3. Representing Geography

Geographic Information Systems and Science

SECOND EDITION

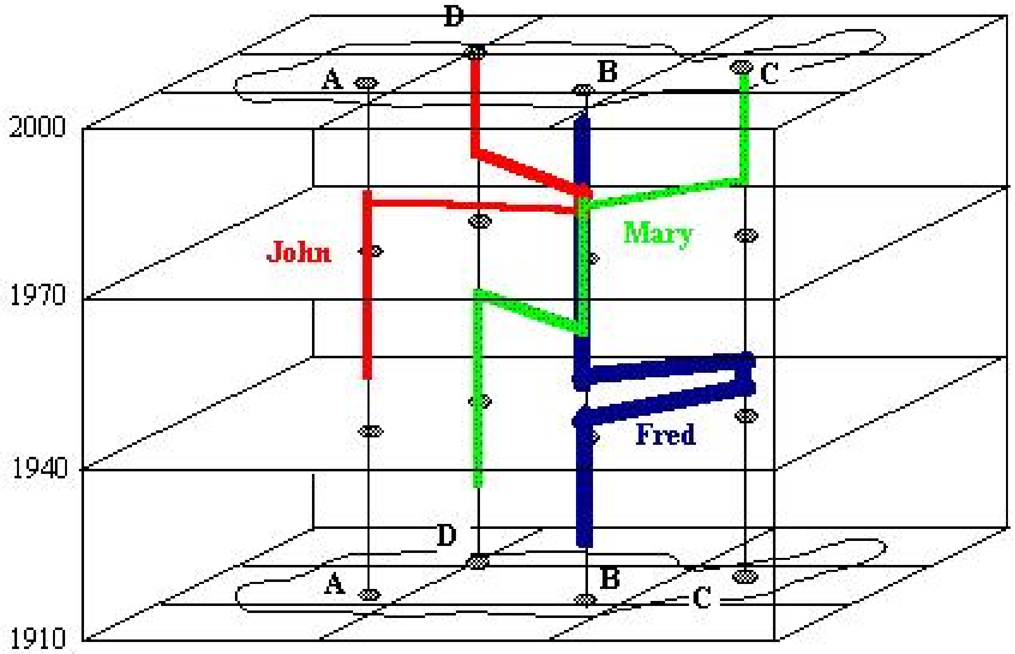
Paul A. Longley, Michael F. Goodchild, David J. Maguire, David W. Rhind

Outline

- What is representation?
- Digital representations
- The fundamental problem
- Discrete objects and continuous fields
- Rasters and vectors
- The paper map
- Generalization

Sensing the World

- Personal experience limited in time and space
  - One human lifetime
  - A small fraction of the planet's surface
- All additional knowledge comes from books, the media, movies, maps, images, and other information sources
  - From indirect or "remote" sensing



Schematic representation of the lives of three US citizens in space (two horizontal axes) and time (vertical axis)

Representations

- Are needed to convey information
- Fit information into a standard form or model
  - In the diagram the colored trajectories consist only of a few straight lines connecting points
- Almost always simplify the truth that is being represented
  - There is no information in the representation about daily journeys to work and shop, or vacation trips out of town

Representations Occur:

- In the human mind, when information is acquired through the senses and stored in memory
- In photographs, which are two-dimensional models of light received by the camera
- In written text, when information is expressed in words

Digital Representation

- Uses only two symbols, 0 and 1, to represent information
- The basis of almost all modern human communication
- Many standards allow various types of information to be expressed in digital form
  - MP3 for music
  - JPEG for images
  - ASCII for text
  - GIS relies on standards for geographic data

Why Digital?

- Economies of scale
  - One type of information technology for all types of information
- Simplicity
- Reliability
  - Systems can be designed to correct errors
- Easily copied and transmitted
  - At close to the speed of light

Accuracy of Representations

- Representations can rarely be perfect
  - Details can be irrelevant, or too expensive and voluminous to record
- It's important to know what is missing in a representation
  - Representations can leave us uncertain about the real world

The Fundamental Problem

- Geographic information links a place, and often a time, with some property of that place (and time)
  - "The temperature at 34 N, 120 W at noon local time on 12/2/99 was 18 Celsius"
- The potential number of properties is vast
  - In GIS we term them *attributes*
  - Attributes can be physical, social, economic, demographic, environmental, etc.

Types of Attributes

- Nominal, e.g. land cover class
- Ordinal, e.g. a ranking
- Interval, e.g. Celsius temperature
  - Differences make sense
- Ratio, e.g. Kelvin temperature
  - Ratios make sense
- Cyclic, e.g. wind direction

Cyclic Attributes

- Do not behave as other attributes
  - What is the average of two compass bearings, e.g. 350 and 10?
- Occur commonly in GIS
  - Wind direction
  - Slope aspect
  - Flow direction
- Special methods are needed to handle and analyze

The Fundamental Problem contd.

- The number of places and times is also vast
  - Potentially infinite
- The more closely we look at the world, the more detail it reveals
  - Potentially *ad infinitum*
  - The geographic world is infinitely complex
- Humans have found ingenious ways of dealing with this problem
  - Many methods are used in GIS to create representations or *data models*

Discrete Objects and Continuous Fields

- Two ways of conceptualizing geographic variation
  - The most fundamental distinction in geographic representation
- Discrete objects
  - The world as a table-top
  - Objects with well-defined boundaries

Discrete Objects

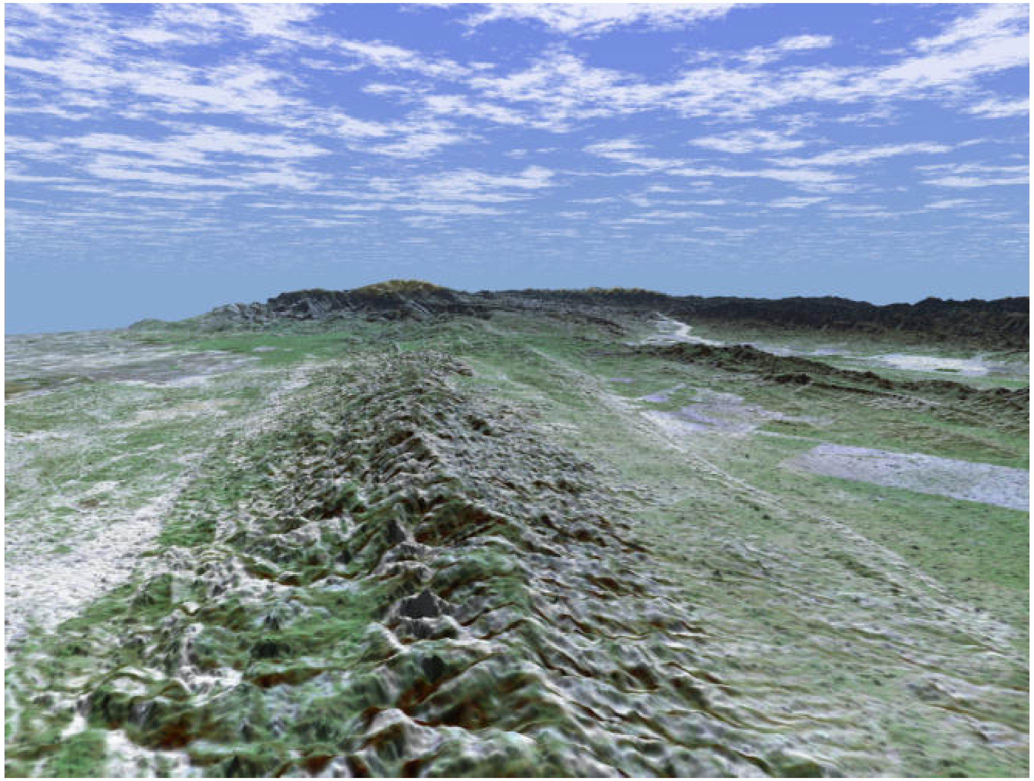
- Points, lines, and areas
- Countable
- Persistent through time, perhaps mobile
- Biological organisms
  - Animals, trees
- Human-made objects
  - Vehicles, houses, fire hydrants

Continuous Fields

- Properties that vary continuously over space
  - Value is a function of location
  - Property can be of any attribute type, including direction
- Elevation as the archetype
  - A single value at every point on the Earth's surface
  - The source of metaphor and language
    - Any field can have slope, gradient, peaks, pits

Examples of Fields

- Soil properties, e.g. pH, soil moisture
- Population density
  - But at fine enough scale the concept breaks down
- Identity of land owner
  - A single value of a nominal property at any point
- Name of county or state or nation
- Atmospheric temperature, pressure



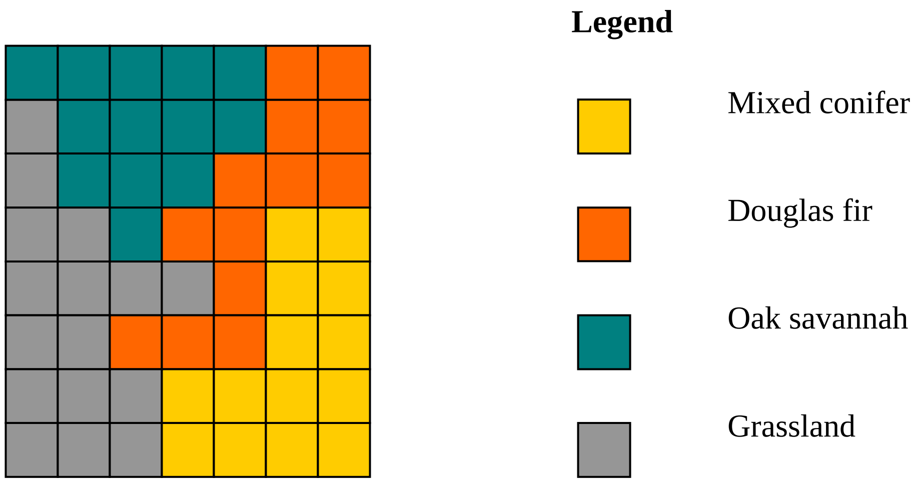
Phenomena conceptualized as fields. The illustration shows elevation data from the Shuttle Radar Topography Mission draped with an image from the Landsat satellite, looking SE along the San Andreas Fault in Southern California, plus a simulated sky

Difficult Cases

- Lakes and other natural phenomena
  - Often conceived as objects, but difficult to define or count precisely
- Weather forecasting
  - Forecasts originate in models of fields, but are presented in terms of discrete objects
    - Highs, lows, fronts

Rasters and Vectors

- How to represent phenomena conceived as fields or discrete objects?
- Raster
  - Divide the world into square cells
  - Register the corners to the Earth
  - Represent discrete objects as collections of one or more cells
  - Represent fields by assigning attribute values to cells
  - More commonly used to represent fields than discrete objects



Raster representation. Each color represents a different value of a nominal-scale field denoting land cover class.

Characteristics of Rasters

- Pixel size
  - The size of the cell or picture element, defining the level of spatial detail
  - All variation within pixels is lost
- Assignment scheme
  - The value of a cell may be an average over the cell, or a total within the cell, or the commonest value in the cell
  - It may also be the value found at the cell's central point

Vector Data

- Used to represent points, lines, and areas
- All are represented using coordinates
  - One per point
  - Areas as polygons
    - Straight lines between points, connecting back to the start
    - Point locations recorded as coordinates
  - Lines as *polylines*
    - Straight lines between points

Raster vs Vector

- Volume of data
  - Raster becomes more voluminous as cell size decreases
- Source of data
  - Remote sensing, elevation data come in raster form
  - Vector favored for administrative data
- Software
  - Some GIS better suited to raster, some to vector

The Paper Map

- A long and rich history
- Has a scale or *representative fraction*
  - The ratio of distance on the map to distance on the ground
- Is a major source of data for GIS
  - Obtained by digitizing or scanning the map and registering it to the Earth's surface
- Digital representations are much more powerful than their paper equivalents

Generalization

- Reducing the level of detail in geographic data
  - By simplifying, weeding, abstracting
  - To reduce the volume of data without adversely affecting its use
  - To "see the wood for the trees"

The Specification

- A set of rules for the construction of a map or database, defining its level of detail
  - A map that is accurate with respect to its specification may be regarded as correct, even though it ignores certain types of detail

Weeding

- The process of removing points in a polygon or polyline while preserving important aspects of shape
  - The Douglas-Poiker algorithm
    - A rigorous process that can be applied to any polygon or polyline
    - Requires the specification of a *tolerance* parameter that defines the allowed deviations between the original feature and its generalized version