Analytical and Computer Cartography Winter 2017

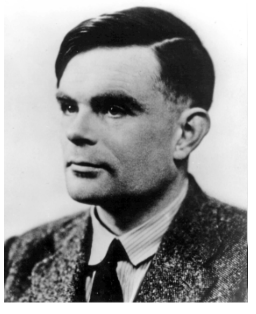
Lecture 8: Algorithms, mosaicing, and conflation

Cartographic Transformations

- Attribute Data (e.g. classification)
- Locational properties (e.g. projection)
- Graphics (e.g. symbolization)
- Information content of maps (e.g. data structure conversion)

Basic Transformation Questions

- Is a transformation quantifiable?
- Can the transformation process be automated? (Alan Turing: Turing Machine and the halting problem, Alonzo Church: Lambda calculus)
- Is a transformation invertible?
- Is a transformation stable?



Types of Transformations

- Map scale
- Dimension
- Symbolic content
- Data structures

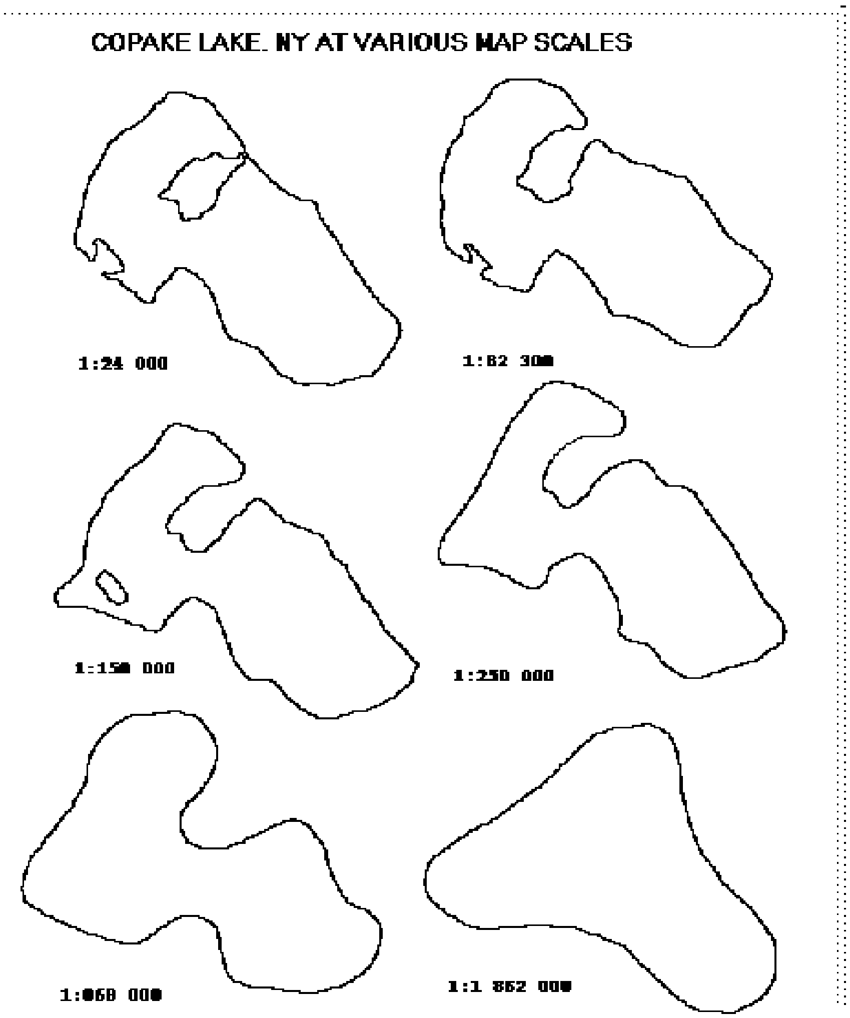


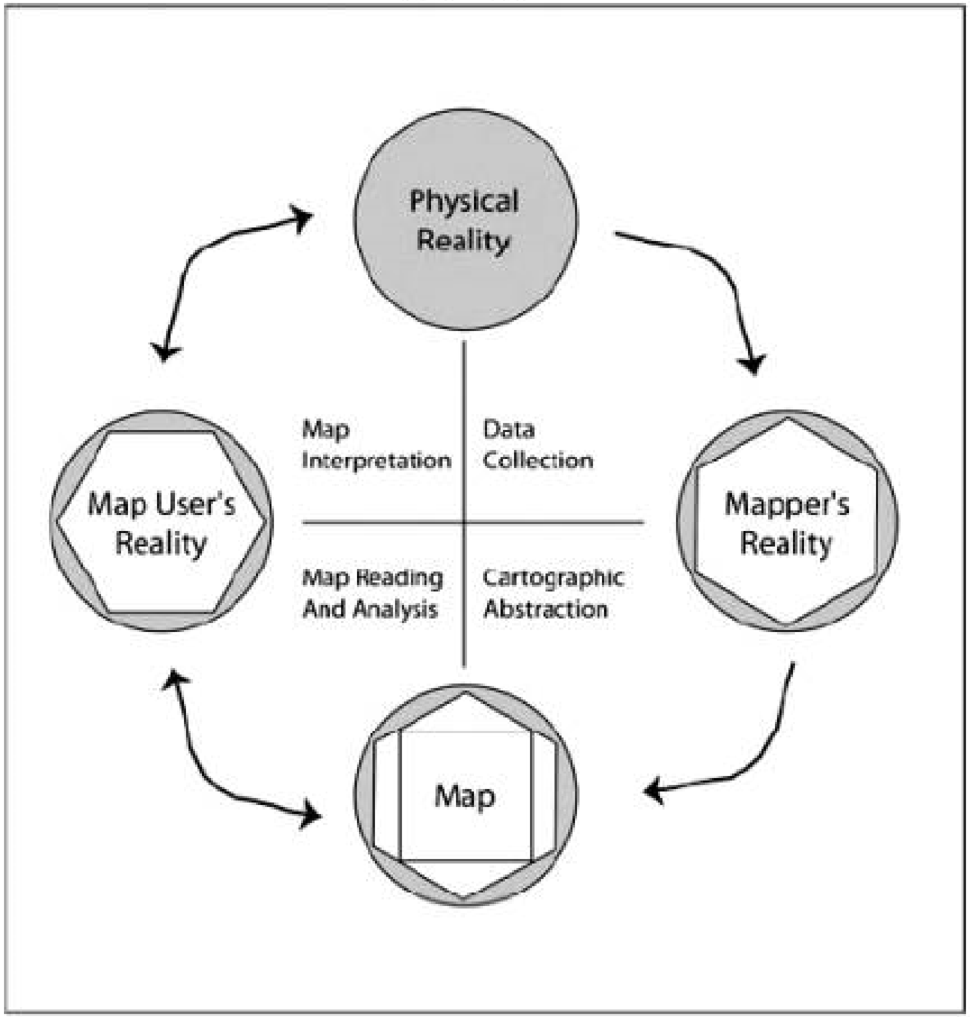
Figure 10.5 Copake Lake (in New York State) at different map scales.§

Why do we need to transform?

- We may wish to compare maps collected at different scales
- We may wish to convert the geometry of the map base
- We may wish or need to change the map data structure
- Almost ALL mapping stages involve transformations!

The mapping process

The Cartographic Process

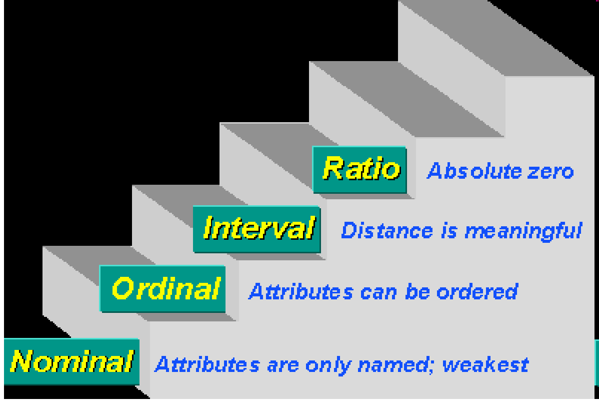


After Muehrcke and Muehrcke, 1992

State Changes and Transformations

- Cartographers are interested in the full set of state transformations
- Each map may have an optimal path through the set
- Design cartography primarily concentrates on the last, or symbolization transformation, and now uses human subjects testing and cognitive engineering
- Four types of transformations shape the mapping process:
  - Geocoding (transforming entities to objects: levels, dimension, data structure)
  - Map Scale
  - Locational Attributes or Map Base
  - Symbolization

Levels of Measurement



- Robinson's Classification was based on dimension and level of measurement
- Level of measurement idea is from Stevens (1946)
- Nominal data assume only existence and type. An example is a text label on a map
- Ordinal data assume only ranking. Relations are like "greater than"
- Interval data have an arbitrary numerical value, with relative value Example: Elevation.
- Ratio data have an absolute zero and scale

Robinson & Sale

Content scaling level	Defining relations	FORM OF CARTOGRAPHIC SYMBOL		
		POINT	LINE	AREA
Nominal	Equivalence	* + + * * + Wholesale and retail establishments	D E A C B Highway connectivity	A B C Land ownership
Ordinal	Equivalence Greater than	Medium Small Large Population centers	Roads by degree of improvement	Low Med High Crop yield
Interval	Equivalence Greater than Ratio of intervals	+ 147 + 210 + 132 + 122 Spot elevations	Graticule	1824 1845 1862 Date of settlement
Ratio	Equivalence Greater than Ratio of intervals Ratio of scale values	Area proportional to population	20 30 20 10 Population density isopleths	Value proportional to population density

Figure 10.1 Classification by scaling and dimension. (After Robinson and Sale, *Elements of Cartography, 3d ed.* ,© 1969, by John Wiley & Sons, Inc. Used with permission.)§

Unwin’s Classification

DATA TYPES

	Point	Line	Area	Volume
Nominal	City	Road	Name of unit	Precipitation or soil type
Ordinal	Large city	Major road	Rich county	Heavy precipitation Good soil
Interval Ratio	Total population	Traffic flow	Per capita income	Precip.in mm Cation exchange

MAP TYPES

	Point	Line	Area	Volume
Nominal	Dot map	Network map	Colored area map	Freely colored map
Ordinal	Symbol map	Ordered network map	Ordered colored map	Ordered chromatic map
Interval Ratio	Graduated symbol map	Flow map	Choropleth map	Contour map

Figure 10.3 Map data and map types. (After Unwin, 1981.)§

Transformations in Choropleth Mapping

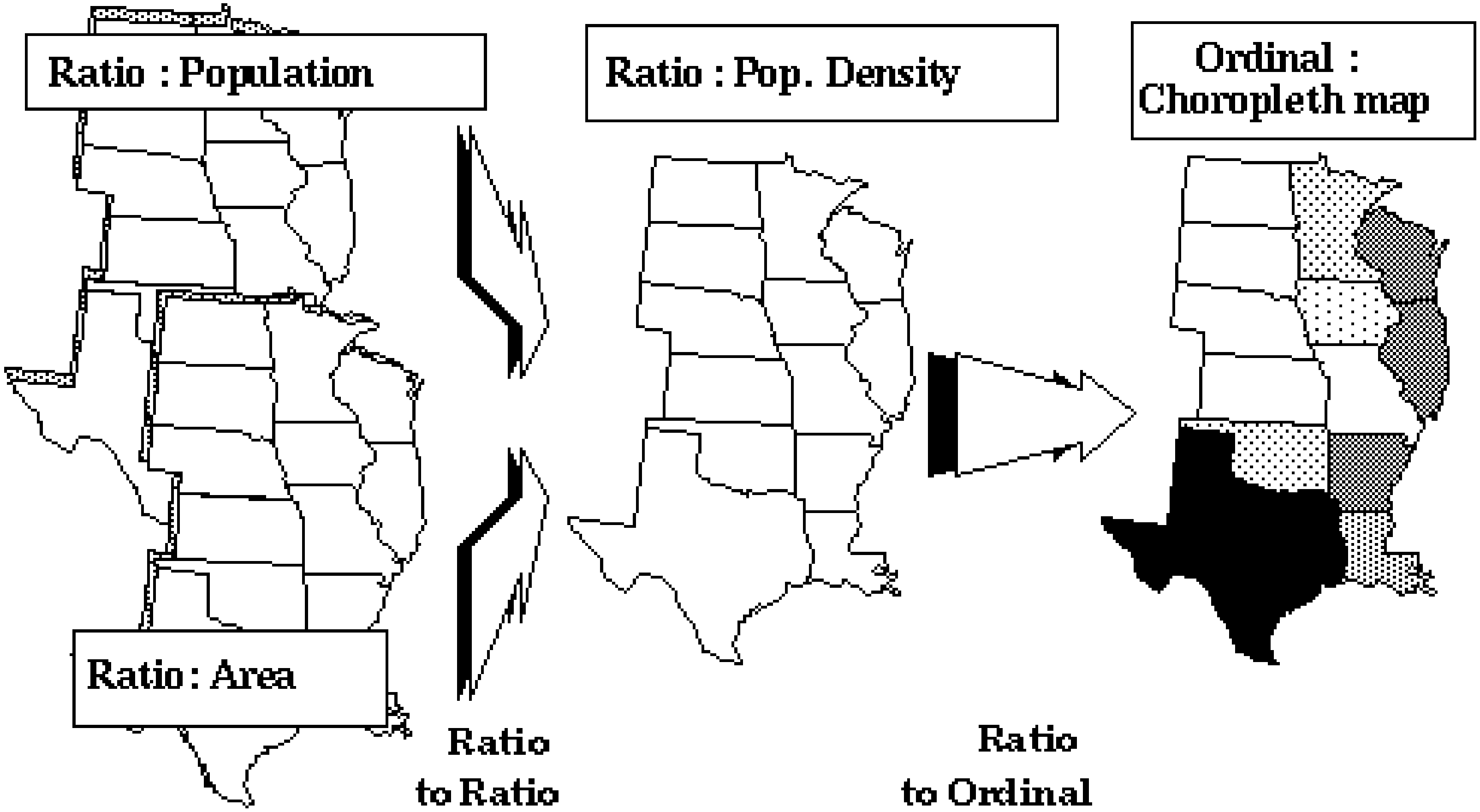
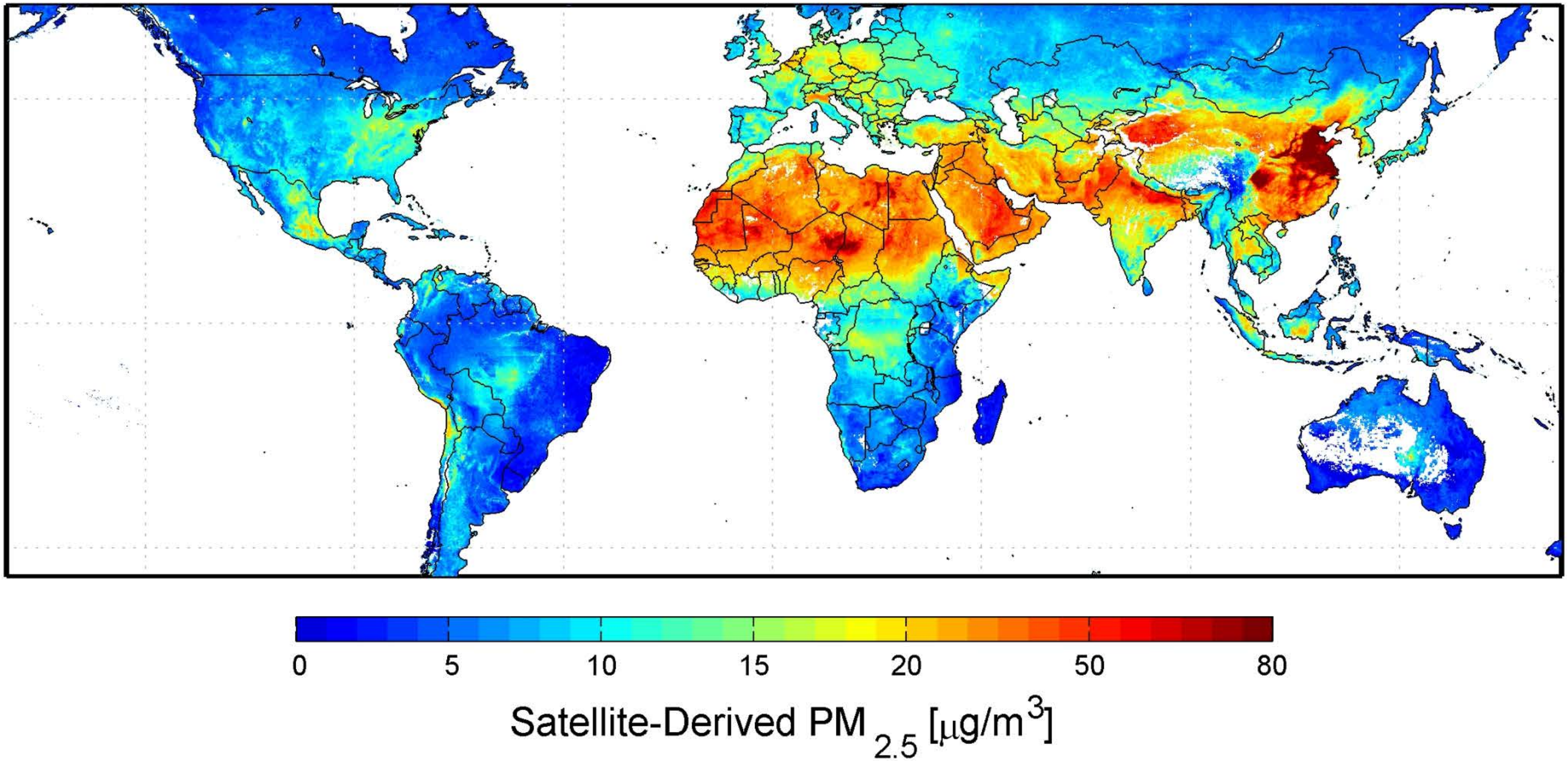


Figure 10.2 Level of measurement transformations for choropleth mapping.$

Air pollution Particulates

Point data (2 instruments and a computer model) -> interpolated to grid > ranked color sequence

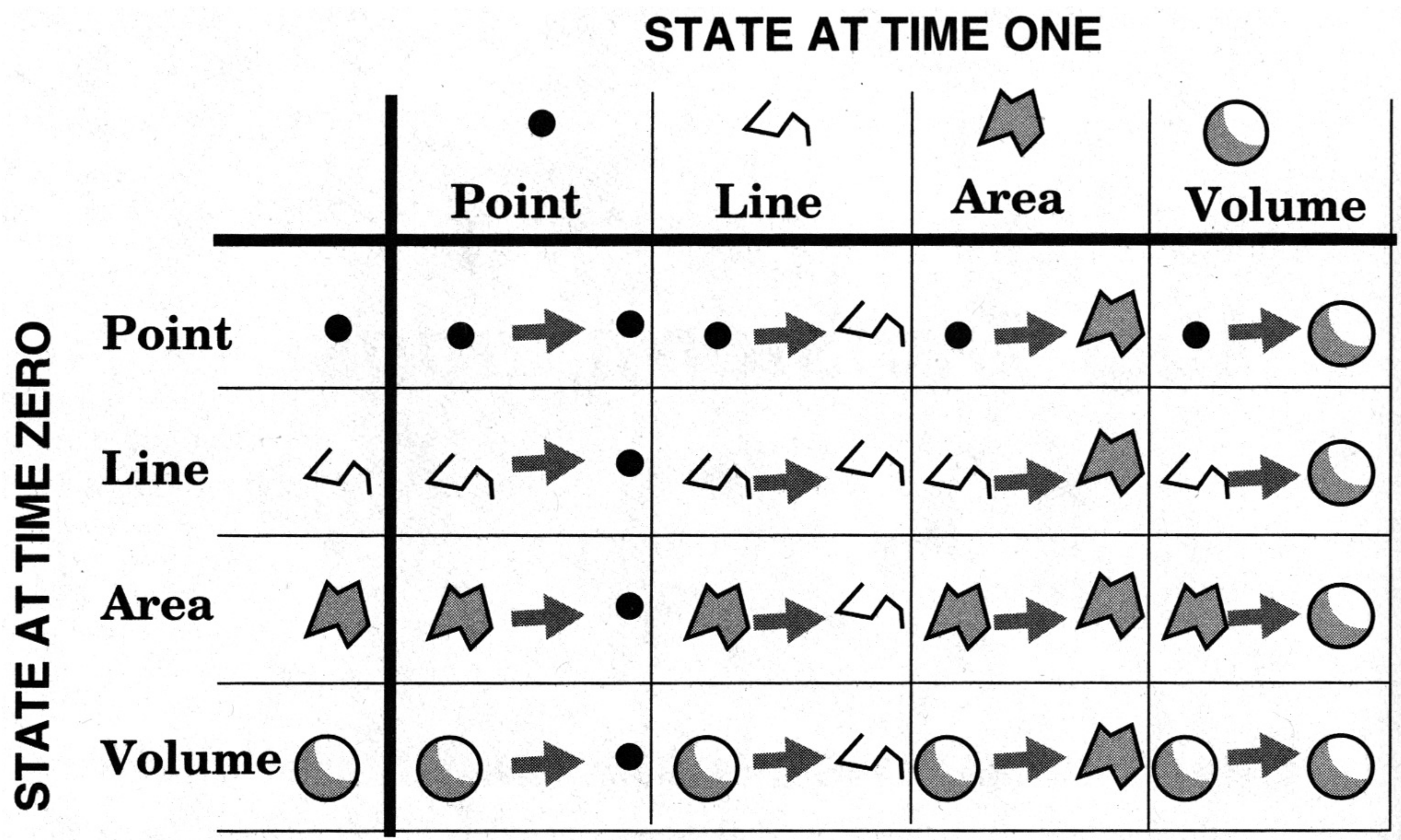
Source: NASA, 2013 and Aaron van Donkelaar and Randall Martin at Dalhousie University



Transformations of Object Dimension

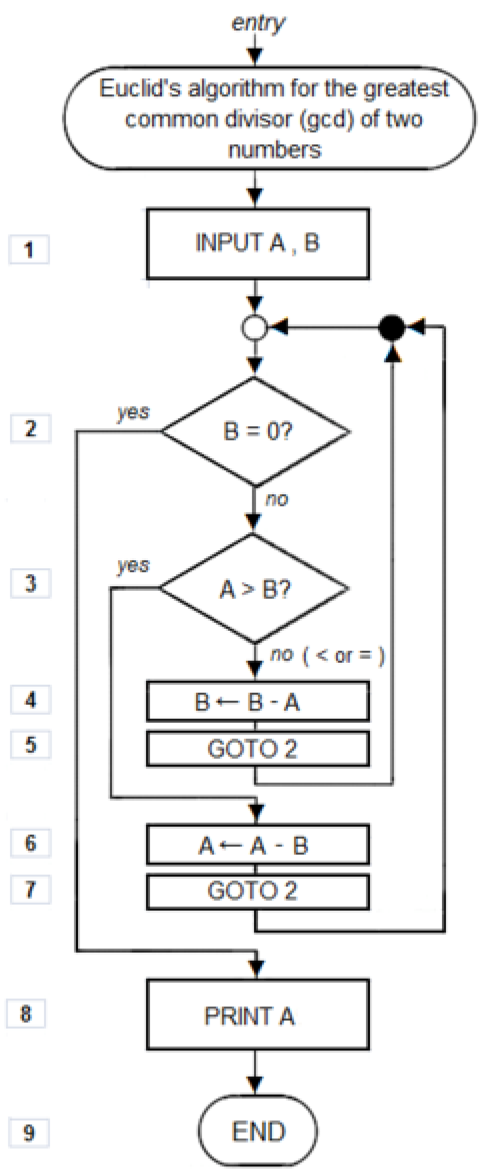
- The four dimension levels of data can be represented at only one level in each state, though a map can contain multiple layer symbols
- Transformations can move data between states
- Full set of state zero to state one transformations is then 16 possible transformations
- Lab exercises fall into several of these
- Dimensional transformations are only one type
- When dimension collapses to "none" result is a measurement

Dimensional Transformations



Algorithm

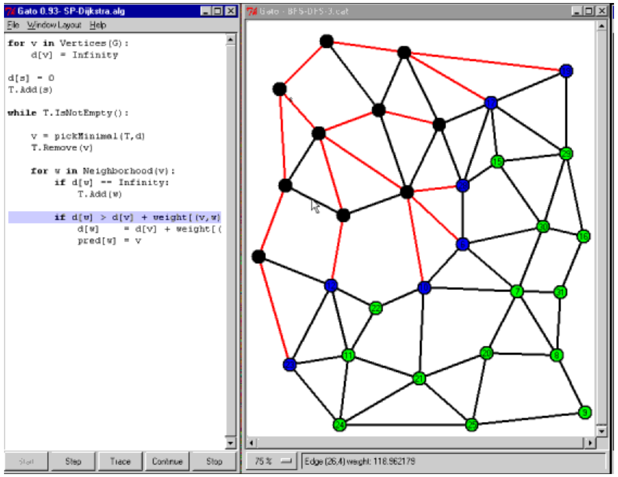
- In mathematics and computer science, an **algorithm** is an effective method expressed as a finite list of well-defined instructions for calculating a function
- Algorithms are used for calculation, data processing, and automated reasoning
- Usually has inputs, result and loops
- Importance of termination
- Divide and conquer



Transformations and Algorithms

- In mathematics, transformations are expressed as equations
- Solutions, inversion and so forth are by algebra, calculus etc.
- In computer science, a set of transformations defining a process is called an algorithm
- Any process that can be reduced to a set of steps can be automated by an algorithm (Church/Turing hypothesis)
- *data structures* + *transformational algorithms* = *maps*

Types of Algorithms in mapping

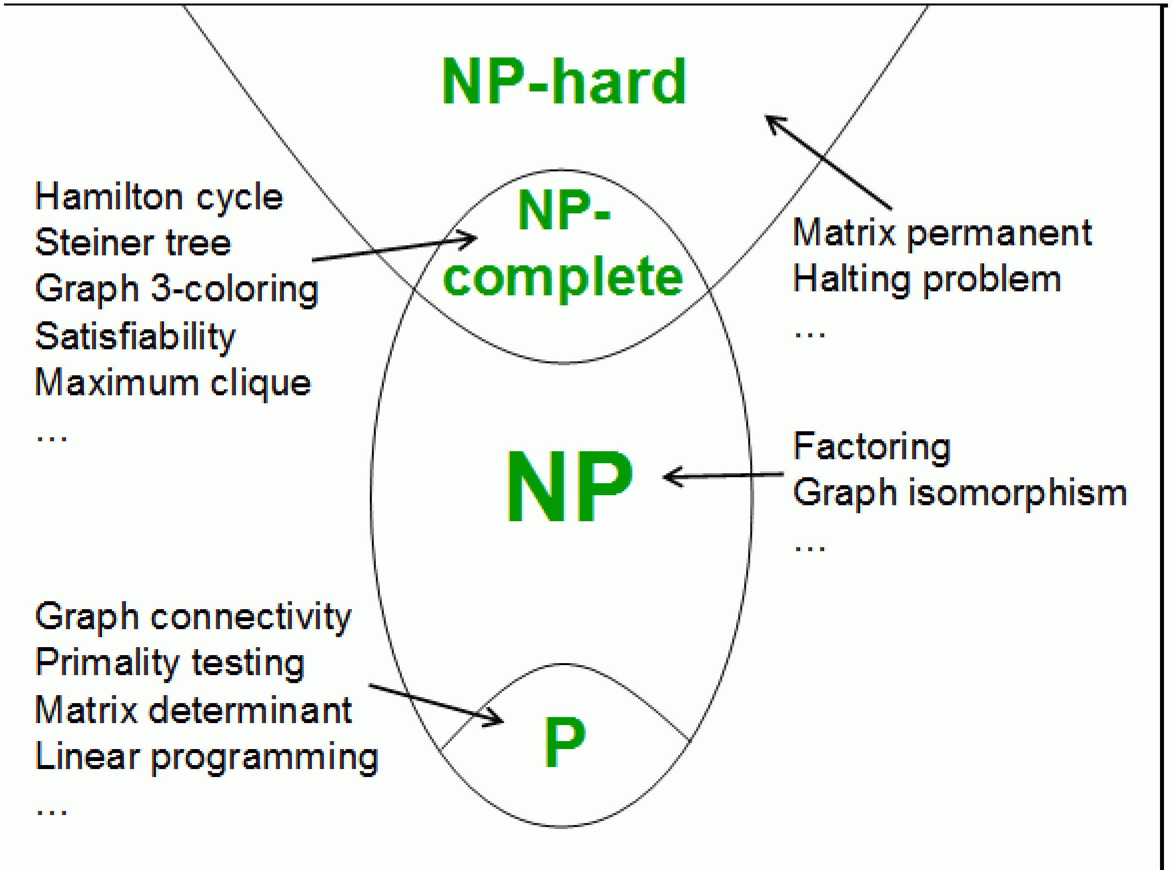


- mathematical
- sorting
- searching
- string processing
- geometrical algorithms (computational geometry)
- graph algorithms
- complex , e.g. decomposition
- In CS, an algorithm is implemented as a function *output* = *f (inputs)*
- **Inputs can be { data, parameters, objects }**

Graphic algorithms

- Algorithm: method for solving problems, suited for computer implementation (Sedgewick, 1984)
- "Most algorithms of interest involve complicated methods of organizing the data involved in the computation. Objects created this way are called data structures.“
- Recursion
- Task decomposition
- Divide and conquer
- Special case vs. Generic solution e.g. vertical lines
- Partitioning: Sequential vs. Parallel (Data and Process)
- Big-O notation and complexity theory
- Solution/Halting problem: Tractability

Complexity Theory



P Polynomial time

N Non-deterministic

T Trivial

Problems needing complex algorithms

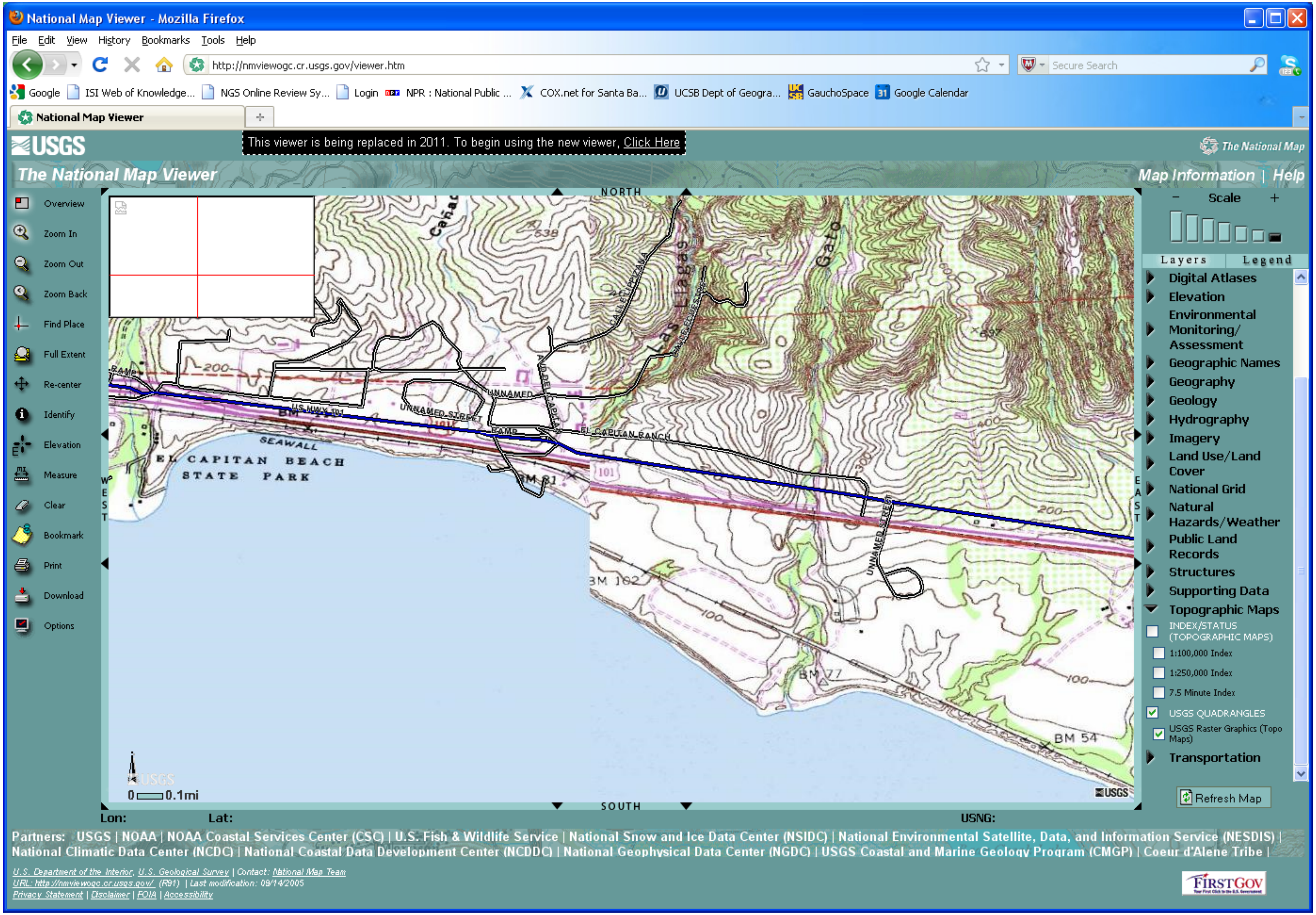
- Mosaicing
  - Forward: Given a large map, divide into regular or uneven tiles in an optimal way
  - Inverse: Given a tiled map, assemble it back into a single network
    - Detect and eliminate errors
    - Adjust geometry
    - Join divided features
- Conflation
  - Given two maps merge their features
    - Geometric error
    - Attribute error
    - Errors of omission and comission

MOSAICING

- **Maps collected separately, or map partitioned for storage/search**
- **Given *n* tiles, with common boundaries, create a new set of common objects that match the geometry in both tiles**
- **Seam stitching**
- **Shift**
- **Mosaicing as conflation**

On the 120th meridian

National Map Viewer: 120th meridian



Dissolve

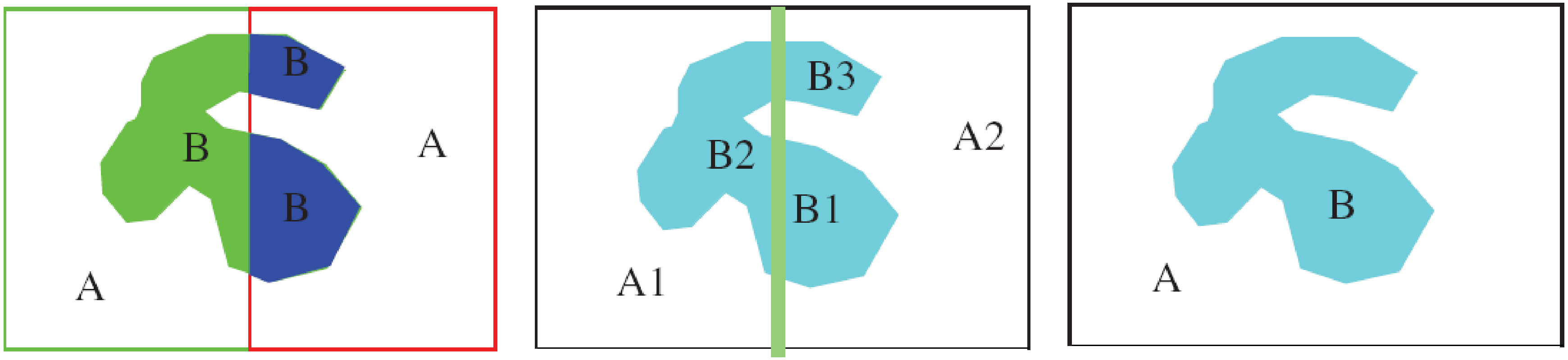
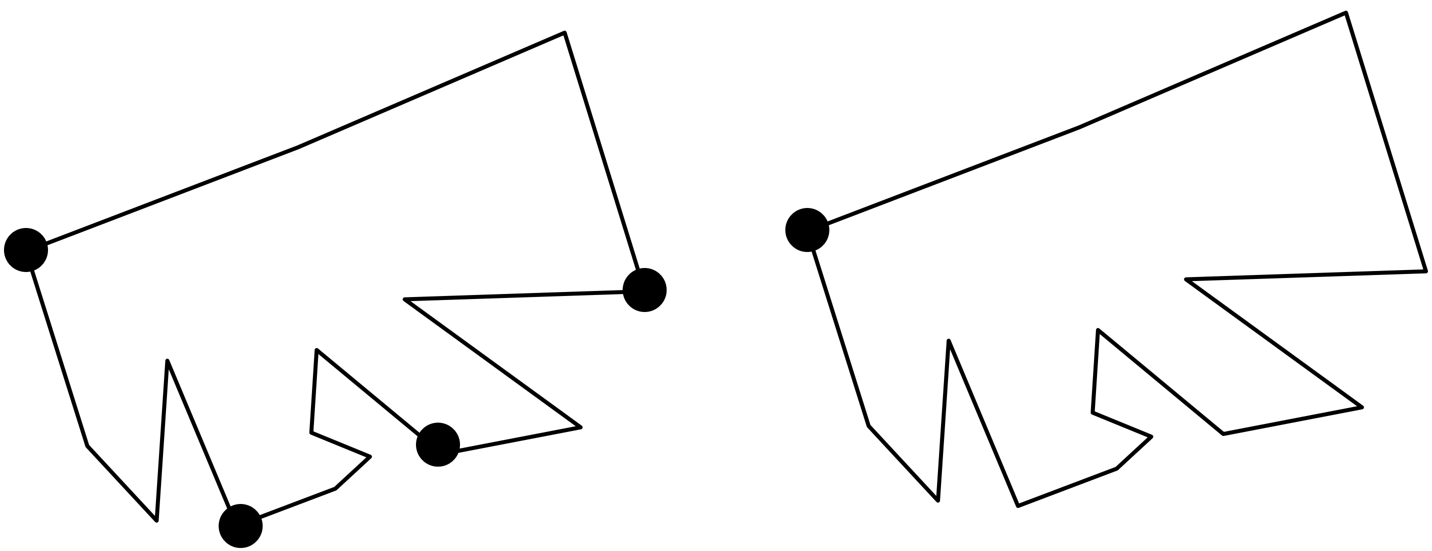
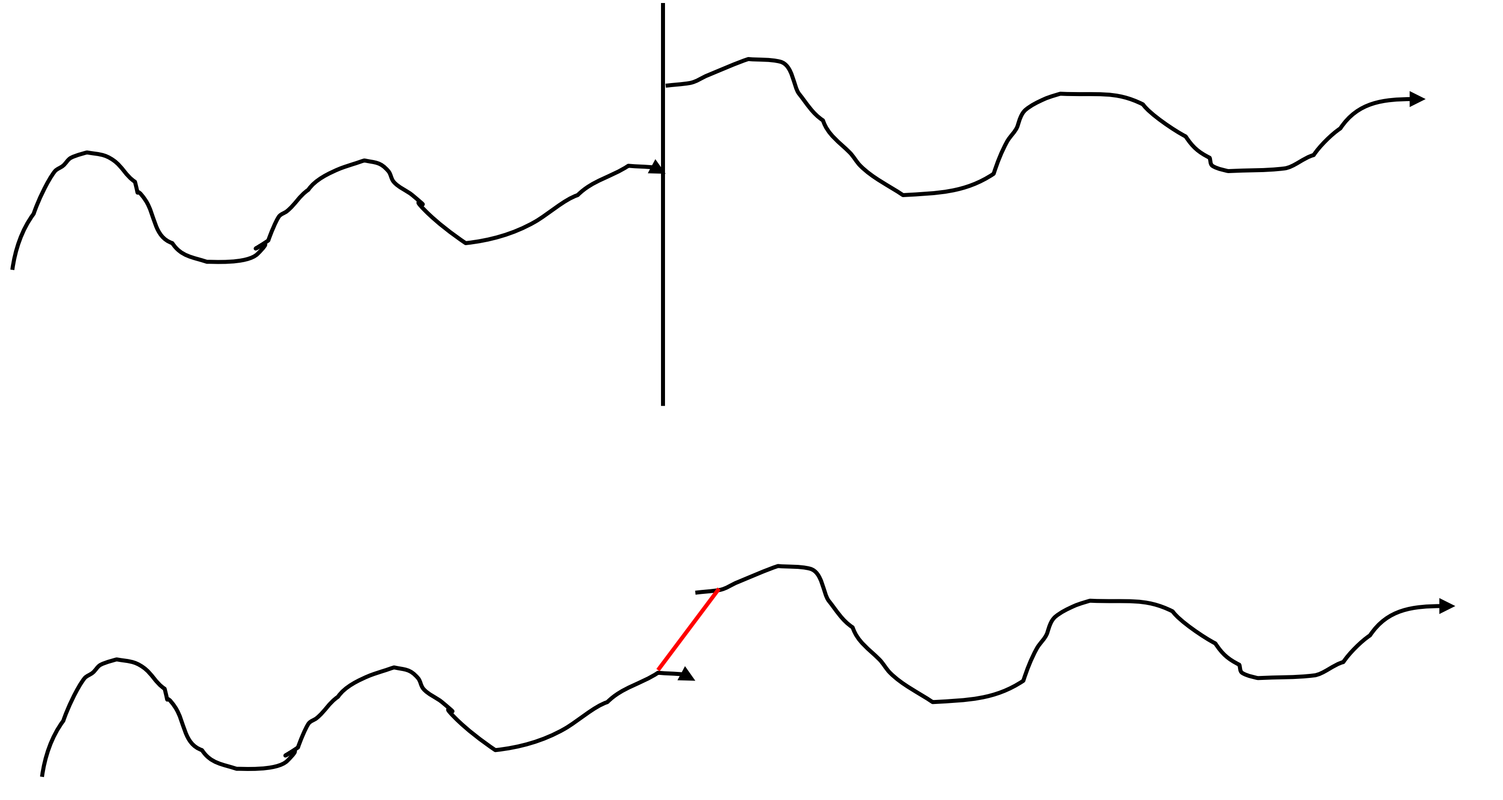


Figure 9.3 Steps in the *dissolve* operation. Left: Two maps show one feature, split across a map edge. Center: Attribute and graphic database have three records for type "B." Right: After dissolve, edge lines are removed and the three type "B" records are amalgamated into a single feature, dissolving the border.



Averaging: Not a solution!



Rubber sheeting

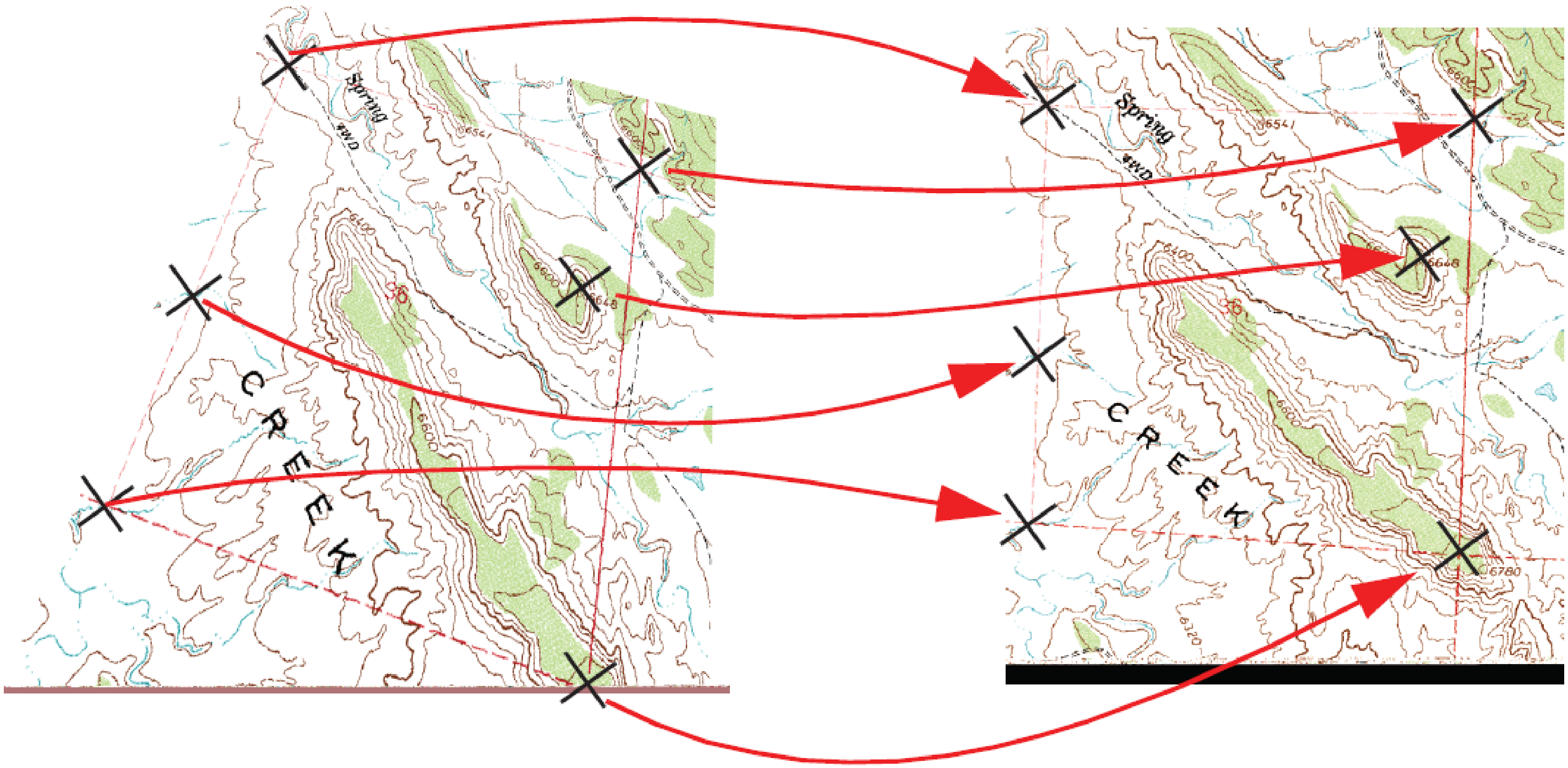
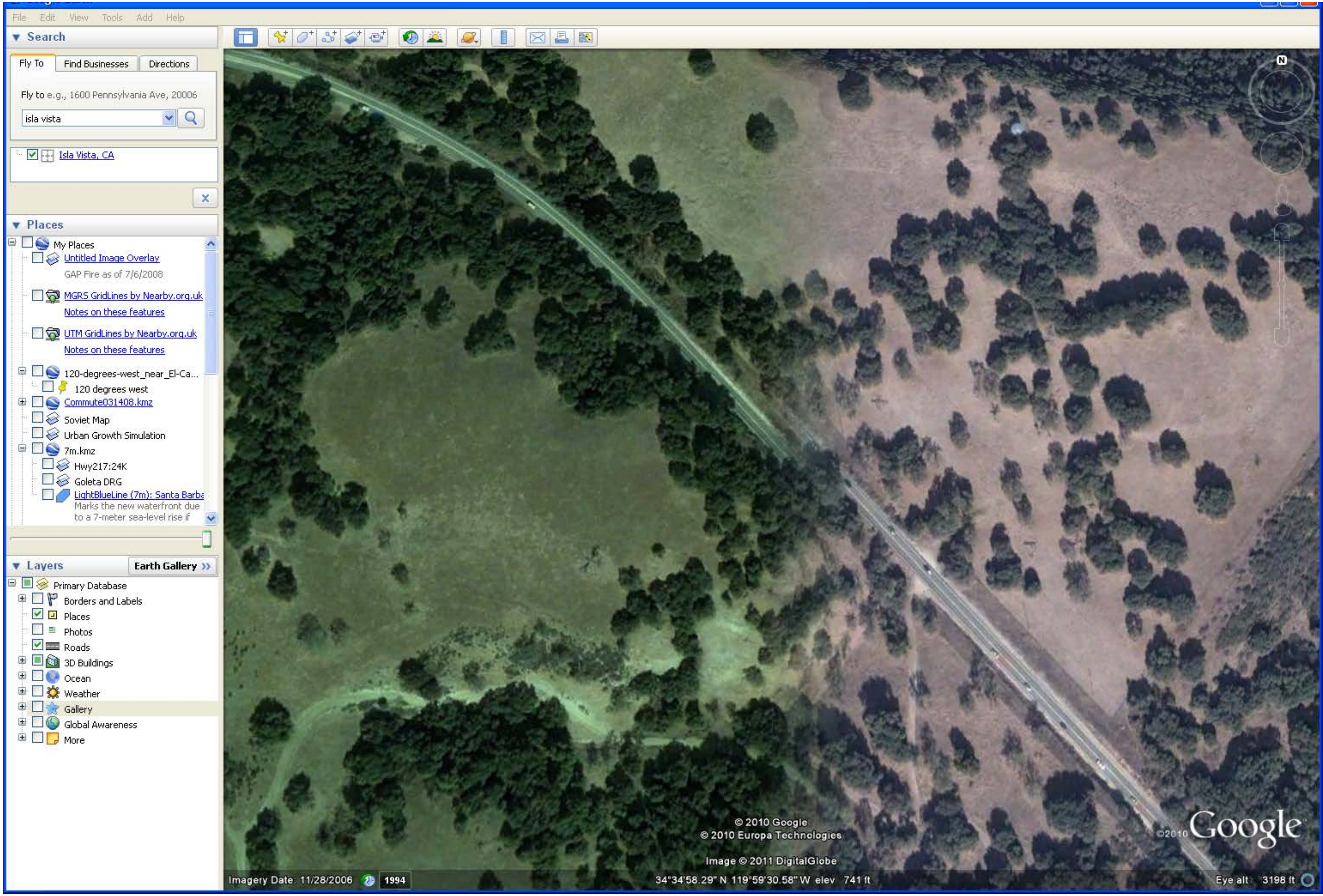


Figure 9.8 The *rubber sheeting method*. A map with unknown geometry (say an air photo taken or scanned map) can be distorted so that its geometry matches that of another map. Pairs of points must be available both on the image and on the map showing the same place or feature location, called control points. Within the GIS, rubber sheeting warps the geometry statistically into that of the map, so that the two geometries match.

Seams

File Edit View Tools Add Help
Search
Fly To
Find Businesses
Directions
Fly to e.g., 1600 Pennsylvania Ave, 20006
isla vista
Isla Vista, CA
Places
My Places
Untitled Image Overlay
GAP Fire as of 7/6/2008
MGRS GridLines by Nearby.org.uk
Notes on these features
UTM GridLines by Nearby.org.uk
Notes on these features
120-degrees-west_near_El-Ca...
120 degrees west
Commute031408.kmz
Soviet Map
Urban Growth Simulation
7m.kmz
Hwy217:24K
Goleta DRG
LightBlueLine (7m): Santa Barba
Marks the new waterfront due
to a 7-meter sea-level rise if
Layers
Earth Gallery
Primary Database
Borders and Labels
Places
Photos
Roads
3D Buildings
Ocean
Weather
Gallery
Global Awareness
More
Calle Real
Gato Canyon Rd
1
101
© 2010 Google
Image © 2011 DigitalGlobe
Google
Imagery Date: 11/28/2006
1994
34°27'31.29" N 119°59'23.68" W elev 185 ft
Eye alt 3266 ft

Near Lake Cachuma



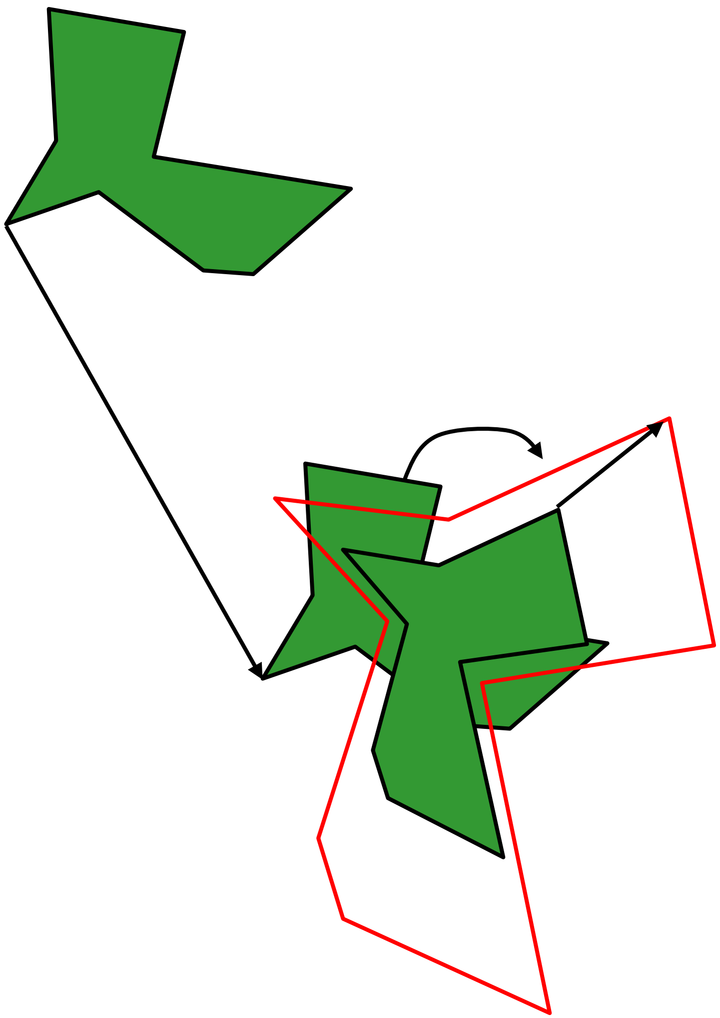
CONFLATION

- **Given two input objects with different (contrary) geometry, generate a single output that conflates the objects**
- **Six-parameter affine**
  (TRS) **Local affine**

$$A_T(f_{x,y}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$

- **Issues:**
  - **Point selection**
  - **Random vs. systematic error**
  - **"truth"**

Affine transformations



$$[p_x' \; p_y' \; 1] = [p_x \; p_y \; 1] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c_x & c_y & 1 \end{bmatrix}$$

Translation

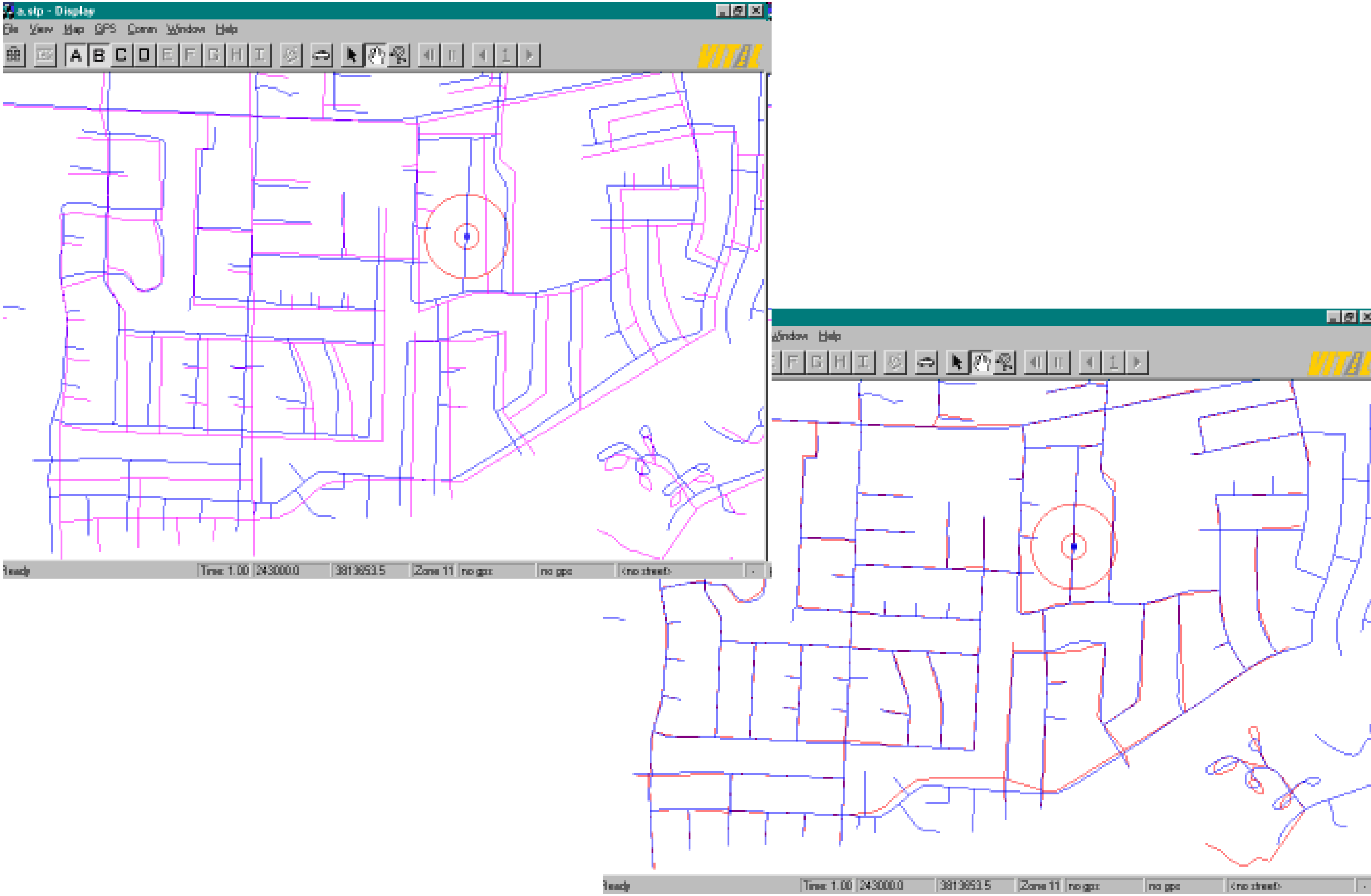
$$[p_x' \; p_y' \; 1] = [p_x \; p_y \; 1] \begin{bmatrix} cos(\theta) & -sin(\theta) & 0 \\ sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Rotation

$$[p_x' \; p_y' \; 1] = [p_x \; p_y \; 1] \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

SCALE

Conflation: Geometry and Attributes



Geometric mismatch

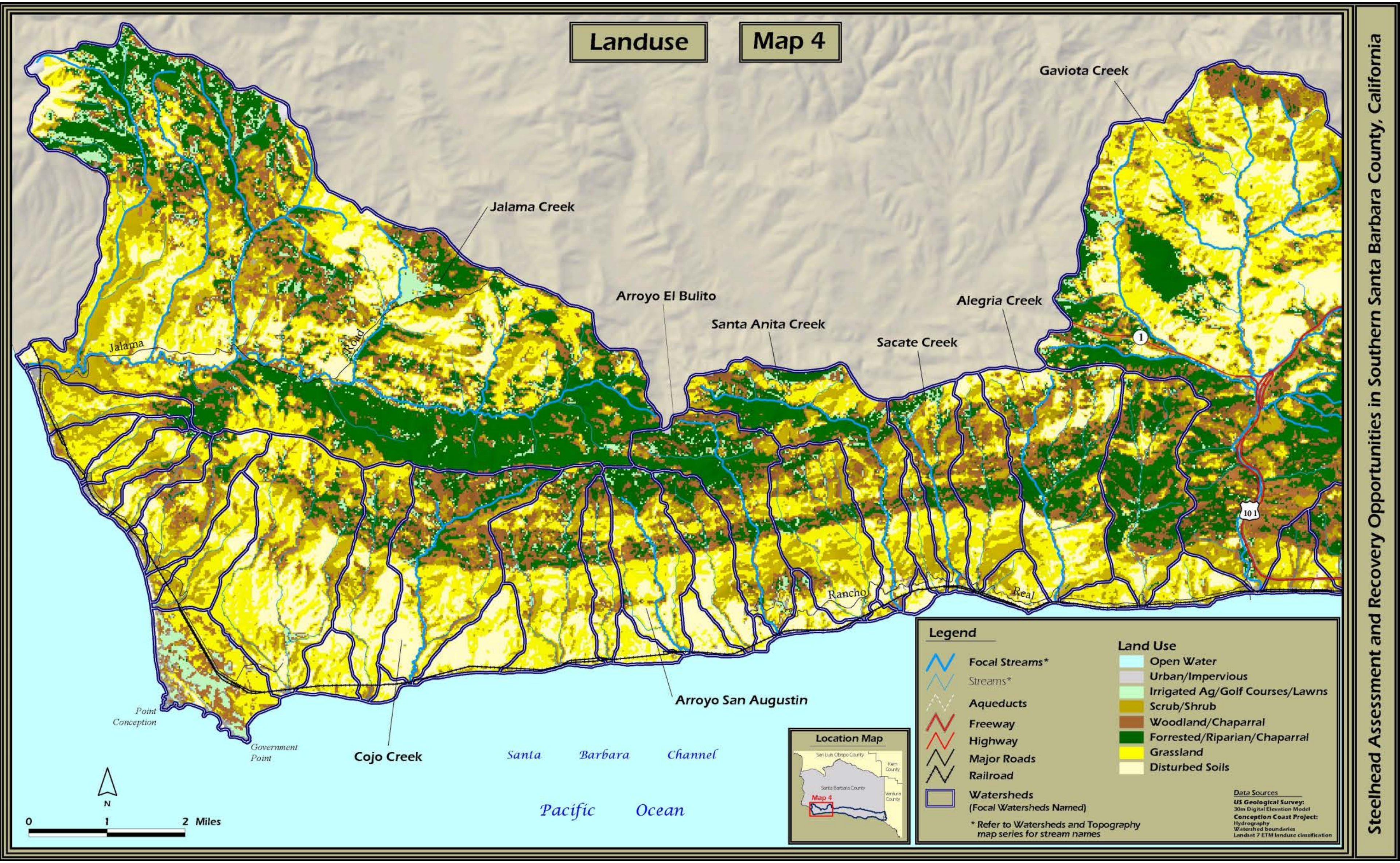
ERROR

- Generic Function Output = f (Input)
- O = f(I)
- $I' = F^{-1}(O)$
- I I' = Identity?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- If not, errors exist.

Categorical map accuracy



Class Accuracy: Ground Truth

Contingency Table: A pixel by pixel comparison of ground reference class to satellite-based map class.

	Cover Type	*Satellite Map Class* (pixel counts) Conifer	Hardwood	Grass	Barren	Total	Producers Accuracy, Percent Correct (1-omission error)
ground reference	Conifer	911	20	1	0	932	97.7
	Hardwood	40	343	72	2	457	75.1
	Grass	0	62	176	14	252	69.8
	Barren	0	0	19	27	46	58.7
	Total	951	425	268	43	1687	
	Users Accuracy, Percent Correct	95.8	80.7	65.7	62.8		86.4

(1-commission error)

Kappa: An accuracy statistic that permits two or more contingency matrices to be compared.

The statistic adjusts overall accuracy to account for chance agreement.

Use kappa to statistically test for agreement between two contingency matrices.

Summary

- Transformations can impact dimension, data level, scale, symbols
- Transformations are chained, and include map reading and interpretation
- Algorithms can make a transformation computable
- *data structures* + *transformational algorithms* = *maps*
- Algorithm computability covered by computational complexity theory
- Examples of hard problems include tiling and conflation
- Methods exist for quantifying and analyzing map error