3

Representing Geography

OVERVIEW

- This chapter introduces the concept of representation, or the construction of a digital model of some aspect of the Earth's surface.
- The geographic world is extremely complex, revealing more detail the closer one looks. So in order to build a representation of any part of it, it is necessary to make choices, about what to represent, at what level of detail, and over what time period.
- Generalization methods are used to remove detail that is unnecessary for an application, in order to reduce data volume and speed up operations.

LEARNING OBJECTIVES

After reading this chapter, your students will be able to explain:

- **The importance of understanding representation in GIS;**
- **The concepts of fields and objects and their fundamental significance;**
- **What raster and vector representation entails and how these data structures affect many GIS principles, techniques, and applications;**
- **The paper map and its role as a GIS product and data source;**
- **The importance of generalization methods and the concept of representational scale;**
- **The art and science of representing real-world phenomena in GIS.**

KEY WORDS AND CONCEPTS

Digital, binary, representation, Tobler's First Law of Geography, attributes, the fundamental problem (the world is infinitely complex), discrete objects, attribute tables, continuous fields, raster, vector, pixel, polygon, polylines, TIN, representative fraction, generalization, simplification, weeding.

OUTLINE

CHAPTER SUMMARY

3.1 Introduction

- This section mentions the notions of
    - *spatial concepts* such as containment and proximity
    - *data models*, the structures and rules that are programmed into a GIS to accommodate data
    - *ontologies*, the frameworks that we use for acquiring knowledge of the world
- Figure 3.1 is a time-space *representation* of the daily journeys of several people
- Representations are reinforced by the rules and laws that we humans have learned to apply to the unobserved world around us, such as using spatial interpolation to guess the conditions that exist in places where no observations were made.
- Introduces *Tobler's First Law of Geography*: Everything is related to everything else, but near things are more related than distant things.

3.2 Digital representation

- This section introduces the notions of *digital* and *binary.*

- Technical Box 3.1 outlines fundamentals about the binary number system, including short (16-bit) and long (32-bit) storage, ASCII, floating point numbers, and BLOBs (binary large object).
- Benefits of digital representations of geography include: can be handled in ways that are independent of meaning; easy to copy and transmit; stored at high density; easy to transform, process, and analyze.

3.3 Representation for what and for whom?

- *Geographic* representation is defined as a representation of some part of the Earth's surface or near surface, at scales ranging from the architectural to the global
- Short history of maps and the significance of their use historically
- Like maps, any application of GIS requires attention to questions of *what* should be represented, and *how*.

3.4 The fundamental problem

- Geographic data are built up from atomic elements, or facts about the geographic world.
- At its most primitive, an atom of geographic data (strictly, a datum) links a *place*, often a *time*, and some descriptive *property*.
- *Attribute* refers to this descriptive property and may:
    - vary slowly or rapidly
    - be physical/environmental or social/economic
    - identify a place or entity
    - provide a measurement of something at that place
- Technical Box 3.3 defines the types of attributes: nominal, ordinal, interval, ratio, and cyclic
- The *fundamental problem* is "the world is infinitely complex, but computer systems are finite".

3.5 Discrete objects and continuous fields

3.5.1 Discrete objects

- In the *discrete object* view, the world is empty, except where it is occupied by objects with well-defined boundaries that are instances of generally recognized categories.
    - Objects can be counted

    - Objects have dimensionality: 0-dimension (points), 1-dimension (lines), 2-dimensions (areas, polygons)
- 3-dimensions are problematic in GIS
- Associating attributes to discrete objects can be expressed in a table (see Table 3.1):
    - each row corresponding to a different discrete object,
    - each column corresponding to an attribute of the object.
- A table does not look like the real world

3.5.2 Continuous fields

- The continuous field view represents the real world as a finite number of variables, each one defined at every possible position.
- Continuous fields, on the other hand, can be distinguished by what varies, and how smoothly.
- The challenge of counting the lakes in Minnesota provides an elegant illustration of various ways to think about the conceptual transition from objects to fields
- Technical Box 3.4 explains the concept of 2.5 dimensions

3.6 Rasters and vectors

- raster and vector are two methods that are used to reduce geographic phenomena to forms that can be coded in computer databases
- In principle, each can be used to code both fields and discrete objects, but in practice there is a strong association between raster and fields, and between vector and discrete objects.

3.6.1 Raster data

- In a raster representation geographic space is divided into an array of cells, each of which is usually square, but sometimes rectangular (Figure 3.9).
- All geographic variation is then expressed by assigning properties or attributes to these cells.
- The cells are sometimes called pixels (short for *picture elements*).
- A few fundamentals of remote sensing are introduced
- When information is represented in raster form all detail about variation within cells is lost, and instead the cell is given a single value based on a rule such as largest share or value at the central point.

3.6.2 Vector data

- Usually, in a vector representation, all lines are captured as points connected by precisely straight lines (some GIS software also allows curves).
- Features are captured as a series of points or *vertices* connected by straight lines
    - areas are often called polygons
    - lines are sometimes called polylines
    - the choice between raster and vector is often complex, as summarized in Table 3.3.

3.6.3 Representing continuous fields

The six approximate representations of a field used in GIS are (Figure 3.12)

- Regularly spaced sample points
- Irregularly spaced sample points.
- Rectangular cells.
- Irregularly shaped polygons.
- Irregular network of triangles
- Polylines representing contours

3.7 The paper map

- Key property of a paper map is its *scale* or *representative fraction*, defined as the ratio of distance on the map to distance on the Earth's surface.
    - when a scale is quoted for a digital database it is usually the scale of the map that formed the source of the data.
    - while the paper map is a useful metaphor for the contents of a geographic database, we must be careful not to let it limit our thinking about what is possible in the way of representation.
- Digital representations can include information that would be very difficult to show on maps.
- Digital databases can represent all three spatial dimensions, including the vertical, whereas maps must always show two-dimensional views.

3.8 Generalization

- ways of simplifying the view of the world include:
    - describe entire areas, attributing uniform characteristics to them, even when areas are not strictly uniform;

    - identify features on the ground and describe their characteristics, again assuming them to be uniform;
    - limit our descriptions to what exists at a finite number of sample points, hoping that these samples will be adequately representative of the whole.
    - some degree of generalization is almost inevitable in all geographic data.
- A map's specification defines how real features on the ground are selected for inclusion on the map.

3.8.1 Generalization about places

- Geometry generalization methods, see Figure 3.14
- Attribute generalization methods include
    - *Simplification*
    - *Smoothing*
    - *Aggregation*
    - *Amalgamation*
    - *Merging*
    - *Collapse*
    - *Refinement*
    - *Exaggeration*
    - *Enhancement*
    - *Displacement*

- Weeding is the process of simplifying a line or area by reducing the number of points in its representation
- The operation of the Douglas-Poiker weeding algorithm is shown in Figure 3.16.

3.8.2 Generalization about properties

- Many representations in GIS bring together multiple properties of places as composite indicators of conditions at particular locations
- Representation may not be fit for its purpose if it is based upon inappropriate indicators of real-world conditions

3.8.3 Representation using VGI

Generalizability of representations based on VGI is questioned as follows:

- Are all places equally accessible to volunteers?
- Are volunteers more likely to provide data on places or properties that interest them than those that do not?

- Where volunteers supply information about themselves, is this representative of other non-volunteers?
- Are volunteers at liberty to collect locationally sensitive data?
- Should volunteers supply data that are open to malevolent use?
- Is it socially acceptable to make available any observations of uniquely identifiable individuals?
- Is the information date-stamped, as an indicator of its provenance and current reliability?

ESSAY TOPICS

Again, these are in rough order of sophistication of necessary response:

1. What do you understand by the terms raster and vector? How would you decide which to use in any specific given project? Describe the situations encountered in this chapter where the distinction between raster and vector becomes blurred.
2. Why is digital map generalization necessary and why does it involve so much more than line simplification?
3. GIS professionals and cartographers represent and understand ‘geography’ in certain ways. Compare and contrast with your own everyday experience of spaces and places.
4. Distinguish between nominal, ordinal, interval, ratio, and cyclic data types, and explain how this classification does not cover all the attribute types that might be recorded in a GIS
5. Table 3.3 summarized some of the arguments between raster and vector representations. Expand on these arguments, providing examples, and add any others that would be relevant in a GIS application.
6. “No representation of geographic phenomena can ever be perfect”. Is this true, are there exceptions, and what implications does this statement have for users of GIS?
7. A map is defined by the International Cartographic Association as ‘A conventionalized image representing selected features or characteristics of geographic reality designed for use when spatial relationships are of primary relevance’ (Board, 1991). Explain the reasoning behind every element of this definition.
8. An argument has been made that representing geography in a GIS can be improved by understanding human perception and cognition. Review and discuss this argument. In what ways do you agree or disagree with this idea?
9. To what extent do you agree with the two propositions expressed by Goodchild in 1995 that 'It is GIS's supreme conceit that one can structure a useful representation of

geographical knowledge in the absurdly primitive domain of the digital computer' and that 'it is geography's conceit that one can accomplish the same with pen and paper'?

10. With reference to sound, maps, and photographic images, write an account of the advantages and disadvantages of digital representations over analog ones.

MULTIPLE-CHOICE QUESTIONS

1. The following lists show a series of numbers and some binary equivalents. Each binary number consists of four binary digits (bits). Match each number to its representation in the binary system by writing its sequence letter in the final column:

Number	**Binary**	**(a)-(e)?**
15	0000	
0	1001	
9	0110	
6	1111	
10	1010	

2. Which form of representation does a paper map use?

    a. analog
    b. digital
    c. binary
    d. decimal

3. Which is **NOT** a commonly used coding scheme for images?

    a. JPEG
    b. GIF
    c. MP3
    d. TIFF

4. The table lists some commonly used attributes. For each write down the measurement scale (nominal, ordinal, interval, ratio) most often used to record it:

Attribute	**Scale**
Building height	
Temperature in ^{o}C	
Land capability	
Soil type	
Atmospheric pressure	

5. Which is **not** characteristic of discrete objects?
    a. they may include points, lines, and areas
    b. they completely cover the space
    c. they can overlap
    d. they can be counted

6. When weather forecasters isolate a cold front on a weather map are they representing this phenomenon as:
    a. A table
    b. A surface
    c. A field
    d. An object?

7. Of the six ways of representing a field in GIS, which is usually used for remotely sensed images?
    a. a raster of regularly spaced sample points
    b. a raster of rectangular cells
    c. irregularly spaced sample points
    d. irregularly shaped polygons

8. Ontology is the study of:
    a. the processes by which knowledge is acquired
    b. the processes by which the world is described

c. properties that are not affected by warping space

d. the naming of places on the Earth's surface?

9. Figure 3.13 is an extract from a 1:24,000 USGS topographic map. Study it carefully and then complete the table below by specifying how it represents each of the named entities in the appropriate blank cells. Note that there may be more than one method for each entity:

Entity	Field	Point object	Line object	Area object
Relief				
Housing				
Wood/forest				
Roads				
Railway				
Public buildings				
Rivers				
Lakes				

CLASS AND INDIVIDUAL ACTIVITIES

1. In class, one of the best ways of stimulating discussion is to look critically at a series of paper maps. Understanding the nature of maps and mapping is a vital step towards understanding GIS. Several of these suggested projects demand access to an extensive map library, or its copyright cleared digital equivalent. Appropriate examples for individual study will be found using the 'image' search at www.google.com.

2. What fraction of the Earth's surface have you experienced in your lifetime? Consider how you would make diagrams at appropriate levels of detail, to show a) where you have lived in your lifetime, b) how you spent last weekend. How would you describe what is missing from each of these diagrams? How could these be shared?

3. The early explorers had limited ways of communicating what they saw, but many were very effective at it. Examine the published diaries, notebooks, or dispatches of one or two early explorers and look at the methods they used to communicate with others. What words did they use to describe unfamiliar landscapes and how did they mix words with sketches?

4. The following exercise can be conducted in class as a sort of collective 'thought experiment'. The 'answers' don't matter. What it does is to look at and discuss a variety of 'paper' maps and 'pre-maps'. It tries to:

- get students into looking at maps and map-like products and so introduce them to the art and science of cartography;
- help put the products that can be obtained from a GIS into a cartographic perspective.

The idea is to find 'sequences' from what all can agree are definitely 'maps' to what are equally definitely images and/or 'plans'.

For example, all maps are scaled representations of the real world but at what point does a map become what we normally think of as a 'plan' as we progress from say, 1:250,000 road map, through 1:50,000 topographic map, 1:10,000 map, to 1:2,500, a map of a room, and even smaller scale? Students should be encouraged to find their own examples at websites, in a local library, or in their city or university map store. For links to URLs concerned with cartography, see http://oddens.geog.uu.nl/index.html.

Similar sequences can be assembled to illustrate how maps distort (is an area cartogram a 'map'? How does this distortion differ from that on a UTM projection? What can be gleaned about how they generalize, the viewpoint and datum they adopt, their subject matter (is a 'map' of DNA still a 'map'?). The book by Dodge, Kitchin and Perkins (2009) Rethinking Maps: New Frontiers in Cartographic Theory (Routledge) also has some good source materials.

5. This is an in-class or project variant on Multiple-Choice Question 10. Take map extracts from the Ordnance Survey of GB or a similar National Mapping Agency, maps at five different scales (1:10,000, 1:25,000, 1:50,000, 1:250,000 & 1:625,000),

and spend, say, 15 minutes making notes and thinking about how these maps become more or less generalized as the scale changes. Individual students can limit what they do to one or two examples of how specific features are represented as point, line, or area objects. They should be encouraged to think about how objects change their representation with scale and the issue of planimetrically 'correct' representation vs the use of symbolism. The implications of these findings for drawing maps from a single digital database should be explored.

6. The word 'mapping' occurs in many contexts. Visit your preferred Internet book vendor (for example www.amazon.com) and do a search for all the books on offer that have the words 'map' or 'mapping' in their titles. How many different contexts can you find, and what does this tell you about geographic maps?

7. Fields are represented in many ways on a typical national-series topographic map. Take a typical map at a scale around 1:20,000 through to 1:50,000 and list and describe all the cartographic methods used. Your list might include, for example, spot heights, contours, hypsometric tinting in which height bands are shown in a graded series of colors, hill shading (shadowing), even hachuring. Having completed this exercise, speculate on how easily each method could be implemented from the six ways of representing a field shown in Figure 3.12

8. Use tracing paper to make a tracing from a map of a limited number of objects, preferably illustrating point, line, and area object classes (see exercise 6). Then overlay on this some transparent graph paper with a regular grid of square cells. Use this grid to hand code the map as a vector data structure and as a raster, both coded according to what is being recorded. This exercise can be done 'on screen', but there is value in using pencil and paper. Its point is to illustrate the nature of a digital caricature of the underlying geography.

In most critiques of GIS, one issue that repeatedly arises is the tendency of GIS to promote or privilege certain types of representation at the expense of alternatives. Organize two teams to debate the motion that 'This house believes that in representing geography in the ways they do, GIS promote a one-eyed view of the world'. A good reference is Schuurman, N. (2004) *GIS: a short introduction,* Oxford: Blackwell.

FURTHER READING

Chrisman N R 2002 Exploring Geographic Information Systems (2nd edn). New York: Wiley.
Fisher P, Unwin D (2005) Re-Presenting Geographical Information Systems. Chichester: Wiley.
McMaster R B, Shea K S 1992 Generalization in Digital Cartography. Washington, DC: Association of American Geographers.
National Research Council 1999 Distributed Geolibraries: Spatial Information Resources. Washington, DC: National Academy Press. Available: www.nap.edu
Gore A 1992 Earth in the Balance: Ecology and the Human Spirit. Boston: Houghton Mifflin
Mardia K V, Jupp P E 2000 Directional Statistics. New York: Wiley
Ryerson R A editor 1998 Manual of Remote Sensing. New York: Wiley.

RELATED READING

Longley P A, Goodchild M F, Maguire D J, Rhind D W, editors 2005 *Geographical Information Systems: Principles, Techniques, Management and Applications* (abridged edition). Hoboken, NJ: Wiley.

- Spatial representation: the scientist's perspective, J F Raper
- Spatial representation: the social scientist's perspective, D J Martin
- Spatial representation: a cognitive view, D M Mark
- Time in GIS and geographical databases, D J Peuquet
- Representation of terrain, M F Hutchinson, J C Gallant
- Digital remotely sensed data and their characteristics, M Barnsley

Maguire D J, Goodchild M F, Rhind D W editors 1991 *Geographical Information Systems: Principles and Applications*. Harlow, UK: LongmanConcepts of space and geographical data, A C Gatrell, pp. 119-34

- GIS and remote sensing, F W Davis and D S Simonett, pp. 191-213
- High-level spatial data structures for GIS, M J Egenhofer and J R Herring, pp. 227-37

ONLINE RESOURCES

- ESRI Virtual Campus course, *Turning Data into Information* by Paul Longley, Michael Goodchild, David Maguire, and David Rhind (training.esri.com)
  - Module 1: Basics of Data and Information, Unit: Representing geography
  - Section 3.1, Module 1: Basics of Data and Information
    Unit: Representing geography

- Section 3.3, Module 1: Basics of Data and Information
  Unit: Representing geography
  Sub-unit: How are geographic data represented?
  Sub-unit: Problems with representing geographic data
- Section 3.4, Module 1: Basics of Data and Information
  Unit: Representing geography
  Sub-unit: What are geographic data?
- Section 3.5, Module 1: Basics of Data and Information
  Unit: Representing geography
  Sub-unit: Discrete objects and fields
  Sub-unit: Comparing the discrete object and field views for lakes in Minnesota
  Section 3.6, Module 1: Basics of Data and Information
  Unit: Representing geography
  Sub-unit: Rasters and vectors

- NCGIA Core Curriculum in GIScience, 2000 (www.ncgia.ucsb.edu/giscc)
    - The World in Spatial Terms (005), ed. Reg Golledge
    - Human Cognition of the Spatial World (006), Dan Montello
    - 1.4.2. Maps as Representations of the World (020), Judy Olson

- NCGIA Core Curriculum in GIS, 1990 (www.ncgia.ucsb.edu/pubs/core.html)
    - 2. Maps and map analysis
    - Models of reality
    - 22. Object/layer debate
    - 73. GIS and spatial cognition