CHAPTER 3.6: SPATIAL ANALYSIS AND MODELING

Michael F. Goodchild
University of California, Santa Barbara

3.6.1 Introduction

In the previous chapters we have seen how a wide variety of types of geographic data can be created and stored. Methods of digitizing and scanning allow geographic data to be created from paper maps and photographs. Powerful computing hardware makes it possible to store large amounts of data in forms that are readily amenable to manipulation and analysis using the routines stored in powerful software. Thus the stage is set for a discussion in this chapter of the real core of GIS, the methods of analysis and modeling that allow us to examine data, to solve specific problems, and to support important decisions, using the capabilities of hardware, software, and data that together compose a GIS.

Chapter 3.5 introduced the various ways in which geographic phenomena can be represented in digital form. A road, for example, can be represented by recording an appropriate value in a swath of cells in a raster representation. With a cell size of 1m, the swath corresponding to a major four-lane highway might be as much as 50 or even 100 cells wide. Alternatively, the road might be represented as a single line, or *centerline*, in a vector database. In this case its location would be recorded by specifying the coordinates of a series of points aligned along the road's center. Such centerline databases are now very commonly used in applications like vehicle routing and scheduling, and by sites such as www.mapquest.com or maps.google.com which offer to find the best routes between pairs of places, and are used daily by millions of people. Finally, the road might be represented as an area, encased by its edges (we use the term *cased* to describe this option, and also call it a *double-line* representation). Figure 3.6.1 shows the three options. Which of these is used in a given case depends on the nature of the application, on the limitations of the available software, on the origins of the data, and on many other factors.

[Figure 3.6.1 about here]

This chapter is about turning data into useful information. Analysis and modeling can make calculations that are too tedious to do by hand, and by doing so provide numbers that are useful for many kinds of applications. They can be used to reveal patterns and trends in data that may not be otherwise apparent. They might also be used by one person or group to draw the attention of another person or group, as might occur in a courtroom where geographic data are used by one side to make a point. Finally, the results of analysis and modeling can be used to provide the information needed to make decisions. In all of these cases GIS is the engine that performs the necessary operations, under the guidance of the user who issues the necessary instructions.

3.6.1.1 Organizing the possibilities

A very large number of methods of analysis and modeling have been devised over the years, many dating from well before the advent of GIS, when calculations and measurements had to be performed by hand. The ability to process large amounts of data quickly has led to a rapid explosion in the list, and today it is virtually impossible to know about every form of spatial analysis. The developers of GIS software often provide thousands, and thousands more are added by specialized companies and individuals (for example, the toolbox provided with ESRI's ArcGIS 9.2 provides over 500 distinct functions, in addition to the hundreds of operations that can be called from its drop-down menus). So one of the most daunting aspects of analysis and modeling with GIS is simply keeping track of the possibilities.

One way to organize methods of analysis and modeling is by the data types on which they operate. There are operations that work on discrete objects -- points, lines, and areas -- and operations that work on phenomena conceptualized as continuous fields. At a different level, and following the concepts introduced in Chapter 3.5, it is possible to separate methods of analysis of vector data sets from those that operate on raster data sets, since there are very few instances of operations that require input of both kinds. Another possibility is to see every operation as a transformation, taking some kind of input and producing some kind of output, and to organize methods on this basis. The structure adopted here is a little of all of these, but is based primarily on popularity: since there are so many possibilities, it is most important to understand the ones that are used most often, and to leave the less popular ones to further study. Each method is described in terms of the problem it attempts to solve, the inputs required, and the outputs that it generates. Both raster and vector operations are covered.

The next section deals primarily with points, and the following section with areas. The fourth section looks in detail at rasters, using the framework of *cartographic modeling*, and includes methods of analyzing digital elevation models. Finally, the last section of the chapter examines methods for optimization and design, where the objective is to use GIS to find the best solution to problems. A more complete discussion is available in the text by Longley *et al.* (2005), and an advanced compilation is provided by De Smith, Goodchild, and Longley (2007).

3.6.2 Methods of analysis for points

We begin with the simplest kinds of geographic objects. Suppose we have records of each of the customers of an insurance agent, including each customer's location. Perhaps these originated in an address list, and were subsequently converted to coordinates using the process known as *geocoding* or *address-matching* (Chapter 3.2). Plotted on a map they might look something like Figure 3.6.2. Several questions might occur to a market researcher hired by the insurance agent to study the state of the business and recommend strategies. For each there is a straightforward type of GIS analysis that can be used to provide the necessary information.

[Figure 3.6.2 about here]

3.6.2.1 Distance

First, we might ask how far each customer is located from the agent's location. The agent might be interested in knowing how many customers do their business with the agent rather than with some other agent located closer to the customer. What is the average distance between each of the customers and the agent, and are there areas near the agent where advertising might increase the agent's market share? All of these questions require the ability to measure the distance between points, a commonly used function of a GIS.

To measure the distance between two points we need to know their locations in (x,y) coordinates. Coordinate systems were introduced in Chapter *, and many of them can be used to measure distances. The UTM coordinate system, for example, is based on measurements of coordinates in meters, so distances calculated in this coordinate system are easily understood. State Plane Coordinates are often expressed in feet. In these cases the distance between two points (x_A,y_A) and (x_C,y_C), representing the locations of the agent and a customer, are given by:

$$D = \sqrt{(x_A - x_C)^2 + (y_A - y_C)^2}$$

This equation is derived from Pythagoras's famous theorem, so it is also known as the Pythagorean distance, or the length of a straight line between the two points.

Pythagorean distance applies to points located on a plane, not on the curved surface of the Earth, and so should be used only when the points are close together. A good rule of thumb is that the study area should be no more than 500km across, because at that distance errors due to ignoring the Earth's curvature begin to approach 0.1%, even with the best-designed coordinate systems.

It is important to recognize that this problem of the Earth's curvature is not resolved by using latitude and longitude coordinates. Suppose we record location using latitude and longitude, and plug these into the Pythagorean equation as if they were y and x respectively (this is often called using *unprojected* coordinates, and is also what happens if we specify the so-called Plate Carrée or Cylindrical Equidistant projection). In this case the coordinates are measured in degrees, so the value of D will also be in degrees. But one degree of latitude is not the same distance on the Earth's surface as one degree of longitude, except exactly on the Equator. At 34 degrees North or South, for example, the lines of longitude are only 85% as far apart as the lines of latitude, and that percentage drops all the way to zero at the poles. So the equation cannot be applied to latitude and longitude. Instead, we should use the equation for distance over the curved surface of the Earth, otherwise known as the Great Circle distance, because the shortest path between two places follows a Great Circle (a Great Circle is defined as the arc formed when the Earth is sliced through the two points and through the center of the Earth). If the point locations are denoted by (ϕ_A,λ_A) and (ϕ_C,λ_C), where ϕ denotes latitude and λ denotes longitude, then the distance between them is given by:

$$D = R\cos^{-1}\left[\sin\phi_A \sin\phi_C + \cos\phi_A \cos\phi_C \cos(\lambda_A - \lambda_C)\right]$$

where R denotes the radius of the Earth, or approximately 6378km. Note, however, that this formula assumes that the Earth is a perfect sphere, ignoring the substantial flattening at the poles and other distortions.

There are many other bases for measuring distance, because it is often necessary to allow for travel that must follow streets, or avoid barriers of one kind or another. So the actual distance traveled between two places may be much more than either of these formulae would predict. Two of these cases are dealt with later in this chapter: when the path followed is the one that minimizes total cost or travel time over a surface (a case of raster analysis, see Section 3.6.*), and where the path follows a network of links with known lengths or travel times (a case of network analysis, see Section 3.6.*).

3.6.2.2 Buffers

Instead of asking how far one point is from another, we might turn the question around and identify all of the points within a certain distance of a reference point. For example, it might be interesting to outline on a map the area that is within 1km of the agent's location, and subsequent sections will cover several applications of this concept. In GIS the term *buffer* is used, and we say that the circle created by this operation constitutes the 1km buffer around the agent. Buffers can be created for any kind of object -- points, lines, or areas -- and are very widely used in GIS analysis. Figure 3.6.3 shows buffers for each of these types of objects.

[Figure 3.6.3 about here]

By finding a buffer, and by combining it with other information using the methods discussed in this chapter, we could answer such questions as: how many customers live within 10km of the agent's location; what is the total population within 10km of the agent's location (based on accurate counts of residents obtained from the census); where are the people with the highest incomes within 10km of the agent's location?

3.6.2.3 Points in polygons

The so-called *point in polygon* operation is another key feature of GIS analysis. Suppose we have a map of the tracts used by the census to publish population statistics, such as the map shown in Figure 3.6.4. In form, the map looks much like a map of states, or counties, or voting districts, or any of a number of types of maps that divide an area into zones for the purpose of counting and reporting. The point in polygon operation allows us to combine our point map with this map of areas, in order to identify the area that contains each point. For example, we could use it to identify the census tract containing each of the customer locations in our point data set. By counting customers in each tract, and comparing the totals to the known populations of each tract, we could get interesting data on market penetration by tract. Alternatively we could add join the attributes of the

containing tract with the attributes of the customer, to see how neighborhood characterists might relate to individual characteristics.

[Figure 3.6.4 about here]

The point in polygon operation is actually very simple to execute, and so the method will be described briefly. One of the points in the figure will be used as an example. A line is drawn from the point, in this case diagonally upwards. The number of intersections between this line and the boundaries of each polygon is counted. In the example, there is one intersection with the boundary of Polygon A, 2 with the boundary of Polygon B, and two with the boundary of Polygon C. The polygon that contains the point is the only one with an odd number of boundary intersections (and it will always be true that exactly one polygon has an odd number -- all other polygons will have an even number).

One useful way to think about the result of the point in polygon operation is in terms of tables. The result of the point in polygon operation could be expressed in terms of an additional column for the census tract table shown in Figure 3.6.4, recording the number of points found to lie in that polygon, as a preliminary to computing the local market penetration. It is applications like these that make the point in polygon operation one of the most valuable in vector GIS.

3.6.3 Analysis of areas

The previous section looked at one method that combines points and areas (or *polygons*, a term commonly used interchangeably with areas in vector GIS, since areas are normally represented as polygons as discussed in Chapter 3.5). This section focuses specifically on operations on areas, again from the perspective of vector GIS.

3.6.3.1 Measurement of area

One of the strongest arguments for the use of GIS, as compared with manual methods of analysis of information shown on maps, is that computers make it easier to take measurements from maps. The measurement of area is in fact the strongest argument of this type. Suppose we need to measure an area from a map, such as the area of a particular class of land use. Perhaps our task is to measure the amount of land being used by industry in a specific city, and we are given a map showing the locations of industrial land use, in the form of appropriately colored areas. In a medium-sized city the number of such areas might be in the thousands, and so there would be thousands of measurements to be made and totaled. But even the measurement of a single area is problematic. Manually, we would have to use one of two traditional methods, known as dot counting and planimetry. Dot counting proceeds by overlaying a prepared transparent sheet covered by dots at a known density, and counting the number of such dots falling within the area of interest. To get a reliable estimate the density of dots needs to be high, and the dots need to be small, and it is easy to see how tedious and inaccurate this task can be. Planimetry proceeds by using a crude mechanical device to trace the outline of the area -- the result is read off a dial on the instrument, and again the process is tedious

and error-prone. In short, manual methods are frustrating, and expensive in terms of the time required to obtain even a poor level of accuracy.

By contrast, measurement of area is extremely simple once the areas are represented in digital form as polygons. A simple calculation is made for each straight edge of the polygon, using the coordinates of its endpoints, and the calculations are summed around the polygon. Any vector GIS is able to do this quickly and accurately, and the accuracy of the final result is limited only by the accuracy of the original digitizing.

The world's first GIS, the Canada Geographic Information System, was developed in the mid 1960s precisely for this reason. A large number of detailed maps had been made of land use and other land properties, with the objective of providing accurate measurements of Canada's land resource. But it would have taken decades, and a large workforce, to produce the promised measurements by hand, and the results would have been of disappointing accuracy. Instead, a GIS was developed, and all of the maps were digitized. In addition to straightforward measurement of area, the system was also used to produce measurements of combined areas from different maps, using the operation of polygon overlay, another important vector operation that is covered in the next section.

3.6.3.2 Polygon overlay

Figure 3.6.5 shows a typical application of polygon overlay. The two input maps might represent a map of land use, and a map of county boundaries. A reasonable question to ask of a GIS in a case like this is: "how much land in County A is in agricultural use?" To answer the question we must somehow compare the two maps, identifying the area that is *both* in County A and in agricultural land use, and measure this area. Polygon overlay allows this to be done. As a method, its history dates back to the 1960s, though efficient software to perform the operation on vector data sets was not developed until the late 1970s.

[Figure 3.6.5 about here]

Figure 3.6.5 shows the result of the overlay operation, in the form of a new data set. Each of the polygons in the new data set represents a unique combination of the inputs, and the boundaries of the new polygons are combinations of the lines from the two inputs. The operation is actually reversible, since it is possible to recover both of the input data sets by deleting some of the attributes of the new polygons and merging polygons with the same remaining attributes. For example, by deleting land use attributes we recover the county data set. In other words, no information is lost or created in polygon overlay -- instead, the information is simply rearranged.

The new data set makes it easy to answer questions like the one we started with. To determine the amount of land in County A in agricultural land use, we simply identify all polygons in the new data set with those attributes, compute their areas, and total. Polygon overlay has a myriad of uses, of which this kind of query is among the simplest. One popular use has to do with population statistics. Suppose a water supply agency wishes to

estimate consumption in one of its service areas, and needs to know the total number of households in the area. The census provides statistics on total numbers of households, but for areas defined by the census, not those used by the water supply agency. So a simple way to estimate households is to overlay the service areas on the census areas, and measure the areas of overlap. Then the household counts from the census are apportioned based on these areas, to obtain estimates that are often remarkably accurate. For example, if one census area has 1,000 households, and 50% of its area lies in Service Area 1, and the remaining 50% lies in Service Area 2, then 500 households are allocated to each of these service areas. When all census counts are allocated in this fashion the service area counts are summed.

Polygon overlay is an important operation for vector GIS, but it has two distinct versions. The examples discussed in this section have been of fields (as defined in Chapter 3.5), such that each point on each map lies in exactly one polygon, and overlay produces a similar map. Overlay in this version is conducted an entire map or *layer* or *coverage* at a time (these terms all have similar meaning in this context). But overlay can also be used for data that represent the *discrete object* view of the world. In this case the input maps can be such that a point can lie in any number of polygons, including zero. This view of the world makes no sense for maps of counties, or land use, but it certainly makes sense for maps of potentially overlapping phenomena like forest fires, or zoning ordinances. In this case, polygon overlay proceeds one polygon at a time, and is used to answer questions about individual polygons, such as "what zoning ordinances affect Parcel 010067?" Some vector GISs support only the field approach (e.g., versions of ArcInfo up to 7), and some (typically those with strong roots in CAD software) support only the discrete object approach -- while others can potentially support both approaches.

Also it is important to recognize the differences between overlay in vector GIS and overlay in raster GIS. The latter is a much simpler operation, and produces very different results, for different purposes. The discussion of raster analysis below includes the essentials of this version of the overlay operation.

3.6.4 Raster analysis

Raster GIS provides a very powerful basis for analysis, that is similar in many respects to the capabilities of other software that also relies on raster representations. For example, some of the raster operations described here will be familiar in concept to people who regularly use software for processing digital photographs or scanned documents, or for processing the images captured by remote sensing satellites. Vector data sets are similar in form to those used in computer-assisted design software (CAD) and in drawing software, but it is unusual to find comparable methods of analysis in these environments.

In raster GIS the world is represented as a series of *layers*, each layer dividing the same project area into the same set of rectangular or square cells or *pixels*. This makes it very easy to compare layers, since the pixels in each layer exactly coincide. But it also means that the ability to represent detail is limited by the pixel size, since any feature on the Earth's surface that is much smaller than a single pixel cannot be represented. Smaller

features can be represented by reducing the pixel size, but only at the cost of rapidly increasing data volume.

Each layer records the values of one variable or attribute, such as land use or county name, for each pixel. The recorded value might be the average value over the pixel, or the value at the exact center of the pixel, or the commonest value found in the pixel. If there are many attributes to record (the census, for example, reports hundreds of attributes for each county in the US), then separate layers must be created for each attribute. However, some raster GISs allow a more efficient solution to this particular problem, in which a single layer is used to record county ID for each pixel, and a related table gives the many attributes corresponding to that ID (see Figure 3.6.6).

[Figure 3.6.6 about here]

An excellent framework for thinking about raster analysis was provided some years ago by Dana Tomlin, who coined the term *cartographic modeling* (Tomlin, 1990). In this scheme there are four basic types of operations (there is some variation in the terms in different versions of the scheme):

- *local* operations, which process the contents of data sets pixel by pixel, performing operations on each pixel, or comparing the contents of the same pixel on each layer;
- *focal* operations, which compare the contents of a pixel with those of neighboring pixels, using a fixed neighborhood (often the pixel's eight immediate neighbors);
- *zonal* operations, which perform operations on zones, or contiguous blocks of pixels having the same values;
- *global* operations, which are performed for all pixels.

Examples of each type of operation are given in the following sections.

3.6.4.1 Local operations

The simplest kind of local operations occur when a single raster layer is processed. This is often done in order to apply a simple reclassification, such as:

> "All areas of Soil Classes 1, 3, and 5 are suitable building sites for residential development; all other areas are not."

This could be operationalized by reclassifying the soil class layer, assigning 1 to all pixels that currently have Soil Class 1, 3, or 5, and 0 to all other pixels. Another might be:

> "Find all areas where the average July temperature is below 25 degrees Celsius."

Again, this could be achieved by reclassifying a layer in which each pixel's attribute is that area's average July temperature.

This kind of reclassification, using simple rules, is very common in GIS. Other local operations on a single layer include the application of simple numerical operations to attributes -- for example, the July temperature layer might be converted to Fahrenheit by applying the formula:

$$F = 9C/5 + 32$$

GISs with well-developed raster capabilities, such as Idrisi, allow the user access to a wide range of operations of this type that create a new layer through a local operation on an existing layer.

Other local operations operate on more than one layer. The raster equivalent of polygon overlay is an operation of this type, taking two or more input layers, and applying a rule based on the contents of a given pixel on each layer to create a new layer. The rules can include arithmetic operations, such as adding or subtracting, as well as logical operations such as:

"If the average July temperature is above 20 degrees celsius and soil type is Class 1 then the pixel is suitable for growing corn, so assign 1, otherwise assign 0."

But notice how different this is from the vector equivalent. The operation is not reversible, since the new layer does not preserve all of the information in all of the input layers. Instead of reorganizing the inputs, raster overlay creates new information from them.

3.6.4.2 Focal operations

Focal operations produce results by analyzing each pixel in relation to its immediate neighbors. Some of these operations work on all eight immediate neighbors, but others focus only on the four neighbors that share a common edge, and ignore the four diagonal neighbors. Sometimes we distinguish these options by referring to moves in the game of chess -- the eight-neighbor case is called the *queen's case*, and the four-neighbor case is called the *rook's case* (see Figure 3.6.7).

[Figure 3.6.7 about here]

Among the most useful focal operations are the so-called *convolutions*, in which the output is similar to an averaging over the immediate neighborhood. For example, we might produce a new layer in which each pixel's value is the average over the values of the pixel and its 8 queen's case neighbors. Figure 3.6.8 shows a simple instance of this. The result of this operation would be to produce an output layer that is smoother than the input, by reducing gradients, lowering peaks, and filling valleys. In effect, the layer has been *filtered* to remove some of the variation between pixels, and to expose more general trends. Repeated application of the convolution will eventually smooth the data completely. Convolutions are very useful in remote sensing, where they are used to remove noise from images (see Chapter *). Of course averaging can only be used if the

values in each pixel are numeric, and measured on continuous scales, so this operation would make no sense if the values represent classes of land, such as soil classes. But in this case it is possible to filter by selecting the commonest class in the neighborhood, rather than by averaging. Note also that special rules have to be adopted at the edge, where cells have fewer than the full complement of neighbors.

[Figure 3.6.8 about here]

If the input layer is a digital elevation model, with pixel values equal to terrain elevation, a form of local operation can be used to calculate slope and aspect. This is normally done by comparing each pixel's elevation with those of its eight neighbors, and applying simple formulae (given, for example, by Burrough and McDonnell, 1998). The result is not the actual slope and aspect at each pixel's central point, but an average over the neighborhood. Because of this, slope and aspect estimates always depend on the pixel size (often called the distance between adjacent *postings*), and will change if the pixel size changes. For this reason it is always best to quote the pixel size when dealing with slope or aspect in a GIS.

Slope and aspect can also be used to estimate the pattern of surface water flow over a DEM, a very useful operation in determining watershed boundaries and other aspects of surface hydrology. Each pixel's elevation is compared to those of its eight neighbors. If at least one of the neighbors is lower, then the GIS infers that water will flow to the lowest neighbor. If no neighbor is lower, the pixel is inferred to be a pit, in which a shallow lake will form. If this rule is applied to every pixel in a DEM, the result is a tree-like network of flow directions (Figure 3.6.9), with associated watersheds. Many more advanced versions of this simple algorithm have been developed, along with a range of sophisticated methods for studying water flow (hydrology) using GIS (see Maidment and Djokic, 2000).

[Figure 3.6.9 about here]

Another powerful form of local operation forms the raster equivalent of the buffer operation, and also supports a range of other operations concerned with finding routes across surfaces. To determine a buffer on a raster, it is necessary only to determine which pixels lie within the buffer distance of the object. Figure 3.6.10 shows how this works in the cases of a point and a line. But suppose we make the problem a little more complex, by asking for pixels that are within a certain travel *time* of a given point, and allowing travel speed to be determined by a new layer. Figure 3.6.11 shows such a layer, and the result of determining how far it is possible to travel in given numbers of minutes. The operation is also known as *spreading*. Of course this assumes that travel is possible in all directions, but this would be appropriate in the case of an aircraft looking for the best route across the Atlantic, or a ship in the open ocean.

[Figure 3.6.10 about here]
[Figure 3.6.11 about here]

This method is used very frequently to find the best routes for power transmission lines, highways, pipelines, and military vehicles such as tanks -- in all of these cases, the assumption of travel in all directions is reasonably accurate. Origin and destination points for the route are defined. A raster GIS is used to create a layer of travel speed, or *friction*. Because this is not uniform, the best route between the origin and destination is not necessarily a straight line. Instead, the best route is modeled as a series of moves in the raster, from each cell to the most appropriate of its eight neighbors, until a complete route of least total time or least total cost is found (Figure 3.6.12). There have been many applications of this GIS method over the past three decades by highway departments, power utilities, and pipeline companies.

[Figure 3.6.12 about here]

The final function considered here is the so-called *viewshed* operation. This also works on a DEM, and is used to compute the area that can be seen from a specified point, by a person at that point. When more than one point is specified, the GIS could be used to determine the area visible from at least one point, or in some military applications it might be useful to determine the area invisible from all points.

The viewshed operation is often used in planning. For example, if a new clearcut is proposed by a forest management agency it might be useful to determine whether the scar will be visible from certain key locations, such as public roads. If a new paper mill is proposed, it might be useful to determine how visible it will be from surrounding areas. One of the complications of the viewshed operation as implemented on a DEM is that it is difficult to incorporate the effects of tree cover on visibility, since a DEM records the elevation of the ground surface, not the top of the tree canopy. Also, it is wise to be careful about the representation of the observer, since far more can often be seen by someone standing on a tower than by someone standing on the ground. Many GIS implementations of the viewshed operation simply assume an observer of average height standing on the ground.

3.6.4.3 Zonal operations

Rasters represent everything in pixels, so line objects and area objects are modeled as collections of pixel values, with no explicit linkage between them. So some operations that are very straightforward in a vector GIS, such as the measurement of area described in Section 3.6.3.1, are quite awkward in raster. If we simply count the pixels assigned to Class 1 on a soil layer, for example, we get the total area of Class 1, not the separate areas of the individual patches of Class 1. Zonal operations attempt to address this, by focusing on operations that examine a layer as a collection of zones, each zone being defined by contiguous pixels of the same class. Zonal operations can be used to measure the areas of patches, or their perimeter lengths, and the results are returned as new layers in which each pixel is given the appropriate measure, evaluated over the zone of which the pixel is a part.

3.6.4.4 Global operations

Finally, global operations apply a simple manipulation to an entire layer, usually returning a single value. Examples include operations to determine the average of all pixel values, or the number of pixels having a specific value, or the numbers of pixels having each unique value, or the total area having each unique value. For example, a global measurement of area on a soil layer with six classes might return six areas, plus the total area. Global operations produce summary statistics, so they are often the last operation in a sequence.

3.6.5 Optimization and design

Many applications of GIS analysis provide information that can be used to make decisions, but some methods come much closer to recommending decisions directly. These are the methods that focus on optimization, by finding the answers that *best* address a problem, and methods that focus on design. Some of the earliest of these were advocated by Ian McHarg, then Professor of Landscape Architecture at the University of Pennsylvania, who developed a method of overlaying maps to find the best locations for new developments (McHarg, 1969). Each map to be overlaid represented some specific issue, such as preservation of agricultural land, or cost of construction, so that when all maps were overlaid the best location would be the one with the least impact or cost. McHarg developed this as a manual method, using transparent sheets, but it is easily automated in a raster GIS. The method of optimum route selection discussed in Section 3.6.4.2 is very similar in concept to McHarg's.

Today, there are many methods of optimization and design available in GIS, in addition to the one already discussed. Many are concerned with optimum routing and scheduling of vehicles operating on a road network, a type of problem that is almost always implemented in vector GIS. In the simplest instance, the GIS is used to find the shortest path through a road and street network between a user-specified origin and destination. Millions of people use this kind of analysis daily when they access Web sites to request driving directions between two street addresses, and the same kind of analysis is possible in the in-vehicle navigation systems that are becoming increasingly common.

Much more sophisticated versions of the same basic idea are used by school bus authorities to design the best routes to pick up children and take them to school, by parcel delivery companies to design routes for their drivers, and by utility companies that schedule daily work orders for their maintenance staff. The term *logistics* is often used to describe this application of GIS.

Another class of optimization problems is termed *location-allocation*. Here the issue is to find sites for activities that serve geographically distributed demand. Examples are retail stores, which must locate centrally with respect to their customers, schools and hospitals, fire stations and ambulance depots, and a host of other services. In all cases the objective is to find locations that are best able to serve a dispersed population. The problem is known as location-allocation because the best solution involves both the *location* of one or more central facilities, and the *allocation* of demand to it, in the form of service areas

(though in some cases this pattern of allocation is controlled by the system designer, as in the case of school districts, but in other cases it is determined by the behavioral choices of customers). In other versions of the problem there is no *location* as such, but instead an area must be divided into the best set of districts, for such purposes as voting, or the assignment of territories to sales staff.

3.6.6 Summary

It should be clear by now that a vast number of options are available in the form of methods of spatial analysis and modeling implemented in GIS. Any good GIS provides many of these, and more are available in the form of add-ons from other vendors, agencies, or individuals. In the case of one popular GIS, ArcGIS from ESRI, the basic package contains many operations in its menus and toolbox; others are available from ESRI in the form of additional special-purpose modules; more are available from other vendors to build on the power of ArcGIS in specific applications; and the package provides tools that allow the sophisticated user to add even more. If your favorite GIS appears not to permit a certain kind of operation, it is certainly worth hunting around, probably on the Web, to see if a suitable add-on is available.

Many GISs also allow operations to be combined into *scripts*, written in special languages such as Visual Basic or Python. Some allow this to be done through a visual interface, with operations represented as icons that can be linked and combined by the user. All of these options have greatly extended the power of GIS to solve problems, and to provide the kinds of information people need to make effective decisions.

Certain aspects should be considered in any application of analysis and modeling using GIS. First, is the analysis more suited to a raster or a vector approach? While many operations are possible in both, this chapter has shown how there can be subtle differences, and how some operations are much more difficult in one form than the other. Second, what level of geographic detail is needed? Most GIS operations give different results depending on the level of geographic detail of the inputs (e.g., the pixel size in a raster operation), so it is important that the level of detail be adequate. Third, what accuracy is achievable? All results from a GIS will be uncertain to a degree, because all inputs are uncertain, so it is important to determine whether accuracy is sufficient. Finally, how should the results be presented? GIS is fundamentally a graphic tool, and the images and pictures it produces can have enormous impact if they are well designed. On the other hand, poor graphic design can easily undermine the best presentation.

References

Burrough P A, McDonnell R A (1998) *Principles of geographical information systems*. New York: Oxford University Press.

De Smith, M J, Goodchild M F, Longley P A (2007) *Geospatial analysis: A comprehensive guide to principles, techniques and software tools*. Winchelsea, UK: Winchelsea Press. Available online at www.spatialanalysisonline.com.

Longley P A, Goodchild M F, Maguire D J, Rhind D W (2005) *Geographic information systems and science*. Chichester, UK: Wiley.

Maidment D R, Djokic D, editors (2000) *Hydrologic and hydraulic modeling support with geographic information systems*. Redlands, CA: ESRI Press.

McHarg I (1969) *Design with nature*. Garden City, NY: Natural History Press.

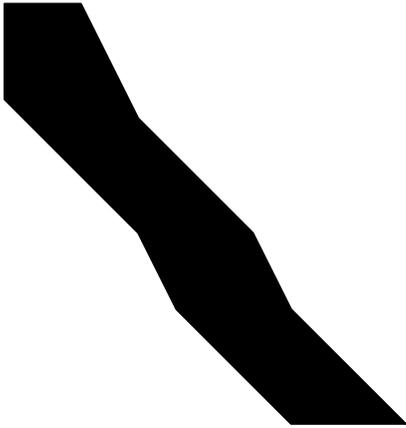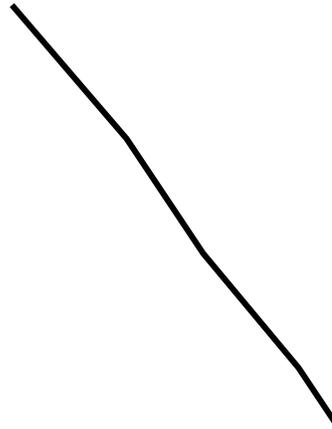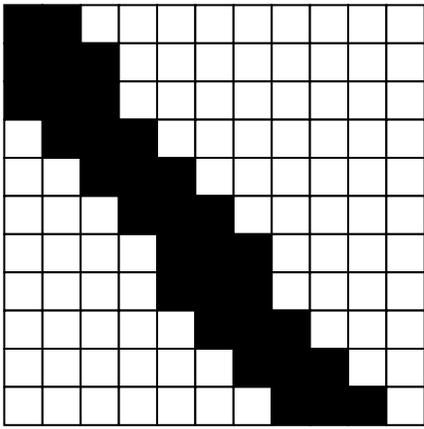Tomlin C D (1990) *Geographic information systems and cartographic modeling*. Englewood Cliffs, NJ: Prentice Hall.

Figure 3.6.1 Three alternative representations of the same road, as (1) a swath of raster pixels, (2) a vector centerline, and (3) a vector area.
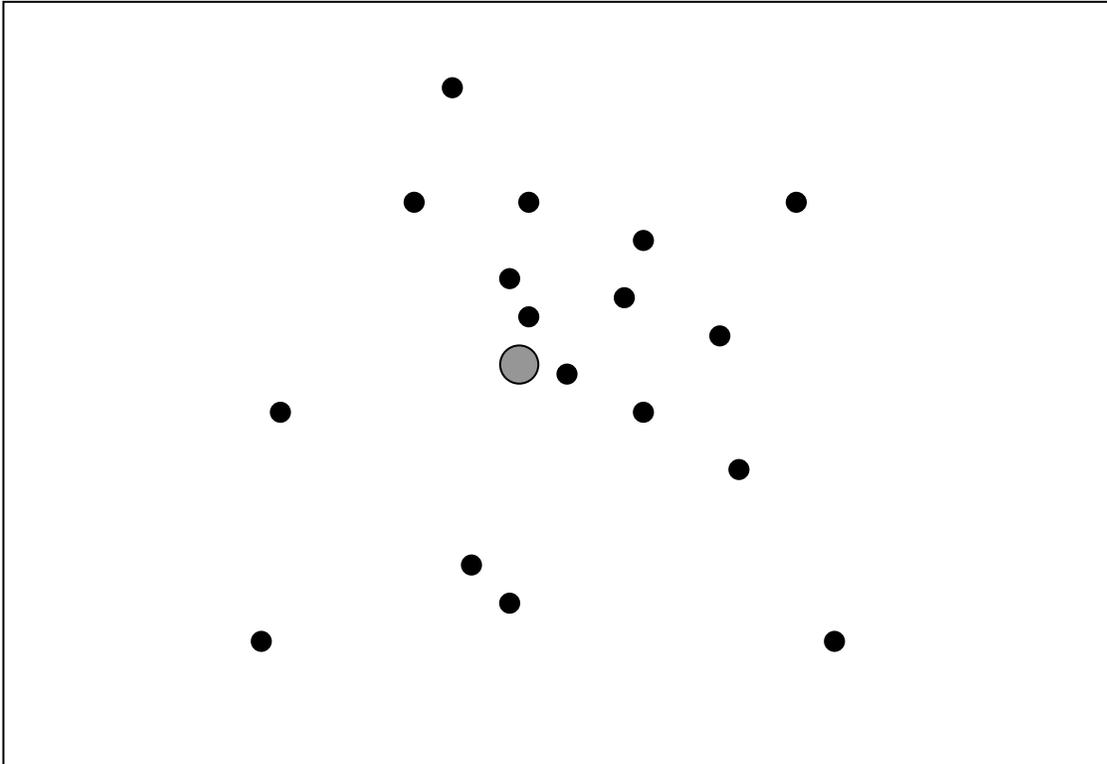
Figure 3.6.2 The locations of the customers of an insurance agent, shown in relation to the agent's office location, and ready for analysis.
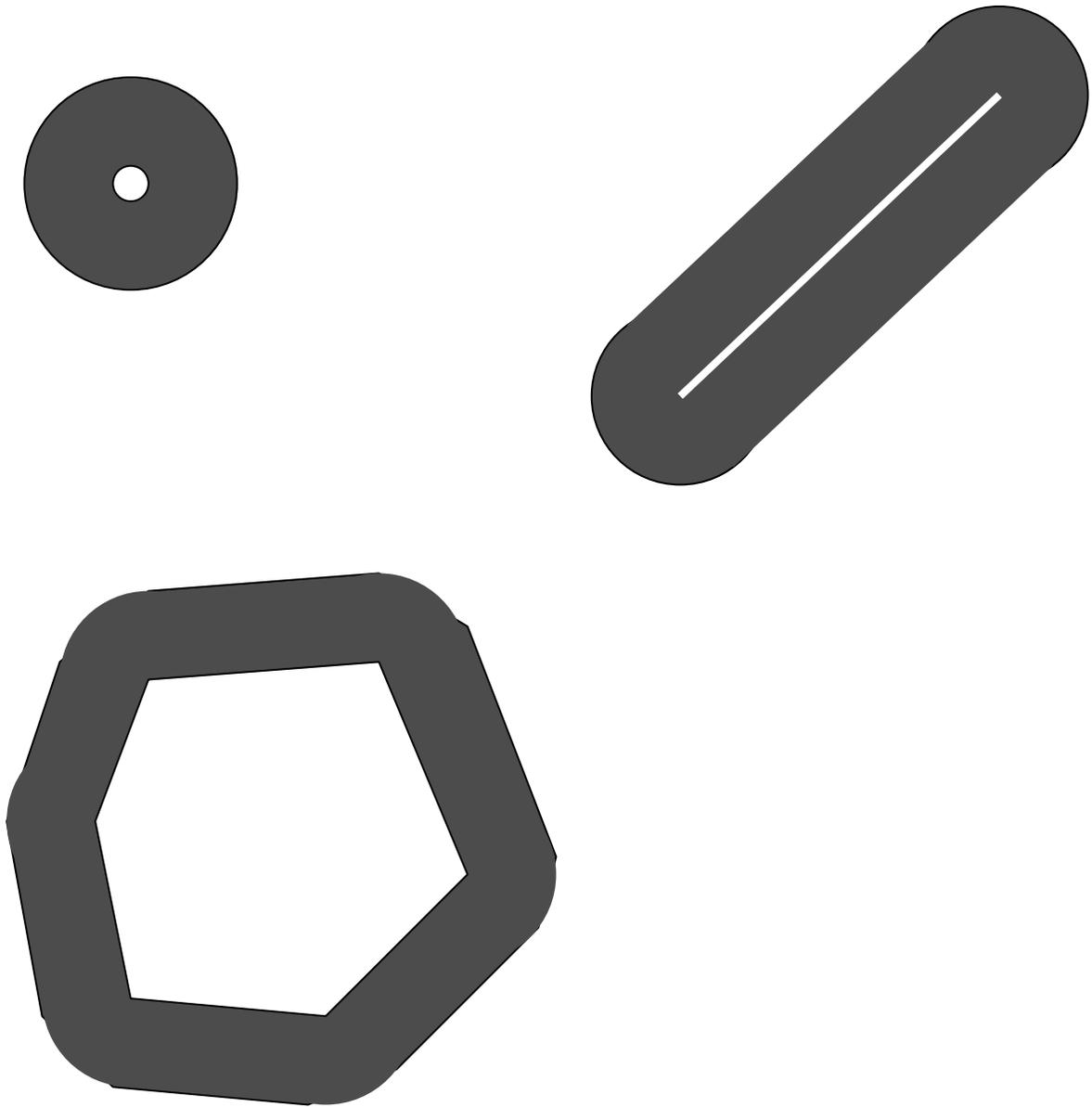
Figure 3.6.3 Buffers for (1) point, (2) line, and (3) area objects.
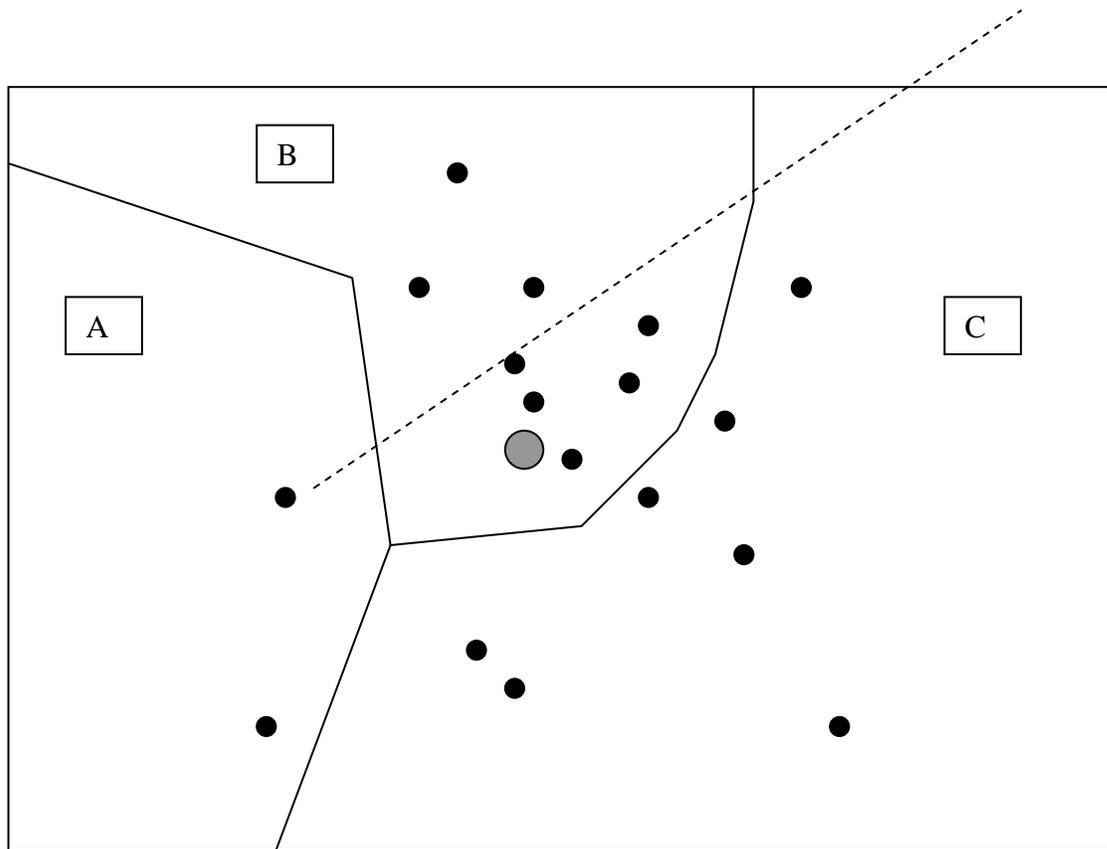
Census Tract	Number of customers	Total households
1	2	249
2	8	377
3	7	401

Figure 3.6.4 By using the point in polygon operation to identify the census tract containing each of the customer locations, it is possible to compare customer counts to other attributes of the tracts, such as the total number of households.
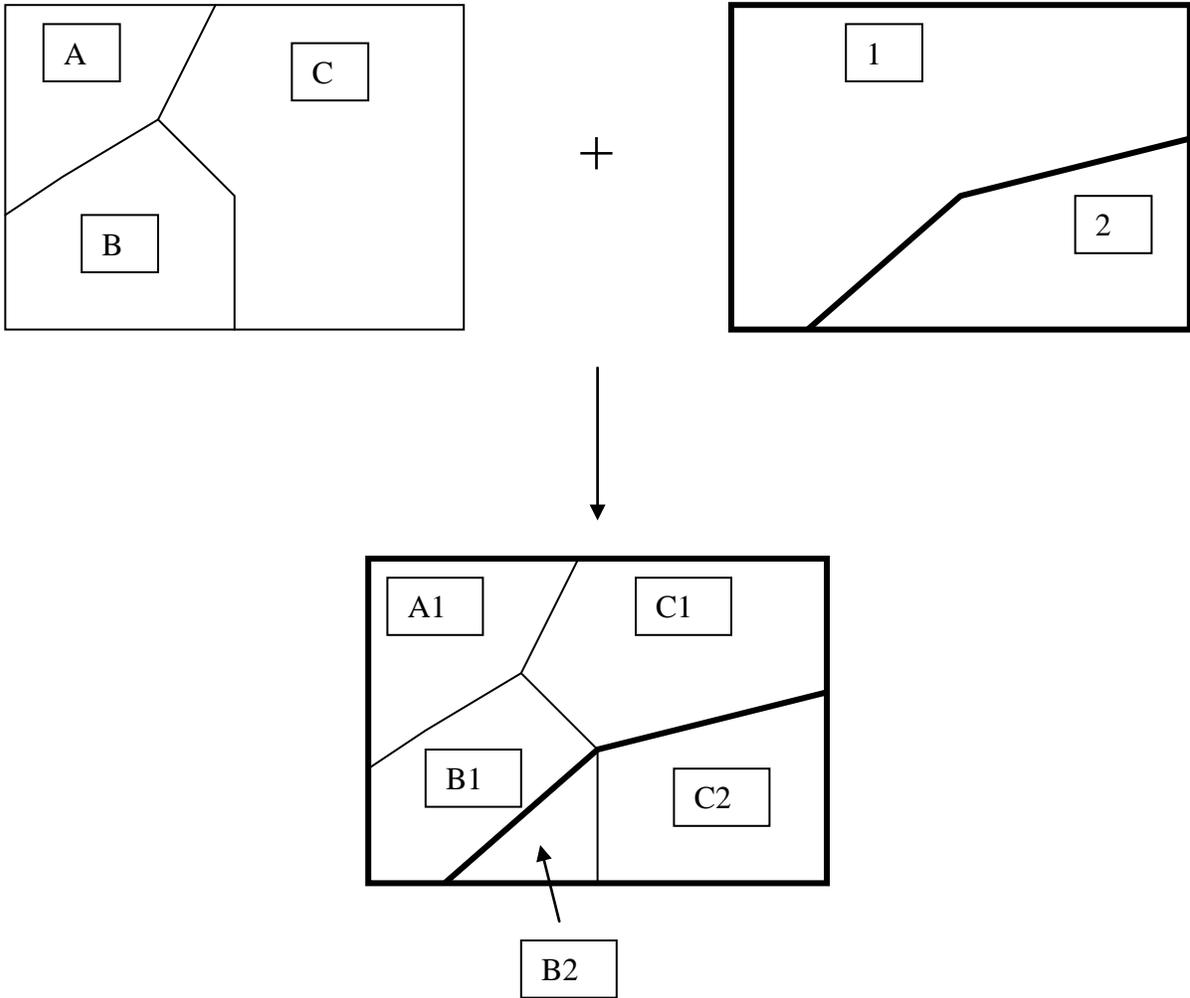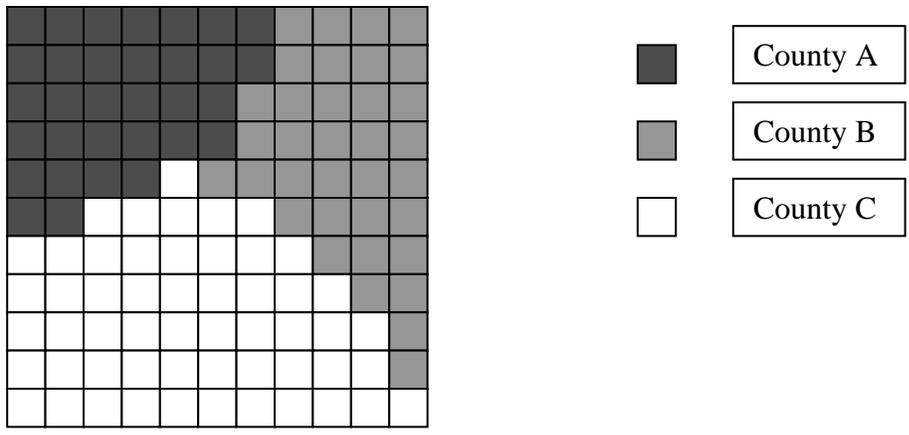
Figure 3.6.5 Polygon overlay, combining a map with three polygons and a map with two polygons to create a map with five polygons. The attributes of the new map combine the relevant attributes of the input maps.

County	Population	Tax rate
A	1,998	5.5%
B	3,941	6.0%
C	15,227	6.75%

Figure 3.6.6 By using raster cells to store the IDs of polygons with multiple attributes in a related table it is possible to store many attributes in a single raster layer.

Figure 3.6.7 Two definitions of a raster cell's neighborhood. In the rook's case only the four cells that share an edge are neighbors, but in the queen's case the neighborhood includes the four diagonal neighbors.

10	9	7	4	5
9	6	5	3	4
6	3	2	2	3

8.5	7.5	5.7	4.7	4.0
7.2	6.3	4.6	3.9	3.5
6.0	5.2	3.5	3.2	3.0

Figure 3.6.8 Application of a convolution filter. Each pixel's value in the new layer is the average of its queen's case neighborhood in the input layer. The result is smoother, picking up more of the general trend. At the edge, where part of the neighborhood is missing, averages are over fewer than nine pixels.
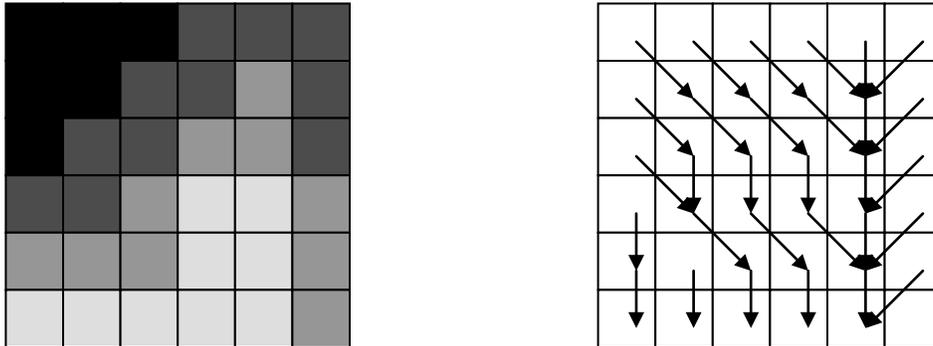
Figure 3.6.9 A digital elevation model, and the result of inferring drainage directions using the simple rule "If at least one neighbor is lower, flow goes to the lowest neighbor".
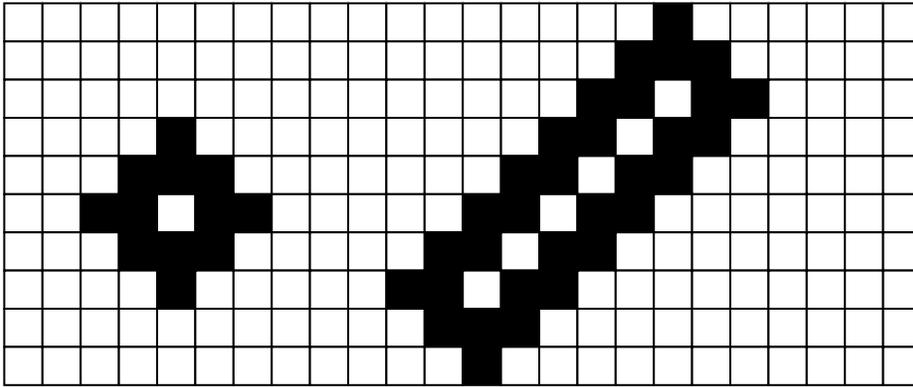
Figure 3.6.10 The buffer operation in its raster form, for a point and a line (indicated by the white cells).
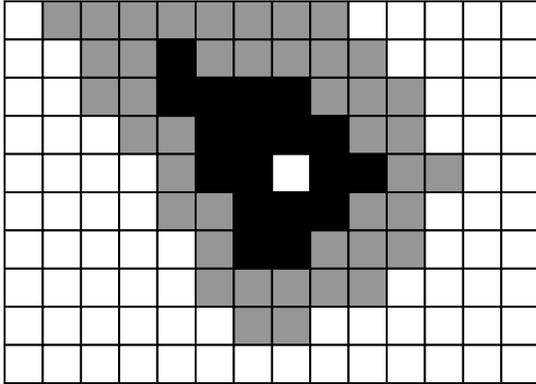
Figure 3.6.11 Illustration of a spreading operation over a variable friction layer. From the central white cell, all cells in the black area can be reached in 10 minutes, and all cells in the grey area can be reached in 20 minutes.
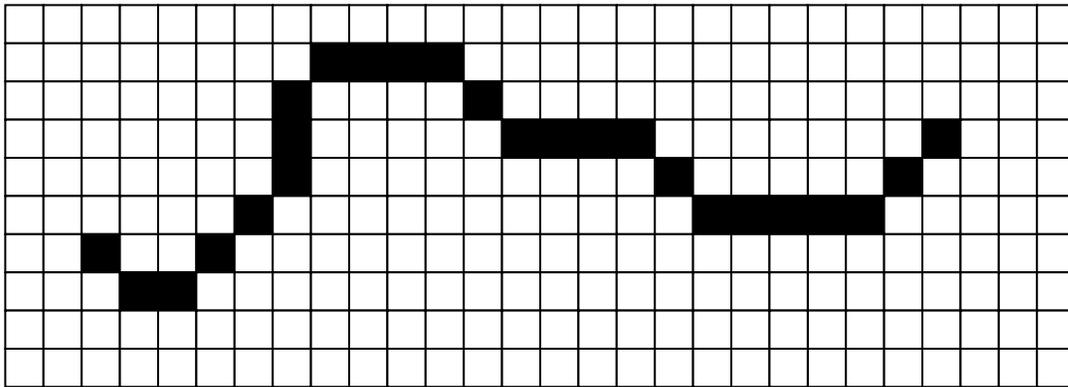
Figure 3.6.12 Finding the least-cost path across a variable-cost layer from a given origin to a given destination, using queen's case moves between raster cells. This form of analysis is often applied to find routes for power lines, pipelines, or new highways.