



Scale and Generalization

This is lecture 12

Different interpretations of scale

- Scales of measurement (nominal, ordinal, interval, and ratio)
- Cartographic scale
- Analysis scale
- Phenomenon scale
- Concepts like detail, granularity, resolution, and generalization

Cognitive perspectives on scale

- Types of scale that humans work with include:
 - a) "minuscule" -- too small to apprehend without technological aid
 - (b) "figural" -- large enough to apprehend without technological aid but smaller than the body
 - (c) "vista" -- significantly larger than the body but visually perceptible from a single vantage point
 - (d) "environmental" -- larger than the body and otherwise obscured so that considerable body locomotion is required for their apprehension, and
 - (e) "gigantic" -- too large to apprehend without technological aid.

Traditional scale/digital scale

- Traditionally scale has been expressed using representative fractions.
- Maps were designated large or small scale depending on their RF.
- But in the digital world, we need more accessible access to scale, from a cognitive perspective.

Children

- Michael Goodchild uses the example of children looking for maps in the web.
- Suppose a child of 10 wants a map of his/her neighborhood.
- What kind of mechanisms can be set up to allow that child to specify the requisite levels of information.

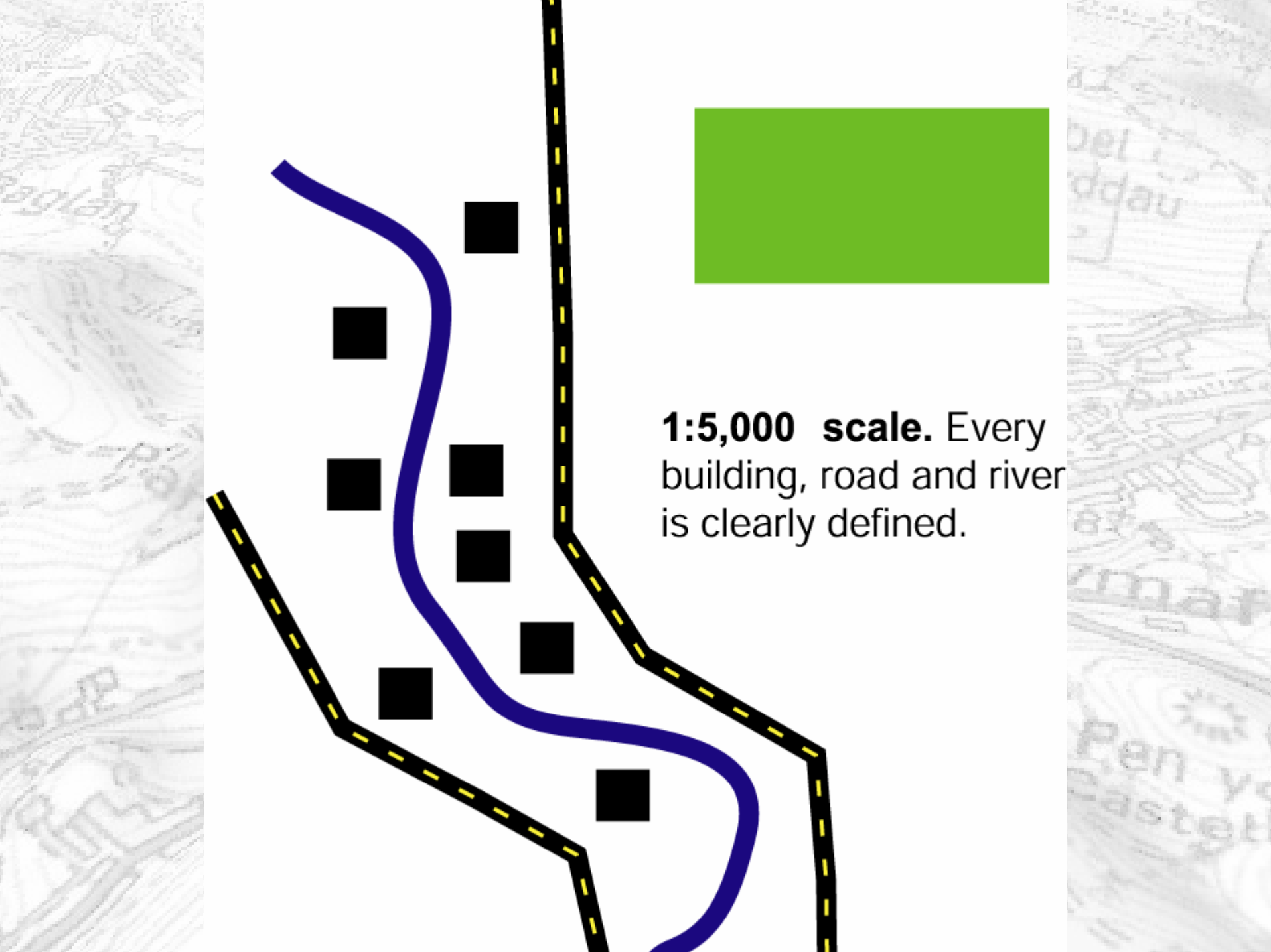
- What happens to the RF when you move into the digital world?
- The notion of RF is unsatisfactory because it is impossible to develop a RF that is constant everywhere.
- A famous example is the airbrushed image of Earth produced by Van Sant at a 4 km resolution. Early editions had on the bottom: scale 1 inch = 720 miles.
- On later editions it had a note underneath it, saying "on the equator".
- In the digital world, the RF is undefined, because there is no single external representation but a digital database.
- Is there scale inside a computer?

Scale and generalization

- Even if there is no scale in a digital world, we use scale as a concept to generalize graphical and model data.
- What is generalization?

Brief review of generalization

- Simple answer: elimination of map detail as scale decreases.
- As scale decreases, geographic objects must be eliminated.



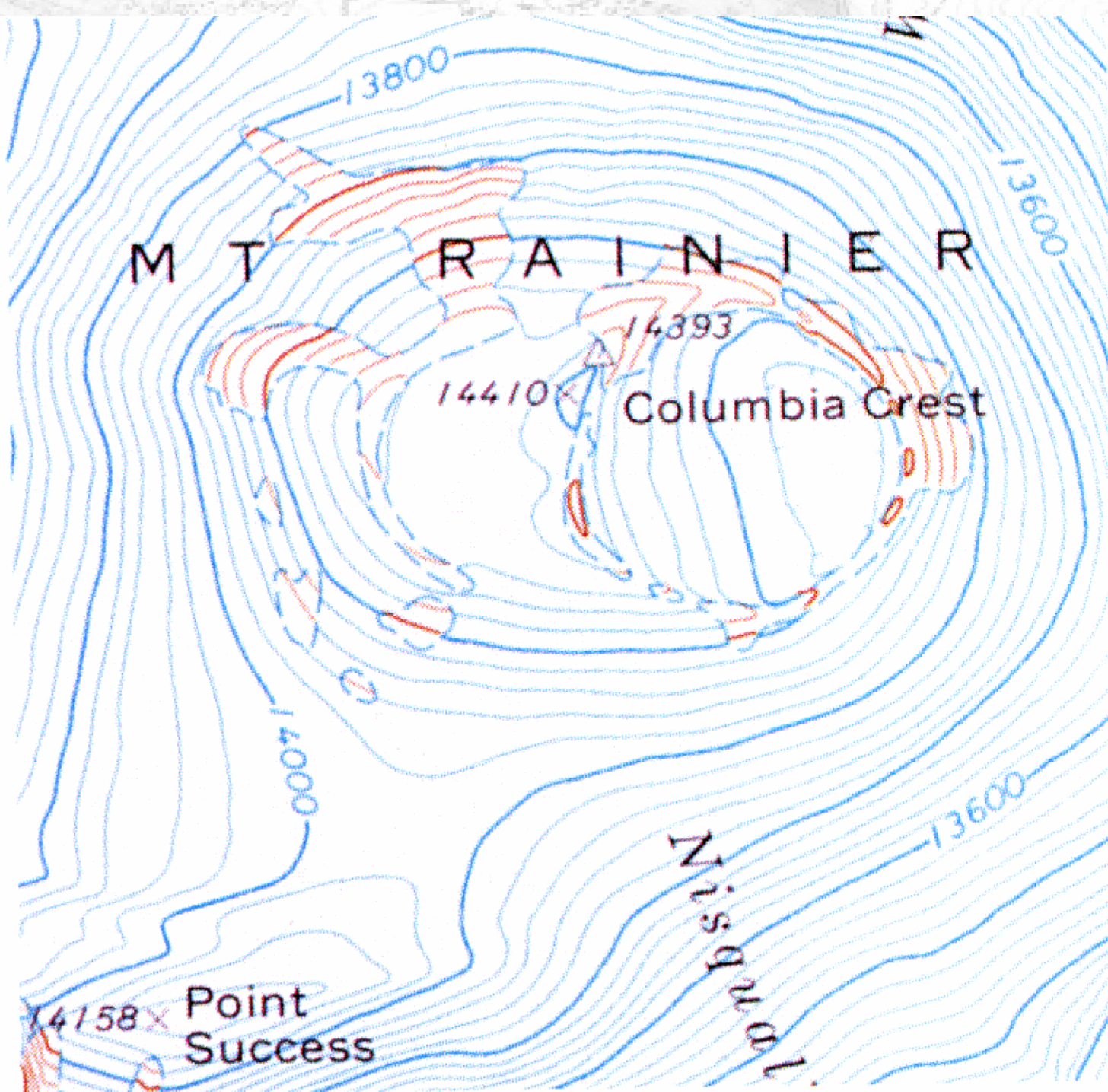
1:5,000 scale. Every building, road and river is clearly defined.



1:500,000 scale. Cities are indicated by coloured areas transected by roads, freeways and rivers.



1:5,000,000 scale. Cities are lucky to be included as coloured dots.

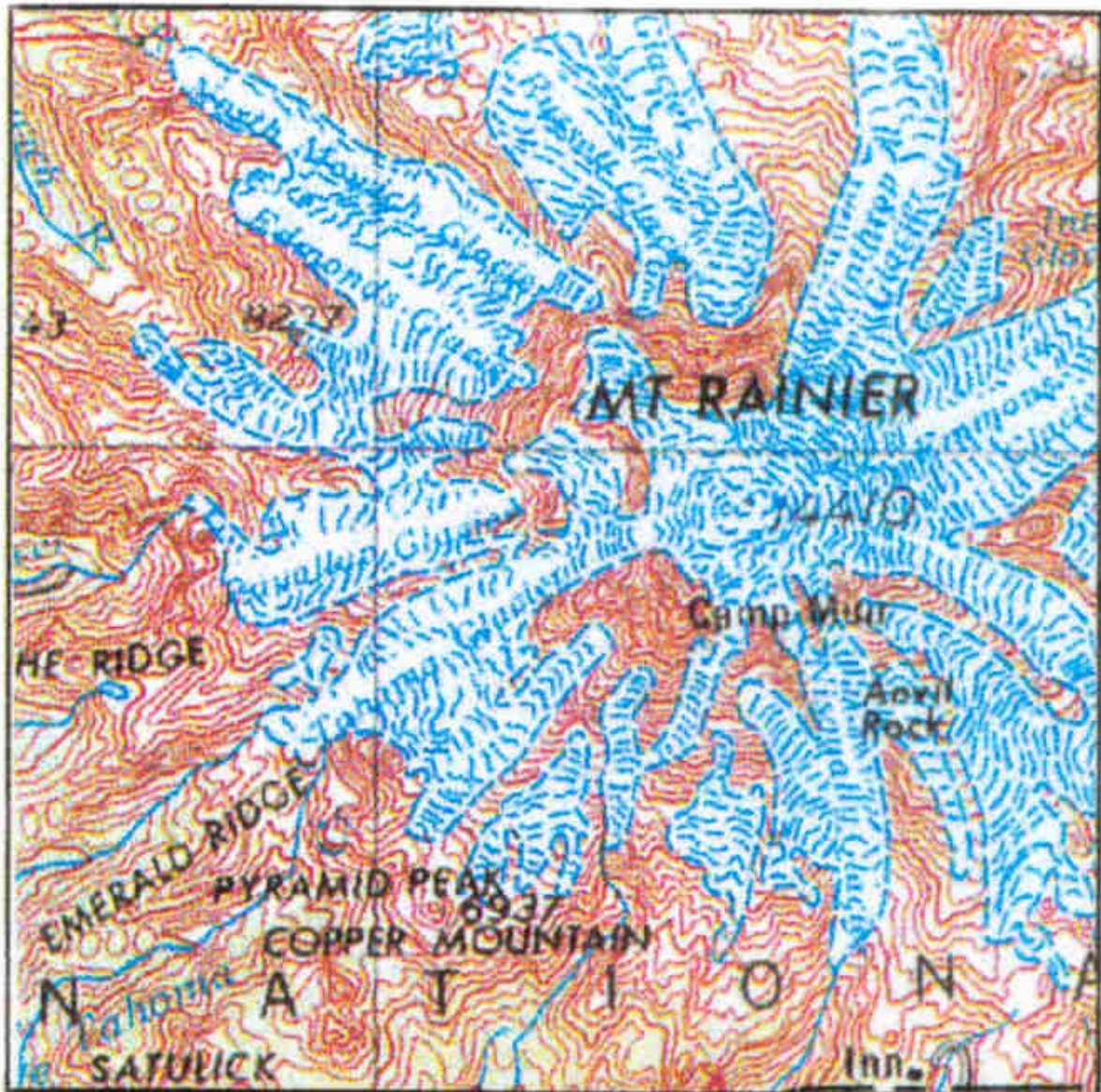


Mount Rainier
At 1:24,000





Mt. Rainier at 1:100,000



Mt. Rainier at 1:250,000

Generalization and context

- Whether a geographic object should be displayed, depends not so much on its size but its value both in the context of the map the me and its surrounding features.
- Context is also important in term of the map's intended use.

Scale transitions

- Some transitions between scale are well worked out in terms of displaying individual objects.
- Transportation routes, waterways and settlements are associated with specific symbols.
- The main impediment to easy generalization is not representation but *conflict* between features and the need to incorporate deep meaning associated with context and structure of geographic objects.

Deep Meaning

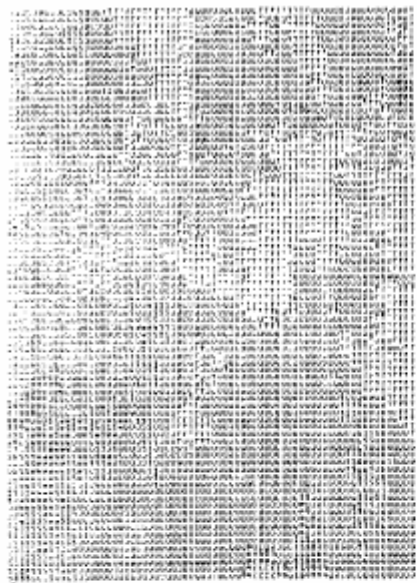
- Maintaining deep meaning implies that map objects will retain logical inter-relationships and symbologies as they move between scales.
- Many phenomena and relationships are difficult to transform between scales in a digital environment because their relationships shift depending on context.

Traditional generalization

- Computerization has resulted in a need for algorithmic solutions to processes which were never explicitly rule-bound though they were subject to convention.
- Digital generalization is, in effect, the search for computerized techniques to do something that was never consistently executed or theorized before the advent of GIS.

A brief history of digital generalization

- In the 1960s and 1970s, geographers' greatest criticism of GIS-generated maps was that they looked *terrible*.
- Geographers were generally slow to embrace GIS; a poor graphical product was seen as a legitimate complaint, especially since the contribution of spatial analysis in GIS had not yet been established.



This map was produced in the early 1980s probably using Fortran-generated output. Each character represents one ground rectangle with different characters signifying different ground cover types.

A vector map from 1984 with polygon identifiers penciled in by hand.



Early GIS output.



Line generalization

- For early GIS, the focus was on geometric simplification.
- Mass digitizing of existing maps was the main source of map data and the expense of storage was prohibitive.
- The Douglas-Peucker algorithm for line simplification was published in 1973 and remains the standard of generalization.

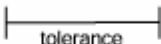
The Douglas-Peucker Line Generalization Algorithm

(adapted from Chrisman, 1997)



1. A tolerance level is established. (Points that exceed the tolerance level will be discarded.)

tolerance
(for deviation from trend line)



A horizontal line with vertical end caps, representing the tolerance level for deviation from the trend line.



4. Distance to point is compared to the pre-determined tolerance
(If smaller, then reject point, else accept point.)



The Douglas-Peucker line generalization algorithm is the most popular of its genre. First published in 1973, it continues to be the workhorse of generalization. It is simple and elegant. But, ironically, its original purpose was data reduction rather than generalization

Limitations of the D-P algo

- Algorithms are designed to work within certain parameters of data quality and these have shifted substantially over the past three decades.
- The D-P algorithm was originally intended for data reduction in the days when hard drive space was at a premium; it was not developed to conform to cartographic principles.

Problems with generalization

- During the **1980s**, point and areal generalization lagged behind line simplification and the inability of automated generalization to deal with context led to despondency about the capabilities of computer in this realm.

Model versus landscape generalization

- Generalization in GIS involves simplification of the database followed by appropriate re-visualization of the information.
- These two steps are distinct yet inseparable; the data are never disassociated from the map but the principles of generalization are very different for both.

Brassel/Weibel paper (1988)

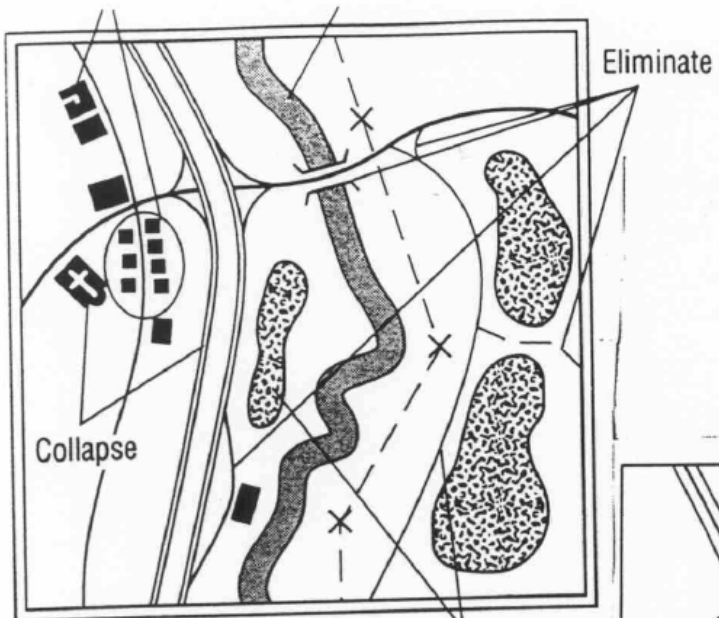
- In 1988 Kurt Brassel and Robert Weibel published a paper in which they made the distinction between model and cartographic processes of generalization.

Model generalization

- Model generalization “is mainly a filtering process” while cartographic generalization is concerned only with the modification of “local structure,” and figures only at the point of display.

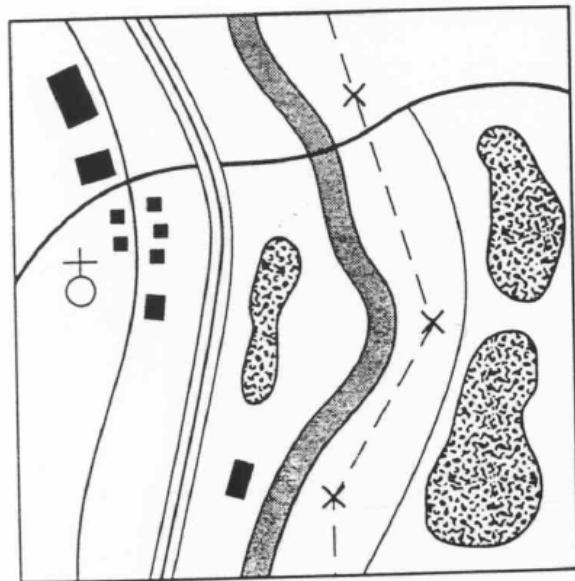
Operators used in generalization

- Basic operators used in landscape generalization are : aggregation, simplification, exaggeration, elimination, displacement, and collapse.

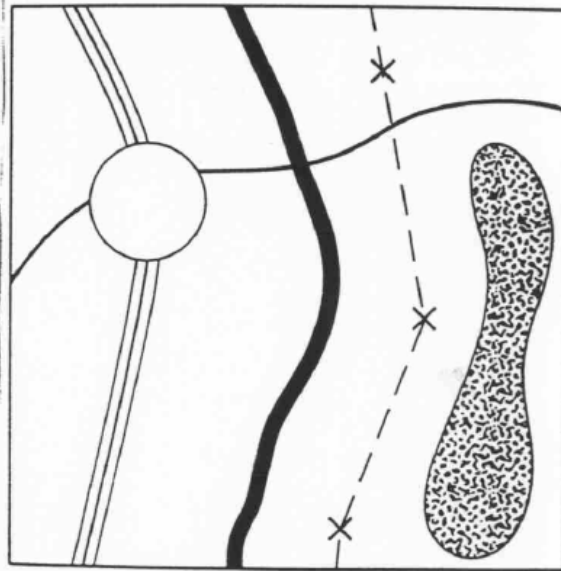


Map at original scale

Displace



Generalised for 50% reduction



Generalised for 25% reduction



Social/physical generalization

- Different operations are used in different contexts with different results.

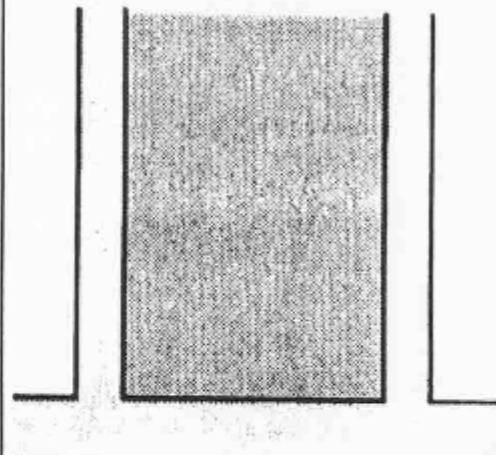
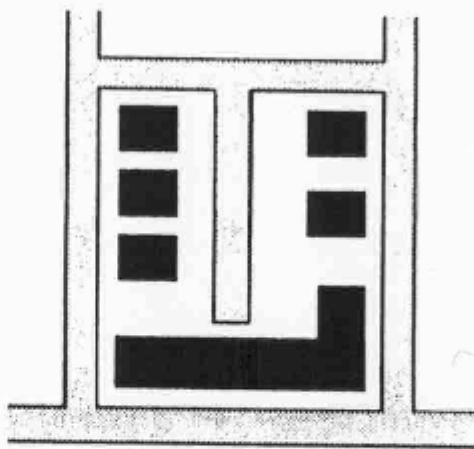
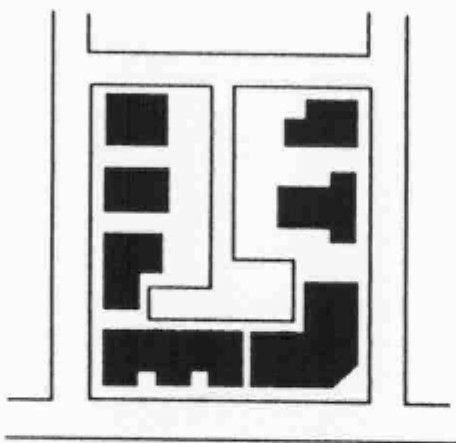
Large scale

Medium scale

Small scale

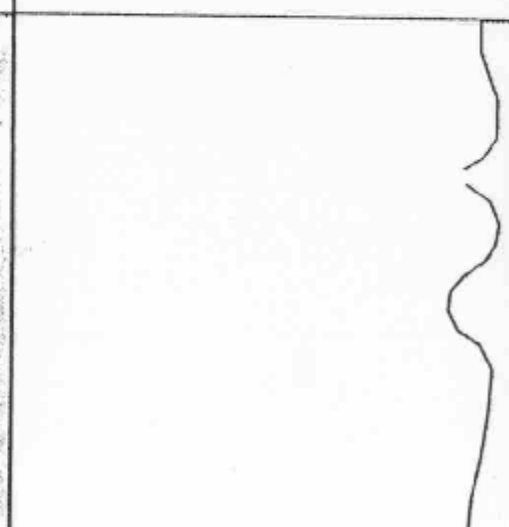
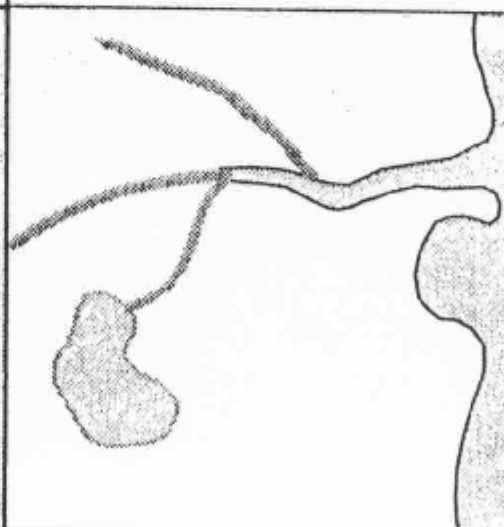
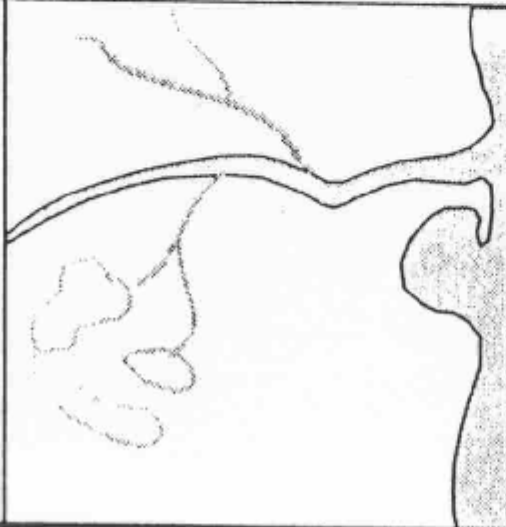
Culture

simplification,
aggregation,
elimination,
collapse



Hydrography

simplification,
exaggeration,
aggregation,
elimination



Rules and artificial intelligence

- AI means different things to different people.
 - It can refer to the IF THEN ELSE rules which are the mainstay of computer programming. A hypothetical algorithmic example for generalization at 1:100,000 looks like this:
 - If [stream=] 0.01 inches < stream width < 0.02 inches for longer than 2.64 inches
then represent stream channel as an area, otherwise a line.
 - If [stream=] 0.02 inches < stream width < 0.03 inches for less than 2.64 inches
then represent as a line, otherwise as an area.
- (Buttenfield 1995, 102).

Formal vs informal environments

- Manual cartographers combined intuitive and formal rules to generalize.
- Computers think sequentially, and automated generalization is necessarily rule-based and procedural.



Problems with expert systems

- Rule-based systems have had trouble dealing with the inter-relatedness of every generalization decision.

Knowledge engineering

- Difficulties in the **knowledge engineering** (KE) department has been one of the chief impediments to developing expert systems.
- Knowledge acquisition now inclines toward machine learning.

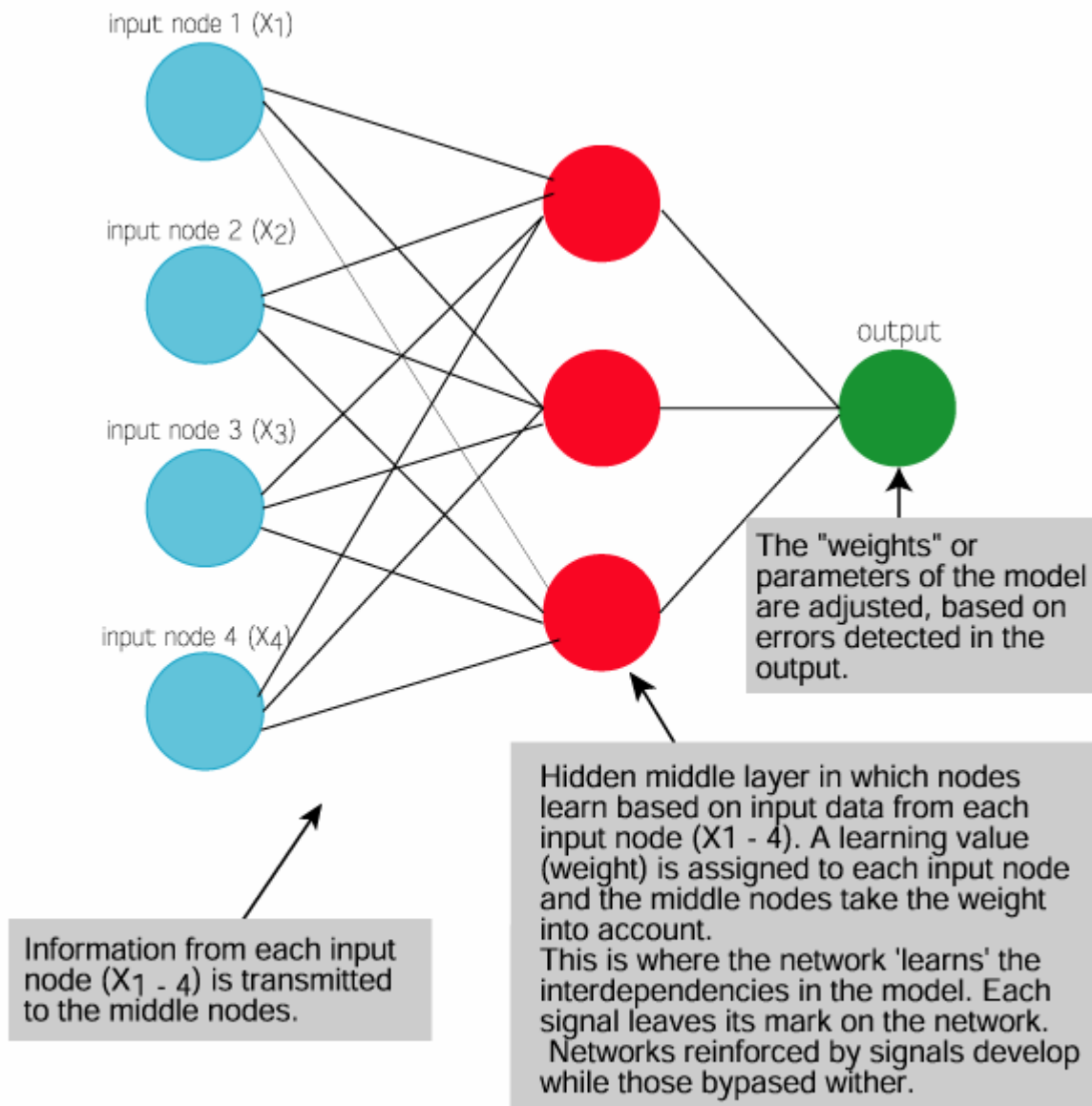
Problems with machine learning

- Regardless of the complexity of their knowledge base, rule-based systems haven't worked.
- They have been “brittle,” lacking the flexibility required for dynamic problem-solving.
- Rules breed exceptions which are numerous and difficult to encode.

The “new” AI

- Two new techniques, in particular, have been investigated since the early 1990s.
- Artificial neural nets (NN) and genetic algorithms (GA) represent possible keys to success for some generalization tasks.
- Computer systems and rule-based systems process information in a linear or serial fashion.

Conceptual structure of an artificial neural network



Neural nets mimic the way that scientists believe the human brain to work. They are a convincing example of diffusion between cultural metaphors in science and technology.

Neural nets

- Neural nets are trained on sets of data which ‘teach’ them how to solve types of problems.
- Information signals are input and then partially processed by single “neurons” which pass them along to other neurons for further processing.
- Signals are then moved between layers (in serial fashion), with processing occurring at every node, until they reach an output layer.
- Links between nodes are weighted according to their connection strength which must be fine-tuned during the learning stage.

Limitations of neural nets

- NN can ‘overlearn’ a certain set of data and consequentially lose their capacity to generalize dissimilar cases.

Uses of NN

- NN **do** have potential use in classification of map features; structure recognition; assessment of a set of different generalization choices; and replacement of algorithms with more holistic solutions.

AGENT

- New research in generalization is focused in Europe.
- AGENT encompasses an object-oriented approach based on entities that have specific, assigned goals.

A cross-section of AGENT

