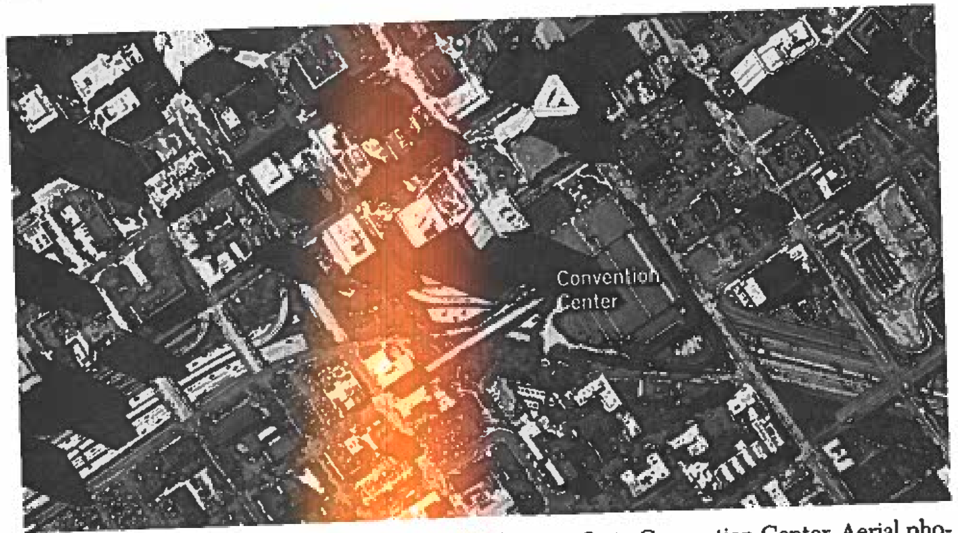

Figure 4-11 Interstate 5 disappears under Washington State Convention Center. Aerial photograph taken April 25, 1995, by Washington Department of Natural Resources.

in measurement framework are particularly apparent in the conversion between raster representations. It makes sense to code an isolated object into a raster representation using a rule such as presence/absence, where a road is shown in all cells it crosses. By contrast, an exhaustive classification is usually controlled by some area rule such as predominant type in the cell. Thus, a selection of one class from an exhaustive coverage may not match the desired object view, particularly for linear features narrower than a pixel.

An analyst must know the limits of the source materials before assuming that any particular mathematics apply directly. On first impression, these problems appear to be inaccuracies in the data, but they arise as much from the assumptions of the model and the measurement framework.

SUMMARY

The simple attribute-based operations introduced in this chapter play a critical, though far from flashy, role in GIS applications. As a general limitation, a derivative representation cannot contain higher resolution (or information content) without having some external source (and the assumptions that go along with incorporating that external information). Although they manipulate the attributes alone, these operations influence the spatial component. In some cases, even simple operations act as transformations between measurement frameworks.

CHAPTER 5

OVERLAY: INTEGRATION OF DISPARATE SOURCES

CHAPTER OVERVIEW

- Review origins of map overlay analysis.
- Describe geometric operations that establish connections between diverse sources.
- Present rules for combining attributes.

The tools presented in Chapter 4 treat attributes related to a single set of geometric objects. Many problems in geographical analysis require integration from a number of sources. The process of overlay discovers the basic spatial relationship between objects using geometric measurements. Then, attributes from the sources can be analyzed or combined. Overlay serves many purposes, but change detection and error investigations provide the simplest examples. Attribute combination also applies to any circumstance with multiple attributes, even if the attributes attach to the same geometric objects. The combination process forms the major part of this chapter.

DEVELOPMENT OF MAP OVERLAY

Geographic understanding has some contrary tendencies; at times it seems important to talk of the unity of the earth as a complete entity, and at times an analyst can cut things apart into small elements treated separately. As the discipline of geography became formalized, the sweeping holism of the nineteenth century was turned into simpler operational techniques (Harvey 1997). The metaphor of a stack of maps came into regular use by the 1930s both in Germany and the United States. In some texts, it is clear that these were transparent maps, and the analyst expected to look through them. Somewhere in this process, map overlay was born.

Connection to Photomechanical Reproduction

The development of the overlay method depended in part on the availability of geometrically stable materials. Stable material, like treated linen, was barely translucent on the best of light tables. The availability of stable photographic material (particularly graphic arts film) contributed directly to the development of cartographic methods based on photographic exposure using multiple negatives. Any graphic to be printed in multiple colors required separate impressions for each ink and thus separate originals for each color. Sherman and Tobler (1957) described a "multipurpose cartography" that broke away from the allocation of a single negative for each color plate. This technique uses dozens of overlays, each one consisting of a group of features selected for having a common attribute that might require distinct graphic representation. Photographic methods served as the model for overlay analysis before the computer database (Finch 1933; Tyrwhitt 1950; Alexander and Manheim 1962; McHarg 1969). To be able to adapt to all potential uses, the photographic process requires one overlay per category (Steinitz et al. 1976), certainly a cumbersome pile of transparencies.

There are many applications that integrate different sources of geographic information. One of these applications, generically termed *site suitability*, played an important role in developing the computer technology. Site suitability examines social, economic, physical, biological, and other criteria to locate potential sites for some purpose. Combining all these factors motivated McHarg's (1969) *Design with Nature*, the book often seen as a harbinger of GIS development. McHarg implemented his vision by combining gray-scaled maps. Each component of the environmental system resulted in a separate "map overlay" in a literal sense—each map was made on a separate transparency. The cartographer used dark shading in the areas considered sensitive and left the rest transparent. All these layers were combined by placing the transparencies on top of each other (in registration) on a light table (Figure 5-1). Once all the transparent overlays were assembled, visual interpretation could distinguish the areas of least sensitivity. McHarg (1969, p. 34) considered and rejected most of the mathematical solutions that his followers later adopted.

McHarg's (1969) book was enormously popular and very quickly influenced many aspects of North American and European politics. In particular, the concept of screening through map overlay was absorbed as the fundamental logic for a whole generation of regulations, beginning with the U.S. National Environmental Policy Act. Of course, McHarg was hardly the first person to use map overlay. Steinitz and his students (1976) recount a rich history of the map overlay method dating back to the late nineteenth century. For example, Manning (1913), a landscape architect, produced a plan for Billerica, Massachusetts, in 1912 that combined four maps: soils,

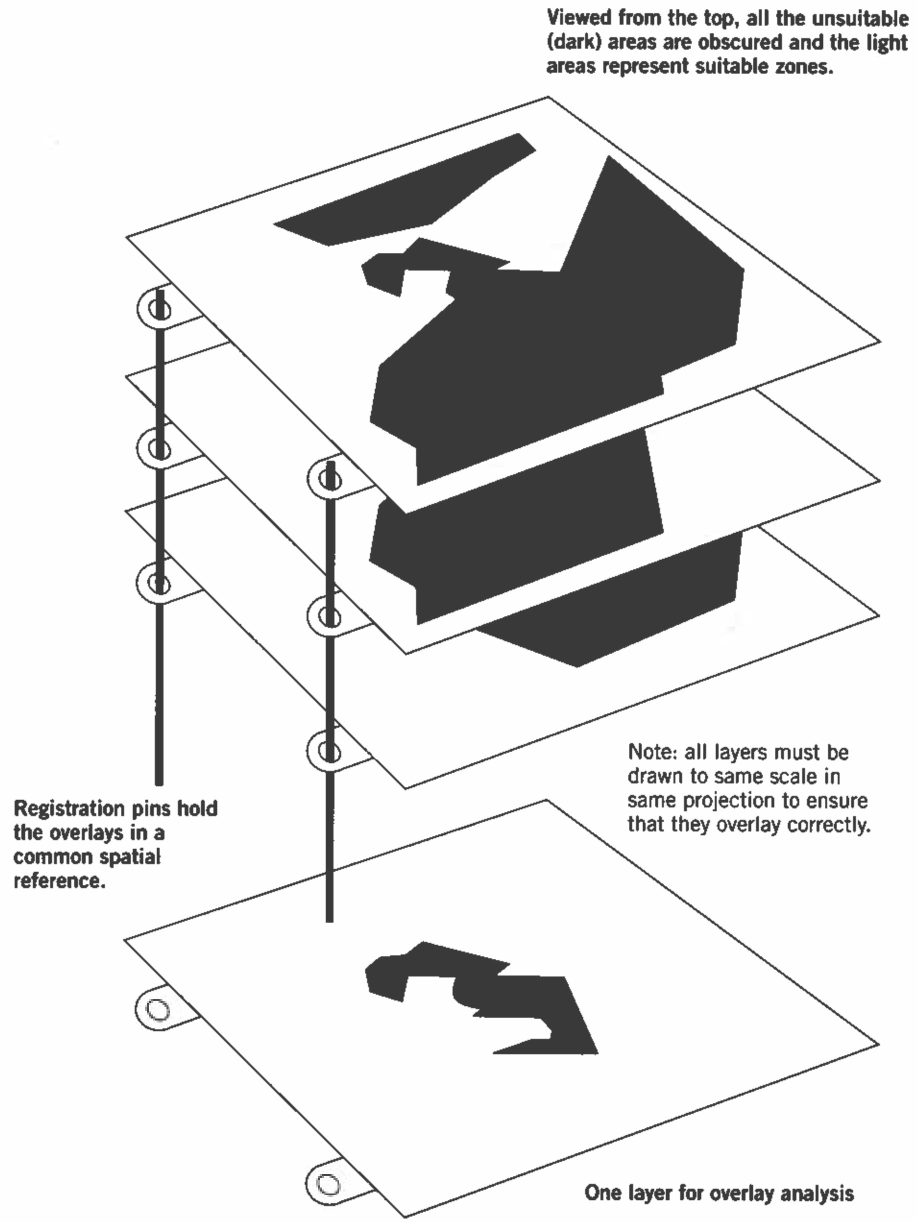

Figure 5-1 Manual overlay process: Transparent "overlays" are registered and combined to make a composite suitability mask.

vegetation, topography, and land use. Lewis (1963) produced recreation plans for Illinois and Wisconsin as the composite of various environmental factors. Thus, map composites were well established, though quite labor intensive and error prone, prior to the analytical solutions possible using a computer-based GIS.

The overlay problem became one of the early objectives for GIS software development. Programs for performing grid-based overlay analysis appeared in the 1960s and came quickly into professional use for the early wave of environmental impact analysis (on nuclear power stations and the like). The Map Analysis Package (developed by Tomlin in the mid-1970s) became a standard for instruction of GIS, and it has been the inspiration for similar packages used in many classrooms (IDRISI, MAP II, ARC/Grid, and others). There were some early algorithms to overlay vector representations (such as MAP/MODEL in 1969, and the original Canada GIS), but not until the late 1970s did prototype systems (such as MOSS and Harvard's ODYSSEY) deliver practical overlay solutions. By the 1980s, commercial software packages built their basic functions around polygon overlay processors (such as ARC/INFO, MGE, DeltaMap, and others). Thus, overlay played a role in developing both raster and vector software strategies.

THE OVERLAY OPERATION

The reason overlay played a key role is that most applications of geographic information must integrate information from different sources. In the terminology of relational databases, map overlay serves as a kind of join, the procedure that links two tables based upon a common key (see Figure 4-9), but the use of the database term may make it seem easier than it really is. Normal joins (as described in Chapter 4) operate through a *foreign key,* a value stored in one table that creates a relationship to a record in another table. Map overlay starts with no such direct correspondence between the layers. Instead it uses the geometric description to discover the connections. The digital map overlay procedure, whether raster or vector, depends on the absolute location of each feature. The spatial reference system provides the geometric basis to connect the two sources. Thus, the first issue in describing the overlay operation involves the spatial reference system, followed by geometric intersection processing, finally leading to attribute combination.

Registration: A Universal Requirement

In the manual graphic form, overlay depends on physical registration of the transparencies. The graphic arts industry uses a range of devices to maintain registration, such as holes punched in the overlay that match a set of pins. These physical techniques work well for multiple overlays in photographic cartography as long as they remain at a consistent scale, projection, and area of coverage. In contrast, digital overlay is not restricted to such consistent sources. Most applications involve sources collected using different methods and at different scales (as described for Dane County in the Introduction). In the era of manual methods, map overlay implied recompiling each overlay onto a common base. The redrafting process, being a visual interpretation, could adjust the layers to become more consistent. Instead of an exterior framework of pin bars, registration came from the content of the map.

Manual integration consumes time, and redrafting introduces additional error. As digital methods replaced manual overlay, the computer offered the chance to escape many of the limitations of the physical media. The external framework of registration moved from the pin bar to the abstract mathematics of a coordinate system. The registration transformations described in Chapter 3 apply to the preparations for overlay. For the La Selva project, the whole reason for the coordinate manipulations was to merge the land use information on the old map with the more recent sources.

Raster Implementations of Overlay

The overlay procedure is actually an inherent feature of a space-controlled representation. Once two maps are rendered into the same grid system, the pixel becomes the base object for both (Figure 5-2). Thus, the methods that operate on pairs of attributes (Figure 4-3) become overlay operations with very little difficulty. The computation of overlay results simply translates into **Boolean** or arithmetic operations on a cell-by-cell basis.

Of course, the trick is to get the geographic information into this common grid reference. First, the operation requires the two grid systems to be identical. A different projection, sampling interval, or even rotation will lead to a mismatch. The process of *resampling* an image (or a cellular database) may be a necessary prerequisite for overlay. In addition, not all geographic information converts from its measurement framework into the grid measurement framework with equal ease. This is particularly true for choropleth tabulations. For instance, a census tract may be assigned a total population, but it is rather confusing to propagate this value to all the cells covered by that census tract without understanding that the measurement is not really located at the cell by itself. Due to these complexities, transformations between measurement frameworks will be discussed in Chapter 9.

Once a raster database organizes a number of layers into the same geometric framework, overlay analysis requires only some relatively simple calculations on the attributes of each cell. Most raster software packages begin with the operators described in Chapter 4 that combine pairs of attribute values. Despite the number of possible operators, they are all rather primitive, so it usually takes a few steps to get from the source layers to the desired result.

Figure 5-3 shows a hypothetical overlay analysis. Three source maps are represented in a common grid by applying one of the rules for spatial control. Attribute processing might include a threshold that classifies the original measurements into two categories, such as suitable and unsuitable. The binary maps can then be combined with Boolean operators, such as AND (set intersection) and OR (set union). These two operations are frequently required by the logic of environmental regula-

Boolean algebra: System of operations applied to sets (and logical propositions); Boolean variables are zero or one, hence strongly connected to modern computing; originated by George Boole in 1847.

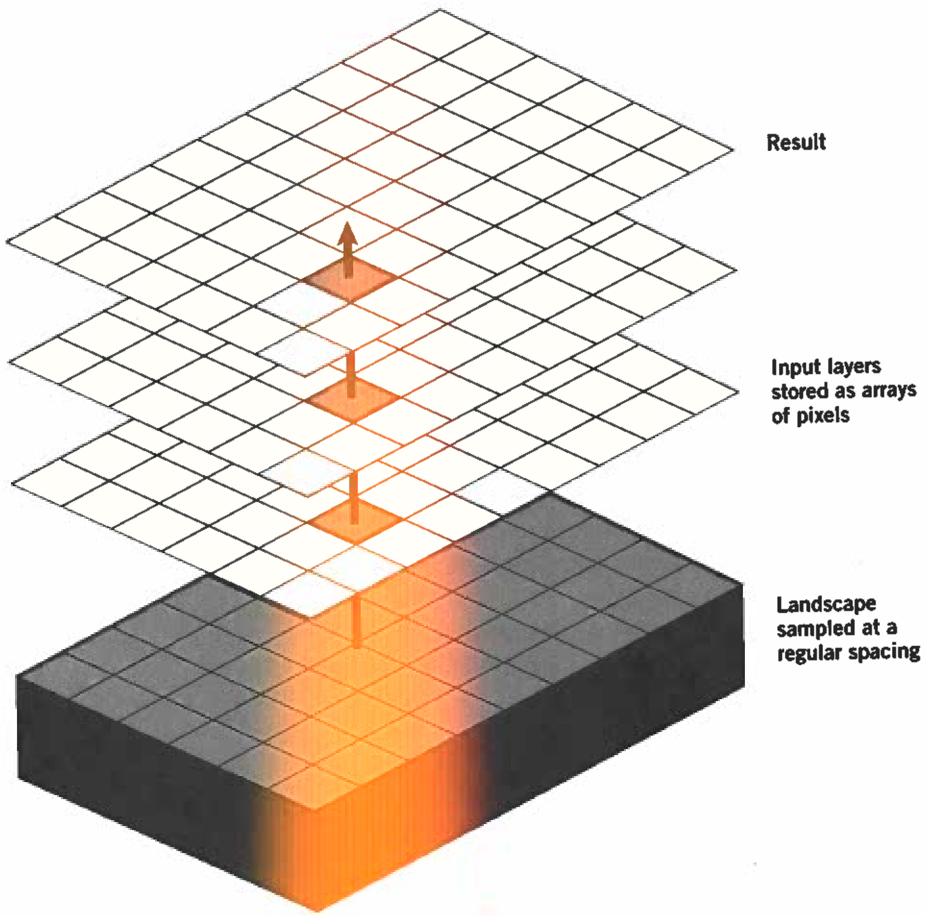

Figure 5-2 Conceptual view of the overlay operation for raster structures. Any mathematical operator can be applied to a pair of values, if justified by the measurements. Results become a new map layer, permitting arbitrarily complex results from simple operations.

tions and suitability analyses. Disqualifying factors often combine with the others using OR because an area is excluded if disqualified by any single factor. If, on the other hand, certain conditions must occur simultaneously, the AND function is the correct choice.

An alternative approach creates a new category for every distinct combination in the pair of maps [LocalCombination in Tomlin (1990); COMBINE in Map II; see Figure 5-4]. This method uses the cross-tabulation operation introduced in Chapter 4. The resultant categories can be grouped into any combinations desired. Finding all the possible combinations involves extra work, and software will limit the number of categories it can handle. However, as an exploratory measure, the complete combination forces the analyst to consider all interactions. The key shown in Figure 5-4 links the new categories back to the categories on the original. Without a rela-

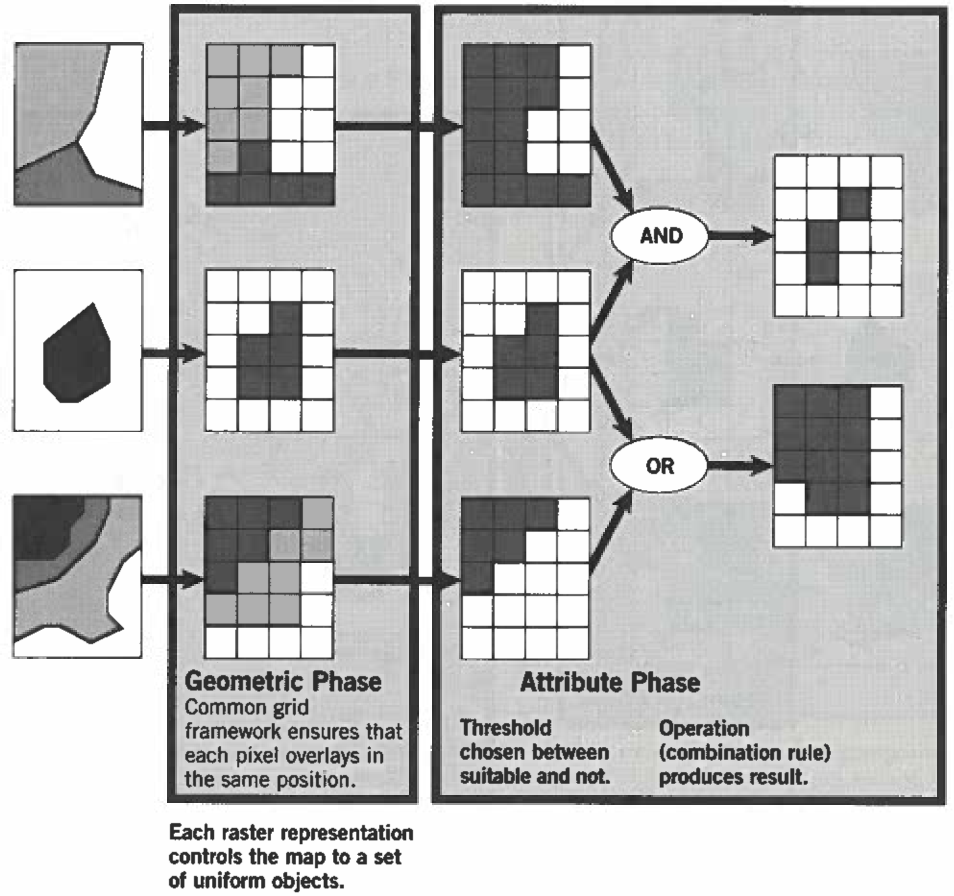

Figure 5-3 Operations for Boolean combination of raster categories. For this illustration, the three sources are converted to raster format using a center-point rule. The output shows two queries, each using two of the input sources.

tional database to manage these critical linkages, the user often has to perform the analytical extractions without any assistance from the software.

These alternative approaches work reasonably well inside the constraints of the raster system. Even a simple analysis can be performed using many different combinations of the basic operations.

Vector Implementations of Overlay

The vector solution to the overlay problem is not as direct as the raster approach. The representation does not guarantee a set of common objects; they must be created geometrically. "Polygon" overlay produces a composite geometric representation where each area has a key to the attribute tables for the two source layers (Figure

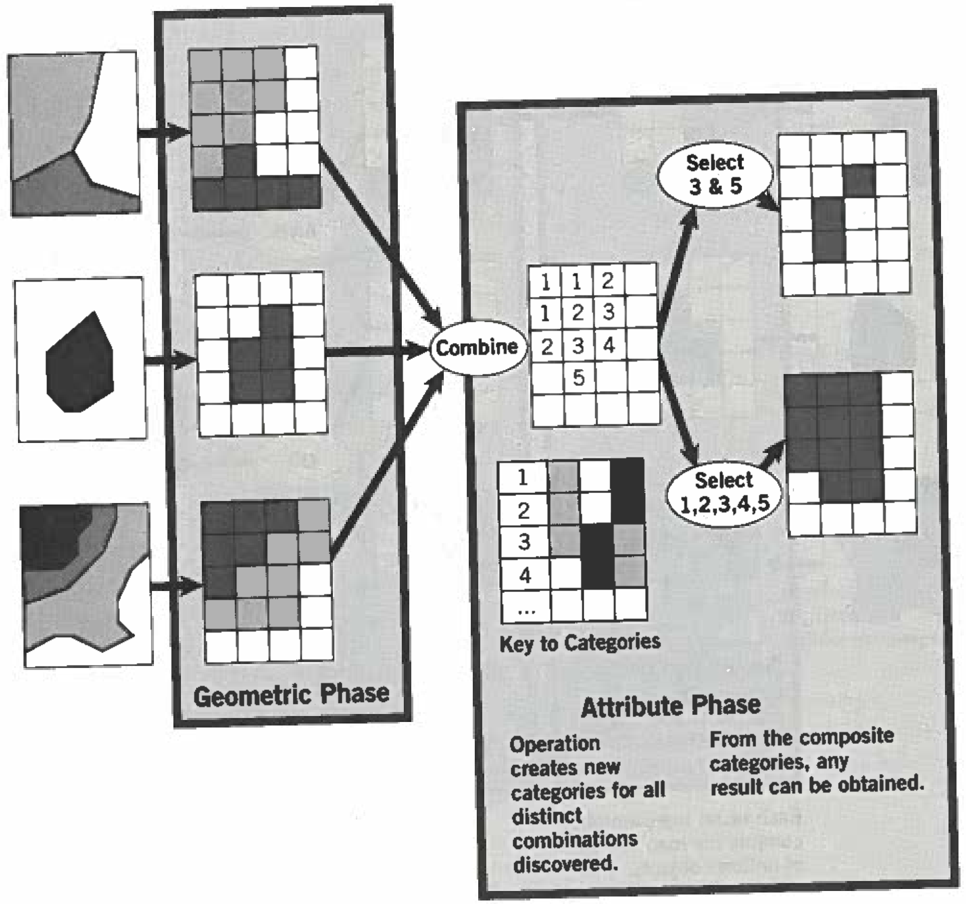

Figure 5-4 Composite combination approach for raster overlay.

5-5). Similar procedures can also combine points and lines with polygons. Following the geometric steps, the attributes of the objects may be combined.

Geometric Intersection Processing Intuitively, the polygon overlay operation involves a pairwise comparison between the objects in the two sources. Polygons that lie outside each other clearly do not require any effort, and polygons totally enclosed by another are not very complicated. Complicated processing is required when they overlap partially (Figure 5-6). The process first finds all intersections between the boundary lines of the polygons. These intersections become new nodes in the composite topology. For this stage, a topological database built of chains eliminates tedious calculations involving shared boundaries of neighboring polygons. With the newly constructed topological network, the procedure must then label all the new objects with a unique polygon identifier. A new attribute table provides access to the

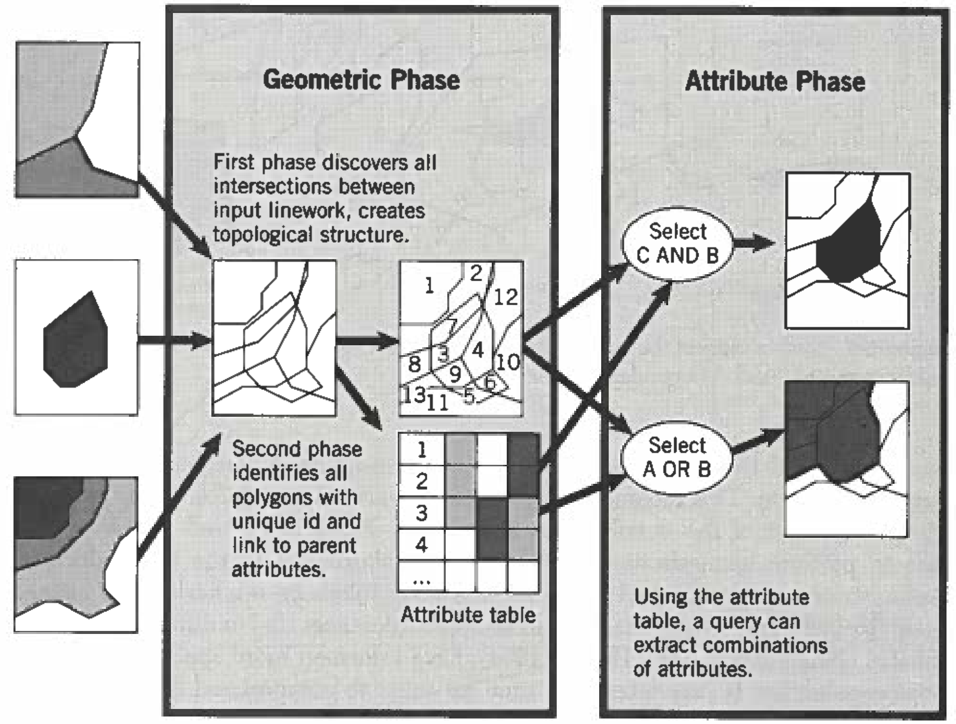

Figure 5-5 Vector overlay: schematic view. A geometric processor constructs a composite topological structure first by finding all intersections, then by labeling all the polygons with a unique identifier and linkage to the source attribute tables. This creates a single coverage with links to all the attributes. Attribute-based operations can produce results from the table and apply to the coverage.

attribute tables from the sources. Efficient overlay processing requires careful use of geometric neighborhoods to limit calculations.

If the overlay calculation had infinite resolution, a single feature appearing as a boundary on two independent sources would be unlikely to coincide perfectly. The result is a flurry of small objects, often called *slivers*, created by slight differences in the representation of boundaries that should have been the same (Figure 5-7a). Slivers have been known to be a major problem ever since the earliest days of the Canada Geographical Information System (CGIS); the vast majority of polygons in the CGIS database were trivial objects created through overlay (Goodchild 1978). The word *sliver* implies thin shapes, though some slivers are quite compact (Lester and Chrisman 1991). Slivers can occur for several reasons, including misregistration of sources, inaccuracy of one source compared to the other, temporal differences between the sources, and more. Most current software offers at least one mechanism to combat slivers caused by overlay. One method was originally termed an *epsilon filter* (Dougenik 1980) to acknowledge the work of Julian Perkal (1956),

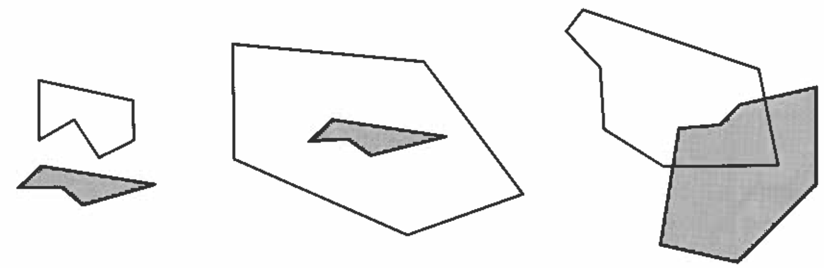

Figure 5-6 Simple cases in the calculation of overlay: (a) Two polygons totally distinct; (b) one encloses another; and (c) boundaries intersect.

who recognized the basic mathematics, but it is now more commonly called the **fuzzy tolerance.** This technique can remove slivers (Figure 5-7b), but it can also filter away most of the detail in the boundaries. Perkal conceived of this operation to perform generalization, not overlay, an alternative to the line reduction demonstrated in Figures 3-11 and 3-12. The fuzzy tolerance is a fairly blunt instrument. Research developments, such as multiple tolerances tied to different error estimates (Pullar 1991, 1993; Harvey 1994), have informed more sophisticated software capabilities. It may take some time for users to comprehend how to control these capabilities.

The geometric integration of polygons overlaid on polygons represents the most complex case. Software built for this job can easily determine *point-in-polygon* or line-in-polygon as well. In the case of points, no new topology is required, but the processor must still use the geometry to determine which points belong in which polygons. If the boundary lines are treated as imprecise using the epsilon model, then some points may fall not inside a polygon but on a boundary line. For example, Blakemore (1984) found that only 55% of industrial locations in Britain could be safely assigned to political jurisdictions, considering the probable error in the polygon boundaries. The other 45% were too close to the boundaries (0.7 km in this case) to be certain.

Attribute Handling Using Results of Overlay Once the geometry has been used to construct a new common framework, attributes are attached to the objects created by the overlay. The attribute handling following point-in-polygon provides a clear example of the role of measurement frameworks in this process. After the geometric

Fuzzy tolerance: A distance within which intersections and points will be treated as coincident. To be processed correctly, the fuzzy tolerance cannot be handled immediately (otherwise a point might be moved twice and beyond its original tolerance). A "cluster" of points must be grouped so that no point is moved more than the tolerance.

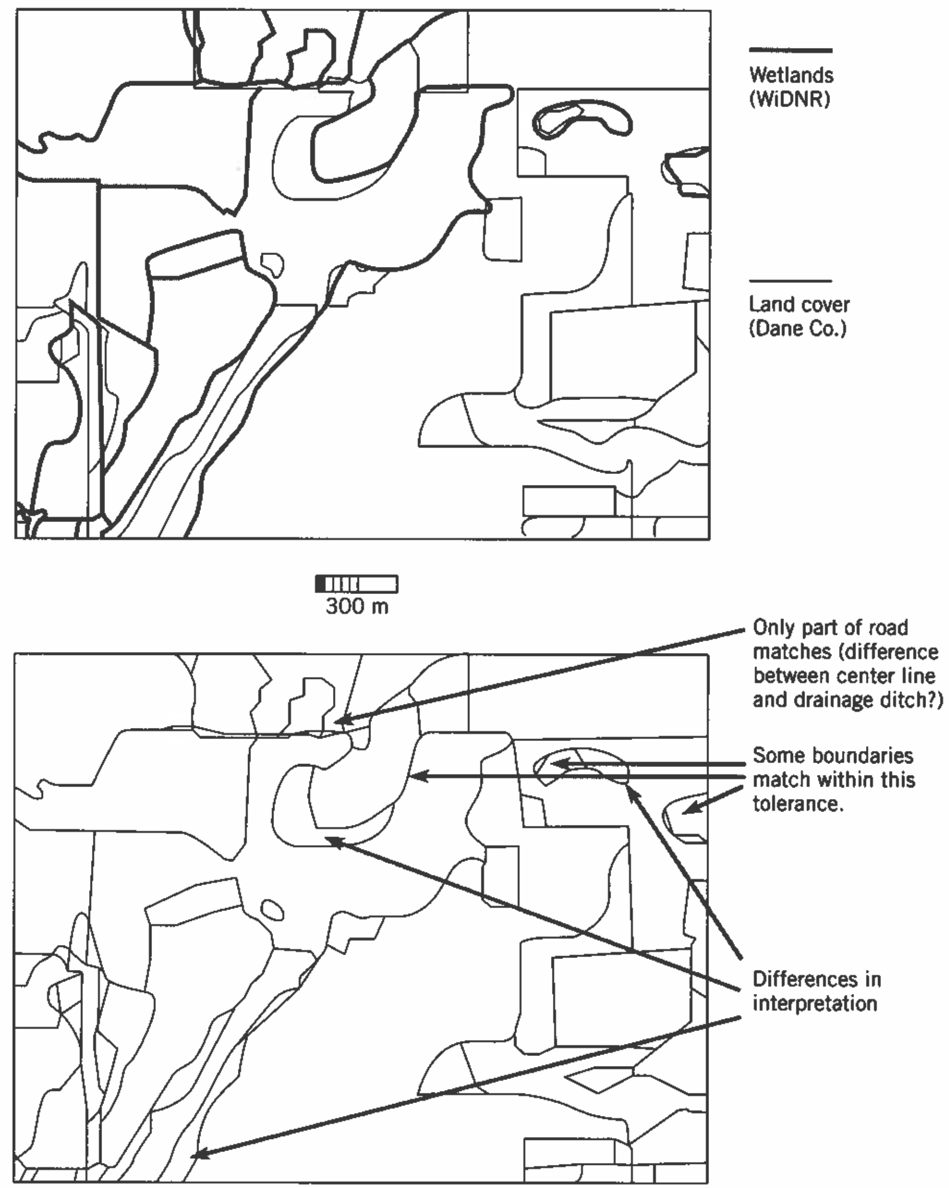

Figure 5-7 Slivers from overlay of land cover and wetlands maps for a portion of Oregon, Wisconsin: (a) Original source layers, both derived from photointerpretation (wetlands interpreted at 1 : 24,000 scale; land cover interpreted at 1 : 15,840; in both classifications, farm fields and wetlands should have similar boundaries); (b) overlay with 30-meter tolerance. (Shaded box in bar scale shows tolerance distance.)

process assigns each point to a polygon, the attributes can be interpreted in different ways. The simplest assigns polygon attributes to the points, as Blakemore assigned the political code (area) to each industrial plant (point). Alternatively, as described in the digitizing procedures (Chapter 3), a single label point located inside each polygon can be used to identify the polygon. The process transfers point attributes to the polygon. However, if multiple points occur inside the polygons, some other rule must apply. Adopting a choropleth framework, a polygon attribute can be obtained by counting the number of points in the polygon or the sum of their attribute values. Using Blakemore's data, the number of plants or the total industrial employment could be tabulated for each political unit.

The fragmentation of polygon overlay makes polygon-on-polygon more complex. Areal attributes must be examined closely to ensure that their measurement framework has not been misinterpreted. Overlay can split areas into many pieces. Typically, the attribute values are simply copied from the parent or referenced indirectly using the parent identifier as a foreign key into the parent table. Either method assumes that the attribute is uniformly distributed in the parent object and that the attribute applies equally to every piece split from the parent. This assumption is defensible, within limits, for most categories and some derived ratios, but it does not apply to extensive measures, such as population. The appropriate procedure for extensive measures requires areal interpolation, discussed in Chapter 9.

Once the attributes are assigned to the composite, the tools used for attributes of a single map (as described in Chapter 4) apply. The new attribute table allows a query to reference characteristics of all the source layers. A new attribute can be calculated using measurements from different layers, choosing appropriate combination techniques from those presented later in this chapter.

The geometric operation is typically the time-consuming part of vector overlay and is done as a preprocessing stage without much operator interaction but at some cost in time and storage. Once the geometric overlay is constructed, many different attribute combinations can be investigated using simple tools and at low cost. This balance of costs would argue in favor of overlaying all layers into one integrated coverage as a part of constructing the database (Chrisman 1975; Frank and Kuhn 1986; Herring 1987). For certain kinds of analytical projects, such as site selection, this may be the appropriate strategy. One integrated coverage does force the recognition of common features to combat slivers. While it may sound efficient, an integrated database makes it much more difficult to maintain each layer's internal consistency as changes occur. In a less centralized institutional arrangement, each custodial organization should maintain its own layer, then overlay them with the current version of other layers as needed for analytical purposes.

If the overlay serves a specific analytical query, the sequence of operations should be reconsidered. Comprehensive overlay produces a complete composite similar to the raster combination operator described in the previous section. Instead of processing all these boundaries in order to ignore them later, it makes sense to aggregate each coverage into the categories of the query before performing the overlay. Simplifying each layer to the boundaries necessary for the specific purpose also has the advantage of reducing the chance that slivers and fuzzy tolerance effects will be caused by the nonessential boundaries. Advanced spatial database systems offer complex intersection processing through the query language, not as distinct processing steps.

Comparisons of Performance and Capabilities

Despite the trenchant rhetoric supporting vector or raster approaches, the analytical method for overlay does not differ dramatically. Both representations use geometric relationships to create a cross reference between attributes. In the raster case, the integration occurs in the creation of the raster, while in the vector case, it has to be discovered for each combination. Folk wisdom contends that "raster is faster, but vector is correcter." This sounds catchy, but it is not necessarily true. This adage came into common usage at a conference in the early 1980s. A raster plotter manufacturer handed out buttons reading "Raster is Faster," probably a true statement when confined to the printing of graphic images compared to a pen plotter. The second part of the phrase was added to homemade buttons by various conference attendees. When applied to analytical operations, raster may seem faster, in that the only step called "overlay" is done after the integrated database is constructed. If the process of creating the raster is included and the resolution is strictly comparable (causing rather large increases in the number of pixels), however, the speed advantage might be diminished. Similarly, the vector method may preserve the cartographic crenulations of boundaries at a level of resolution far beyond the reliability of the boundaries. All the effort and expense may be driven by cartographic convention, not a careful analysis of the accuracy of the information. Thus, a vector overlay may not be any more correct. Both raster and vector analysis can be used properly or improperly.

COMBINING ATTRIBUTES AFTER OVERLAY

The overlay process, whether performed in a raster or vector representation, provides access to all the attribute values that occur at one location. For most applications, the geometric processing is simply a mechanism to allow attribute comparisons. A number of different rules can be applied to combine the attribute values placed in contact by overlay. It is common to divide techniques based on the type of mathematics used, and thus to talk about Boolean operations, arithmetic, and so on. These groupings may reflect programming issues, but they obscure differences in information content.

At a generalized level, the possible forms of attribute combination can be organized into four groups (Figure 5-8). The simplest method preserves all the source material; it supports direct analysis of the results by *enumerating* all the combinations discovered. The next method selects one value from those available. This will be termed a *dominance* rule because other attributes are ignored. The next group uses

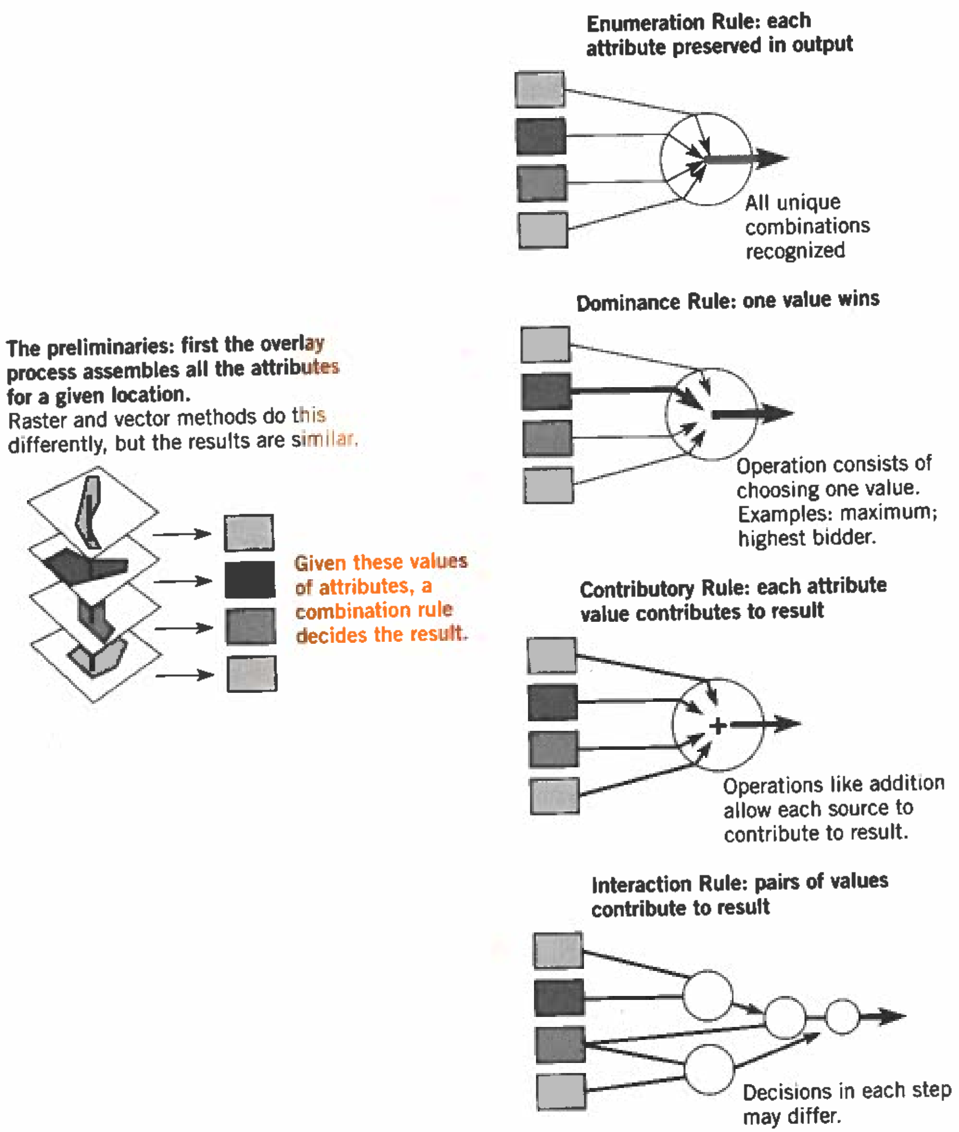

Figure 5-8 Combination rules classified in terms of information use.

each layer's attribute value to create a composite result, often using a mathematical operation such as addition. Because each value contributes to the combination without regard for the others, the rule will be termed *contributory*. The fourth group exploits the interaction between pairs of values. These will be termed *interaction* rules. Each of these four groups can have many variants; those listed in Table 5-1 are intended as illustrations, not an exhaustive set. In practice, combination methods may borrow from a number of groups to solve a problem.

TABLE 5-1 Attribute Combination Methods Following Overlay

Enumeration Rules	
Cross-tabulation	All combinations preserved
Change detection	Special case for repeated measurement
Dominance Rules	
Exclusionary screening	One strike and you're out
Exclusionary ranking	Extreme value from rankings
Highest bid	Extreme value from continuous data
Highest bidder	Records identity of extreme value
	Contributory Rules
Boolean operations	Logical combinations of binary variables and sets (queries)
Voting tabulation	Sum of binary exclusions
Weighted voting	Weighted sum of binary factors
Linear combination	Sum of "ratings" (mean, etc.)
Weighted linear combination	Weighting and rating game
Product	Multiplication of factors
	Interaction Rules
Contingent weighting	Linear combination weighted by other factors
Integrated survey	Informal judgment, "gestalt" interpretation
Rules of combination	Formal interaction tables

Previous Descriptions of Overlay Operations

In the landscape planning literature, various procedures to combine suitability information are listed without an organizing principle. In the early period of GIS development, Hopkins (1977) presented a rather complete consideration of the possibilities for map combination, guided by Stevens' levels of measurement. At the time, this helped organize the possibilities, but the level of attribute measurement is not enough to understand the choices.

Another, more recent stream of research has relied on the "multicriteria decision-making" literature. This interdisciplinary field has worked on problems of rendering a decision for environmental problems that involve more complex criteria than traditional optimization methods from operations research (Voogd 1983; Hobbs 1985). There have been some examples of applications of these multicriteria methods in GIS (Carver 1991; McCartney and Thrall 1991). The complex measurement basis of geographic information requires some care in adapting multicriteria techniques to the overlay combination problem.

Enumeration Rules

Presented with a richness of multiple attributes, one reaction is to preserve them all. In one sense, this combines attributes by respecting all the combinations. Because they list all the combinations, they are termed *enumeration* rules. In most situations, a complete enumeration simply postpones the hard choices. The methods in Chapter 4 will be required to extract a specific result.

In describing the vector technique of performing overlays, the process acts like a join that combines attributes from both sources in a combined table. It takes very little additional work to construct an attribute that combines both sources. In an earlier age these were often multidigit composite codes. In the raster approach, there is a similar tool that gives a distinct identifier to each pair of values discovered. This kind of code is harder to read, but no different in information content. In either case, the clear distinguishing feature is that the source attributes can be reconstructed without loss.

The number of potential categories does increase rather rapidly with more factors. In the worst case, an uncorrelated environment, the number of overlaid categories will be the product of the numbers of categories in the source layers. After a few overlays, such a product is frighteningly large. Due to associations in the environment, the number of categories is much less likely to rise that quickly. For example, in a four-factor inventory of forest characteristics in northern California, the product of the categories was 12 × 5 × 6 × 3 = 1080 potential categories. The actual 5-acre minimum mapping unit coverage only used 105 of these possibilities. Smoothed to 40 acres, only 64 survived. While this is still a large number, it shows that enumerating the results does not expand quite as fast in practice.

Detecting Differences If two categorical coverages of the same region are available for two different times, an overlay can detect the differences by enumerating all the attribute combinations. If the coverages used identical categories, the analysis is particularly easy. Those areas with the same category did not change, while those with different classifications appear to have changed. For example, a land use map for Cwmbran (on the edge of Wales in the United Kingdom) was created in 1967 as part of the Second Land Use Survey (Coleman 1961). I created another land use map in 1981 using similar categories and on the same base (Figure 2-9). The specific transitions from category to category can be presented as a transition matrix (Table 5-2). Note that 20 of the 81 cells in this table are zero; those particular changes did not occur in this time period. The exact same procedure can be used for error testing (Chrisman and Lester 1991) by overlaying two sources obtained at the same time, instead of reporting changes. In this case, the matrix is usually called a *misclassification matrix*.

The results of the overlay provide the basis for analysis. It may be important to know that residential uses expanded into 324 hectares of pasture land. Of course, neither map is perfectly error free, so the apparent change includes some amount of error mixed in with the actual change. For example, the water category was the same reservoir in both time periods, but about 0.7 hectares of slivers accumulated from two separate attempts to delineate the same feature. The overlay applied a 20-meter tol-

TABLE 5-2 Transition Matrix: Land Use Change 1967–1981 Cwmbran, Wales

1967/81	Resid.	Indust.	Open	Pasture	Crops	Woods	Heath	Water	Trans.
Residential	689.2	23.0	20.8	7.7	0.0	1.2	0.8	0.0	0.7
Industrial	30.3	215.0	3.1	5.7	0.0	10.2	9.1	0.0	32.4
Open Urban	17.6	13.9	157.9	9.4	0.0	2.4	1.6	0.0	0.0
Pasture	324.2	98.6	167.4	2715.0	150.1	101.6	23.3	0.1	75.9
Crops	2.1	2.7	1.5	197.7	20.3	7.3	0.0	0.0	1.6
Woods	13.9	7.0	19.3	46.5	1.5	337.8	12.8	0.5	3.6
Heath	17.0	2.9	1.0	50.4	0.0	7.9	490.3	0.1	0.0
Water	0.0	0.0	0.0	0.0	0.0	0.6	0.1	134.5	0.0
Transition	28.9	38.8	26.1	0.3	0.0	0.0	1.6	0.0	1.4

Figures in hectares tabulated from results of polygon overlay of land-use maps compiled at 1:25,000 for an 8 × 8 km around Cwmbran, Gwent, Wales. Categories grouped from classes used in Coleman survey (1967) and Anderson codes (1981). Residential includes commercial and retail, industrial includes transportation and mining (Cwmbran had steel mills and coal for the steel facilities). Heath includes gorse, bracken, and tundra vegetation not of great value for pasture.

Source: Chrisman (1982a).

erance to avoid slivers, but this does not cover all the sources of error. Despite the potential for error, change detection remains a major application for overlay processing.

Interpreting Differences In both change detection and overlay accuracy testing, overlay creates a cross-tabulation as the primary raw result (for example Table 5-2). Shown in tabular form, these matrices summarize the area in all possible combinations of the categories. A number of techniques can analyze these tables. The **diagonal** of the matrix is the focus for some of the simpler techniques. The change matrix can be simplified into categories of change and no change, based on the diagonal or off-diagonal elements. In Cwmbran, these two categories provide the first result desired in change detection (Figure 5-9). This new attribute is a kind of interaction rule because each resulting value depends on pairs of values.

In other situations, the analyst does not want an attribute for each object, but an overall summary of the comparison. The percentage correct (applied to error matrices) is simply the sum of the diagonal divided by the total frequency (in the overlay case, the area of the region). This percentage is easy to understand, but it may not represent the only (or even the best) measure of the matrix. Differences in the number of categories and in the distribution among these categories can change the meaning of this single figure. As a remedy, some suggest the use of **Cohen's kappa,** a measure that deflates the percentage correct by the amount that would be expected

Diagonal: Cells in a square matrix whose row and column indices are the same. In a transition matrix, they represent no change; in an error matrix, they represent no error.

Cohen's kappa: A measure of agreement between two classifications. Defined as (observed accuracy − change agreement)/(1 − chance agreement) where the chance agreement is estimated by the cross product of marginal frequencies (statistical independence model).

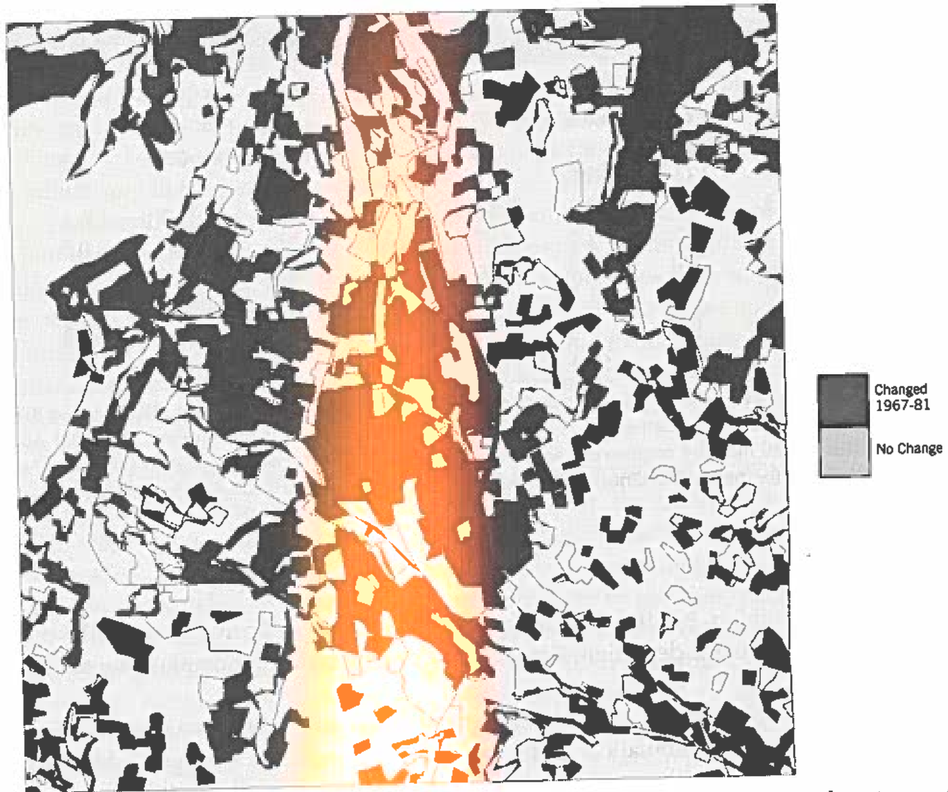

Figure 5-9 Change detection for Cwmbran, Gwent (UK). Area covered is 8 × 8 km. Areas of change between 1967 and 1981 land-use maps.

by chance (Rosenfield and Fitzpatrick-Lins 1986; Congalton 1991). This measure represents a simple case of applying statistical methods, such as discrete multivariate analysis, to the matrices that summarize the results of overlay. Unfortunately, there are some large differences between regular statistical measurement frameworks and the circumstances that produce these matrices. The model for chance agreement is statistical independence, which is not really the baseline expected either within or between coverages. Both change detection and error analysis begin with a presumption of no change or no error. In addition, the quantities in the matrix represent areas, a ratio measure with an arbitrary unit of measure, not counts of discrete independent objects. By definition, categorical coverages expect extremely strong correlation between adjacent points, so the model behind kappa (and other discrete multivariate techniques) does not fully apply.

While percent correct or kappa provide a single index, other techniques produce two measures to describe dependencies. Each entry in a change or error matrix can be compared to its row and column total, giving a proportion. In the remote-sensing accuracy literature, the proportion comparing a cell to the total ground truth for that class is called the *producer's accuracy*, while the figure divided by the total from the satellite source is called the *user's accuracy*. In a change matrix, these have to be interpreted, respectively, as the proportion unchanged at the earlier time period (row) and the proportion unchanged at the later time period (column). Zaslavsky (1995) has adopted similar measures inside a technique called determinacy analysis.

Explaining Selective Harvest Patterns at La Selva

The analytical goal of the La Selva project was to study the areas of the Sarapiqui Annex that had been selectively logged before they were purchased by the Reserve. These areas should not be considered as undisturbed primary forest in the various scientific studies under way and may require restoration efforts. One overlay analysis combined the soils coverage with the land use layer. Of the area that was in the T/H/L Tropaquept soil class, 82% had been selectively logged, while only about 40% of the other soils categories had been selectively logged (Figure 5-10). This association is helpful, but it only accounts for 4% of the total area that was selectively logged. A comprehensive explanation needs to be both accurate and complete (a high proportion in both row and column directions). The factors that provided the best explanation for logging were generated by a distance calculation, so they will be described in the next chapter.

By adding explanatory categories, determinacy analysis tries to obtain the most accurate and complete statements to describe the results in a matrix. Because it does not use the condition of independence as a benchmark, it can be applied to error matrices and change matrices without requiring statistical assumptions.

Dominance Rules

A dominance rule determines the result of combination by selecting a single value from those found at the same place. The selection is governed by a set of external rules, not the combination of values available. The one value selected then dominates all the others. Various possibilities arise based on the rules that choose the single value from those available.

Exclusionary Screening The simplest form of dominance is exclusionary screening, a rule that could be summarized: "One strike and you're out." Screening imposes a binary vision of the world, assigning priority to any detrimental attribute. Rephrased in terms of Boolean logic, exclusion uses an OR operator to combine categories lumped into an excluded class from a series of maps. This rule may be appropriate for factors that by themselves are so damaging that there is no need for fur-

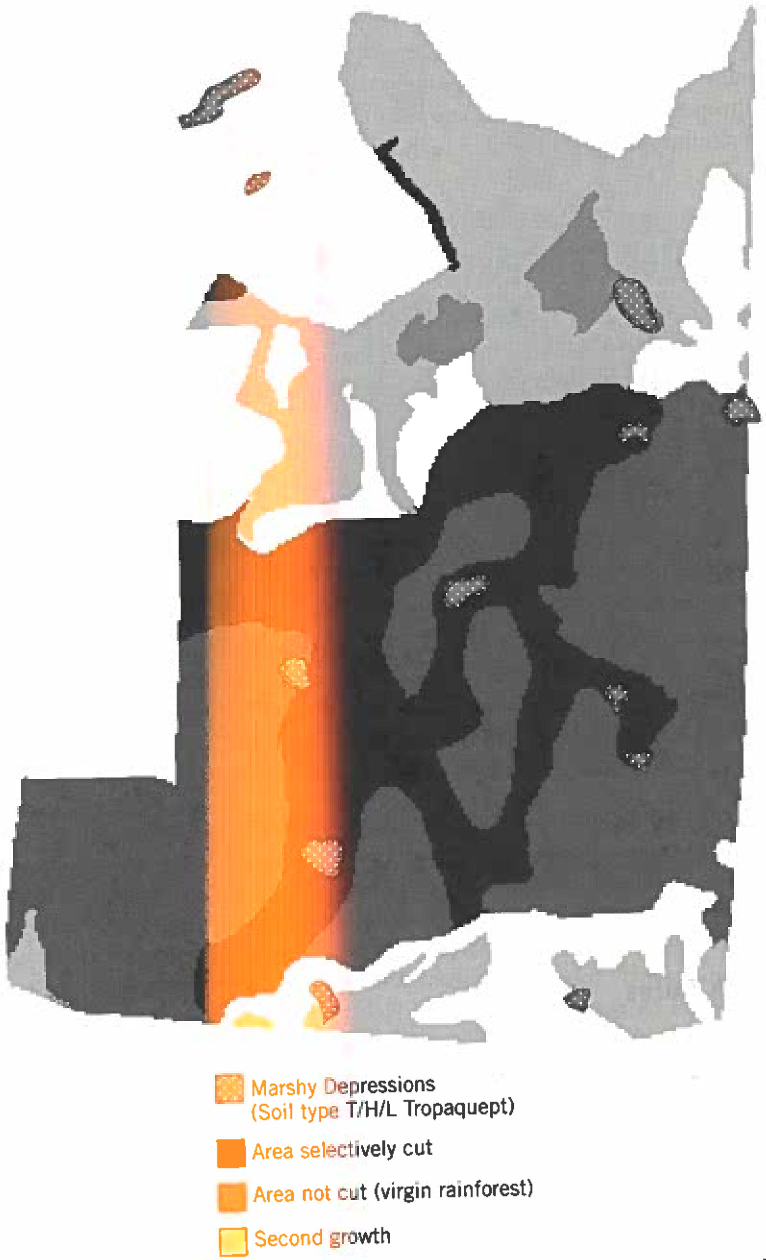

Figure 5-10 Overlay of selected soils and land uses, Sarapuiqui Annex, La Selva Reserve.

ther study. If a factor could be mitigated by some technology or could be neutralized by some other factor, then the method may exclude some areas falsely. This dominance rule also gives the same result for those areas with a single detrimental factor and for those areas with all the possible negative factors.

The Pennsylvania project to site a disposal site for low-level radioactive waste **(LLRW)** provides one example of a siting project centered around a GIS and the overlay tool. The first phase of this project disqualifies any portion of the state that exhibits one of 18 characteristics (Table 5-3). Some criteria are simple, and some are

TABLE 5-3 Disqualifying Criteria for a Low-Level Radioactive Waste Disposal Facility

	Stage Implemented		
Criteria	*One*	*Two*	*Three*
Masking facilities			X
Active faults		X	x
Geologic stability			X
Slope			X
Carbonate lithology			
Outcrop at surface	X	x	x
Within 50 feet of surface, >5 feet thick		X	x
Potential for subsidence		X	x
Evidence of subsidence			X
River floodplains			X
Coastal floodplains	X	x	x
Important wetland			X
Dam inundation			X
Public water supply			X
Surface water intake			X
Wildlife area boundaries (many sources)	X	X	x
State forests and gamelands	X	x	x
Watersheds	X	x	x
Oil and gas areas	X	X	X
Agricultural land			X
Mines		X	X
Protected area boundaries (various)	X	X	X

X, first stage identified.

x, previous criteria updated by more detailed material.

Source: Simplified from Chem-Nuclear Systems Inc. (1994) Table 4-1 and 4-2, pp. 7–10.

LLRW: Low-level radioactive waste; low-level waste is generated (on the order of 160,000 m^3 for the United States per year) by nuclear power plants, hospitals, and various other industries; excludes waste from weapons construction and the spent fuel from nuclear reactors (high-level waste).

much more complex. Any one disqualifying factor excludes the area as a potential site. Knowing this logic, the analyst could allocate effort somewhat strategically. The 18 factors were not equally well established or equally expensive to measure. There was no need to spend time and money creating a digital coverage to exclude an area that was already excluded. The screening phase was conducted in three stages of increasing spatial resolution. Stage One excluded about 23% of Pennsylvania using components of 7 of the factors. Stage Two refined the layers used in Stage One and added 3 more to exclude some 46% of the state. Stage Three (Chem-Nuclear Systems, Inc. 1994) included all 18 factors and excludes 75% of the state (Figure 5-11).

The disqualification stages relied on many sources. Some disqualifying criteria came from a single source, as slope gradient was derived from digital elevation data generated by the U.S. Geological Survey. Other criteria, such as protected areas, required information from all levels of government. Some sources, like the National

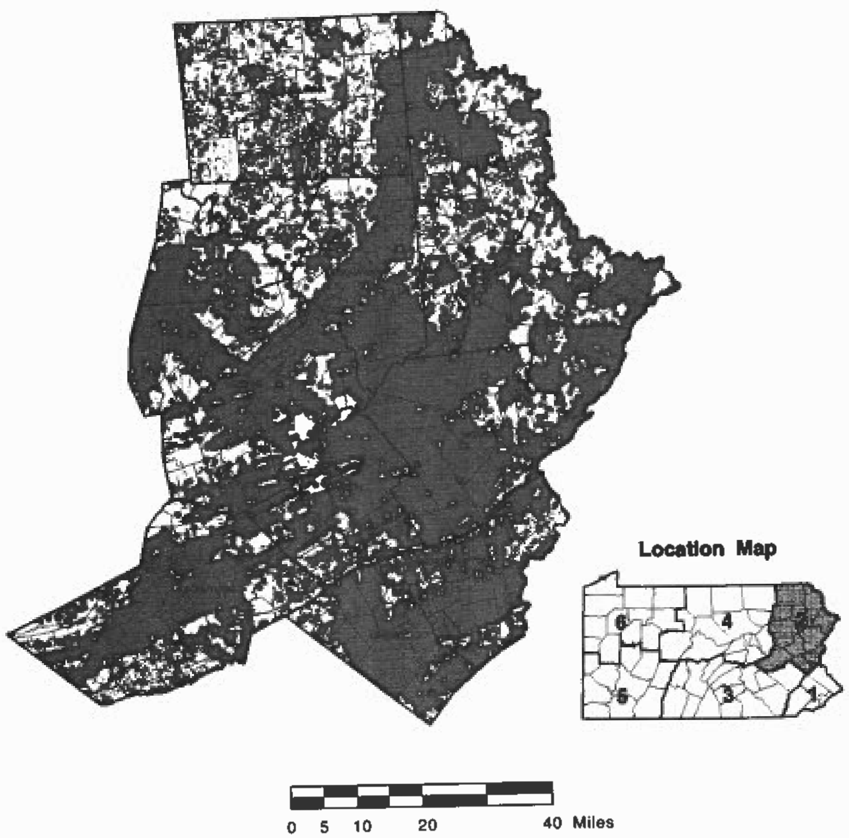

Figure 5-11 Portion of disqualifying coverage from Stage Three for northeast portion of Pennsylvania. Scale 1:1,000,000; most elements derived from 1:24,000. This region included large regions not disqualified in previous stages. (Source: Chem-Nuclear Systems, Inc. 1994; Appendix A, Plate 3)

Wetland Inventory, did not use quite the same set of categories, requiring some interpretive reclassification. Other disqualifying factors required compilation and digitizing efforts by the contractors. For example, coal mines have been recorded individually on surveys of diverse scale, some dating back to the Works Progress Administration in the 1930s. This project paid for the conversion of these records, which are seemingly unrelated to their mission. Such a requirement creates a large barrier to the use of GIS for less well-funded concerns. It is ironic that a search for a 500-acre site requires such massive information processing.

Under the best conditions, an exclusionary screening analysis will remove the bulk of the study area, leaving just a few alternatives for further examination. However, there is no guarantee that it will work out that well. One risk is that the criteria exclude only a small portion, leaving a huge area for future study (and the methods developed below). Another more troubling risk is that the whole study area might be ruled out (this was the result of the LLRW siting process in Michigan). External constraints (such as a decision that the facility must be accommodated) may require some relaxation of the criteria to find a possible site. The dominance approach provides no hint regarding the most prudent way to relax the restrictions.

Exclusionary Ranking A somewhat more complex dominance rule relies on ordinal measurement. If the attributes in the various layers are ranked according to a common scheme, then a rule for combination would select the most extreme (highest or lowest) value at each location. This exclusionary ranking is simply an evolution of exclusionary screening. The concept of limitations used by the U.S. Department of Agriculture and the UN Food and Agriculture Organization (FAO 1976) ranks soils and other resources according to limitations from severe to moderate to slight. These rankings were mentioned in Chapter 1 as a prototype for ordinal measurement, but the concept is not limited to a single measure. It also provides a rule to combine multiple factors. These rankings imply a dominance rule for the various ingredients (slope, drainage, etc.) that contribute to each evaluation. When combining various factors (such as availability of water, oxygen, or nutrients), the FAO method selects the most limiting ranking. This logic is in common use for land resource assessment (Kiefer 1967; Burrough 1986, p. 95). It makes sense as long as there is no interaction between factors; for example, no way for a "severe" limitation on one factor to be mitigated by some other factor's value.

Highest Bid/Highest Bidder Dominance rules can also apply to interval or ratio attributes, as long as each is expressed in the same attribute reference system. The highest bid rule differs from exclusionary ranking only in allowing numerical results in place of the ranked categories. When the highest value overrides the others, it simulates bidding, a basic rule of a market economy, and many forms of rational decision making. For example, in von Thünen's location theory, purchasers offer a price at a market center for agricultural produce, but progressively higher transportation costs reduce the net profit to producers that are located farther from the market. Thus, a producer grows the crop that offers the highest profit at that distance from the market.

If highest bid records the maximum value, then another related dominance method could be devised that assigns the location to the highest bidder, the source that contributed that highest value. This rule processes ordinal, interval, or ratio measures, producing an identifier for the extreme value. For example, Young and Goldsmith (1977) prescribed land use alternatives for an area in Malawi based on the most suitable of six alternative crops and competing uses.

Dominance rules are perhaps the lowest form of combination since they are all or nothing. Yet, this logic does play an important role in many applications. It is particularly important to use a dominance rule for absolute exclusions that precede other steps. Some of the GIS literature belittles these kinds of operations as being too simple, but they are useful exactly because they are simple and understood by all parties. Environmental regulations are often simplified into binary criteria for simplicity and clarity. The concept of "highest and best use" is basic to land economics and property appraisal. These culturally sanctioned decision rules provide the basic rationale for many GIS applications.

Contributory Rules

The next group of operations uses all the information from the various sources. The values from one layer contribute to the result without regard for attribute values from other layers. The combination process can use any arithmetic operation, but the most common example is addition. Addition is, after all, considered the basic rule for extensive measurement scales. As long as the measurements extend along the same axis and in the same units, addition provides the appropriate combination. Each contributory rule depends on certain assumptions about the source materials.

Voting Tabulation The simplest form of voting applies to nominal categories. By examining all the values present at one location, the procedure can select the category that is most frequent (the mode of the distribution) or least frequent or other variations. This differs from the dominance rules because the choice of the value comes from the set of values available, not from some external rule. This procedure can produce either the value selected or the number of votes cast for the winner.

Voting can also be applied to the layers used in an exclusionary screening. If the positive factors are scored as one and the negative as zero, then the sum of the values can tell how many positive factors occur at the location. Or the scores could be turned around to tabulate the number of negative factors. This method, which might be best termed a voting tabulation, provides more information for site suitability or environmental regulation. Compared to exclusionary screening, it helps to know that an area has been excluded for multiple reasons. This result can be obtained by using addition on the binary layers instead of Boolean operations.

The process of counting exclusions seems quite reasonable, but it leads directly to a troubling question. Is an area excluded on two criteria twice as excluded as an area only excluded for one reason? For example, following the Washington State (1990) critical area designations created by the Growth Management Act, are wetlands and geologically hazardous areas of equal sensitivity? Should the components of the geological hazards (seismic risk, slope stability, erosion) be given separate status? Should each source of landslide information be counted separately or together? These questions simply cannot be answered in the general case. There is simply no universal calculus that tells that one floodplain equals two aquifer recharge areas. In the colloquial, this is the problem of apples and oranges. At one level it is possible to generalize and to call it all "fruit," but the unit of measure is not as stable or as comparable. In some rough sense, the count of critical areas may act as an ordinal scale, though even that requires a stretch of measurement rules.

Weighted Voting Another form of voting tabulation weights each factor by a number. The weight is supposed to express the relative importance of the factor. The attribute values are still scored zero or one. Carver (1991, p. 324) adopts a seven-point ordinal scale (from 7 = very important to 1 = unimportant) as the weight. While this allows finer gradations of preference, there is little logic to suggest that "very important" is exactly seven times larger than "unimportant" or that "important" belongs at three or four. The wording of such ordinal scales conflicts with their use as numeric values on an interval scale.

For the Pennsylvania LLRW facility, Chem-Nuclear Systems (1992) applied a variant of a weighted voting called rank-sum for the phase designed to follow the exclusionary screening. The votes consist of zero–one thresholds of nine factors. These nine factors are ordered according to importance, from nine down to one. This rank is used as the weight. The score for a site is the sum of the weighted votes, so that the factor chosen as most important contributes a value of nine if present, zero if not, and the lowest factor contributes one or zero. This technique should only be used if the top factors are much more important than the bottom ones.

These forms of voting tabulation approximate popularity of each alternative based on presence or absence of certain factors. Voting makes sense if the votes come from totally distinct sources (not alternative sources for a composite environmental sensitivity). Each factor contributes its weight to the result democratically, but the result depends on how votes are counted.

Linear Combination The more common form of a contributory rule involves a more general use of addition. A linear combination adds up the values for a particular location. Compared to the voting rules, the measurement scale becomes even more important.

A simple example of linear combination shows many of the hidden assumptions. In 1982, the Dane County (Wisconsin) Regional Planning Commission (1980) was charged with selecting a site for solid waste disposal. Six site factors ranked on the scale of slight, moderate, and severe were represented by the numbers 1, 2, and 3 (Figure 5-12). Once assembled, the ratings were added up for each grid cell. The minimum possible value was 6, from a ranking of "slight" on all factors. If the landfill site could have been found with this score, the method would have replicated exclusionary screening. As noted above, it is not always possible to find a site that will pass

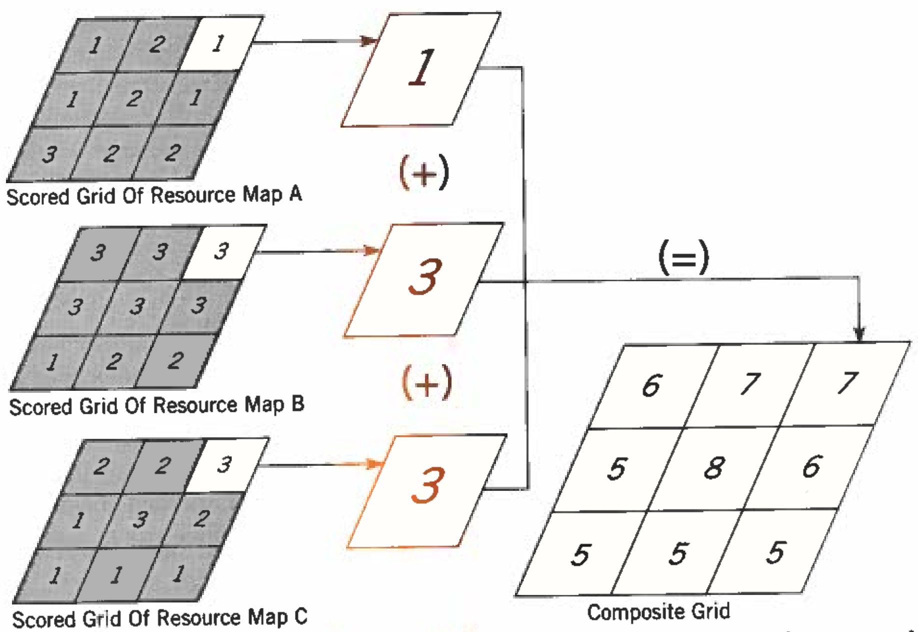

Figure 5-12 "Linear combination" in its simplest form simply adds up the values. Example shows the ratings assigned to categories for the Dane County Solid Waste Plan (Dane County Regional Planning Commission, 1980). Ratings for grid cells were added to obtain a composite ranking.

all the exclusionary filters. The proposed landfill had some cells with composite ratings of 8 and 9, values that can each occur with two possible combinations. For example, a value of 8 implies either two values of "moderate" (1 + 1 + 1 + 1 + 2 + 2 = 8) or one "severe" (1 + 1 + 1 + 1 + 1 + 3 = 8). Does one "severe" (and one "slight") equal two "moderates"? In this case, the factors may not hold equal importance to a landfill. Some factors can be remedied with engineering efforts, while others cannot. The relative environmental risk is not strictly equivalent. A result of all "moderate" scores (2 + 2 + 2 + 2 + 2 + 2 = 12) may rank more favorably than (1 + 1 + 1 + 1 + 1 + 3 = 8), if the one "severe" rating totally prohibits the landfill's economics. In conclusion, simple addition does not provide an unbiased site selection method for this landfill siting database. In its defense, the Regional Planning Commission did this study using manual methods in a predigital era, but the technology is not really the issue. Rules for combining scores must be justified, whatever technology is applied. Using addition imports a set of axioms that may not apply to the measurements available.

Despite years of criticism from sceptics such as Hopkins (1977), many practitioners of GIS still use linear combination for ordinal scale data. For example, Cornwell and Rohardt (1983) report a site selection study for a steel plant in Nigeria. The study rated 17 factors on a common ordinal scale of suitability. The overall suitability

was then calculated by the sum of all the rankings. In their study, the authors recognize some drawbacks in adding ordinal scores. They seek a solution in rescaling the numbers assigned to the ordinal scores, not in some other method of combination.

Various methods have been proposed to convert ordinal rankings to a continuous scale. Carver (1991) suggests a proportion of the range for each variable between minimum and maximum. Eastman and others (1993) describe standard deviations from the mean and histogram equalization. Each of these techniques depends on the distribution within the study area. The mean and standard deviation make sense only for interval scales. Histogram equalization converts all distributions to rankings, which removes any sense of an extensive measurement. The median will always be given the midpoint value, no matter the distribution. Rankings remain orderings, not interval or ratio measures.

As presented in Chapter 4, new information can evaluate each ordinal category on a true ratio scale. For example, categories from a number of factors can be rated in terms of dollar cost to a developer (Table 5-4) (Fabos and Caswell 1977). As a

TABLE 5-4 Assigning a Ratio-scaled Measure to Ordinal Categories

Factors and Dimensions	*Estimated Added Costs (in dollars)*
Depth to bedrock	
0–2 feet	$20,000
2–5 feet	5,000
5 feet +	200
Depth to water table	
0–3 feet	$5,000
3–5 feet	1,400
5 feet +	0
Drainage	
Poorly to very poorly drained	$5,000
Moderately poorly drained with hardpan	1,400
Other	0
Slope	
15% +	$*****
8–15%	1,300
0–8%	0
Topsoil	
Poor (0–4 inches)	$1,500
Fair (4–6 inches)	600
Good (6 inches +)	0
Bearing capacity	
Plastic and nonplastic; silts and clays; peat; muck	$1,500
Other	0

Source: Fabos and Caswell (1977) *METLAND Landscape Planning Research*, Research Bulletin 637, Massachusetts Experiment Station.

scale, money has a strong attraction, though the 1977 dollars in Table 5-4 certainly would need adjustment to be taken seriously by current developers. The particular values shown here would certainly not apply in all surroundings. In earthquake country, the cost of building on peat or muck should be substantially higher. In Hong Kong and other places, a 15% slope gradient is certainly not an absolute restriction on construction. Each marketplace and physical environment values these categories differently.

As much as one can argue with the specific values assigned, there are also problems that arise simply from trying to monetize all the factors. Why should the cost to the developer be the only cost considered? How can we include benefits to society that are not traded in a market of any kind (like scenic beauty)? On the more technical level, the sharp breaks between categories point to the weakness of assigning ratio values to measurements that are just categories. This project expected to obtain most of those physical factors from soils map and similar sources, using indirect procedures. Consequently, there will be very sharp differences in development cost along boundaries that may represent rather gradual transitions between soil properties.

Weighting and Rating The full form of the linear combination method requires a slightly more general form (Equation 5-1) to include a weighting factor for each attribute rating. The insertion of a weight simply recognizes that the simple sum assigns an implicit weight of one to each factor.

Equation 5-1

$$V_j = \frac{\sum_i w_i r_{ij}}{\sum_i w_i}$$

where

V_j refers to the resultant value for each object j

r_{ij} refers to the rating of each object j for each attribute i

w_i refers to the weight assigned to attribute i

Weighted addition is essentially identical to the procedure termed *ideal point analysis* in the multicriteria literature (Hobbs 1985). A more neutral label for this method might be weighted linear combination, but in many applications it turns into "The Weighting and Rating Game." The selection of weights and ratings is often established iteratively based on their results, without sufficient attention to the assumptions involved.

The U.S. Department of Agriculture's Land Evaluation and Site Assessment (LESA) methodology uses weighted sums of many factors (Table 5-5). As applied in a county in Kansas (Williams 1985), the land evaluation component of the score is based on a single index crop, grain sorghum. This figure is similar to the corn grain yield shown in Figure 2-17, an indirect attribute of eight soil productivity classes (in

TABLE 5-5 Implementation of Weighted Linear Combination: Land Evaluation and Site Assessment for Douglas County, Kansas

Land Evaluation

Grain Sorghum Yield		*Rescaled to Set Highest Value at 100*	
Site Assessment		*Source*	*Weight*
1	% of area agricultural within 1.5 miles	Land use, buffered	10
2	Adjacent agricultural land	Land use, buffered	7
3	Farm size	Parcel area	2
4	Average parcel size within 1 mile	3, neighbor sums	4
5	Agrivestment in area		3
6	% of area zoned ag within 1.5 miles	Zoning, buffered	8
7	Zoning of site (for agriculture)	Zoning, rescored	6
8	Zoning for alternative (residential use)	Zoning, rescored	6
9	Nonfarm AND poor soil	Zoning, soil overlay	6
10	Need for urban land	Land use/city limits	8
11	Compatibility of adjacent uses	Land use, neighbor	7
12	Unique features		3
13	Adjacent to unique features	12, buffered	2
14	Floodplain or drainage way	Hydrology	8
15	Suitability for on-site septic	Soils, rated	5
16	Compatibility with comp. plan	Plan	5
17	Designated growth area	Plan	5
18	Distance from city limits	City, buffered	6
19	Distance from transportation	Roads, buffered	5
20	Distance from water service	Water, buffered	4
21	Distance from sewer lines	Sewers, buffered	4
	(reweighted to sum to 200 points)	Total	114

Overall score = Land evaluation (100) + Site assessment (200).

Source: Williams (1985), p. 1928.

turn a grouping of soil mapping units). Thus, the largest contributor to the score has only eight possible values, not a continuous measure from 0 to 100. On the land evaluation scale, an increment of one represents 0.6 bushels of sorghum per acre per year. The site assessment scores are measured and reweighted into 200 points on the same scale. The 21 site assessment factors are scored on a scale from 0 to 10, then weighted. A few of these factors are based on distance measurements from lines or areas. The tools to produce these buffers are described in the next chapter. The distance measures in this study were rated on a logarithmic scale, for some reason. For example, the distance to water lines (factor 20) was scored 0 in the same 100-meter cell as the water line, 2 for less than $\frac{1}{8}$ mile, 4 for $\frac{1}{8}$ to $\frac{1}{4}$ mile, 6 for $\frac{1}{4}$ to $\frac{1}{2}$ mile, 8 for $\frac{1}{2}$

to 1 mile, and 10 for more than 1 mile. These ratings do not predict the costs of installing pipes very closely.

This study targets the conversion of agricultural land to residential use. Another set of factors and weights would be required for another application. As the author points out (Williams 1985, p. 1930), the LESA linear combination assumes independence of the factors, but suitability for septic fields (factor 15) is irrelevant if the new residential development is attached to the sewer system (factor 21). Similarly, the highly weighted factors for percent agricultural (1) and percent zoned for agriculture (6) may simply provide double weight for strongly related maps.

Specialists in multicriteria methods react with suspicion to the choice of weights using a "magic number" method. A system of rating scores and weights can be justified if the measurements turn into a common ratio scale of measurement. The magnitude of a weight must express the amount of one quantity that one is willing to trade for another. In the LESA case, the weight of 4 assigned to distance from water lines means a specific relationship between bushels of sorghum and each increment away from water lines. Moving $\frac{1}{8}$ mile away from the water lines converts to 0.8 bushels of sorghum; beyond 1 mile to 4.2 bushels. It is doubtful if the value of 4.2 bushels of crop productivity can compensate for the cost of installing 1 mile of water line over any reasonable budgeting period.

One method to prepare a defensible system moves beyond the technical level to recognize the institutional and social context. Any procedure, no matter how boneheaded, is adequate if all participants accept it. Many environmental decisions become heated public issues, and the technical components can become the overt battlefield for disagreements with other origins. Faced with this situation, there have been many attempts to use conflict resolution methods in the technical process. Methods of compromise and consensus building, such as Delphi panels, have been used in various site selection processes since the origins of GIS (see, for instance, Dames and Moore 1975; McCartney and Thrall 1991; Eastman et al. 1993). Bringing the group to agreement on a set of weights is a fine goal, but it will not remove the mathematical objections to weighting as a technique. The process of social agreement may often find a compromise by taking a middle value, thus importing the same set of measurement assumptions.

Nonlinear Combinations Beyond the linear combinations, there are many alternative mathematical functions to combine variables. As with addition-based methods, all the available values contribute to the result according to the rules of algebra. Most of the criticism applied to the linear methods becomes even stronger as the functions become more powerful.

One common method involves a multiplicative rule in the place of addition. Sometimes the unsuspecting user is not aware of the switch to multiplication. If the values for the variables are transformed by a logarithmic functions [one of the suggestions in Cornwell and Rowhardt (1983)], then the addition produces a result that is essentially multiplicative. Other adoptions of the multiplicative form are more conscious, such as Storie's (1933) index of crop productivity. One of the most common examples of an overtly multiplicative model applied in GIS is the universal soil loss equation (USLE) (Wischmeier and Smith 1978) introduced in the description of the Dane County Project.

A multiplicative rule is particularly sensitive to errors in the original sources since small changes in one factor can influence the whole result. Multiplication is particularly sensitive to zero values, which can dictate the result despite the contributions from other values. With a zero value, multiplication acts as a dominance rule, though with nonzero values, each value contributes to the result.

Interaction Rules

The contributory rules assume that each attribute should contribute to the result without regard for the specific level of some other attribute. Effectively, this rules out interactions between factors. However, interactions between attributes should be considered to be the norm, not the exception. In some situations, such as change detection or error testing, the specific interactions are easy to choose, so long as each source used the same set of categories (Figure 5-9). In other cases, simple Boolean operators (like AND and Exclusive OR) are actually a very simple form of interaction. It is less easy to include interactions for situations with more complex attributes. The research community has struggled with interrelated factors for a number of years. Few of the methods below are in common use, perhaps due to the complexities in implementation.

Contingent Weighting Methods that include interactions between factors can be developed from any of the contributory methods. One example involves the Wisconsin Groundwater Contamination Susceptibility Map (Riggle and Schmidt 1991). Five statewide source maps were overlaid to create a composite of some 54,000 polygons. The categories on four maps were scored in the common approach (termed evaluation in Chapter 4) in which ordinal categories are given status as continuous measures (Table 5-6). The high values in this scoring system denote less chance of groundwater contamination. Thus, carbonate rocks have greater risk than shale; shallow groundwater is easier to contaminate; and so on. The treatment of interaction comes from the table that gives a weight for these four factors based on the value of a fifth factor, depth to bedrock. The weight given to the type of bedrock declines as the depth to bedrock increases. Thus, a shallow situation will be more strongly influenced by the type of bedrock. The zero value for depths over 50 feet mean that the factor has no influence on the result. Here the weighting serves to remove a factor from the contributory equation. The other three factors have weights that increase with depth to bedrock. Thus the importance of soil characteristics is greater for deep soils.

TABLE 5-6 Formula for Groundwater Contamination Susceptibility Score

Attributes	*Value Assigned*
Type bedrock (TBV)	
Carbonate	1
Sandstone	5
Igneous/metamorphic	6
Shale	10
Depth to water table (DWV)	
0–20 feet	1
20–50 feet	5
>50 feet	10
Surficial deposits (SDV)	
No material	0
Sand and gravel	1
Sandy	2
Peat	5
Loamy	6
Clayey	10
Soil characteristics (SCV)	
Coarse texture/high permeability	1
Medium coarse texture/high–medium permeability	3
Medium texture/medium permeability	6
Fine texture/low permeability	10

	Weights Based on Value of Depth to Bedrock			
Depth to Bedrock	*TBW*	*DWW*	*SDW*	*SCW*
0–5 feet (>70%)	13	1	0	1
0–5 feet (<70%)	11	1	1	2
5–50 feet	6	2	4	3
50–100 feet	0	3	8	4
>100 feet	0	3	8	4

GCSS = (TBW × TBV) + (DWW × DWV) + (SDW × SDV) + (SCW × SCV).

Note: High scores are less susceptible.

Source: Riggle and Schmidt (1991).

The values and weights in the Wisconsin groundwater study are difficult to justify as pure extensive measures. The source materials available statewide were not sufficient for anything but the crudest determination of depth to bedrock (50-foot contours). The use of a linear equation may not be properly justified for the ordinal categories on the original maps. Nevertheless, this study did wrestle with the interaction between some of its factors. This procedure is an example of how interactions can be inserted into contributory numerical procedures.

Evaluating Interactions A rule based more completely on interactions involves **integrated survey,** a landscape evaluation approach that makes multicriteria evaluations of suitability a part of the fieldwork process (Mabbutt 1968). Integrated surveys uncover associations by discovering the process that causes factors to combine to make the landscape suitable for a given purpose. Examining the interrelationships in the field is a demanding process that is not easy to replicate, even with the same team. The ultimate weakness is that the field team must be sent out with the appropriate mission. Evaluation is conditional on a purpose, for instance, the original surveys in Australia were looking for expanded agriculture (not ecotourism or biodiversity). This limits the use of the survey in a multipurpose database for unknown future requirements. Still, this method remains one of the best designed to create scientific consensus about a landscape.

While the integrated survey method relies on fieldwork to handle all the interactions, there are other methods better connected to current technology. The method that Hopkins (1977) called *factor combination* involves creating an overlay of all possible factors (the basic enumeration step), then rating all the combinations for their suitability. It differs from integrated survey in that various factors are mapped separately and then combined by overlay. *Integrated terrain unit mapping* (Dangermond 1979) is another term for an automated form of factor combination.

Most of the literature on map combination throws up its hands at the prospect of ranking all the combinations of categories that arise from an overlay. As described above, these numbers do not expand quite as rapidly as the worst case. In the multicriteria decision-making literature, the procedures of *conjoint measurement* (Keeney and Raifa 1976) have been developed for exactly this purpose. Instead of trusting some linear combination of individual ratings, a conjoint method evaluates the result in the full multidimensional space including the interactions. Thus, these various terms boil down to the same basic step of evaluation.

Rules of Combination The integrated survey approach challenges an analyst to consider all the possible interactions between factors, while the contributory methods offer the simplicity of performing the combinations automatically using an algebraic analog. There is a compromise between these extremes, involving some of the simple rules of the contributory methods, modified by interaction tables. An analyst will decompose the problem into subsets on which simpler rules apply, then combine the subsets according to some other rule.

At an early stage in the development of geographic information processing techniques, the Honey Hill project (Murray et al. 1971) developed a simplified method to accommodate a certain amount of interaction between the factors of an overlay. In this somewhat artificial prototype, each factor had three categories. A three-by-three table combined two factors into a three-category result, which was combined with the third factor in another three-by-three table. Certainly, the restriction to a three-level

Integrated survey: An approach to land evaluation that combines the opinions of many disciplines in producing a common representation of the processes that form a landscape.

ranking limits this particular method, but the underlying logic can recognize interrelationships not treated by an arithmetical procedure.

A generalized set of rules provide a way to fill in the complex multidimensional interactions of a complete overlay (as in the factor combination or integrated terrain unit approach) with a bit more clarity. There seem to be few published examples of such an approach, however.

Summary of Rules

This presentation of overlay combination distinguishes methods based on the amount of information used. In all cases, a number of map layers have been combined using overlay. Each location on the map has an attribute on each of the input layers. The enumeration rules preserve all this detail for later consideration. The dominance rules pick one of the values as the result; the exclusionary screening rule is used quite frequently in applying regulations because of its simplicity. The contributory rules use all the attributes available to create some kind of composite attribute value. The limitation is that each layer contributes its value without regard for the others. These methods are relatively easy to implement but full of assumptions. Interaction rules try to provide a method to treat the complexities of the environment. The result depends on the specific combination of attribute values for some layers taken together. It is not surprising that these methods are underdeveloped in practical cases. Truly defensible applications of GIS will require continued development of the procedures for map combination. Many models of environmental interactions exist, but they are not always implemented in their full complexity.

SUMMARY

Map overlay had its origins in a simpler world of physical maps drafted on transparent material and examined on a light table. In a GIS, overlay constructs the linkage between different sources of information using the geometry as a key. Attributes attached to the sources can then be combined using four general classes of rules: enumeration, dominance, contributory, and interaction. Chosen with care, the composite information forms a strong base for geographic analysis because it can integrate diverse sources. By itself, overlay provides direct answers to certain analytical needs such as change detection. However, it only addresses phenomena that are collocated. The next step is to broaden the analysis to larger neighborhoods.

CHAPTER 6

DISTANCE RELATIONSHIPS

CHAPTER OVERVIEW

- Construct vector buffers around isolated objects.
- Perform distance measurements and construct distance fields for rasters.
- Introduce extended Voronoi representations of distance relationships.

Chapter 5 concentrated on the overlay method, a technique to integrate different sources. As with any tool kit, there is a risk of seeing the world through the viewpoint of that one method. To the hammer salesperson, the whole world is full of nails (that need hammers). With geographic information, there is a similar risk. Regulations and specifications are usually written to reflect what seems technically feasible. Scientists, naturally, tend to use the techniques with which they are most familiar, sometimes missing important insights. There are many tools beyond overlay. This chapter will introduce one such tool in some depth that will lead to a more general set of neighborhood tools in the next chapter.

The next step in expanding the range of tools from the strict locality of overlay is adding tools that discover distance relationships implicit in a spatial representation. These operations can construct a new representation using distance measurements from some existing representation, thus qualifying as a limited form of transformation. The nature of these transformations depends on the assumptions implicit in the measurement framework. Chapter 7 will continue in the same direction with more generalized methods that operate on neighborhoods.

EXAMPLES OF DISTANCE SPECIFICATIONS

Before considering the procedures used to implement distance measurement, it is useful to consider some of the motivations for using distance in geographical analy-