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 ele-

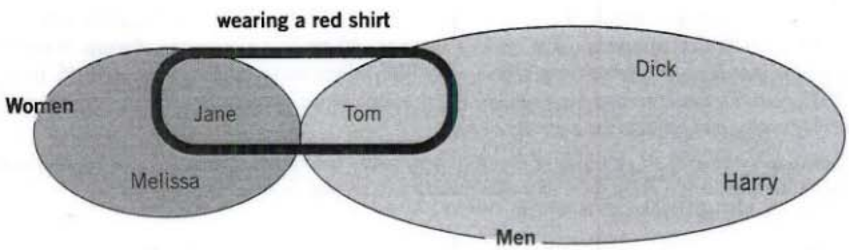

Figure 1-4 Nominal measures are not on scales at all. They create categories that can be treated as sets.

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.

ment 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.

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-

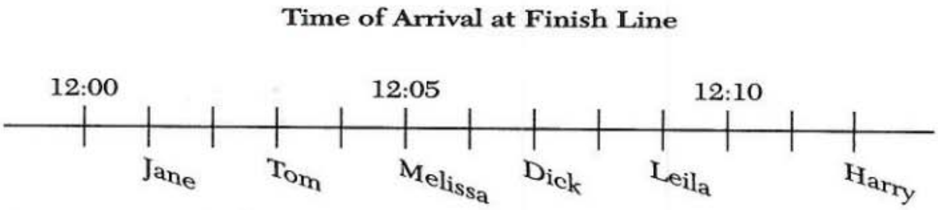

Figure 1-6 Interval scales mobilize a number line, but the origin and the unit are arbitrary.

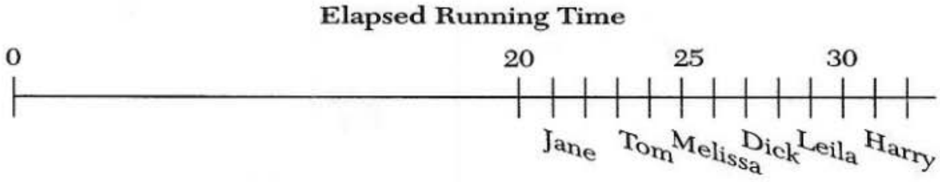

Figure 1-7 Ratio scales, the classical ideal for physical measurement, have a true origin and an arbitrary unit of measure.

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)].

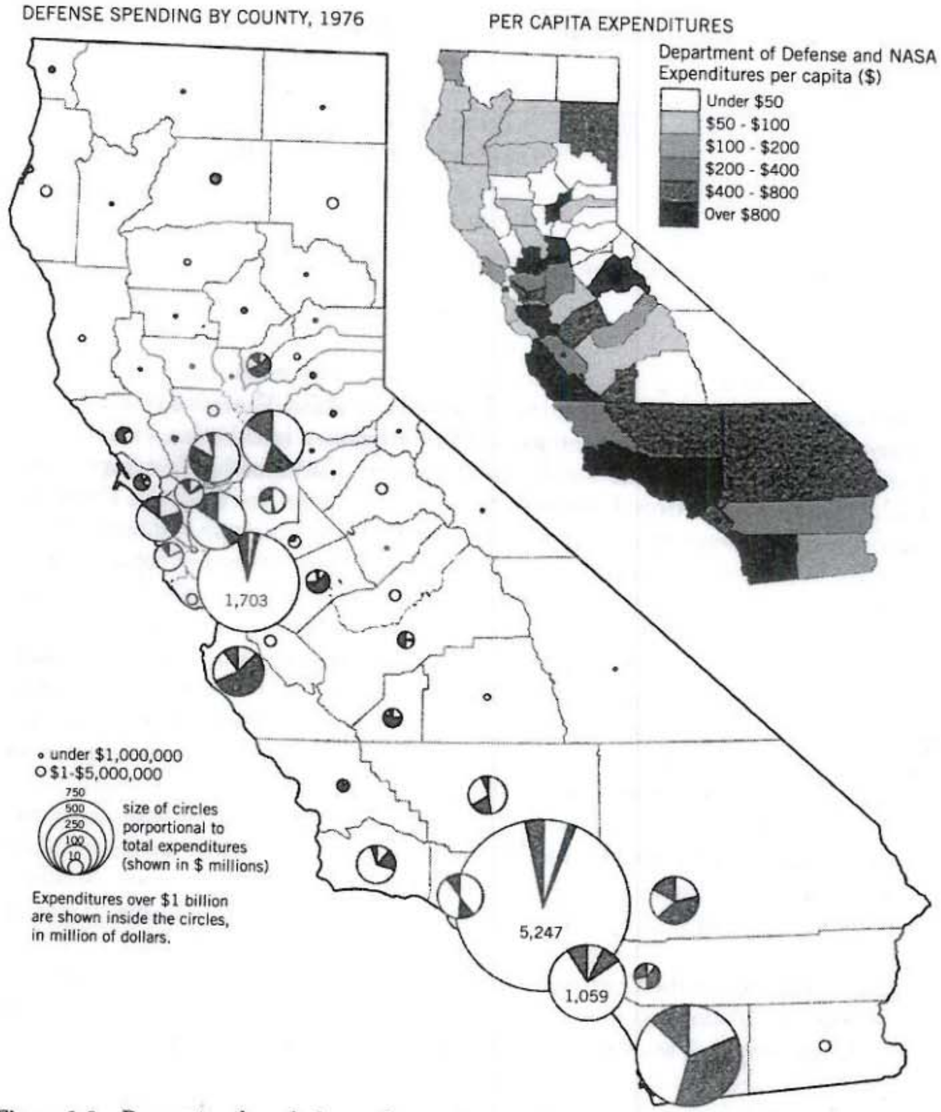

Figure 1-8 Proportional symbols use the graphic variable of size for a simple geometric symbol such as a circle. They are considered appropriate for raw measures, such as total population or total economic output. This map portrays defense spending for each county by scaling the area of the circle to 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.

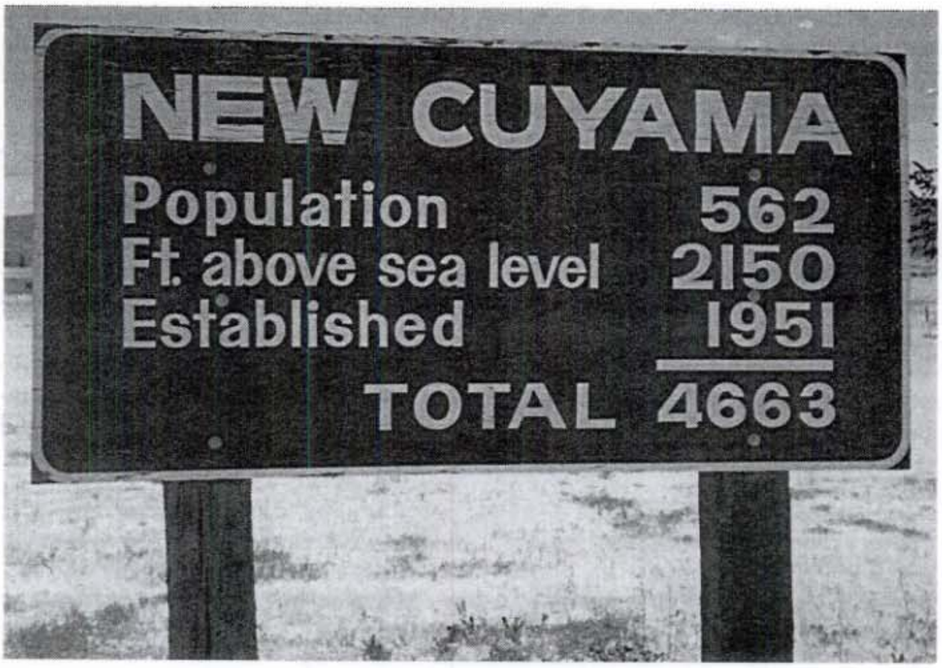

Figure 1-10 Sign posted at the entry to New Cuyama, California.

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.