4

Geocoding

4.1 GEOCODING AND COMPUTER CARTOGRAPHY

Geocoding is the conversion of spatial information into computer-readable form. As such, geocoding, both the process and the concepts involved, determines the type, scale, accuracy, and precision of digital maps. An important aspect of geocoding is that effective geocoding requires an understanding of some basic geographic properties underlying geographic data. Unfortunately, it is quite easy to convert information into computer-readable form as data, but not all data are information, and all too frequently analytical cartography drowns in a sea of meaningless and unusable numbers. Most essential for analytical and computer cartography is that the information should be about geographic phenomena. The role of the cartographer is to subdivide the broad landscape elements of geographic interest into smaller units more suitable for mapping, termed here *cartographic entities*. The cartographic entities can then be photographed, measured, sampled, or surveyed and then entered into computer mapping systems via *geocoding* methods to become *cartographic objects*.

Geocoding is the first stage of computer cartography and is done for two reasons. First, we can geocode with a specific mapping purpose in mind, for example, capturing lines that are going to be parts of polygons to be used to produce a choropleth map of Western Europe. Second, we can do general purpose geocoding; in other words, we can collect every piece of information we can about a specific area and assemble it for service in many possible cartographic contexts. The geocoding of general purpose cartographic data has great potential importance, just as accurate mapping of the country by topographic survey was important. The agencies performing this type of data collection and encoding have traditionally been government agencies, and we are fortunate that these agencies have often been at the forefront in making geocoding effective and efficient.

We can also geocode at two different levels. First, we can simply convert the graphic elements of a map into numbers so that we can reproduce the map using the methods of computer cartography. Many geographic databases consist of geocoded data in this form. Alternatively, we can encode important topological information about the data we are geocoding. Increasingly, as GISs become the principal users of geocoded data, the encoding of topology becomes essential to the use and survival of cartographic data sets.

To understand this distinction, a good example is the symbolization of roads and rivers. If the digital map consists of a road and a river that cross, a nontopological approach would be to plot the two lines one on top of the other. If the topology is encoded, we could recognize that either the road crosses the river on a bridge, in which case we draw a road symbol only and break the river symbol, perhaps adding the symbol for a bridge. Alternatively, the river could cross the road, implying an aqueduct or tunnel. Critical to this instance is that both lines, the river and the bridge, must be broken to include the feature at the intersection as part of their geometry. The importance of topological geocoding is considered later. First, however, we will return to the idea mentioned above, that effective geocoding comes from a clear understanding of the fundamental geographic properties and their manifestations.

4.2 CHARACTERISTICS OF GEOGRAPHIC DATA

If the purpose of geocoding is to encode the fundamental characteristics of geographic data digitally, then we must understand what those characteristics are before we start designing strategies for capturing them. The characteristics of geographic data are summarized in Figure 4.1.

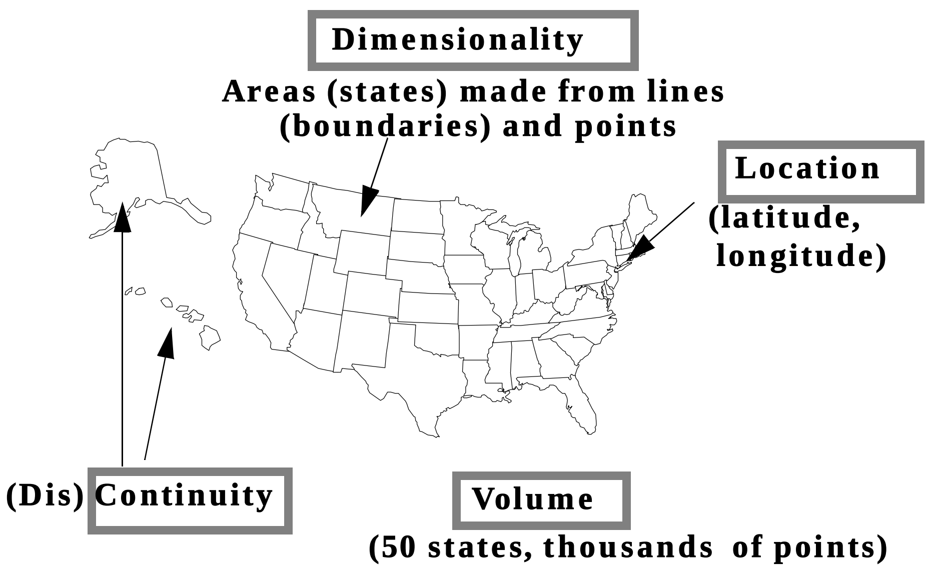


Figure 4.1 Some characteristics of geographic data.

4.2.1 Location

Fundamental to geographic data is the attribute of *location* on the earth's surface. Although we have to use a third dimension to describe elevation on the earth, the two dimensions of location on the plane or sphere are probably *the* basic geographical property. Normally, x and y values represent latitudes and longitudes, but often in geocoding we assume a map projection, such as a transverse Mercator, and we use coordinate systems such as Universal Transverse Mercator or State Plane to give locations. Occasionally, polar coordinates are used, giving an angle sometimes clockwise from north, and a distance.

Most coordinate systems for use with computer mapping are based on cartesian coordinates. This implies that the axes of the two directions, such as eastings and northings, are orthogonal, that is, they are at right angles to each other. This allows us to specify a location in space by referring to a pair of coordinates (x, y), or an easting and a northing. Leaving out one or both of these means that a location is simply undefined. Location is therefore the most fundamental characteristic of both cartographic and geographic data.

4.2.2 Data Volume

A second fundamental characteristic of geographic data is *high volume.* Computer science traditionally has dealt with small databases by cartographic standards. Cartographic and geographic databases contain thousands or sometimes millions of data elements, an amount that is directly related to the scale of the data and the extent of the map area.

A typical application in remote sensing uses seven one-byte bands of attributes and image arrays typically 512 by 512 in size. This represents a quarter of a million points for which we have multiple pieces of information all required for a single display. Many cartographic data processing problems are generic problems of large data sets. As a result, in computer cartography we often have to deal with memory constraints and the efficiency of our data structure.

Fortunately, the cost of storing data has decreased dramatically. Even on small computers, new storage methods have increased available memory from kilobytes to gigabytes within only a decade. The effect has been to change the emphasis from simple storage volume to storage access time as the primary volume-related consideration, although memory constraints will remain a factor.

4.2.3 Dimension

A third fundamental characteristic of geographic data is *dimensionality.* Traditionally, cartography has divided data into *points, lines,* and *areas.* Somewhat related is the concept of level of measurement. Levels of measurement are divided into *nominal, ordinal, interval,* and *ratio.* These two divisions have formed the basis of at least two classifications of mapping methods and cartographic data and are considered in more detail in Chapter 8.

4.2.4 Continuity

A fourth major characteristic of geographic data is *continuity*. Some map types, such as contour maps, assume a continuous distribution, while others, such as choropleth maps, assume a discontinuous distribution. Continuity is an important geographical property. The best example of a continuous variable is probably surface elevation. As we walk around on the earth's upper surface, we always have an elevation. There is no point where elevation is undetermined.

In real terrain, there are very few exceptions. Vertical overhangs and cliffs do indeed have areas where elevation is locally undetermined, but on the whole, elevation as a geographic distribution is continuous. Continuity does not always apply to statistical distributions. For example, tax rates are a discontinuous geographic variable. A resident of New York has to pay the state personal income tax, but by living just 1 meter inside Connecticut, a person will not pay that tax. In such an example, the tax rate is a discontinuous geographic variable because on the boundary line, the tax rate is undefined. It was once somewhat facetiously suggested that the only truly discontinuous geographic variables were tax rates and road surfaces, as anyone who has paid New York taxes or driven into the city of New York knows.

In addition, geographic continuity is an important property. Space classifications by areas must be exhaustive for continuity, that is, there should be no holes or unclassified areas. Similarly, for a set of categorical attributes reflecting a map, the set should contain all the objects found on the map, without any "other" or "miscellaneous" categories.

4.3 FUNDAMENTAL PROPERTIES OF GEOGRAPHIC OBJECTS

So far we have discussed the characteristics of geographic data. We should now consider some of the fundamental underlying properties that shape geographic and cartographic objects. Although characteristics of geographic data influence computer cartography, the implications of the fundamental properties are closer to the concerns of analytical cartography. A full understanding of the properties of geographic objects allows more effective geocoding, provides for the correct use of cartographic data structures, and facilitates the use of cartographic data transformations. The properties are illustrated in summary form in Figure 4.2.

4.3.1 Size

A basic underlying property is *size* and its characterization in measurement. Most geographic phenomena can be measured directly, for example by survey or air photo. A point has the measured aspects of location (x, y), adjacency, and elevation. A line has length, direction, connectivity, and "wigglyness." A polygon has topology (whether there are holes or outliers), area, shape, and boundary length, as well as location and orientation. A volume has topology, continuity, surface slope, surface aspect, surface trend, structure, location, and elevation. Most of these properties are comparatively simple to measure if the cartographic data are geocoded. Some are extremely difficult to measure in the real world and as such can only be analyzed using the data abstractions of mapping and rep-

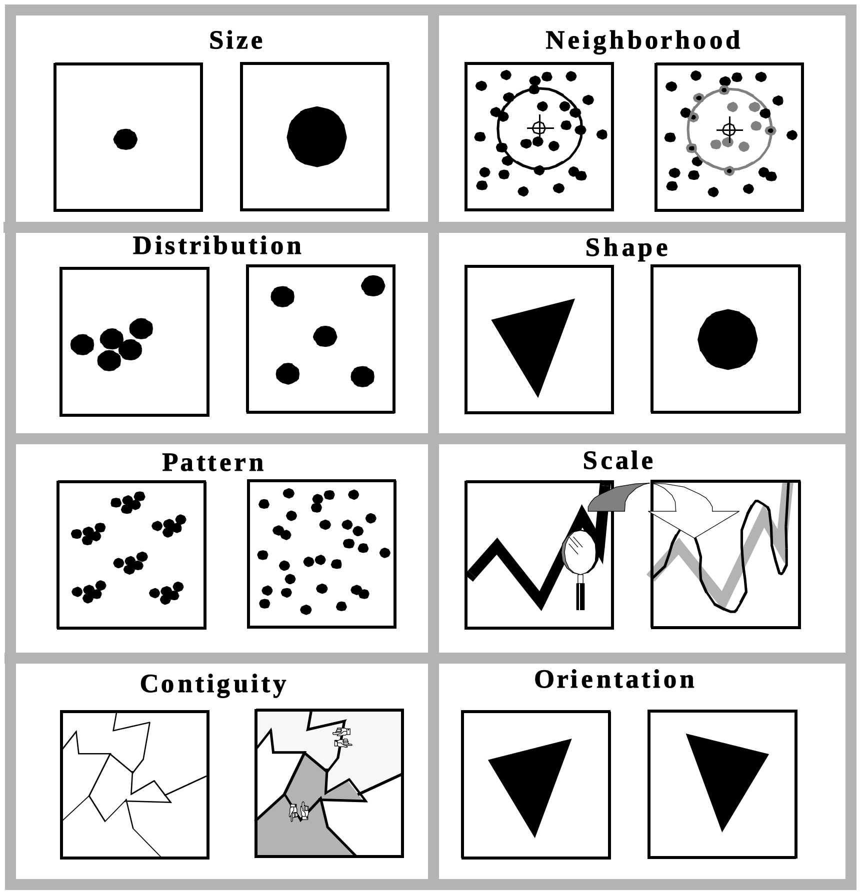


Figure 4.2 Fundamental properties of geographic objects.

resentation as cartographic objects. These measurements can usually be implemented by simple algorithms, some of which are given in the following chapters

4.3.2 Distribution

Another fundamental property is *distribution*. Density is a measure of the distribution of a phenomenon across space. Density can be computed by counting cartographic objects or their attributes over a set of geographic units, such as a grid or a set of regions. The density of a geographic phenomenon has a great many implications not only for how we measure and geocode it, but also how we can generalize it. This influences our decision on map coverage, and how we symbolize it on a map.

4.3.3 Pattern and Orientation

Another fundamental property of geographic objects is *pattern*. Pattern is actually a characteristic of distributions and is a description of their structure. Pattern can be thought of as a lack of randomness. A "first law of geography" is that anything that is of geographic interest lies at the intersection of between two and four maps, photos, or images. This relates to distribution via the sampling theorem.

The "second law of geography," with apologies to Tobler, is that everything is related to everything else, but *near* things are more related than others. In other words, if two things are close to each other, they are more likely to be similar than if they are separated by a long distance. This implies that proximity and adjacency imply a stronger interrelation between geographic objects, and that the strength of the relationship can be measured.

The simplest way in which this "nearness" phenomenon can be seen and measured is by repetition. Those relationships that repeat themselves at distances of less than half the size of the map result in patterns, implying that repetition is very important for pattern. For simple point distributions we can measure randomness, or lack of randomness, by using the nearest-neighbor statistic.

A test for pattern goes beyond simple measurement, however, because to observe pattern we need a model or description of the pattern we wish to find. In remote sensing and image processing, we can pass "templates" over a continuous distribution looking for a match of the discontinuous feature we wish to detect. This approach works even if the image is obscured by error or atmospheric effects. Patterns repeated through scales are also of interest and can be called "self-similar."

Probably the simplest pattern is direct repetition. Another basic property of objects distributed in space, however, is their *orientation*. In point distributions, elongated distributions or asymmetrical patterns show orientation, lines have obvious directions, areas are sometimes rotated and scaled equivalents of each other, and surfaces have dip and aspect, all implying orientation.

4.3.4 Neighborhood

The *neighborhood property* is, along with pattern, one of the most definitive of geographic properties. If pattern is the repetition of an attribute over space, the neighborhood property defines how the property varies over space. A key aspect of the neighborhood property is that variation takes place with distance, so little separations mean similarity, and big separations mean dissimilarity. Geographic studies often examine the relationship between some geographic phenomenon and distance.

Usually we see the neighborhood property as a distance function, a mathematical expression of the relationship between a geographic property and distance. Within geography, this distance function has been characterized and measured using tools such as the autocorrelation function, spatial interaction models, distance decay models, and the variogram. These functions express mathematically exactly what the second law of geography states.

4.3.5 Contiguity

Contiguity is another important geographic property. Contiguity is the property of being related by juxtaposition, that is by a sharing of a common boundary. Contiguity, therefore, is one of the geographic expressions of topology. The best example of contiguity is the sharing of a common boundary. Political geographers may be interested in the length of the border of Poland, perhaps even the geometric shape of Poland, but the main item of geographic interest would be a list of nations with a common boundary. Similarly, in land cover terms, we may be interested in the land uses that are most likely to be found around lakes. Whether the answer is beach cottages or swamps, we have a distribution of direct geographic interest

Contiguity is expressed in many ways. We can define it in terms of shared boundaries, in which case we can measure the number and lengths of shared boundaries. Within networks, contiguity is referred to as connectivity. For example, a network may contain connected links between nodes. In the sense of the network, the nodes are contiguous; that is, spatially you can get there from here in one step. Also, often we think of contiguity in terms of pixels or a grid structure. In fact, we do this so frequently that we even have special terms for it: four-cell and eight-cell contiguity (Figure 4.3). The center pixel for a neighborhood is called the kernel, and the contiguous pixels are those that share a direct common boundary or a corner.

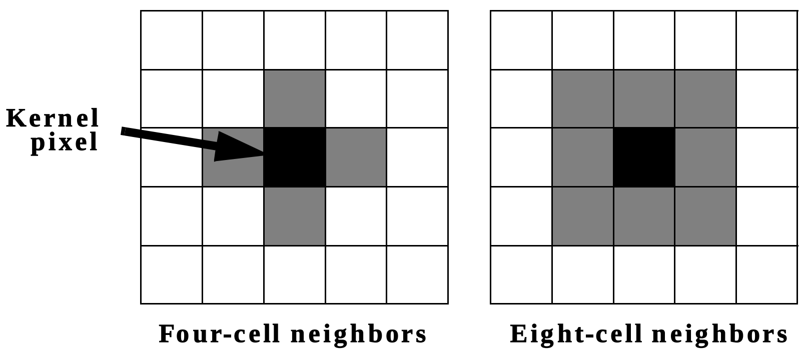


Figure 4.3 Contiguity of pixels.

4.3.6 Shape

Shape is another very geographic property. Shape is a difficult property to measure directly, so most shape measures are really measures of the level of correspondence between shapes. Some, however, are graphic with fewer dimensions. An excellent review of shape measurement is that by Pavlidis (1978). Scientists have measured the shapes of many phenomena. For example, geologists have measured the shape of Pacific atolls in the ocean, and biologists have measured the shape of cells and the wings of butterflies. Many geographers are interested in the shape of cartographic objects, such as the shapes of congressional districts or geomorphological features.

It is best to illustrate the property of shape by example. A particularly simple shape

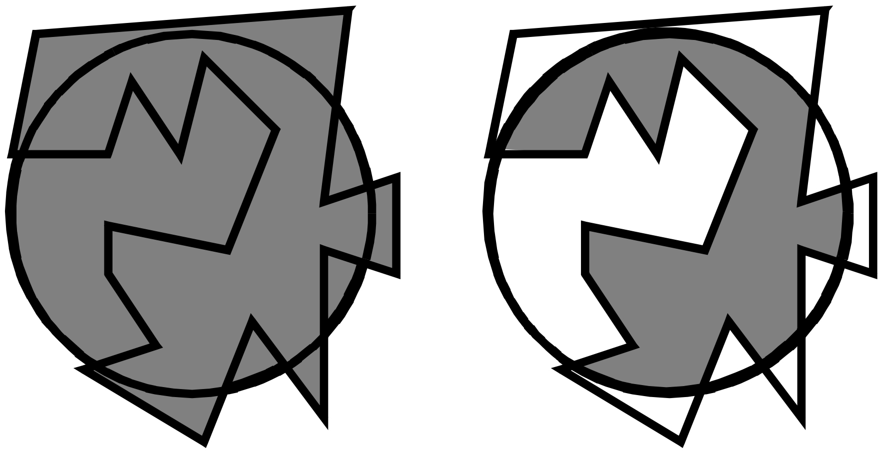


Figure 4.4 The Lee and Sallee shape measure.

measure is that of Lee and Sallee (1970), illustrated in Figure 4.4. This measure chooses a point somewhere inside the shape and draws a circle with the same area as the shape. The two figures are then overlain, forming three types of regions, an overlap or intersection, and then remaining parts of each of the two shapes. In set theory, the shape and the circle can be termed *A* and *B*. Then the Lee and Sallee shape measure is expressed as

$$s = 1 - \frac{A \cap B}{A \cup B}$$

Notice that for a circle, $A \cap B$ and $A \cup B$ are both equal to 1. Subtracting $\frac{1}{1}$ from 1 gives zero. So the base measurement is zero, and the shape measure compares shapes with a circle. Note, however, that the shape number depends on the area of the chosen "comparison shape," the point at which the two figures are overlain, and the orientation. The same measure could be used to compute the resemblance between any set of shapes, such as the resemblance of the remaining 49 state boundaries to the shape of New York. There are many other shape measures. Shape, however, is a very complex property. For example, the measure described above could give the same number to an infinite number of different shapes. An inversion of the method, for example, to perform recognition, would be impossible.

4.3.7 Scale

The final fundamental geographic property is *scale*. As far as cartography is concerned, the property of scale is the most distinctive feature of things that are geographic. The simplest expression of scale is the representative fraction, the ratio of distances on a map to the same distances in that part of the world shown on the map. There is a limited range of scales over which cartography has an interest in a phenomenon. If instead of thinking about things at a particular scale we think about things through scales, then one of two things can happen. Either objects become more clear at certain scales, or they never change with scale. These two subproperties are called scale dependence and scale independence.

If things do not change with scale, they can be modeled using fractal geometry. Fractal geometry has dimensions that are not whole numbers. If we go back to our original geographic characteristic of dimension, we can consider the point, line, area, and volume as special geometric cases. In fractal geometry, lines have dimensions between 1.0 and 2.0, and areas and surfaces have volumes between 2.0 and 3.0. For example, a line on a map is represented by a geometrically infinitely thin cartographic object, yet it can only be symbolized by covering an area on the map with color. Scale dependence means that there is an appropriate scale for making a map. Analytical cartography can contribute by determining this scale objectively.

4.3.8 Measurement and Fundamental Properties

The purpose of geocoding from an analytical cartographic standpoint is to encode digitally the fundamental properties of geographic entities with the objects so that their later analysis is possible. Encoding these properties, usually as geographic relationships between objects, can sometimes be performed during the direct data capture process, but this encoding is most often captured by processing geocoded data using computer algorithms. Examples of quantitative measures used to represent each of the fundamental properties are summarized in Figure 4.5. Many of these measures and the algorithms for their computation require, in addition to explicitly locational data, the geocoding or the computation of the topological relationships between objects.

4.4 GOALS OF GEOCODING METHODS

Geocoding seeks many conflicting goals. Each should be kept in mind before geocoding cartographic data, and the computer cartographer should be realistic about which goals are at odds with each other.

4.4.1 Minimize Labor Input

Because one of the major sources of error in the geocoding process is human error, and because labor costs are usually high, especially for semiautomatic digitizing, an important goal for geocoding is to minimize the amount of manual labor involved. The manual component for converting existing paper maps is high and can multiply unexpectedly. Chrisman (1987) reported manual digitizing times of eight hours per 300 polygon soil map sheet, with an additional four hours of editing, even using software designed explic-

	POINT	LINE	AREA	VOLUME
SHAPE	Feature type	Curvature	Shape measure	Dimension Resemblance to figure (e.g. cone)
CONTIGUITY	Link	Intersection	Shared boundary	Shared face
ORIENTATION	Of cluster or pattern	Bearing Trend	Of axis Of pattern	Dip, drift, trend, aspect
SIZE	Number	Length	Area	Volume Surface area
SCALE	Range at which object is a point	Range at which object is a Line	Range at which object is an area	Range at which object is a volume
NEIGHBORHOOD	Set of nearby points	Connected lines. Lines within a range	Contiguous areas Area within a range Connected areas	Adjacent voxels Overlapping volumes Shared faces
PATTERN	Pattern matching Fourier analysis	Curve measures Fractal dimension	Shape distribution Description	Fourier power spectrum Trend surface
DISTRIBUTION	Standard distance Nearest neighbor number Autocorrelation	Line density Length, Intersection frequency	Coverage Autocorrelation	Variogram

Figure 4.5 Quantitative descriptors of the fundamental properties of geographic objects.

itly for the reduction and detection of digitizing errors, with automatic topological correction and automatic end-node snapping. Chrisman pointed out that careful attention to the editing capabilities of digitizing software, coupled with consistency checking, can allow even a low-cost microcomputer workstation to produce high-quality, accurate, digital cartographic data.

For example, we may seek a digital land-cover map at 1:500,000. One approach would be to draw a grid with squares of 20 mm, representing 5 km on the final grid after scaling, over four 1:250,000 land-use and land-cover maps from the U.S. Geological Survey. We could then go grid cell by grid cell writing down the land-use category at each intersection. We could then type the land-use category numbers one by one into a computer file. Assuming that we can draw the grids and write down the numbers at the rate of 200 per hour, and can enter the data at the same rate, we would need about 400 hours, or 10 weeks of full-time labor, to geocode the map, without any checking.

Although in actuality geocoders become faster at a task with experience, anything we can do that makes data entry easier, we should do. In many cases, simple steps can be taken. Digitizing tablets that are adequately lit, and are repositionable, especially so that the user can sit down, can save substantial amounts of labor time, because much work is necessary in going back over errors and in cross checking. Immediate video and audible feedback for errors can save starting over after a string of errors, as also can direct editing, off-line control, and a quiet, nondistracting workplace.

More substantial steps can also be taken to reduce labor. Many steps are totally unnecessary, such as the writing down on paper in the example above. If in doubt, the best way to estimate the time for each step of the digitizing process is to conduct a small test on a pilot data set, taking it through the entire procedure. The time can then be multiplied up to get overall estimates.

4.4.2 Detect and Eliminate Errors

Many early geocoding systems had only limited editing capabilities. They allowed data entry, but error detection was by batch processing and correction was by deletion and reentry. Anything we can do in the geocoding process that reduces errors, or that makes errors easily detectable, we should indeed do. As an absolute minimum, data for lines and areas can be processed automatically for topological consistency, and any unconnected lines or unclosed polygons can be detected and signaled to the user.

The easiest way to avoid errors in geocoding is to make sure that they are detected as soon as possible and then to make their correction easy. Video display during digitizing and audio feedback for error messages are essential. Software should spell out exactly what will happen in the case of an error. A common geocoding error is to overflow a hard or floppy disk while digitizing. Some software continues to accept data as if nothing is wrong, and gives a “disk full” message only when you exit from the digitizer software.

Some easy-to-detect errors are slivers, spikes, inversions, and disappearing nodes (Figure 4.6). Scaling and inversion errors are usually due to an incorrect digitizer set-up procedure, that is, they are systematic errors caused by incorrectly entering the control points for establishing the map geometry. Spikes are random hardware or software errors in which a zero or extremely large data value erroneously replaces the real value in one of the coordinates. Spikes are also sometimes known as zingers. Errors in topology, missing or duplicate lines, and unsnapped nodes are operator errors (Ward and Phillips, 1987).

Errors which are more difficult to detect take more effort. Some rules of thumb for the most difficult errors are as follows:

- If something looks wrong on the display, it is probably wrong in the data.

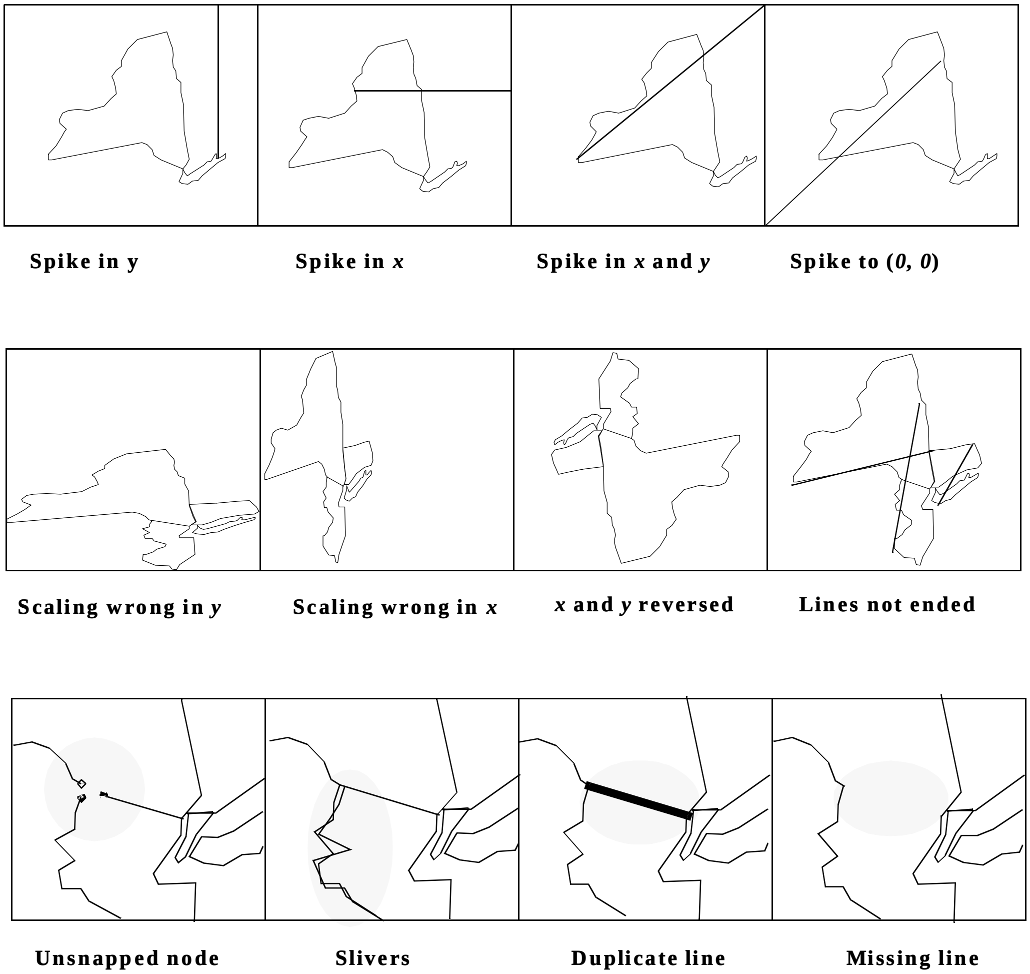


Figure 4.6 Some easy-to-detect geocoding errors.

- Believe the locational data files before the display.
- Believe the source map before your version of it.
- Stopping or having another person work on geocoding often is a source of errors.
- Postponing a correction gives the correction only a 50% chance of ever being made unless postprocessing for errors is available.
- Sloppy geocoding makes unreliable and erroneous maps.

In the Spatial Data Transfer Standards, a distinction is made between positional and attribute accuracy and logical consistency. Positional and attribute accuracy can be tested and measured by direct comparison with the source. Very useful aids in detecting errors are overlay plots at the same size and scale as the source map, especially if they can be superimposed on the original. There is no substitute, however, for systematically parsing through the geocoded data looking for discrepancies. Often, plotting the data becomes a useful aid because unplottable data often have bad geocodes. Similarly, attempting to fill polygons with color often detects gaps and slivers not visible in busy polygon networks. The best check for positional accuracy is a check against an independent source map of higher accuracy.

A data set that is correctly geocoded both positionally and with attributes is not necessarily logically consistent. Logical consistency can be checked most easily for topological data. Topologically, data can be checked to see that all chains intersect at nodes, that chains cycle correctly in a ring around polygons, and that inner rings are fully enclosed within their surrounding polygons. Otherwise, attributes can be checked to ensure that they fall within the correct range and that no feature has become too small to be represented accurately.

4.4.3 Optimize Storage Efficiency

Data contains two parts: volume and information. Geocoding often seeks to eliminate volume while retaining information. We need to know what parts of the data are redundant, what can be left out, and what must be kept.

The ideal storage method for geocoded data saves only the minimum, the distilled information. Redundant information, however, is often a necessary part of geocoding. A straight state boundary on one map projection may become curved on another and would need many points to make it look curved, many more than the “optimal” two necessary for a straight line. In addition, storage efficiency depends entirely on which data structure we will be using to map the digital data or how we will move the raw geocoded information between systems or applications.

4.4.4 Maximize Flexibility

A critical goal is to optimize flexibility. The most elegant data structure in the world is worthless if the author is the only one who can read it or if wants to read it because it is so obscure. We need data structures and storage methods that allow analyses we did not anticipate when we geocoded the data originally. These unanticipated data uses are not the exception, they are the norm. This flexibility is the very aspect of computer cartography that has allowed the development of geographic information systems.

As soon as data exists in digital form new uses for the data will be discovered, each of which will place a new set of demands on the chosen data structure. A measure of the flexibility of a data structure is how well it holds up to these unanticipated demands, rather than the demands of the original uses for which it was designed. Geocoding, therefore, should not impose upon the data restrictions on accuracy, precision, and reliability.

4.5 LOCATIONAL GEOCODES

Geocoding uses a map coordinate system. Coordinate systems can be standard or arbitrary. There are many reasons to avoid arbitrary referencing systems, among them incompatibility, lack of ability to document the alternative systems, inability to adapt to new levels of precision, or the assumption of a specific map projection. An important factor in choosing a coordinate system with which to geocode a map is its universality. A good coordinate system works worldwide; it is simple, accurate, precise, terse, and adaptable. Unfortunately, the problems of global coordinate systems are much like those of map projections; that is, not all goals can be served at once. As a result, we have several established systems for specifying locations on the earth's surface.

4.5.1 Geographic Coordinates

Many global data-bases record locations using latitude and longitude or geographic coordinates. Latitude and longitude are almost always geocoded in one of two ways. Latitudes go from 90 degrees south −90 to 90 degrees north (+ 90). Precision below a degree is geocoded as minutes and seconds, and decimals of seconds, in one of two formats: either plus or minus DD.MMSS.XX, where DD are degrees, MM are minutes, and SS.XX are decimal seconds; or alternatively, as DD.XXXX, or decimal degrees. Longitudes are the same, with the exception that the range is −180 to +180 degrees. In the second format, degrees are converted to radians and stored as floating point numbers with decimal places.

It is especially important that we record how each map has been geocoded if geographic coordinates are used. Maps look particularly strange when decimal degrees are taken for the degree-minute-second format. The relatively open exchange of geocoded data makes this recording of the system even more important. The simple rectangular projection, in which latitude and longitude are simply drawn without projection at all, may get the map done, but denies a cartographic tradition going back over 2,000 years and may give incorrect results if used for computations.

4.5.2 The UTM Coordinate System

If we use the ability to georeference other planets as a measure of success, a successful geocoding system for cartography is the Universal Transverse Mercator (UTM) coordinate system. The equatorial Mercator projection, which distorts areas so much at the poles, nevertheless produces minimal distortion along the equator. Lambert modified the Mercator projection into its transverse form in 1772, in which the "equator" instead runs north-south. The effect is to minimize distortion in a narrow strip running from pole to pole.

UTM capitalizes on this fact by dividing the earth up into pole-to-pole strips or zones, each 6 degrees of longitude wide, running from pole to pole. The first zone starts at 180

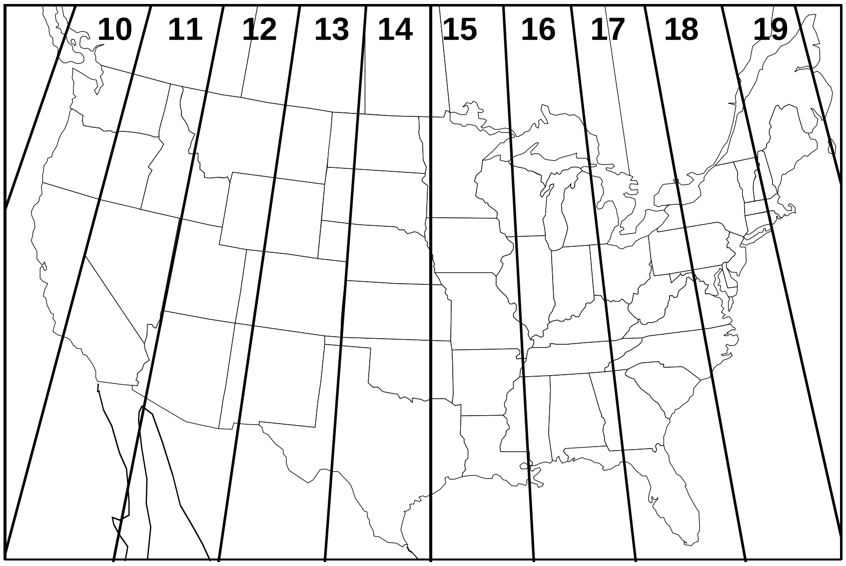


Figure 4.7 Universal Transverse Mercator zones in the 48 contiguous states.

degrees west (or east), at the international date line, and runs east, that is, from 180 degrees west to 174 degrees west. The final zone, zone 60, starts at 174 degrees east and extends east to the date line. The zones therefore increase in number from west to east. For the United States, California falls into zones 10 and 11, while Maine falls into zone 19 (Figure 4.7). Within each zone we draw a transverse Mercator projection centered on the middle of the zone. Thus for zone 1, with longitudes ranging from 180 degrees west to 174 degrees west, the central meridian for the transverse Mercator projection is 177 degrees west. Because the equator meets the central meridian of the system at right angles, we use this point to orient the grid system (Figure 4.8). In reality, the central meridian is sometimes set to a map scale of slightly less than one, making the projection for each zone secant along two lines at true scale parallel to the central meridian.

Two forms of the UTM system are in common use. The first, used for civilian applications, sets up a single grid for each zone. To establish an origin for the zone, we work separately for the two hemispheres. For the southern hemisphere, the zero northing is the South Pole, and we give northings in meters north of this reference point. Fortunately, the meter was originally defined as one–ten millionth of the distance from the pole to the equator, actually measured on the meridian passing through Paris.

Although the distance varies according to which meridian is measured, the value 10 million is sufficient for most cartographic applications. Although the meter has been redefined in a more precise way, the student may wish to compare the utility of its origin with the origin of the foot, first standardized as one third of the distance from King Henry I's nose to the tip of his fingers.

The numbering of northings starts again at the equator, which is either 10 million meters north in southern hemisphere coordinates or 0 meters north in northern hemi-

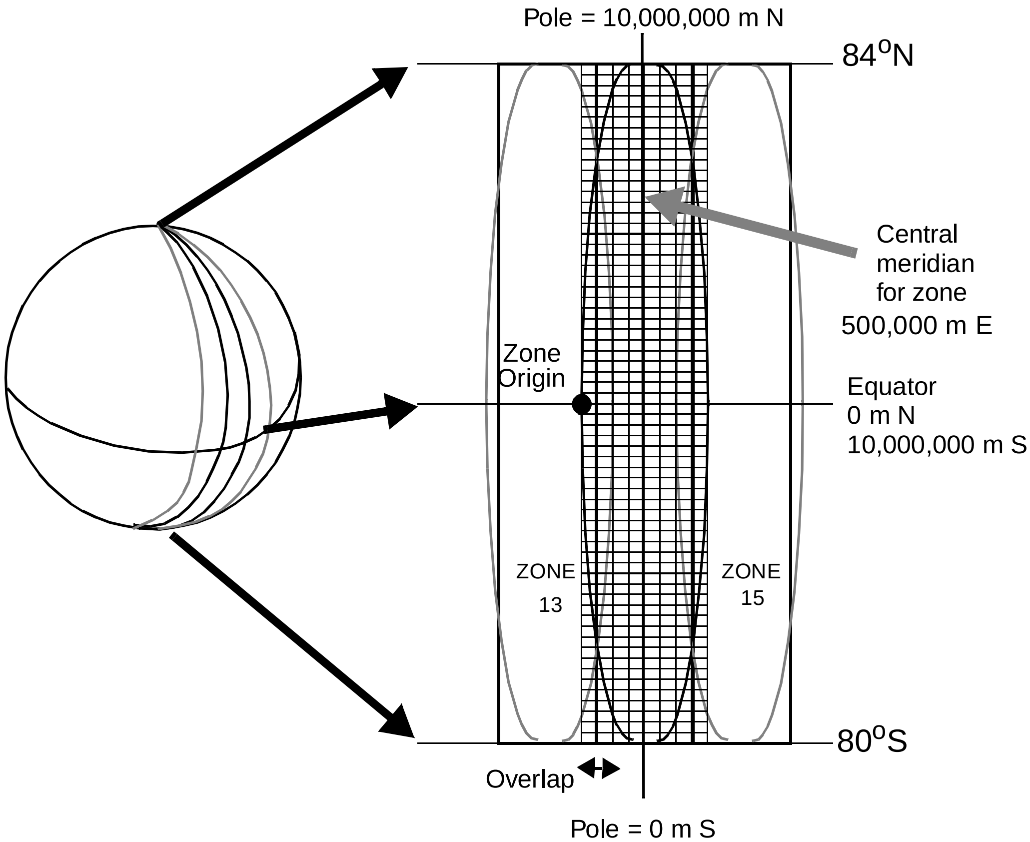


Figure 4.8 The Universal Transverse Mercator coordinate system.

sphere coordinates. Northings then increase to 10 million at the North Pole. Note that as we approach the poles, the distortions of the latitude-longitude grid drift farther and farther from the UTM grid. It is customary, therefore, to use the UTM system neither beyond the land limits of North America, nor for the continent of Antarctica. This means that the limits are 84 degrees north and 80 degrees south. For the polar regions, the Universal Polar Stereographic coordinate system is used.

For eastings a false origin is established beyond the westerly limit of each zone. The actual distance is about half a degree, but the numbering is chosen so that the central meridian has an easting of 500,000 meters. This has the dual advantage of allowing overlap between zones for mapping purposes and of giving all eastings positive numbers. We can tell from our easting if we are east or west of the central meridian, and so the relationship between true north and grid north at any point is known. To give a specific example, Hunter College is located at UTM coordinate 4,513,410 meters north; 587,310 meters east; zone 18, northern hemisphere. This tells us that we are about four-tenths of the way up from the equator to the North Pole, and are west of the central meridian for our zone, which is centered on 75 degrees west of Greenwich. On a map showing Hunter College, UTM grid north would therefore appear to be east of true north.

For geocoding with the UTM system, 16 digits is enough to store the location to a precision of 1 meter, with one digit restricted to a binary (northern or southern hemisphere) and the first digit of the zone restricted to 0 to 6 (60 is the largest zone number). This coordinate system has two real cartographic advantages. First, geometric computations can be performed on geographic data as if they were located not on the surface of a sphere but on a plane. Over small distances, the errors in doing so are minimal, although it should be noted that area computations over large regions are especially cartographically dangerous. Distances and bearings can similarly be computed over small areas.

The second advantage is that the level of precision can be adapted to the application. For many purposes, especially at small scales, the last UTM digit can be dropped, decreasing the resolution to 10 meters. This strategy is often used at scales of 1:250,000 and smaller. Similarly, submeter resolution can be added simply by using decimals in the eastings and northings. In practice, few applications except for precision surveying and geodesy need precision of less than 1 meter, although it is often used to prevent computer rounding error.

4.5.3 The Military Grid Coordinate System

The second form of the UTM coordinate system is the military grid, adopted for use by the U.S. Army (Department of the Army, 1983) and many other organizations. The military grid uses a lettering system to reduce the number of digits needed to isolate a location, since letters can be spelled out using simple words over radio broadcasts. Zones are numbered as before, from 1 to 60 west to east. Within zones, however, 8 degree strips of latitude are lettered from C (80 to 72 degrees south) to X (72 to 84 degrees north: an extended-width strip). The letter designations A, B, Y, and Z are reserved for Universal Polar Stereographic designations. A single rectangle, 6 by 8 degrees, generally falls within about a 1,000 kilometer square on the ground. These grids are referenced by numbers and letters; for example, Hunter College falls into grid cell 18T (Figure 4.9).

Each grid cell is then further subdivided into squares 100,000 meters on a side. Each cell is assigned two additional letter identifiers (Figure 4.10). In the east-west (x) direction, the 100,000-meter squares are lettered starting with A, up to Z, and then repeating around the world, with the exception that the letters I and O are excluded, because they could be confused with numbers. Thus the first column, A, is 100,000 meters wide and starts at 180 degrees west. The alphabet recycles about every 18 degrees and includes about six full-width columns per UTM zone. Several partial columns are given designations nevertheless, so that overlap is possible, and some disappear as the poles are approached.

In the north-south (y) direction, the letters A through V are used (again omitting I and O), starting at the equator and increasing north, and again cycling through the letters as needed. The reverse sequence, starting at V and cycling backwards to A, then back to V, and so forth, is used for the southern hemisphere. Thus a single 100,000-meter grid square can be isolated using a sequence such as 18TWC. Within this area, successively

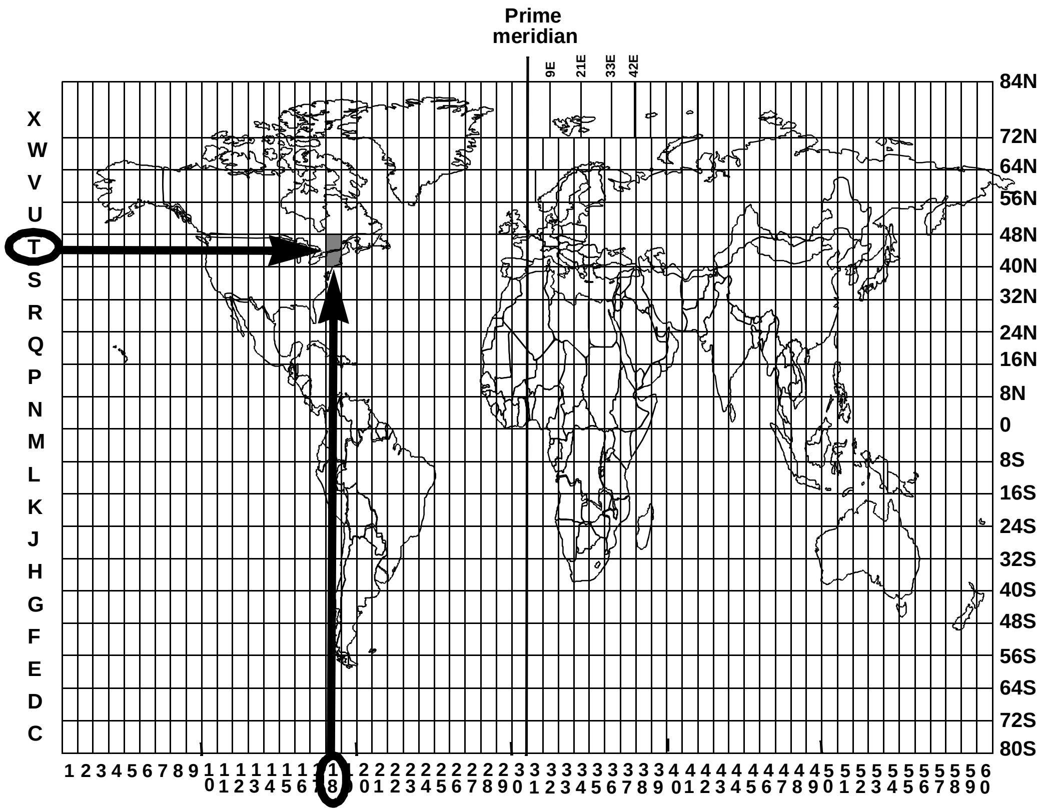


Figure 4.9 Six-by-eight degree cells on the UTM military grid.

accurate locations can be given by more and more pairs of *x* and *y* digits. For example, 18TWC 81 isolates a 10,000-meter square, 18TWC 8713 a 1,000 meter square, and 18TWC 873134 a 100-meter square. These numbers are frequently stored without the global cell designation, especially for small countries or limited areas of interest. Thus WC873134, two letters and six numbers, would give a location to within 100-meter ground accuracy.

The Universal Polar Stereographic (UPS) coordinate system, also part of the military grid, is based on a stereographic map projection centered on each on the poles. The two projections are centered so that the western hemisphere is to the left. For the north polar region, the western zone is designated Y, and the eastern Z, each extending from 84 to 90 degrees north. The prime meridian is used as the right angle for the grid, and 100,000-meter grid cells are simply lettered from A to P running from the bottom (the prime meridian) to the top.

The eastings for the letter designations Y and Z are chosen so that the first column left of the pole is Z and the first to the right is A. For cell Y, R is the first column, while for cell Z, J is the last. For the South Pole, the situation is identical, except that the cells are inverted, that is, the Greenwich meridian is to the top. The cells are designated A to the

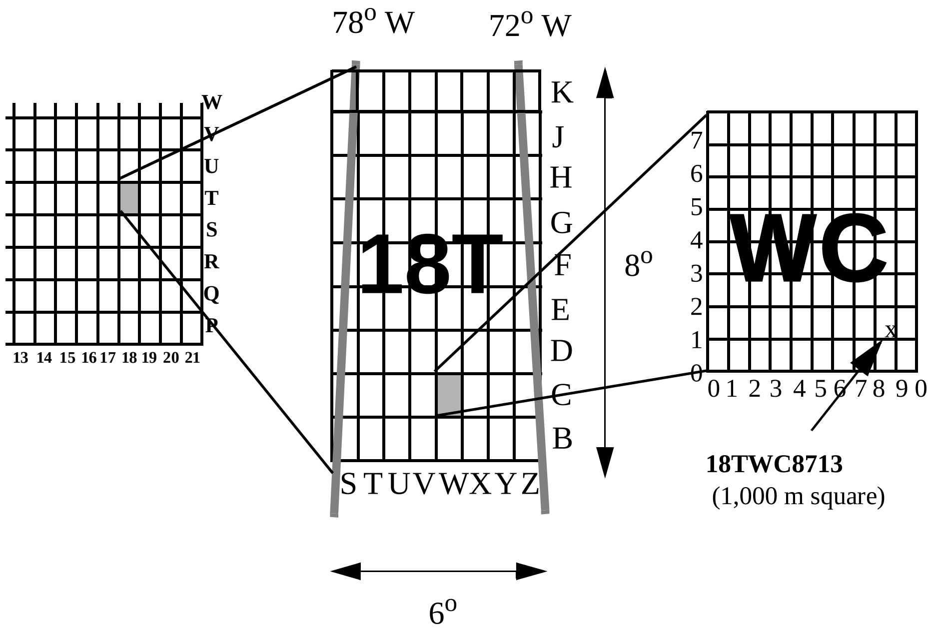


Figure 4.10 Military grid cell letters.

west and B to the east, and since the zone is larger (80 to 90 degrees south), the letters go from A to Z as northings and J to Z and A to R as eastings, respectively (Figure 4.11).

4.5.4 The State Plane Coordinate System

Much georeferencing in the United States uses a system called the State Plane Coordinate System (SPCS). The SPCS is based on feet and has been used for decades to write legal descriptions of properties and engineering projects. Legal documents are probably the least modifiable descriptions on Earth. The SPCS is based upon a different map of each state, except Alaska. States that are elongated north to south, such as California, are drawn on a Lambert Conformal Conic projection. States that are elongated east to west, such as New York, are drawn on a transverse Mercator projection, because the zones are divided into north-south strips. The state is then divided up into zones, the number of which varies from small states, such as Rhode Island with one to as many as five. Some zones have no apparent logic; for example, the state of California has one zone that consists of Los Angeles County alone. Some have more logic, so, for example, Long Island has its own zone for the state of New York. Because there are so many projections to cover the land area, generally the distortion attributable to the map projection is very small, much less than in UTM, where it can approach 1 part in 2,000.

Each zone then has an arbitrarily determined origin that is usually some given number of feet west and south of the southwestern-most point on the map. This again means that

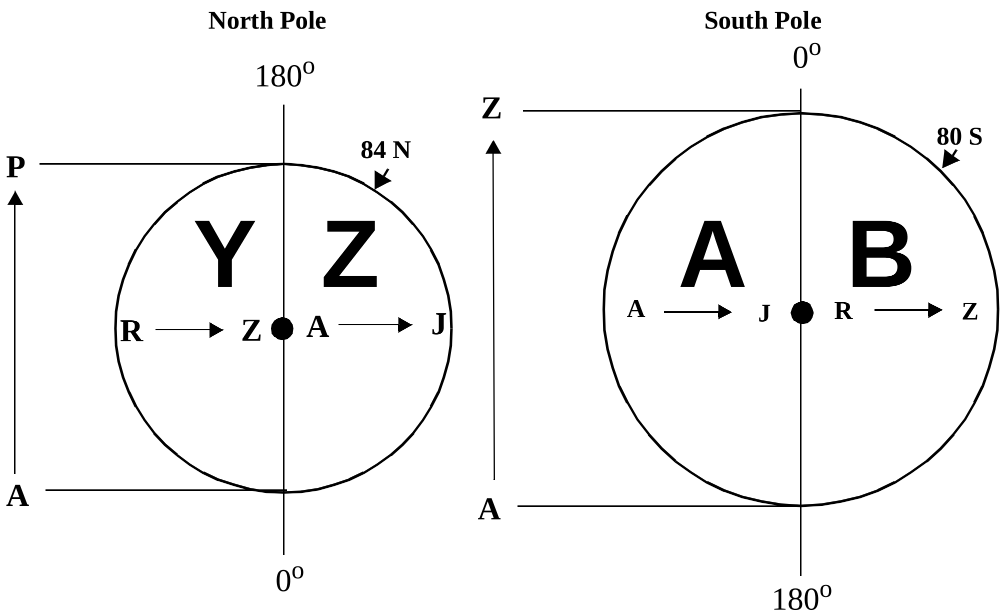


Figure 4.11 100,000 meter cells on the UPS grid.

the eastings and northings all come out as positive numbers. The system then simply gives eastings and northings in feet, often ending up with millions of feet, with no rounding up to miles. The system is slightly more precise than UTM because coordinates are to within a foot rather than a meter, and it can be more accurate over small areas. A disadvantage is the lack of universality. Imagine mapping an area covering the boundary between not only two zones, but two states. This means that you could be working with data that fall into four coordinate systems on two projections. Calculating areas on that basis becomes a set of special case solutions. On the other hand, SPCS is used universally by surveyors all over the United States.

4.5.5 Other Systems

There are, in addition, many other georeferencing systems. Most countries have their own, although many use UTM or the military grid. The National Grid of the United Kingdom uses the lettering system of the Military Grid. In a few cases, particularly Sweden, the national census and other data are directly tied into the coordinates. Within the United States, many private companies and public services use unique systems, usually tied to specific functions such as power lines, or a specific region such as a municipality, or even a single construction project. When using a georeferencing system for geocoding, we should be sure to remain consistent within that system and to record the relationship between the system and latitude and longitude or some other recognized system. Also, we should be sure to use precision and numbers of significant figures that make sense. Can we really measure distances over entire states down to the micrometer or less? And even if we can, is this storage efficient? On the other hand, there is also a tendency to throw away precision needlessly. If we round all cartographic data up to 78 meters, a convenient number determined by an arbitrary orbiting satellite system, how will we deal with newer, higher-resolution systems?

In summary, the universality of the UTM system makes it attractive for world-scale geocoding, while for small areas other systems may be more accurate. Care should be taken to be consistent and accurate, and to use an appropriate level of precision. Fortunately, computer software for mapping allows data to be used from more than one referencing system.

4.6 GEOCODING METHODS

Historically, many different means have been used to geocode. At first, some computer cartographic software actually required maps to be encoded and entered by hand. The hours of monotonous work required for this task made errors common and their correction difficult. Since special purpose digitizing hardware became available, and especially since the cost of this hardware fell substantially, virtually all geocoding has been performed by computer.

4.6.1 Semiautomated Digitizing

There are two distinct technologies for geocoding. The first assists a person in performing the digitizing task; the second performs the task in a completely automated fashion. The first type, *semiautomated digitizing* involves the use of a digitizer or digitizing tablet (Figure 2.2). This technology has developed as computer mapping and computer-aided design have grown and placed new demands on computer hardware.

The digitizing tablet is a digital and electronic equivalent of the drafting table. The major components are a flat surface, onto which a map is usually taped, and a stylus or cursor, which has the capability of signaling to a computer that a point has been selected. As we saw in Chapter 2, the mechanism to capture the location of the point differs. Many systems have connected arms, but most have embedded active wires in the tablet surface which receive an electrical impulse sent by a coil in the cursor. In some rare cases, the cursor transmits a sound, which is picked up and recorded by an array of microphones.

The actual process of digitizing a map proceeds as follows (Figure 4.12). First, the paper map is tailored or preprocessed. If the map is multiple sheets, then the separate sheets should be digitized independently and digitally merged (zipped) later. The next major step, unless annotations have to be made onto the map to assist the geocoding, is to derive a coordinate system for the map. Most applications use UTM, the Military Grid, or latitude and longitude, but many cartographers ignore these standard systems and use hardware coordinates or map inches or millimeters. Map units are sometimes used when precise matching between the digitized map and its source is required. The map is then transformed into geographic coordinates when the editing and proofing is complete.

As a minimum, the coordinate locations of three points are required, usually the upper right easting and northing and the lower left easting and northing, and one other corner. From these points, with their given (user or world) coordinates, and their raw digitizer coordinates, all the parameters can be computed for affine transformations.

This means that the orientation of the map on the tablet need not be exact as is required of some digitizing packages. Many software packages require four of these control points for computing the affine transformation parameters, and it is advisable to repeat digitize points and average coordinates to achieve higher accuracy (see Chapter 9).

The beginning of the digitizing sequence involves selecting the control points and interactively entering their world coordinates. This is a very important step, because an error at this stage would lead to a complex systematic error in every pair of coordinates. After the map is taped to the tablet, it should not be moved without reregistration, and it is preferable to perform the registration only once per map. Ideally, the entire digitizing process should be finished at one sitting, although this is often impossible.

Tape should be placed at each map corner after smoothing the map, and care should be taken to deal with folds and the crinkles that develop during periods of high humidity with certain papers. A stable base product such as Mylar is preferable for digitizing. The lower edge, which will have the cursor and your right sleeve (if you are right-handed) dragged over it many times, should be taped over its entire length. Always permanently record the x and y values of the map control points, ideally digitally and with the geocoded data set. This may allow later recovery of lost resolution or systematic errors.

Digitizing then proceeds with the selection of points. The cursor may have multiple buttons and may be capable of entering text and data without using the keyboard. Voice data entry and commands are also sometimes used. On specialized workstations, there may even be a second tablet with its own mouse or cursor for commands. Errors can be reduced during this process by reading the documentation in advance and by occasionally stopping to review the actual data being generated on the screen.

Points are usually entered one at a time, with a pause after each to enter attributes such as labels or elevations. Lines are entered as strings of points and must be terminated with an end-of-chain signal to determine which point forms the node at the end of the chain. This signal must come from the cursor in some way, either by digitizing a point on a preset menu-area or by hitting a preset key. Unless the chains are to be software-processed for topology, chain-linkage information may need to be entered also, or a point within each polygon digitized and tagged.

Polygons are usually digitized as chains, although sometimes an automatic closure for the last point (snapping) can be performed. Finally, the points should be checked and edited. The digitizing software may contain editing features, such as delete and add a chain or move and snap a node. The software may also support multiple collection modes. *Point mode* simply digitizes one point each time the button on the cursor is pressed. *Stream mode* generates points automatically as the cursor is moved, either one point per unit of time or distance. This mode can easily generate very large data volumes and should be avoided in most cases. Error correction is especially difficult in this mode. *Point select mode* allows switching between point and stream mode. This mode is sometimes used when lines are both geometric and natural, such as when following a straight road and then a river.

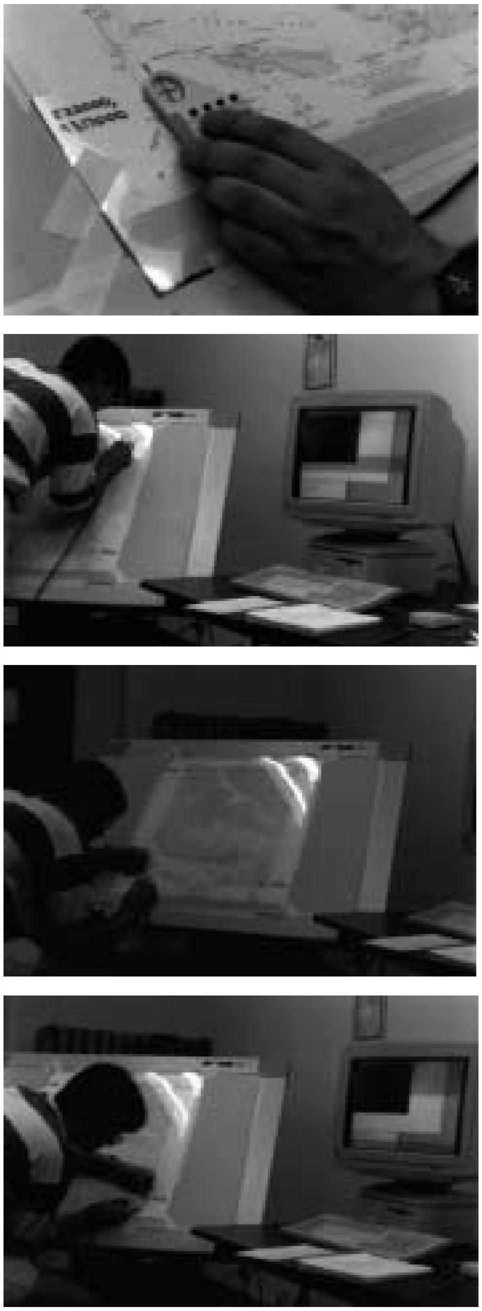
Stable base copy of map is prepared by choosing digitizing control points at known locations. Any features to be selected should be marked in advance.

Control points and window limits to be used are marked clearly for use and reuse. Cursor cross hairs should be used at 45 degrees to grid, which should be as fine as possible. Repeat digitizing of controls is advisable for precision.

Control point coordinates are labeled. These should be checked at least twice during tablet setup. Systematic errors will result otherwise. These points can also be used to register test plots.

Map is firmly taped or fixed down to tablet. No movement of the map should be possible. Surface should be flat and free of folds, bubbles, and so forth. Double tape over edges that will be rubbed by elbows and forearms.

Figure 4.12a The semiautomated digitizing process (Part 1).



Control points to be used in the affine transformation are entered one at a time, along with their map coordinates in geocoded coordinate space such as latitude and longitude or UTM meters, or in map inches or millimeters if the affine transformation is to be performed after map editing.

Control point at upper right is entered.

Control point at lower left is entered.

Third control point at lower right is orthogonal, allowing computation of map rotation angle, scaling, and location of tablet versus map original. Map setup is now complete. The setup should be tested by digitizing several points and checking the display and the computed coordinates.

Figure 4.12b The semiautomated digitizing process (Part 2).

Digitizing begins. Map features are traced out using the cursor. Care is taken to capture features accurately, with a suitable level of detail. Points can be selected one at a time or in a stream turned on and off from the cursor. Attributes can be entered as features are completed.

At various stages, and at completion, the map is plotted from software. The plot should be at the exact same scale as the original, allowing overlay. Features should be edited, deleted, or added as appropriate.

The edit plot should be compared with the original, and any necessary final edits should be made.

Final map should be saved carefully. The user can now move ahead to process topology or any other necessary stage before data can be used.

Figure 4.12c The semiautomated digitizing process (Part 3).

At this point the data are ready either for direct integration into the computer mapping software or are ready to be used as input for cartographic transformations to change data structure, map base, or scale. As a human-machine interactive process, the digitizing process has not frequently been studied in detail by cartographers. The process is important to understand, because most of the errors in digitizing can be reduced or eliminated using some simple ergonomic principles (Jenks, 1981).

4.6.2 Automated Digitizing

The second digitizing process is *automated digitizing,* of which there are many types, and, indeed, it is rapidly broadening in scope as a means of data capture. The earliest device for automatic digitizing was the scanner, which receives a sheet map, sometimes clamped to a rotating drum, and which scans the map with very fine increments of distance measuring the radiance, or sometimes transmission of the map when it is illuminated, either with a spot light-source or a laser (Figure 4.13). The finer the resolution, the higher the cost and the larger the data sets. A major difference with this type of digitizing is that lines, features, text, and so forth are scanned at their actual width and must be preprocessed for the computer to recognize specific cartographic objects. Some plotters can double as scanners, and vice versa.

For scanning, maps should be clean and free of folds and marks. Usually, the scanned maps are not the paper products but the film negatives, mylar separations, or the scribed materials that were used in the map production. An alternative scanner is the automatic line follower, a scanner that is manually moved to a line and then left to follow the line automatically. The Altek Apache digitizing cursor is halfway between this and semiautomated digitizing, because it is a manual cursor that has a small scanning window on the cursor itself. Automatic line followers are used primarily for continuous lines, such as contours. These and other scanners are very useful in CADD systems, where input from engineering drawings and sketches is common.

Increasingly, video scanners are becoming important geocoding devices. These scanners are simply television cameras, sometimes with highly enhanced resolution, that can be mounted on a stand and pointed at a stationary image, air photo, or map. Early versions were monochrome only and had limited gray levels. More recently, color versions with many gray levels, and even color look-up tables, have been used. Map separations can be scanned and entered as separate data layers, and data can be sent directly to a microcomputer with local editing, storage, and even image-processing capabilities. Even complex maps such as topographic quadrangles can be scanned by these devices with suitable results, although for a whole quadrangle multiple scans must be used for a reasonable resolution. Again it should be noted that the scanner sees folds, pencil lines, erasures, correction fluid, and coffee stains as easily as cartographic entities. Great care should be taken with scanning not to introduce complex geometric relationships between the map and the image, such as the effects of using different lenses.

Finally, low cost scanners are now available that can read and interpret both documents and text. This is important, because typed and other text can be entered directly from a map and then manipulated and plotted within a mapping system. Simple graphics

Figure 4.13 The automated digitizing process (scanning).
(Photo by the U.S. Geological Survey. Used with permission.)

scanning is rarely adequate for cartographic purposes, but can be used to put a rough sketch into a CADD system for reworking. In this way, first-draft or worksheet sketches can be used as the primary source of information for developing the final map design. Most of the graphics in this book were produced using a combination of scanning maps and photographs and graphics editing using a CADD-like package.

4.7 TOPOLOGICAL GEOCODING

Bearing the goals of geocoding in mind, let's look at a real geocoding method. More important, let's look at how it has evolved over time. The example we use is the geocoding associated with the mapping needs of the U.S. Bureau of the Census, part of the Department of Commerce. This geocoding system is particularly good for seeing how topological geocoding became important over time, and introduces the topic of data structures taken up in the following chapters.

The specific mapping need is to support the decennial census effort (as required by the U.S. Constitution), by generating street-level address maps for use by the thousands of census enumerators. Fairly early on, the use of the computer was recognized as critical. An early system, the address coding guide, was largely a text database, both computerized and manual, that listed all of the street addresses within census enumeration districts.

Included as part of the record was a designation listing the block side of the address, each one given a unique number, as well as geographic information such as census tract, block, ward, and post office codes.

Eighty-eight of the standard metropolitan statistical areas (SMSAs) for the 19th census (1970) were entered manually, with the rest coming from existing computerized mailing lists and post office checks. All 233 SMSAs were covered by this system, with over 40 million addresses.

A need was identified to link this address information with the hand-drawn enumeration maps used in the field. These maps were compiled by asking local communities for all available maps, which were then used as a base for designating the enumeration districts. Understandably, the lack of standardization and variation in accuracy and age of the maps was considerable.

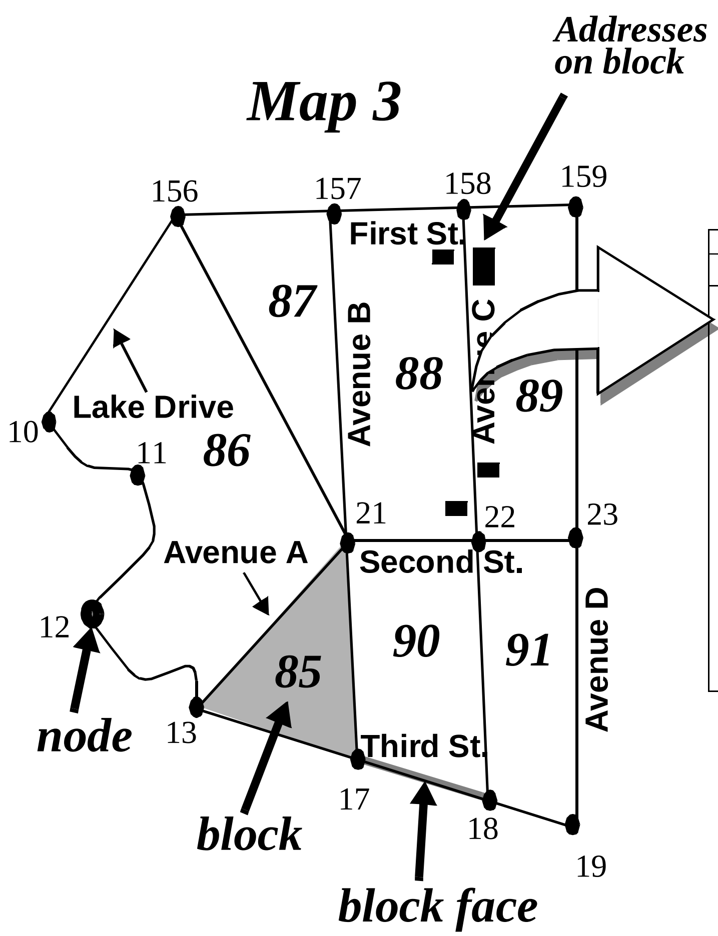
In planning for the 1970 census, a computer-based address coding guide (ACG) was tested in 1967 for New Haven, Connecticut. because a major area of interest in this pilot study was computer mapping, an attempt was made to add geocodes to the ACG files. At that time, digitizing tablets were rare, and the Census Bureau had to design and build its own prototype in-house system. Many maps were digitized using light tables and graph paper. Geocodes with state plane coordinates were added to the ACG records, although some records used latitude and longitude and “map miles.” This is one of the earliest cases of using digitally encoded topology to supplement location geocodes. Further use of ACG in this context, however, was abandoned. The process of using the guide had proven too difficult, with too many conflicts, errors and inefficiencies in the digitizing methods used.

The resultant technical steering group, which was overseeing the census use study in New Haven, recommended developing new methods for future censuses. A proposal was made to build geographic base files for the census separately, using graph theory as the underlying concept. Each street, river, railroad track, municipal boundary, or other map feature was considered as a straight line segment. Curved streets were constructed from multiple straight segments. Each node, line segment, and enclosed area was uniquely identifiable within the full network. Line segments were labeled with street names from the base maps, and nodes were numbered sequentially.

The entire system was called dual independent map encoding (DIME), because the basic file was created by computing two independent incidence matrices from the source map, line segment/node, and line segment/enclosed area (Department of Commerce, 1970). DIME, therefore, built substantially upon ACG in that it allowed external checking of the logical consistency of the data by performing topological checks. A DIME file was constructed for New Haven. The major difference between this file and the ACG file was that DIME records contained codes for both sides of a street; that is, the records were segment—rather than block side—based (Figure 4.14). The first few DIME files were used for experimental computer mapping, using such software as MAP01 and SYMAP.

Between the 1970 and 1980 censuses, DIME was updated and extended in scope. The geographic coverage was also extended so that a geographic base file of many SMSAs became available to census users. Automated error-detecting methods were devised that used the topological geocodes to locate and correct errors. Names were standardized and header records were devised for the computer tapes on which the GBF-DIME data sets were distributed. For the 1980 census, virtually all the major SMSAs were covered, with a total of 300,000 enumeration districts. The 1980 GBF-DIME files were used commercially and have remained as important sources of data for a large number of applications.

Analytical flexibility was a key to DIME. Although the geocodes were designed as a way of producing enumeration maps, it was soon realized that DIME was suitable for more general computer mapping. DIME was used for collection of statistical data, automatic generation of centroids, thematic mapping, and automatic address matching, a form of geocoding in itself. Much of this success is because DIME is georeferenced into the census data itself, allowing automatic thematic cartography with thousands of attributes to choose from.



From Node		To Node		Blocks		Left		Right	
Map	Node	Map	Node	Left	Right	Low	High	Low	High
3	**158**	**3**	**22**	**89**	**88**	**1**	**99**	**2**	**98**
3	159	3	23	00	89	1	99	2	98
	E T C								

Figure 4.14 Geocoding topologically using the DIME system.

After the use of GBF/DIME for the 1980 census, people started looking at automated thematic mapping as an effective means of understanding the mass of data the census provided. One of the problems for mapping directly from DIME, however, was that most of the segments were street or political boundaries. In addition, they are highly generalized. The demands of detailed computer cartography derived from census data called for a revision of the DIME system. As noted above, an important measure of a geocoding strategy is the ability to adapt to demands previously unforeseen. DIME gave a certain storage efficiency and more analytical flexibility than was at first realized. It produced a lot of labor in geocoding, and there were errors in the original GBF files, but in the long run the system proved invaluable, succeeding in its original goal and going on to different applications.

For the 1990 census, the Census Bureau developed a system called TIGER, for topologically integrated geographic encoding and referencing. TIGER came with a refinement of the DIME terminology (Figure 4.15). Instead of using the block face or street

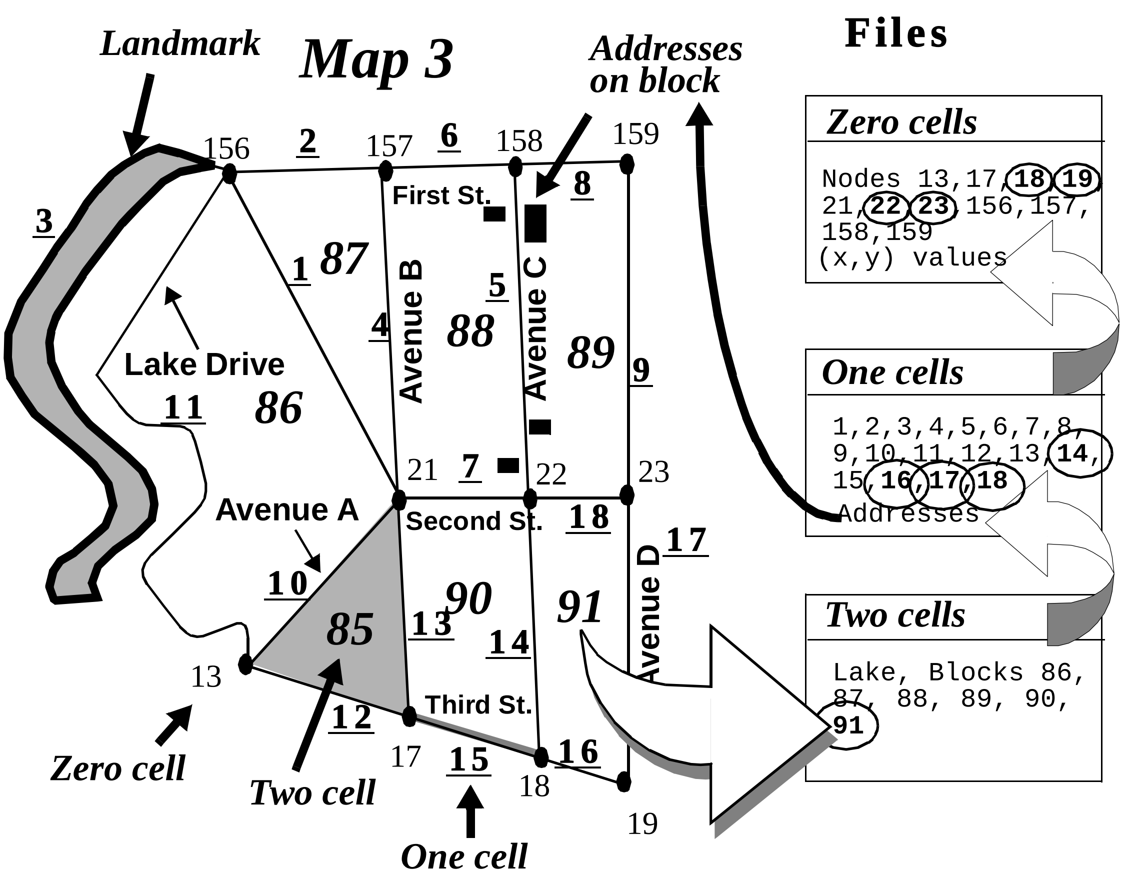


Figure 4.15 Geocoding using the TIGER system.

segment as the basic entity, the TIGER system recognizes cartographic objects of different dimensions. The objects are points (nodes), lines (segments), and areas (blocks, census tracts, or enumeration districts). In the TIGER terminology, points are zero cells, lines are one cells, and areas are two cells.

Under the TIGER system, the digital cartographic boundaries of census tracts follow real things on the earth's surface. Cartographic elements that were digitized included roads, hydrography, political boundaries, railroads, and some miscellaneous features, such as commercial buildings. Each was digitized as a point or a line and assembled to enclose areas. Interactive checking and maintenance software allowed manipulation of these objects, individually and collectively, as well as labeling and consistency checking. The new files were structured topologically and linked together by cross-referencing rather than by having all data in a single file. Thus linkage with the address records is possible, as is the isolation of the data and map base required for a particular thematic map. Choropleth maps, for example, need only the attributes and the two cells, while detailed maps need mostly one cells and labels.

This required a whole new series of maps. A large-scale cooperative effort prepared these maps for the 1990 census. Focal to the effort was having part of the digitizing phase performed in collaboration with the United States Geological Survey. As a result, the Geological Survey was able to convert to a digital basis many of the 1:100,000 series and some of the 1:24,000 series topographic maps.

In the evolution of the use of digital data at the Bureau of the Census, foremost is the collection of massive amounts of digital map data. By changing the basic geocoding method, going from a system based on the block face with a sequential topology to a system based on entities of different dimensions, there has been a substantial increase in analytical flexibility. At the same time, intergovernmental cooperation has reduced the labor in geocoding. The incorporation of topology has allowed a substantial amount of automated error checking, which has improved the quality of the data. The evolution from GBF/DIME to TIGER, therefore, has achieved many of the goals for improving geocoding. In addition, the TIGER data are likely to prove an important archive for future map production of many different types and by many different groups, proving once again that success can be measured by the ability to plan for the unanticipated future applications of geocoding.

4.8 REFERENCES

Chrisman, N. R. (1987). "Efficient Digitizing Through the Combination of Appropriate Hardware and Software for Error Detection." *International Journal of Geographical Information Systems*, vol. 1, no. 3, pp. 265–277.

Department of Commerce (1970). *The DIME Geocoding System*, Report Number 4, Census Use Study. Washington, DC: U.S. Bureau of the Census.

Department of the Army (1973). *Universal Transverse Mercator Grid*, TM 5-241-8, Headquarters, Department of the Army. Washington, DC: U.S. Government Printing Office.

Department of the Army (1983). *Grids and References*, TM 5-241-1, Headquarters, Department of the Army. Washington, DC: U.S. Government Printing Office.

Jenks, G. F. (1981). "Lines, Computers, and Human Frailties." *Annals, Association of American Geographers*, vol. 71, no. 1, pp. 1–10.

Lee, D. and T. Sallee (1970). "A Method of Measuring Shape." *Geographical Review*, vol. 60, pp. 555–563.

Pavlidis, T. (1978). "A Review of Algorithms for Shape Analysis." *Computer Graphics and Image Processing*, vol. 7, pp. 243–258.

Ward, J. R., and M. J. Phillips (1987). "Digitizer Technology: Performance Characteristics and the Effects on the User Interface." *IEEE Computer Graphics and Applications*, April, pp. 31–44.