CHAPTER 1

REFERENCE SYSTEMS FOR MEASUREMENT

CHAPTER OVERVIEW

- Introduce reference systems for time, space, and attributes.
- Extend Stevens' levels of measurement to provide a richer basis for attribute measurement.

HOW INFORMATION WORKS

An exploration of geographic information systems must begin with an operational understanding of information, then tighten the focus on what is specifically geographic, and finally examine the system components. The word *information* appears nearly everywhere in current life. We are said to be in an Age of Information, much as earlier and not so distant periods were termed the Age of the Airplane and the Age of the Atom. Such trendy catchphrases reflect what is economically important at a given time, but the glare of the spotlight may flatten out important details. **Information** occupies a middle stage in a process modeled on the scientific method. The starting point involves data—raw observations, that have no particular value by themselves. Somehow, through procedures not often totally explained, these raw data acquire value when placed in a frame of reference—a system of relationships among objects and assumptions about those relationships. For example, the digits 8.5 do not mean much unless you know they measure a water level in meters vertically from extreme low tide. This process also implies that information leads to higher levels of knowledge, through further refinement and interpretation.

Information: Data (observations, measurements, etc.) placed in context of a system of meaning (a set of relationships and assumptions about those relationships). Information, built into larger context, constructs knowledge.

This sequential flow from data to information and eventually to knowledge does not occur without human beings actively engaged. Gigabytes of databases do not produce refined information products unless someone does a lot of work. In addition, there is no guarantee that everyone will extract exactly the same results when confronted with the same data. Knowledge is not simply a passive result of assembling the data. It takes special talent to be willing to abandon your assumptions and to be open to surprise.

This exploration will not reveal a specific boundary where the data instantly become useful information. The whole book is about the process of providing context and of discovering relationships. Constructing information requires experimentation and exploration. It is not some smooth, guaranteed progression as on an assembly line. Understanding often comes with a spark of recognition—a realization that some particular fact jolted you to recognize a new relationship. Though this moment may seem distinctly personal, the practice of measurement provides guidance from the shared experience of communities of scientists, government officials, and business people. Measurement techniques ensure that separate facts relate to a common reference system. Geographic information depends on common forms of measurement, although it raises some particular issues and special problems.

BASIC COMPONENTS OF GEOGRAPHIC INFORMATION

Geographic information is commonly broken into the components of *space, time,* and *attribute*. Space, although it is an obvious component of geographic information, can be understood from a number of different perspectives. At the simplest, a space can consist of distinct "places" that are only different from each other. However, it does not require much observation to figure out that certain places are nearer than others and that some align in the same direction. Building up these relationships leads to the basics of geometry. The world of sensory experience is basically three dimensional. Objects have length, width, and height, and each is located at some distance and direction from the others. When dealing with larger geographic regions, attention can be limited to a thin shell of the earth's surface. For much mapping (and GIS), this space is predominantly two dimensional. Even though it is often convenient to use **Cartesian** geometry within confined regions, the surface of the earth is not a limitless plane. The most common spatial reference systems for GIS provide a translation between the three-dimensional shape of the earth and the flat spaces traditionally used for maps.

Time often plays a silent role in maps, though there is always some implicit or explicit temporal reference. The most common map works like a snapshot—valid for a specific moment in time. Of course, time is more than a collection of points. The inexorable progression of time and the inability to turn back the clock lead to the most

Cartesian: Geometry of unbounded spaces, particularly in two dimensions; the geometry of a flat plane treated by the tools of analytical geometry (cartesian coordinates) developed by René Descartes in the seventeenth century.

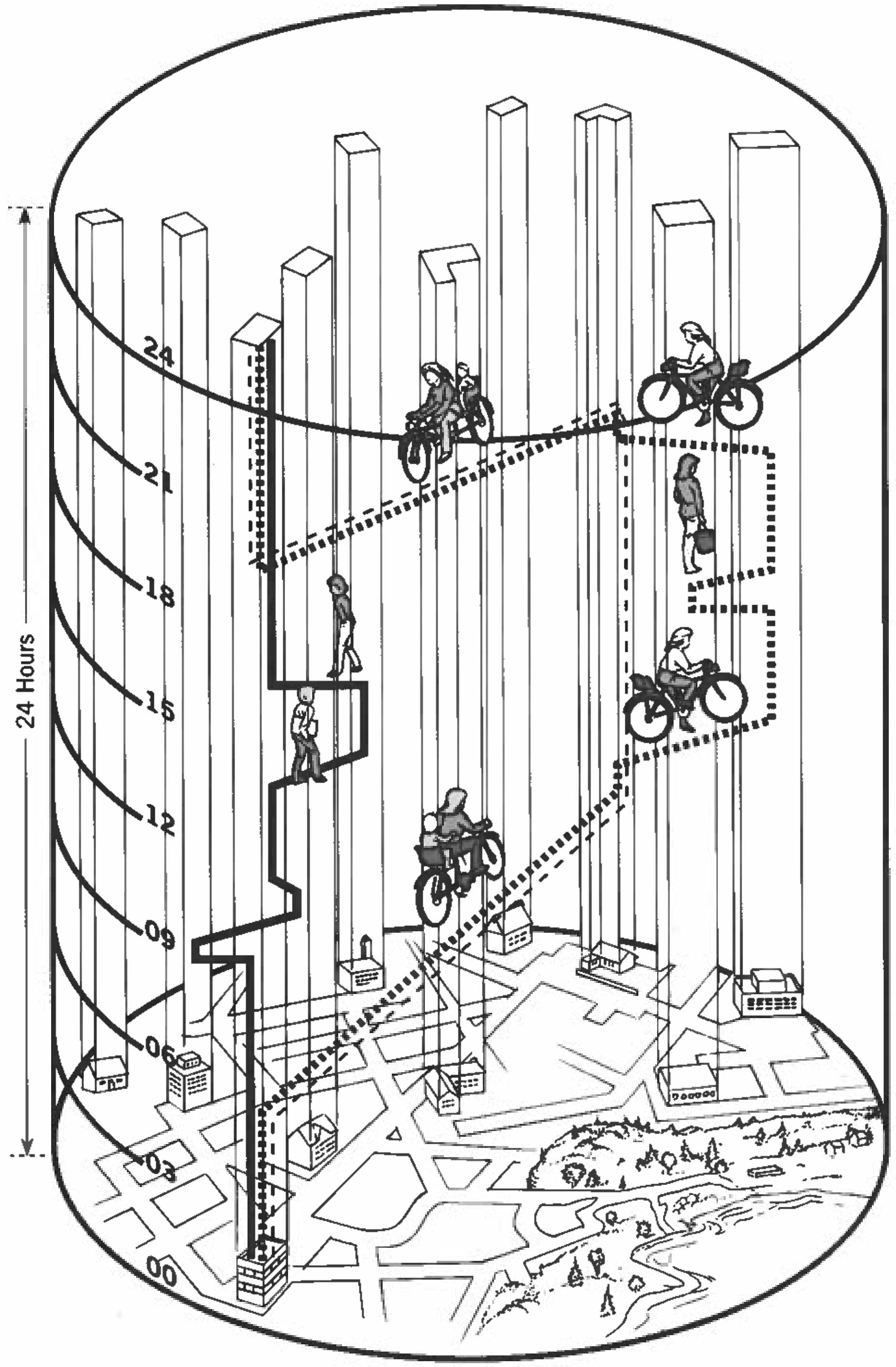

Figure 1-1 Space–time diagram showing people in their daily lives. Redrawn from Parkes and Thrift (1980).

prevalent view of time as an arrow—a line with one direction. Sometimes, however, it makes sense to treat time as cyclical, such as when astronomical and climatic events repeat. For example, when the earth returns to a particular position in its orbit, the stars return to the positions observed the previous year. In these cases, the sense of cycles comes from the connection between the spatial realm and the temporal, not because any span of time actually repeats.

The interplay of time and space in daily life can be read from the time–space diagrams produced by time geographers [following the lead of Hägerstrand (1970)]. Figure 1-1 shows a day (0 to 24 hours vertically) in a small town. The buildings remain in place, appearing as pillars (unless a new one appears or an old one gets demolished on this particular day). Members of one family start their day together, their movements charted as diagonal lines starting from their apartment building. Quite early, the mother takes off with the younger child on her bicycle, leaves the child at a day-care center (light dashed line), and continues on with her daily activities (thick dashed line) until she returns to the day-care center and then home. The other child (dark line) stays closer to home, going to school and to some afternoon lessons, returning home before the mother returns. The social interaction of this town is strongly synchronized by clocks. The mother's life is a dance of social expectations to be in certain places at particular times. The movement between places and times creates the social web of the town. This diagram illustrates how time and space provide a frame of reference for all activities.

The third component of geographic information, the **attribute,** can range from observable physical properties to aesthetic judgments. In Figure 1-1, there are dozens of possible attributes, including the width of streets, the names of the people, and the scenic beauty from each viewpoint. Information extracted from the time and space dimensions, such as a rate of speed, can also be treated as attributes. The attributes attached to a GIS are certainly the most varied of the three components. While this chapter can give technical details for the measurement of time and space, it can suggest only general classes for attributes.

Reference Systems

Each of the three components of geographic information is measured with respect to some particular **reference system.** Such a system provides rules to interpret individual observations with respect to others and to document the rules so that results can be repeated and compared. The technology of temporal reference systems is quite ancient, since calendars and clocks have existed for millennia. The other forms of reference systems are more recently established and less universal. The clearest reference systems to use are those established by explicit standards, though workable results can come from less formalized procedures as long as they are shared by all users.

Attribute: The range of possible values of a characteristic; an attribute value is a specific instance of the characteristic associated with a geographic feature.

Reference system: An established set of rules for measurement. Provides a means to compare a particular measurement to others performed with reference to the same set of rules. Geographic information requires reference systems for time, space, and attributes.

Temporal Reference Systems Time, with its strong sense of a linear order, is simpler than space. The linear axis of time—measured in units such as seconds, hours, and years—orders our lives in many ways (Figure 1-2). A simple temporal reference system merely requires an origin (a time to call zero) and a unit of measurement, such as a second. A stopwatch starts from its own zero each time you start it. A more complex system, such as a calendar, requires rules for counting days in months, and as long as others follow the same rules it remains reliable. Living together in a civilized society creates the need for a shared temporal reference system. Each ancient civilization created its own calendar, some based on the lunar cycle, others based on the solar year. During the past century, global activity became synchronized by a common reference time (Greenwich Mean Time) and a common calendar. Each time zone around the world sets its clocks so that solar noon corresponds roughly to 12 o'clock. A time zone is a subsidiary reference system based on an offset from Greenwich Mean Time. Similarly, certain countries and religious communities retain alternative calendars, but the correspondence is well understood. By adopting a common reference system, time measurements can be compared and mathematical operations such as subtraction produce useful results.

Some aspects of time are cyclical. In environmental studies of all kinds, the seasons play an important role. A "growing year" can be thought to start in spring while a "water year" might start in fall. There is no necessary starting point to a yearly cycle.

Another kind of reference system consists of ordered periods. For example, administrative procedures often specify a sequence of events, perhaps with some guidelines for duration, without respect to any particular starting point. Thus, an environmental impact statement might have a series of planned phases such as a scoping process, public comment, the analysis phase, a draft report, more public comment, then a final report. Any particular project can be located along this time sequence without needing a numerical measure.

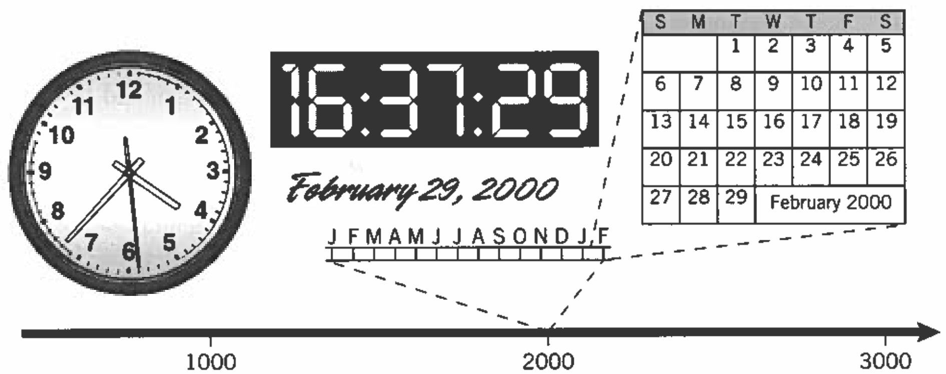

Figure 1-2 Temporal reference systems, lines and cycles. Clocks with rotary hands or digital displays repeat a cycle of measurements, while a time line continues indefinitely. The conventional calendar is a complex combination of cyclical elements at different frequencies (weeks, months, years, leap years, etc.).

Spatial Reference Systems Perhaps the family members in Figure 1-1 navigate through their town as a network of named places: buildings, roads, parks, and neighborhoods. They may never see it from above or draw a map of it, but still they would eventually learn certain geometric rules (such as the triangular inequality—the straight line home is shorter than combining it with an errand). The relationships between places create a geometry, even if it is not formalized and measured. A **spatial reference system** provides a refined tool by formalizing these relationships using analytical geometry.

Beyond specifying an origin and a unit of measure, spatial measurement requires geometric assumptions because an object cannot be located in a two-dimensional space by a single number along one axis. Spatial measurement for a limited site can use a local grid laid out in the field as long as the two axes are strictly laid out at right angles. Here the geometric model is a flat plane. The spatial reference for the La Selva Biosphere Reserve project introduced earlier was about as simple as it could be. Like many scientific sites, the researchers are more interested in local coordinates and the ability to find things in the field, not the relationships to places outside the Reserve. At La Selva during the 1980s, ecologists placed pieces of plastic pipe in the ground on a 200-meter grid. The idea was that researchers could locate themselves using a compass and a tape measure relative to the nearest pipe. The map the students had to digitize was created with this 200-meter grid as its reference system.

In 1991, the Reserve adopted a new spatial reference system and installed another set of pipes spaced 50 meters apart along one axis and 100 meters on the other. This new reference system provided a much denser grid of plastic pipes in the field and was laid out using improved surveying instruments. Due to imperfections in the original survey, the 200-meter grid was not on exactly the same axes as the successor; it was rotated at a small angle. The 1980s map could not be incorporated directly into the Reserve's GIS because of the difference in the spatial reference system. Because both systems were totally local, the reference systems could only be related by measuring some objects in both systems to derive the geometric relationships between the two planes. A student did this work in the field by measuring the markers of the 1980s system according to the 1991 reference system.

Local coordinate systems disconnected from a geodetic framework will probably die out as positioning technology becomes more common and the demand for GIS integration increases. La Selva with its deep "cloud forest" canopy may be one of the last places to make the conversion. In any case, GIS workers who must salvage records from the past will have to cope with obsolete systems for a long time to come.

While some local mapping projects can still be performed using an isolated, planar reference system, a GIS must usually mobilize a more complex process with a series of geometric steps that mobilize the science of **geodesy** [see Defense Mapping Agency (1984) for a comprehensive review] and analytical cartography. The actual shape of the earth, the **geoid,** is too lumpy to use as a reference surface. The first step adopts a model of the earth, usually in the form of a reference **ellipsoid.** There are dozens of ellipsoids in use, each chosen to fit the apparent shape of the earth in the regions surveyed. Each of these reflects the transitional period when local surveys were connected together in larger networks, but, until recently, it was technically challenging to connect all these networks into a global system. Once connected through global geodesy, the variations since the 1970s have become effectively insignificant for mapping purposes (Table 1-1).

While the ellipsoid provides a smooth surface, it is featureless and thus not sufficient as a reference system. A **geodetic datum** populates an ellipsoid with specific points whose locations have been established through astronomical surveying and careful adjustment to compensate for errors. It is a long-standing practice to specify positions on an ellipsoid using coordinates of latitude and longitude given as angles (Figure 1-3a). This geometric model uses the axis of the earth's rotation as its north–south axis. The plane of the equator—at right angles to this axis—provides the origin for angles north and south (latitude). Longitude measures the angles east and

Spatial reference system: A mechanism to situate measurements on a geometric body, such as the earth; establishes a point of origin, orientation of reference axes, and geometric meaning of measurements, as well as units of measure.

Geodesy: Science of measuring the shape of the earth and establishing positions on it. Involves study of geophysical properties such as variations in gravitational field. Adjective form: geodetic.

TABLE 1-1 World Geodetic Standards: Reference Ellipsoids[a]

Name	*Equatorial (Major) Axis in Meters*	*Flattening (1/f)*	*Region*
Airy 1830	6377563	299.325	Great Britain
Bessel 1841	6377397.2	299.153	Central Europe
Everest 1830	6377276.3	300.80	Indian subcontinent
Clarke's 1866	6378206.4	294.98	North America
Clarke's 1880	6378249.2	293.47	Africa; France
Krasovsky 1940	6378245	298.2	Former Soviet Union
World Geodetic System 1972	6378135	298.26	NASA, U.S. military
GRS 1980/ WGS 84[b]	6378137	298.257	GPS, new systems

[a]These reference ellipsoids may serve as the best fit to the actual geoid in different parts of the world. A horizontal "datum" adopts a reference ellipsoid and locates geodetically surveyed points on that ellipsoid.
[b]At the resolution shown, Geodetic Reference System 80 and World Geodetic System 84 are the same.
Source: Snyder (1987), Table 1, p. 12.

Geoid: Three-dimensional shape of the earth defined by the surface where gravity has the value associated with mean sea level.

Ellipsoid: Three-dimensional object formed by rotating an ellipse around its minor axis; an oblate ellipsoid approximates the shape of the earth (geoid), computed by the best fit to geodetic observations. See Table 1-1.

Geodetic datum: A geodetic reference system, usually divided into vertical and horizontal standards. A horizontal datum is based on a given ellipsoid and specified latitude–longitude coordinates for certain points. The plural of datum is *datums*, despite the word's Latin origins. (Not to be confused with the usage of *datum* as a single value, plural *data*.

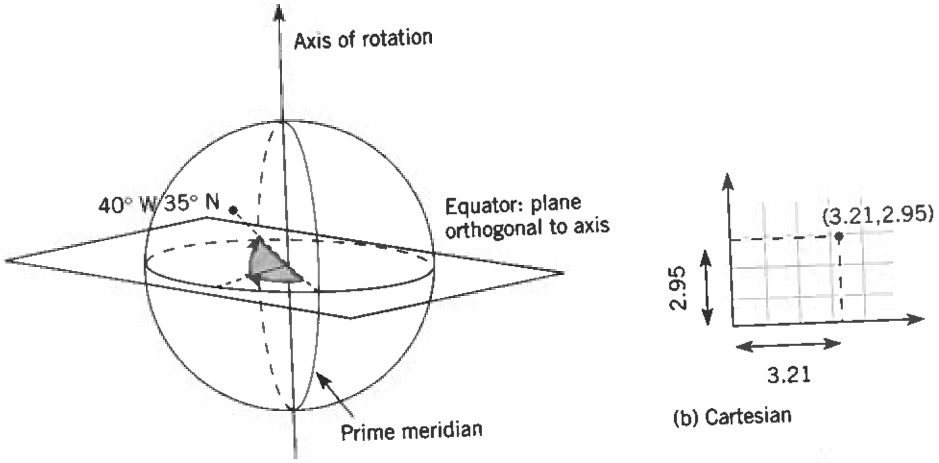

Figure 1-3 Spatial reference systems: Latitude–longitude pairs measure angles from the plane of the equator (orthogonal to the rotational axis) and the prime meridian (the meridian of Greenwich, England). The World Geodetic Reference System specifies a particular ellipsoid that provides the size of the earth to use with this model of angular measurement. Other reference systems may use a plane or a cone as a projection surface, connected to the geodetic reference system. Measurements on these flat projection surfaces are usually distances from a pair of axes (cartesian coordinates).

west around the equator using an arbitrary origin set by international convention along the meridian through Greenwich, England.

A particular place on the earth will have different values of latitude–longitude, depending on the choice of datum and ellipsoid. One common source of incompatibility between spatial reference systems in the United States comes from the transition between the **North American Datums (NAD)** of 1927 and 1983. This includes a change in reference ellipsoids plus the more local effects of readjusting the geodetic survey data. For example, the first-order **geodetic control** station Northwest Blake (PID SY5304) has been located on the same boulder on the beach of Blake Island in Puget Sound since it was first established in 1857. The coordinates assigned to this point have changed due to surveying technology and the development of more comprehensive geodetic models. Using the most current geodetic adjustments, this marker has a coordinate 4.46 seconds of longitude farther

North American Datum (NAD): An adjustment of geodetic measurements that provides the accepted horizontal reference system for North America. The 1927 Datum held Mead's Ranch, Kansas, as a fixed point, while the 1983 Datum performed a simultaneous adjustment of all measurements. NAD 1927 uses Clarke's 1866 ellipsoid, while NAD 1983 uses the 1980 Geodetic Reference System.

Geodetic control: Reference marks used to establish a spatial reference system for a specific project; under ideal circumstances these are geodetic survey points established as part of the National Geodetic Reference Network.

west of Greenwich than it had using NAD27. Such a difference means about one kilometer at this latitude, although the marker has not moved at all. For this reason, knowing the full spatial reference system can be very critical when registering sources compiled at different dates. Since ellipsoids are mathematically derived, they can be converted from one to another, but two datums require further measurements for accurate conversion. Similar differences occur between some adjacent European countries.

While geodetic reference systems are most appropriate for geographic information, maps are not constructed on ellipsoidal geometry. Another step is required. Most maps adopt a simpler geometric model (such as a cylinder or a cone) positioned with respect to the geodetic model—a process called **projection.** This technique was developed to permit map construction on two-dimensional media, but the simplicity of cartesian coordinates has been retained for many software packages. Cartesian coordinates measure distances along two axes at right angles (Figure 1-3b), but when defined through a projection, these positions retain their link to the global geodetic reference system. There are hundreds of possible projections and infinite ways to position them on the ellipsoid (Snyder 1987). In practice, for a given area a small number of projections will be used. Projections can be classified by the properties they preserve as well as by their geometry (whether based on a **developable surface** such as a cone or cylinder).

Selection of a map projection is often a matter of tradition, linked to the historical development of a given set of map users. Nautical charts, for example, are nearly always drawn on the Mercator projection, with the equator as line of tangency. This cylindric projection is **conformal,** and it has a special property, namely, that corrected compass bearings are straight lines. Manual plotting of courses and positions requires these particular properties. For topographic mapping, most of the world is covered at some scale in either a **transverse** Mercator such as Universal Transverse Mercator (**UTM**) or a conformal conic such as Lambert Conformal Conic (Table 1-2). This choice was based on the needs of artillery officers and engineers, who originally developed these maps. Equal-area projections might be more useful for many kinds of GIS analysis.

Projection: Geometric transformation that converts latitude–longitude coordinates into planar coordinates. Projections can be based on a developable surface (such as a plane, cylinder, or cone) or on a mathematical function.

Developable surface: Three-dimensional object that can be flattened into a plane without scale distortions. Cylinders and cones, since they curve in only one axis, can be converted into planes by making a single cut. Common projections use developable surfaces but some projections use more complex functions.

Conformal: Property of a projection that preserves the shape of geographic features. Within an immediate vicinity of a point, angles are preserved, but linear and areal scales have to be sacrificed to obtain this property.

Transverse: A projection oriented at right angles to the equator. A transverse cylindric projection uses a meridian of longitude as its central meridian.

UTM: Universal Transverse Mercator; a spatial reference system using a set of transverse Mercator projections 6° wide that cover the earth (except for polar regions covered by two polar stereographic projections).

TABLE 1-2 Common Projections Used as Spatial Reference Systems

Transverse Mercator Systems	
Worldwide	Universal Transverse Mercator (6° strips)
United Kingdom	National Grid
Germany	Gauss-Kruger (3° strips)
State Plane System[a] (USA)	North–south states, e.g., Illinois in two zones
Lambert Conformal Conics	
France	Grille Lambert, three zones
State Plane System[a] (USA)	East–west states, e.g., Washington in two zones
Other Projections	
Malaysia (peninsular)	Malayan skew orthomorphic (oblique Mercator)
Switzerland	National Grid (Gauss-Kruger)
State Plane System[a] (USA)	Alaskan Panhandle; oblique Mercator

[a]The State Plane System for the United States defines 125 zones, some as small as single counties.

A common system of spatial reference is a critical element of a GIS, since it brings different map layers into correspondence (Figure I-1). The Dane County project (see Introduction) discovered a half-dozen projections in active use, plus many maps without any known projection (much like the La Selva local coordinate systems). The project performed geodetic surveying to relate the different reference systems, then transformed each source following procedures described in Chapter 3.

Some countries try to limit incompatibilities by ordaining a particular projection as a common spatial reference system. All maps in Switzerland and the United Kingdom are referenced to their respective national grid systems. Larger countries and most American states cannot cover their territory with a single projection zone without distortions beyond the tolerances required for many applications. Most projects require translation between multiple spatial reference systems.

With computer representation, converting between geodetically referenced systems poses no real difficulty. A map in one projection can be converted to latitude–longitude (through the inverse of the projection function) then projected into some other datum and projection. Some calculation time may be required, but on current computers it is no great burden. The U.S. government has placed the General Cartographic Transformation Package (GCTP) in the public domain, and many commercial packages have incorporated this software.

Attribute Reference Systems Just as reference systems apply to time and space, they also apply to attributes. Unlike the common approaches that apply to time and space, each particular attribute scale requires its own reference system. There are some general rules for attribute measurement that are widely used in cartography and social science statistics. The next section reviews some of these concepts, with some additional rules related to common geographic attributes.

LEVELS OF MEASUREMENT

The wide diversity of attributes indicates a huge number of techniques for measurement. To a purist, like a classical physicist, measurement provides a numerical relationship between some standard object and the object being measured. Consider the attribute *length*. Every entity in space can be measured by comparing its length to some other length. The procedure begins by placing a standard measuring rod alongside the object to be measured, marking where the end of the rod falls, then placing the rod again beginning at the mark, until the end of the object is reached. This procedure implements a physical form of addition. The number of times the rod is placed represents the ratio of the length of the object to the length of the rod. Using similar comparisons, physicists developed procedures to measure temperature, mass, electrical charge, and more.

In nineteenth-century physics, fundamental physical properties were considered *extensive* because they *extended* in some way as length does in space. Other properties, like density, were built up as ratios of the extensive properties and were thus *derived*. The fundamental physical properties form the basis for the international standards comprising the metric system or **SI.** But the attributes used for geographic information reach far beyond the SI measures and ratios derived from them.

To provide a framework for a broader range of measurement types, Stanley Stevens (1946), a psychologist at Harvard University, published an article in *Science* proposing a framework based on what he called **levels of measurement.** Stevens adopted a very simplified definition of measurement as the "assignment of numbers to objects according to a rule" (Stevens 1946, p. 677). Stevens' schema has become a basis for social science methods and a framework for cartography and GIS. Because Stevens' classification is often misapplied and misinterpreted, the levels of measurement deserve careful scrutiny. Some revisions and extensions must be considered to accommodate geographic information.

Stevens used the concept of **invariance** under transformations. Invariance considers the degree that a **scale** can retain its essential information content even if it is not identical to some other scale. A level groups the scales that share a set of possible transformations. One example of invariance involves temperature. A measurement in °F can be transformed into °C without loss of information; thus, these mea-

SI: Système International d'Unités; the system of weights and measures established by international agreement in 1875. The International Bureau of Weights and Measures in Sèvres, France, oversees the measurement standards. SI defines seven base units from which many others can be derived: meter for length, kilogram for mass, second for time, kelvin for temperature, ampere for electric current, mole for chemical quantity, and candela for intensity of light.

Level of measurement: A grouping of measurement scales based on the invariance to transformations. A measurement scale at a given level of measurement can be transformed into another scale at the same level without losing information.

Invariance: Properties that remain unchanged despite transformations of the numbers used to represent the measurement.

Scale: When applied to a scale of measurement, a system used to encode the results of a measurement; typically a number line, but generalized to include a list of categories.

surements are at the same level. So, despite having different zero values and different units of measurement, the two scales can be related to each other. By contrast, a scale consisting of {cold, warm, hot} cannot retain all the information recorded in either Fahrenheit or Celsius scales.

The following sections explain Stevens' four levels of measurement and illustrate them with a common example. Imagine a marathon in which contestants become associated with certain attributes or measurements.

Nominal

At the most basic level, Stevens described a nominal "scale" in which objects are classified into groups. Any assignment of symbols can be used, so long as the distinct nature of each group is maintained. A nominal measure is based on set theory. The use of the word *scale* for a nominal measurement may evoke the traditional number line, but there is no such ordering implied.

In the marathon example, each contestant gets a number to wear. What does this number mean? Is it a measurement? If the number is simply pulled randomly out of a box, it has to be considered an arbitrary symbol (like a word or an icon). Other nominal attributes could be determined, such as the set of contestants wearing red shirts (Figure 1-4). Another nominal grouping might allocate contestants into either the women's event or the men's event. Any numerical symbol for these two categories (0 and 1, or 1 and 2, or 359 and 213) would be totally arbitrary. In this sense, nominal data remains invariant under the most extreme alterations; any symbol can be converted to another symbol, as long as they remain distinct from each other.

Ordinal

The ordinal level introduces the concept of an ordering. An ordinal scale applies when objects can be sorted in some manner; such a scale can exist in many forms. The most exhaustive form orders all objects completely without any ties (Figure 1-5). An example is the order of runners finishing the race (first, second, third, . . .). It makes sense to use the word *scale* for such an ordering because each successive element follows in the same direction. In another example, the number on the shirt of each contestant in the race might also represent an ordinal measure, if the numbers are handed out of the box sequentially. Thus a seemingly nominal identifier might hide an ordering based on time or on some other property.

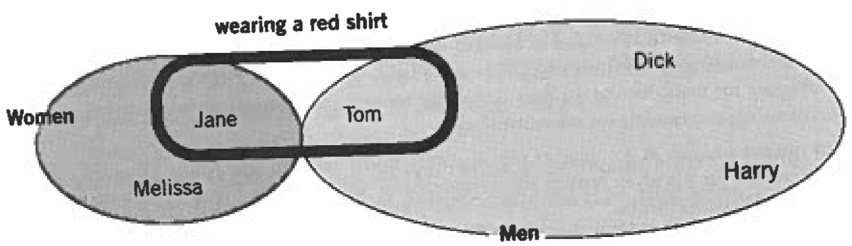

Figure 1-4 Nominal measures are not on scales at all. They create categories that can be treated as sets.

Ordinal

Order of arrival of contestants	*Women's race*	*Men's race*
First	Jane	Tom
Second	Melissa	Dick
Third	Leila	Harry

Figure 1-5 Strictly ordinal scales can arise from a total ordering, but ordinal scales may also arise from partial orderings.

Not all orderings are as well behaved as this ideal model without ties. The more ties there are, the less scalelike the ordering becomes. For example, some ordinal categories use a semantic scale of words. Soils are ordered from "poorly drained" through "somewhat poorly drained" to "well drained" and "excessively drained." Opinion polls use orderings such as "strongly disagree" to "strongly agree." The ordinal level covers a wide range of possibilities. Some scales behave in a nearly numerical way, whereas others are barely evolved from a nominal level.

Ordinal measurement does not constrain numeric representation very much. It may be conventional to give out the numbers 1, 2, 3, . . . to finishers in a race, but the numbers could be any increasing sequence (0.5, 0.66, 0.75, 0.8, . . .) or (1, 3, 597, 6667, . . .) because the order is all that matters. In the example of soil drainage, we do not know if the step from "poorly drained" to "somewhat poorly drained" is identical in magnitude to the step between "well drained" and "excessively drained." For different analytical purposes, the importance of each step in the scale might vary. Some orderings may relate to an underlying numerical scale, and others may not. For example, it is hard to assume that "good" means the same thing for all respondents in an opinion poll. Ordinal values are essentially categories without the arithmetic properties usually ascribed to numbers. Hence, nominal and ordinal measures are sometimes grouped together as *categorical* measurements. It is important to remember these limitations when some GIS user wants to standardize rankings on a scale from 1 to 9. Encoding with numbers does not automatically make arithmetic valid.

Interval

In Stevens' scheme, the quantitative realm begins with interval scales that give numbers algebraic meaning. An interval scale involves a number line with an arbitrary zero point and an arbitrary interval (the unit of measurement). Thus, interval scales can be shifted by changing the zero without changing the meaning of the measurement. For example, years can be recorded on the Gregorian calendar (A.D.), the Islamic calendar (1 A.H. is A.D. 622), or the geologists' calendar [0 Before Present (B.P.)

is A.D. 1950]. In all these systems, the numerical value of a year has no particular significance. The year 2000 is not twice the year 1000 in some magnitude.

In the case of the marathon, we could assign arrival times to runners by simply noting the clock time for each arrival (Figure 1-6). As long as some basic assumptions are valid, particularly that all runners departed at the same time, then these numbers capture all the ordinal results. In addition, the differences between arrival times can be interpreted. Some of the arrivals are closer to the next arrival than others, establishing a truly numerical measure of difference between values. An elapsed running time can be twice as long as another, for example.

Ratio

Arrival times for a race provide raw results awaiting further processing. Contestants would obtain a more useful measure by subtracting the time at the start from the time at their finish. In fact, a difference between two interval measures becomes a measure on Stevens' next level: ratio.

In measurement theory, the ratio level gets the most attention. Ratio measures retain the arbitrary unit of measure from the interval scale but substitute a true origin (zero value). These properties support the arithmetic operations of addition, subtraction, multiplication, and division. On a ratio scale, if a value is twice that of another, then it represents a doubling of the quantity. The easiest ratio measures to visualize are classical extensive quantities. In the race example, the elapsed time in running the race is a ratio measure obtained by subtracting two interval measures for the start and finish (Figure 1-7). The ratio measure of elapsed running time contains all the information of the ordinal scale for ranking winners, plus it adds the numerical properties that measure how fast each contestant ran. It is clear that these ratio measures convey more information and permit more analytical treatment.

Extensive and Derived Scales

Stevens tried to combine extensive and derived measurement into one level. He defined the ratio level based on the invariance related to the arbitrary *unit of measure*. Thus, a length in feet can be converted to meters with no real change in length. But the invariance group misses some important distinctions between geographic mea-

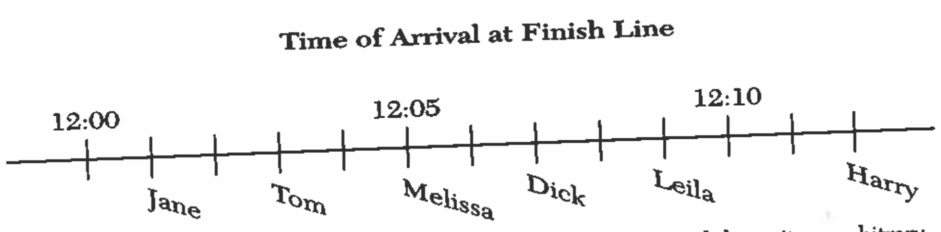

Figure 1-6 Interval scales mobilize a number line, but the origin and the unit are arbitrary.

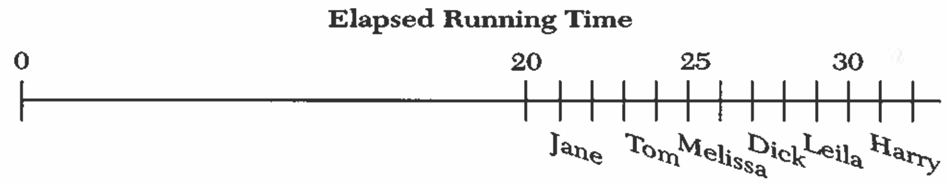

Figure 1-7 Ratio scales, the classical ideal for physical measurement, have a true origin and an arbitrary unit of measure.

surements. Consider some economic activity in a county measured in dollars. This scale is arbitrary because yen would work as well as dollars. The value of zero applies to any scale. The value for the county could be combined by addition with the value of others to obtain a measure for the state economy. Cartographic design principles would suggest **proportional symbols** for such raw values (for example, Figure 1-8 shows the total expenditure by the Department of Defense for each California county in 1976). In place of total expenditure, the county can be given a per capita measure by dividing the dollar figure by the number of persons in the county. Such a transformation removes the influence of population and concentrates on relative expenditure. This value is just as much a "ratio" as is the total expenditure figure, but cartographic rules suggest a **choropleth map** presentation for these per capita figures (Figure 1-9). Per capita values of two counties cannot really be added together because the denominators (populations) might be totally different. Notice that some counties with low total figures can have high per capita figures.

Stevens' system has been used to suggest which statistical methods apply to a given measurement. Many introductory statistics books, particularly for the social sciences, connect the levels of measurement to a group of appropriate tools. Similarly, cartography texts connect the cartographic tool kit of **graphic elements** to specific levels of measurement. The connection is not entirely straightforward, as demonstrated by the case of the California county data. For geographic data, it would make more sense to divide ratio measures into the invariance classes applied in selecting thematic map types. This would also separate those measures that are aggregated by

Proportional symbols: A thematic mapping technique that displays a quantitative attribute by varying the size of a symbol. Typically, proportional symbols use simple shapes such as circles and are scaled so that the area of the symbol is proportional to the attribute value. Each symbol is located at a point, even if it represents data collected for an area.

Choropleth map: A thematic mapping product that displays a quantitative attribute using ordinal classes applied as uniform symbolism over a whole areal feature. Sometimes extended to include any thematic map symbolized using areal objects.

Graphic elements: The characteristics of a symbol system that can be manipulated to encode information. For cartography, these include size, shape, hue, saturation, brightness, orientation, and pattern. [See Robinson et al. (1995)].

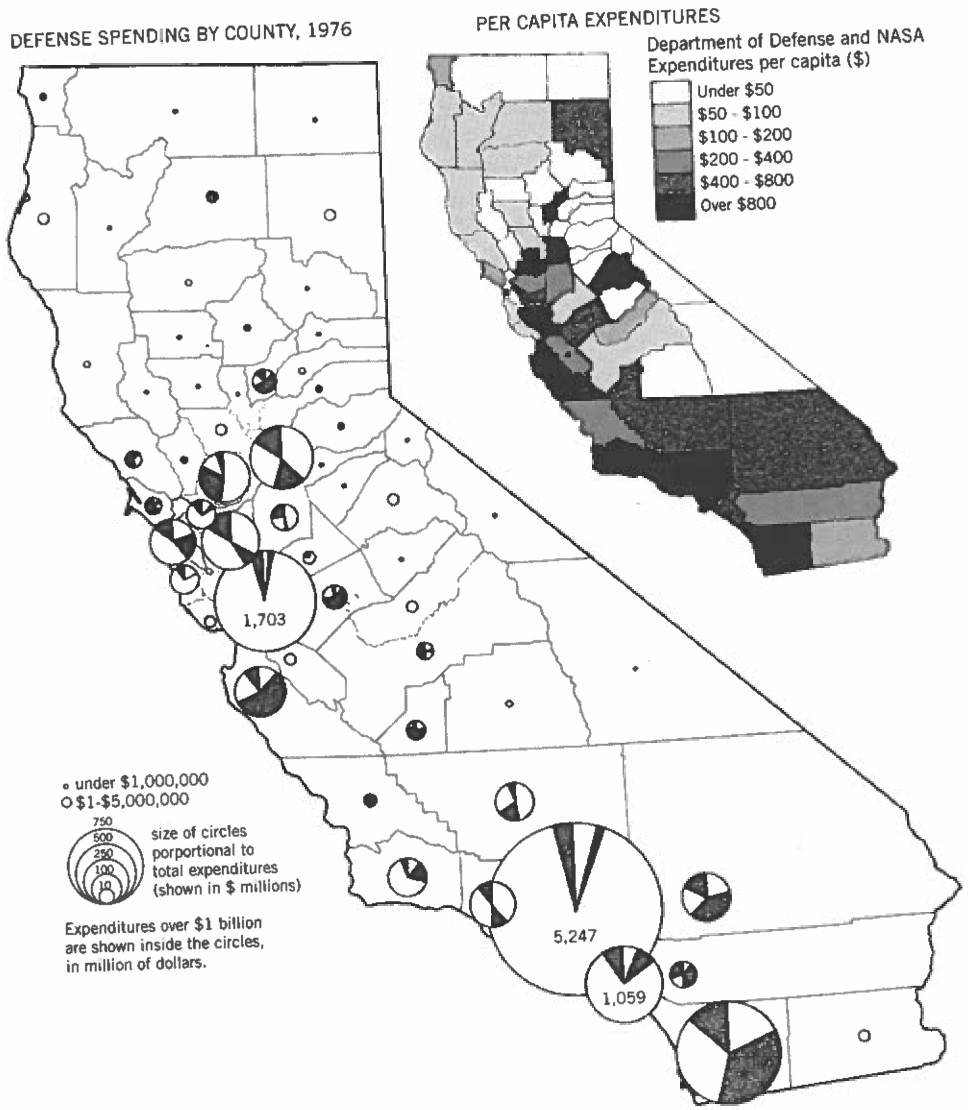

Figure 1-8 Proportional symbols use the graphic variable of size for a simple geometric symbol such as a circle. They are considered appropriate for raw measures, such as total population or total economic output. This map portrays defense spending for each county by scaling the area of the circle to be proportional to the dollar figure. (Source: Donley et al. 1979, p. 46.)

Figure 1-9 Choropleth maps use the spatial object as the symbol. The graphic variable size cannot be used without a cartogram. Here, as in many cases, the graphic variable value (a gradation from light to dark) shows the range of the attribute. Since the area of the object is a part of the symbol already, this method is most appropriate for density measures, or a derived ratio such as dollars per capita. (Source: Donley et al. 1979, p. 46.)

addition (extensive) from those that must be weighted (such as derived ratios). These are just a few examples of the operations that form the main objective of this book in later chapters.

Perhaps the best way to explain why attribute reference systems matter is by a counterexample. Along highways, it is common to announce the towns and villages through which the road passes. Most highway signs announce the name with some extra information, such as population or elevation. One town in California has a sign that takes the spirit of local pride to an extreme (Figure 1-10). Adding these three numbers is clearly a joke; yet, professionals who work with geographic information often commit equally meaningless combinations with no humorous intent. The number 4663 measures nothing about New Cuyama because it combines the count of people, the elevation (in feet above sea level), and the year the town was established (on a certain calendar). Having three numbers does not ensure that addition will produce any sensible result.

What Is Missing from Stevens

Stevens' four levels are usually presented in the geographic literature as a complete set, but they are not enough for practical applications of geographic information. There are at least four revisions that need to be added.

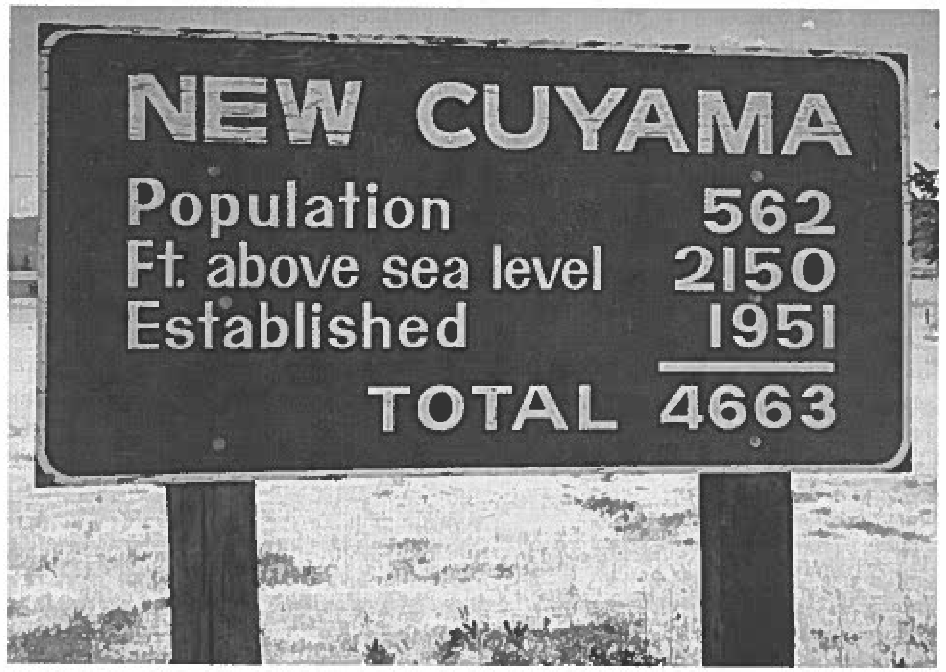

Figure 1-10 Sign posted at the entry to New Cuyama, California.

Absolute Scales Ratio is not actually the highest level of measurement. A ratio scale is higher than an interval scale because the value of zero is no longer arbitrary. A higher level of measurement would be achieved if the unit of measure were not arbitrary. When the whole scale is predetermined or *absolute* (Ellis 1966), no transformations that preserve the meaning of the measurement can be made. One example of such an absolute scale is *probability*, where the meaning of zero and one are given. Even though it is common to report probabilities as percentages, the relationships of probability (such as Bayes' law of conditional probability) operate correctly only when scaled from zero to one.

Cyclical Measures While Stevens' levels deal with an unbounded number line, there are many measures that are bounded within a range and repeat in some cyclical manner. Angles seem to be ratio, in the sense that there is a zero and an arbitrary unit of measure (degrees, grads, or radians); however, angles return to their origin. The direction 359° is as far from 0° as 1° is. Any general measurement scheme needs to recognize the existence of such cyclical measures.

Counts Another class of geographic measurements deals with counting objects aggregated over some region in space, such as a human population. The objects counted are discrete; there is no half person. Yet, unlike the other discrete levels of measurement (nominal and ordinal), the result of a count is a number. Since the zero is a fixed value, counts may seem to be ratios, but the units of a count are not arbitrary, so they cannot be rescaled as freely. Counts are more similar to absolute scales, with a restriction to discrete integers. They become ratios when the unit of measure is rescaled to "population in millions" or something that loses the discrete identity of the objects.

Graded Membership in Categories A further criticism of Stevens' system is that nominal categories are not always as simple as portrayed. Nominal measures apply the strict rules of classical set theory. All members of a set are meant to belong in that category equally; these sets are called "sharp." Many classifications, however, adopt more flexible rules; they involve some kind of graded memberships as formalized in **fuzzy set theory** or they involve comparison to a **prototype** member of the class. In both these situations, an object will have some degree of membership (represented by a proportion or percentage), rather than just belonging or not. Some members of the group are just more typical than others. Accommodating a more nuanced

Fuzzy set theory: An extension to set theory that permits an object to have a degree of membership (usually represented as a number between 0 and 1). Fuzzy membership values do not have to follow the rules of probability.

Prototype: An approach to categorization that defines a category by identifying a particular object as the typical example. Other objects assigned to this category may not share all characteristics with the prototype object. The degree of resemblance represents graded membership.

approach to categories remains a research frontier in GIS (Burrough and Frank 1996).

Thus, Stevens' four levels of measurement are not the end of the story. A closed list of levels arranged on a progression from simple to more complex does not cover the diversity of geographic measurement. Still, Stevens' terminology provides a starting point for the bulk of common situations.

Applying Levels of Measurement to Attribute Reference Systems

The previous section ended with the conclusion that attribute reference systems seemed too varied for standardized treatment. Though they do not specify all details, Stevens' four levels of measurement (with extensions) do prescribe the information required for an attribute reference system (Table 1-3). Absolute measurements can simply state what they measure because the whole scheme is implicit. For a count, the reference system is the kind of object counted. For a ratio scale, the unit of measure must be given, along with some additional information to sort out the subcases of cyclical scales and derived ratios. For interval measures, the units and the zero point are required. The categorical levels require more information because each category has its own definition. Perhaps the ordinal categories of "somewhat poorly drained" and "poorly drained," for example, are divided at a specific threshold on a ratio measure of permeability. Other ordinal values may not have explicit links to a numerical scale, just a ranking. With nominal categories, each category needs to be described. Some category systems are simply lists, as in the Anderson land use codes (Anderson et al. 1976). Another way to present categories uses a series of structured questions. For example, does the tree have leaves or needles? Are the needles in groups or singly attached? Are the groups of five, three, or two? Such a key can emphasize different characteristics in the various paths, each leading to a particular category.

TABLE 1-3 Information Content for Attribute Reference Systems

Level of Measurement	*Information Required*
Nominal	Definitions of categories
Graded membership	Definition of categories plus degree of membership or distance from prototype
Ordinal	Definitions of categories plus ordering
Interval	Unit of measure plus zero point
Extensive ratio	Unit of measure (additive rule applies)
Cyclic ratio	Unit of measure plus length of cycle
Derived ratio	Unit of measure (ratio of units; weighting rule)
Counts	Definition of objects counted
Absolute	Type (probability, proportion, etc.)

Attribute Reference Systems for the La Selva Project

The La Selva project demonstrates a diversity of attribute reference systems. As in many GIS projects, the bulk of the data sources available to the students for the La Selva project were in categories. The project focused on the Sarapiqui Annex, a parcel of land. Inside this boundary, the map showed some points representing the pipes that demarcate the spatial reference system and other points depicting stumps (signs of logging activity). A tree stump is a member of a very simple nominal category. There were also roads, trails, and streams. To some extent, the roads and trails are a part of an ordinal set of classes. The primary content of the map was a land use delineation in which the categories were ordered along a gradient of human disturbance. "Primary Forest" included the rain forest with the least human influence, followed by "Selectively Cut," "Cleared Land," and so on. The land use mappers did not document the rules that were applied in deciding how many trees had to be removed to qualify for each category. Presumably, these categories made sense considering the local economy. Some forest was completely cleared to create agriculture, while other operations targeted specific kinds of trees.

The La Selva project also used some information from the Reserve's existing database. Elevation data appeared in standard meters. The zero was mean sea level, which was not particularly relevant in the cloud forest. For all practical purposes, the elevation data were interval measures, largely useful when compared to each other to construct measures (to be introduced in Chapter 7) such as slope gradient (an absolute scale) or slope aspect (a cyclical ratio). The soils data were classified according to an international nomenclature for soils. These classes were divided from each other according to multidimensional thresholds: permeability, organic content, and grain sizes. Analytically they were treated as sharply distinct sets, although they probably represent a series of complex gradients.

SUMMARY

Geographic information must be embedded in a reference system for time, space, and attribute. Time and space have fairly standardized systems in common use. Attributes, by contrast, come in all flavors. The generic typology of Stevens' levels of measurement provide a starting point to develop the information content required. Numerous additions and special cases must be recognized.

To use measurements effectively, additional distinctions must be made. These distinctions do not come from the numbers but from a larger framework surrounding the measurements. A spatial reference system provides a mechanism to construct a more integrated structure, but coordinates by themselves do not ensure compatibility of diverse information. Time and attribute also have reference systems, but these three systems just provide the basic axes. A more comprehensive framework must include the interactions of these three components. The next chapter develops such a framework for geographic information.