Spatial Data Transfer Standard: Guide for Technical Managers". This paper is an excellent review of the system but is not required reading for this particular lesson. The paper can be accessed using the following address: ftp://sdts.er.usgs.gov/pub/sdts/articles/pdf/mgrs.pdf.

Review Questions

1. Non-graphic data can be divided into four types. What are they? Give an example that is different from those presented in the lesson.

2. What is the difference between a point and a node? Explain your answer.

3. All map projections distort data. How are distance and angular distortions minimized in the State Plane Coordinate System?

4. Describe the different data models used to manage graphic data.

5. What is topology? Why is topology important in a GIS?

6. Why is it necessary to have a standard exchange format for spatial data (see SDTS Senior Management Overview paper)?

7. What is the difference between an intermediate exchange and direct exchange format (see SDTS Senior Management Overview paper)?

8. SDTS base specifications are described in three parts. Describe them (see SDTS Senior Management Overview paper).

9. What is a profile (see SDTS Senior Management Overview paper)?

Appendix
Excerpts from SDTS

1 Introduction

This Spatial Data Transfer Standard (SDTS) provides a solution to the problem of spatial (i.e., geographic and cartographic) data transfer from the conceptual level to the details of physical file encoding. Transfer of spatial data involves modeling spatial data concepts, data structures, and logical and physical file structures. To be useful, the data to be transferred must also be meaningful in terms of data content and data quality. SDTS addresses all of these aspects for both vector and raster data structures.

The standard is in three parts. Part 1 addresses the logical specifications in terms of conformance requirements, a conceptual model, quality specifications, the data structure model, and the transfer format. Part 2 addresses data content by providing a standard list and definitions of spatial features and their attributes. Part 3 specifies the implementation of SDTS in terms of the International Organization for Standardization for a Data Descriptive File for Information Interchange.

Section 1 of part 1 includes a statement of scope and conformance requirements for SDTS. It also includes normative references to other standards and definitions of terms.

The conceptual model of spatial data is presented in section 2 to provide a framework for defining spatial features and a context for the definition of a set of spatial objects. This conceptual model supports the translation of user data models to and from the SDTS model. Within section 2 is a defined set of spatial objects, for zero, one, and two dimensions, used in spatial data systems to represent real-world spatial phenomena. Three-dimensional spatial objects have not been specified. The defined set of objects will support the three major types of spatial data operations: (1) geometry only, (2) geometry and topology, and (3) topology only. These objects have been specified in a modular fashion so that more elaborate composite objects can be constructed from them.

Section 3 includes specifications for a quality report concerning the objects in a transfer and their attributes. The purpose of the quality report is to provide detailed information for a user to evaluate the fitness for a particular use. This style of standard can be characterized as "truth in labeling," rather than fixing arbitrary numerical thresholds of quality. To implement this portion of the standard, a producer is urged to include the most rigorous and quantitative information available on the components of data quality described in this section.

Sections 4 and 5 present specifications for the transfer of spatial data. Section 4 contains general concepts and specifications, the underlying models that pertain to the transfer module specifications of section 5. Section 4 also specifies the general elements of an implementation, the relationships of the logical constructs of the data models to the general elements of a detailed implementation, and general constraints on the implementation. Finally, section 4 presents the transfer module specification conventions used in section 5. Logical modules consisting of detailed record, field and subfield specifications are presented in section 5.

1.4 Definitions

The following terms are used in the definitions of concepts essential to this standard. The concepts for which this standard provides normative definitions appear in section 2, Spatial Data Concepts; section 3, Spatial Data Quality; and section 4, General Specification. An informative list of all defined terms is contained in annex E.

Accuracy - The closeness of results of observations, computations, or estimates to the true values or the values accepted as being true.

Altitude - Elevation above or below a reference datum, as defined in FIPSPUB 70-1; the z-value in a spatial address. See also elevation.

Control (mapping) - A system of points with established horizontal and vertical positions that are used as fixed references in positioning and relating map features.

Coordinates - Pairs of numbers expressing horizontal distances along orthogonal axes; alternatively, triplets of numbers measuring horizontal and vertical distances.

Data base - Related subject information stored as a volume set, volume, file set, or file.

Data element - A logically primitive item of data.

Digital encoding - To convert to a form that can be operated upon by electronic computer as binary digits.

Elevation - Conforming to FIPSPUB 70-1, the term "altitude" is used in this standard, rather than the common term elevation, for the z-value in a spatial address.

Field - Consists of one or more related subfields. It may contain part or all of a module field. It does not contain parts of two or more module fields.

Field name - A name associated with a field.

File - An identifiable collection of zero or more related records. It may contain part of, or all of one or many modules.

File set - An identifiable collection of zero or more related files.

Geocodes - A system of encoding used to represent an exhaustive list of a class of spatial features (usually applied to political units).

Implementation method - A method of encoding data content and data structure to accomplish a transfer between dissimilar computer systems without loss of content, meaning, or structure.

Map - A spatial representation, usually graphic on a flat surface, of spatial phenomena.

Media - The physical devices used to record, store, and (or) transmit data.

Media record - A physical unit of data. The characteristics of a record and its means of delimitation are defined by standards specific to each given medium.

Misclassification Matrix - Results of an attribute accuracy test given in the form of a row by column contingency table (crosstabulation) sometimes called a classification error matrix. The rows represent the interpretation tested and the columns represent the verification assumed to be correct. The diagonal elements represent the correct classifications when the matrix is square and the rows and columns are strictly comparable. The remaining elements can be treated row-wise as errors of commission, and column-wise as errors of omission.

Module - A logical collection of module records.

Module field - A defined set of one or more module subfields in a Spatial Data Transfer.

Module record - A defined set of one or more module fields in a Spatial Data Transfer.

Module specification - The meaning, identification, order requirements, and data structure requirements for data belonging to the module.

Module subfield - A logical construct defining a single data element in a Spatial Data Transfer.

Primitive - The quality of not being subdivided; atomic.

Quality - An essential or distinguishing characteristic necessary for spatial data to be fit for use.

Quality overlay - A collection of points, lines, and areas organized to represent quality information for another set of map information. An overlay describing lineage may be termed a source data index; a positional accuracy overlay may be termed a reliability diagram.

Record - An implementation-dependent construct that consists of an identifiable collection of one or more related fields.

Representation - Graphical symbolization of a spatial object.

Resolution - The minimum difference between two independently measured or computed values that can be distinguished by the measurement or analytical method being considered or used.

Spatial data transfer - A collection of related modules.

Subfield - A physical area containing, or logical construct defining, a single data element.

Transfer construct - A volume set, volume, file set, file, module, module record, module field, module subfield, field, subfield, record, or media record.

Transformation - A computational process of converting a position from one coordinate system to another.

Volume - A media-dependent construct consisting of an identifiable collection of part of or all of one or more files. A volume is a discrete interchange construct such as a unit of dismountable media or a single online session.

Volume set - A media-dependent construct consisting of an identifiable collection of one or more volumes containing a single file set.