TITLE: FIELD-BASED SPATIAL MODELING

BYLINE: Michael F. Goodchild, University of California, Santa Barbara, www.geog.ucsb.edu/~good

SYNONYMS: None

DEFINITION:
A field (or continuous field) is defined as a mapping from location **x** to a function *f*. In modeling geographic phenomena the domain of **x** is most often the two dimensions of geographic space, but may include the third spatial dimension for applications that extend above or below the Earth's surface, and may include time for dynamic phenomena. Fields can also be defined on one-dimensional networks embedded in two- or three-dimensional space. Moreover most applications are limited to a specified sub-domain of geographic space, such as the limits of a country or county, or of a map sheet or arbitrarily defined study area. The domain of *f* includes scalar measurements on interval and ratio scales, nominal and ordinal classifications, and vectors describing such directional phenomena as wind or topographic gradient. Field-based spatial modeling can in principle be employed in the representation of any space, including the spaces of the human brain, the surfaces of other planets, or complex buildings.

Fields are one of two ways of conceptualizing the geographic world. By contrast, the discrete-object conceptualization imagines a world that is empty except where it is occupied by discrete, countable objects that maintain identity and geometric form through time. This conceptualization is more often adopted in the representation of biological organisms, manufactured objects, and man-made structures, whereas the continuous-field conceptualization is more appropriate for variables that can be defined at every location in space, such as air temperature, terrain height, soil moisture content, or wind speed.

HISTORICAL BACKGROUND:
The field/object distinction has ancient roots. Its significance for geographic information science was first recognized in the late 1980s and early 1990s [1] [2], and it has come to be acknowledged as the most important distinction underpinning the entire field of geographic data modeling [3] [4]. Consider, for example, the section of the U.S. county boundary map shown in Figure 1. If counties are regarded as discrete objects then one should be able to move them around as pieces of a jigsaw, possibly overlapping and modifying the coastline of the U.S., as shown on the left. On the other hand if the variable *county* is regarded as a single-valued function of location **x**, defined everywhere inside the boundary of the U.S., then the boundaries merely indicate where the value of *county* changes. Edits that move common boundaries are feasible, but not edits that move the coastline, or edits such as that shown, which attempts to move vertices of the boundary outside the coastline. Conceptualizing the phenomenon as a field clearly results in a more consistent representation that is more appropriate to the nature of the phenomenon.

[Figure 1 about here]

Figure 2 shows an area-class map, depicting the variation in vegetation cover class across a geographic domain. If such a map were conceptualized as a collection of discrete objects, then the issue of accuracy would come down to such questions as “Is the number of areas correct?”; “Are the boundaries in the correct places?”, and “Are the classes assigned to each area correct?” However uncertainty would be addressed very differently under a field conceptualization; one could ask only one form of question, “Is the value recorded at **x** correct?” It turns out that this second conceptualization is far more productive and tractable than the first [5].

[Figure 2 about here]

Strong associations have been recognized between the two alternative conceptualizations and patterns of human thought. Many would argue that the discrete-object view is more compatible with human cognition – that our brains are in effect hard-wired to segment any visual image into a collection of discrete objects, and to track their movements through time. On the other hand the continuous-field view underlies many of the most significant advances in science, including electro-magnetism (the Maxwell equations), hydrodynamics (the Navier-Stokes equation), and quantum mechanics (the Schrödinger equation). Some of the most challenging problems in science are described in partial differential equations defined on fields, and solved using a variety of computational methods. The distinction is strongly linked to concepts of scale, as for example when the behavior of a group of ants is modeled at very detailed scale as interactions between discrete objects, or at a coarser scale as modifications to a continuous field of ant density. Very crudely, continuous-field conceptualizations tend to be more common in the natural and physical sciences, while discrete-object conceptualizations are more often found with reference to social phenomena.

SCIENTIFIC FUNDAMENTALS:

Computers are fundamentally discrete machines, founded on the representation of information in a two-valued alphabet, and as a result discrete-object conceptualizations are more readily implemented in spatial databases. There is little ambiguity, for example, in the representation of the current location of every aircraft in an airline’s fleet as a point in space. However other phenomena are inherently continuous across space, including terrain, rivers, roads, and the tracks of moving objects, and their representation necessarily involves some form of discretization. A river, for example, may be broken into reaches, either at junctions or at points where the direction of the river changes significantly. Each reach will be represented using a simple mathematical function, most often a straight line but sometimes as an arc of a circle or a spline function. Roads in a road network may be broken into segments at intersections, or at other points where direction changes. These stages necessarily modify the phenomena they are used to represent, since the discretized object will in most cases differ geometrically from the original, and discretization is essential if a geometric shape that is potentially infinitely complex is to be represented in a digital store of finite capacity. Thus in almost all cases the representation of real-world geometry requires the loss of some degree of detail.

This stage of discretization is necessary whenever an arbitrarily shaped geographic feature must be represented in digital form. But a further stage is needed when the characteristics of the phenomenon also vary continuously over space, in other words in the representation of phenomena conceptualized as continuous fields. Figure 3 shows the six methods most commonly employed in spatial databases and geographic information systems. Two are based on a raster, in other words the discretization of space into a regular array with finite spacing, while the remaining four use vector methods, specifying the position of each element of the discretization as coordinates and representing volumes as collections of polygonal faces, areas as polygons, and lines as polylines.

[Figure 3 about here]

Figure 3a shows one of the raster options, the representation of a spatially continuous phenomenon as a collection of sample values regularly spaced over a rectangular array. This is the method most commonly employed in the representation of terrain, in what are known as digital elevation models (DEMs). The spacing of sample heights implies a well-defined level of spatial resolution. Note, however, that since it is impossible to lay a rectangular grid over a curved surface, it is similarly impossible to construct a DEM of a significant part of the Earth's surface with a precisely constant spatial resolution.

Figure 3d shows the other raster option. Here the study area has been divided into rectangular cells, and a single measurement provided in each cell. This is often the mean value in the case of measurements on interval and ratio scales, but in other cases may be a modal value or some more complex function of the values of the field within the cell. Again spatial resolution is constant and well defined. This approach is most commonly encountered in the discretization of images, such as those obtained from satellites.

Figure 3b shows a vector option. In this case the field has been sampled at a set of irregularly spaced locations, and the value of *f* reported. Sample locations may be selected using some specified rule, or may be a historic artifact, as they are in the case of weather data, when the points correspond to weather-observation stations. Nothing is known about the values of *f* between sample locations, though a wide range of techniques exist for making intelligent guesses under the rubric of spatial interpolation.

Figure 3c shows another vector option involving irregularly spaced sample points, though in this case a set of triangles have been created by connecting them. The field variable is assumed to vary linearly between points, ensuring continuity of value across triangle edges. In essence this option, which is generally known as a triangulated irregular network (TIN) or triangular mesh, adds a specific method of spatial interpolation to Option 3b. It is commonly used as a method for representing terrain, where it is particularly efficient for terrains characterized by long uniform slopes and sharp ridges; and as an internal representation in contouring algorithms.

Figure 3e shows the discretization of a field as a collection of space-exhausting, non-overlapping areas, that will themselves be discretized geometrically as polygons. The field variable is assumed uniform within each area. This is the commonest discretization when the field variable is nominal or ordinal, as it is for maps of soil class and land cover class for example. It is also commonly used to represent data collected by statistical agencies and aggregated by reporting zones such as counties or census tracts. In such cases the number of variables recorded for each reporting zone may be very large. For some of these variables, termed spatially intensive, the value reported will represent the mean within the zone of a variable such as average income, or a proportion such as percent black, or a density per unit area. In other cases the reported value will be a total, such as total population or total income, corresponding to the integration of a density over the reporting zone's area; this type of value is termed spatially extensive.

Finally, Figure 3f shows the last example, in which the field is represented as a collection of digitized isolines. This method is commonly used to capture the contours shown on topographic maps. While it is effective for visual purposes, its value from an analytic perspective is far less than the DEM or TIN because of the very uneven sampling that is achieved.

KEY APPLICATIONS

Many examples of phenomena conceptualized as fields have already been cited; this section addresses the functions that are commonly applied to field representations, and the software that implements them. The functions described in this section are commonly found in GIS packages, and they and many others are reviewed by De Smith, Goodchild, and Longley [6].

A variety of functions are available to manipulate representations of topography. Most use the DEM option (3a) though very powerful algorithms have been described for similar operations on TINs. Visualization is a common requirement, and functions have been developed to compute surface gradient and hence simulated illumination; to compute and plot isolines; and to compute solar insolation as a key variable in understanding vegetation patterns on rugged topography. Another class of algorithms concern visibility, and can be used to compute the area visible by an observer positioned a given height above the terrain. Algorithms have been described for computing the most exposed and most concealed points on a landscape, as well as most exposed and most concealed routes; and to position a minimal number of observers such that the entire landscape can be observed.

Another important collection of algorithms concerns drainage. Starting with a DEM, it is possible to compute drainage directions as a vector field, and to integrate these into catchments and stream channels. DEMs are used to predict and manage flooding, and to plan modifications to the landscape such as the construction of levees.

Reference was made earlier to methods of spatial interpolation, which address the task of predicting the value of a field at locations where it has not been measured. Most often these methods are applied to the representations illustrated in Figures 3a and 3b, but a related technique known as areal interpolation has been devised for the task of predicting the values associated with areas that do not match (cut across the boundaries of) the reported areas of 3d and 3e.

Of particular interest are algorithms that produce representations of fields from collections of discrete objects. They include density estimation, which produces a field of feature density, most often of points; and calculation of the distance from any point in the plane to the nearest of a collection of discrete objects.

The most powerful collections of field-based manipulation functions clearly exist for raster data, because of the possibility that a collection of fields can be represented using a set of co-registered rasters. This is the principle of raster GIS, most clearly illustrated by packages with firm roots in raster-based discretizations, such as Idrisi and GRASS. Most have adopted a common language for expressing instructions known as map algebra [7]. A powerful alternative geared particularly to simulation is PCRaster (http://pcraster.geo.uu.nl), developed at the University of Utrecht and having its own manipulation language that is considerably more succinct and powerful than map algebra.

FUTURE DIRECTIONS

The comparative importance of continuous fields and discrete objects is a matter of continual debate. On the one hand, the majority of GIS applications occur in worlds where the objects of interest are well-defined and often man-made. On the other hand, many phenomena in the natural world are essentially continuous, and representing them as collections of discrete objects invites error and misuse. Cognitive scientists would argue that humans are hard-wired to see the world as a collection of discrete objects, while environmental scientists might argue that continuous fields are one of the most significant breakthroughs in the history of science, lying at the root of such developments as the calculus and hydrodynamics. Many of the forms of analysis commonly used in the environmental sciences are based on fields; while most of those commonly used in the social sciences are based on discrete objects.

GIS software today embraces both conceptualizations. But it does so in a somewhat unsatisfactory manner, in which representations of continuous fields must be reduced to collections of discrete objects, without any record of that process of reduction. Thus it is impossible for the user of a GIS database to enquire whether a collection of points in the database represents a set of points sampling a field, or a set of isolated point-like objects in an otherwise empty space. Unfortunately this means that inappropriate operations can easily be performed. Nothing prevents the GIS user from applying spatial interpolation to points conceptualized as discrete objects, or from computing a density field of sample points; the former is of course more disastrous than the latter. Similarly nothing prevents a GIS user from moving an isoline so that it crosses another, or from making two polygons in a Figure 3e representation overlap.

Several interesting and potentially powerful research directions have been pursued in the hope of eventually improving this situation. One approach has been to argue that the user should be able to interact with the concept of a field directly, rather than with the elements of one of the six representations of Figure 3 as at present. Another has been to search for a visual paradigm for handling fields that matches and has similar power to the visual representation of sets of discrete objects that is found in UML (Unified Modeling

Language) and related methods. Now that GIS technology includes the capability to design databases in UML and to create and populate the necessary tables automatically, it would make good sense to develop parallel methods for handling fields. Finally, much recent effort has gone into finding ways of reconciling and bridging the field/object dichotomy [8].

CROSS REFERENCES: Digital Elevation Models, Geographic Information System, Raster Data Management and Multidimensional Arrays, Spatial Data Analysis, Triangulated Irregular Networks, Unified Modeling Language.

RECOMMENDED READING

1. Goodchild M. F. (1989): Modeling error in objects and fields. In Goodchild M. F., Gopal S., eds. Accuracy of Spatial Databases. Taylor and Francis, 107-114.
2. Couclelis H. (1992): People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In Frank A. U., Campari I., Formentini U., eds. Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, Springer-Verlag, 65-77.
3. Worboys M. F., Duckham M. (2004): GIS: A Computing Perspective. CRC Press.
4. Longley P. A., Goodchild M. F., Maguire D. J., Rhind D. W. (2005): Geographic Information Systems and Science. Wiley.
5. Zhang J. X., Goodchild M. F. (2002): Uncertainty in Geographical Information. Taylor and Francis.
6. De Smith M. J., Goodchild M. F., Longley P. A. (2007): Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools. Winchelsea Press.
7. Tomlin C. D. (1990): Geographic Information Systems and Cartographic Modeling. Prentice Hall.
8. Goodchild M. F., Yuan M., Cova T. J. (2007): Towards a general theory of geographic representation in GIS. International Journal of Geographical Information Science 21(3): 239-260.

FIGURE CAPTIONS

1. Counties conceptualized as (a) discrete objects and (b) a continuous field. As a discrete object, any county is free to move independently of its neighbors and the coastline. As a continuous field, however, only the common boundaries where county values change can be moved, and not so far as to intersect the coastline (the edit shown would not be accepted).
2. A map of vegetation cover class conceptualized as a nominal field, superimposed on the Santa Barbara area. Each area denotes a particular type of vegetation.
3. The six common methods of discretizing a field. See text for explanation.