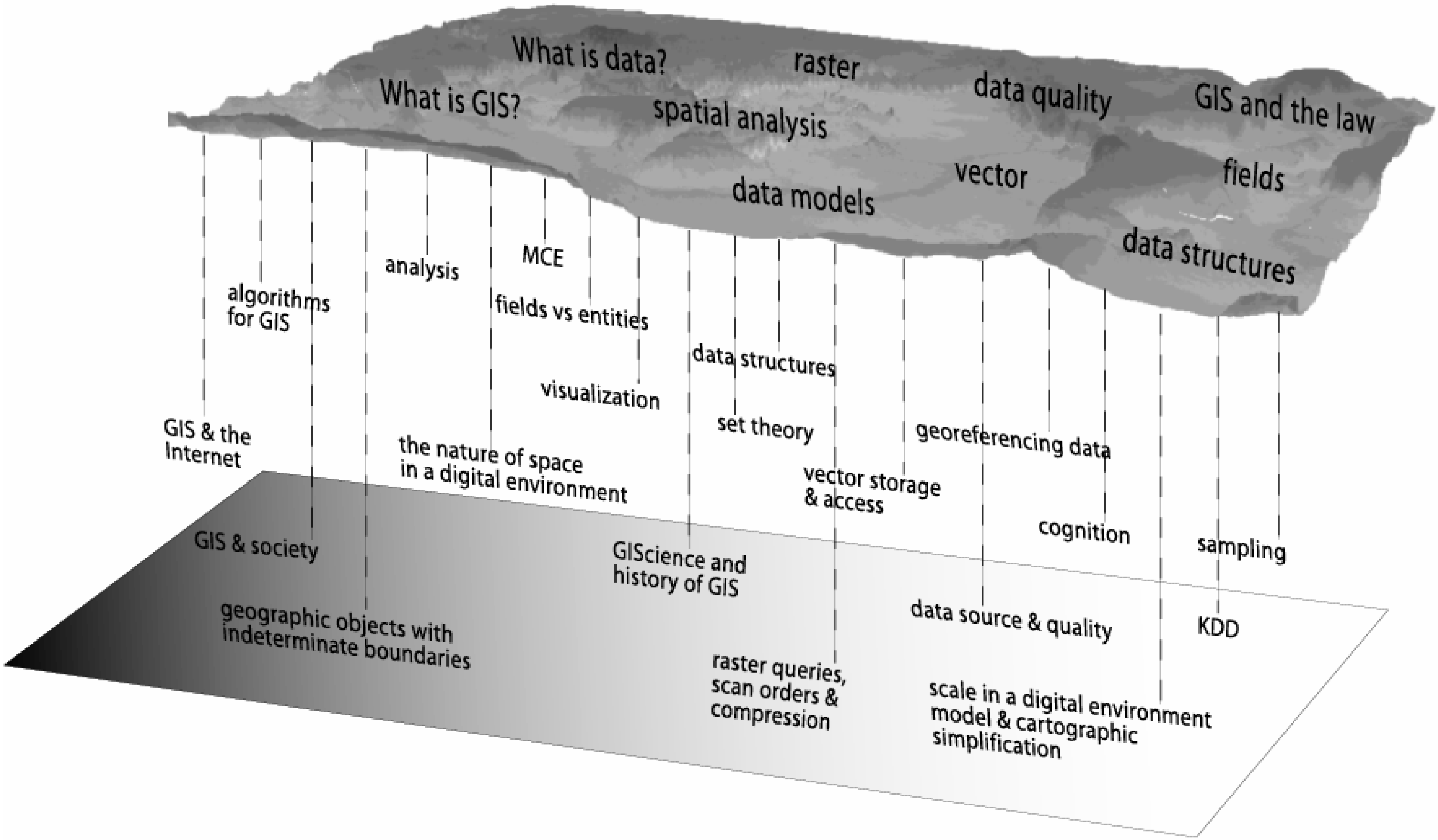
Representing the Earth

GIScience and the nature
of digital space

Structure of the course

- Think of 354/255 (your introduction to GIS) as being a surface.
- This course digs below the surface to discover what the underlying mechanisms are.


What is data?
raster
What is GIS?
spatial analysis
data quality
GIS and the law
vector
fields
data models
data structures
analysis
MCE
algorithms
for GIS
fields vs entities
data structures
visualization
set theory
GIS & the
Internet
georeferencing data
the nature of space
in a digital environment
vector storage
& access
cognition
sampling
GIS & society
GIScience and
history of GIS
data source & quality
KDD
geographic objects with
indeterminate boundaries
raster queries,
scan orders &
compression
scale in a digital environment
model & cartographic
simplification

Geographic information systems versus science

- This class is more about geographic information science than systems. I want to start by differentiating the two.

- Question: How would you describe geographic information *systems*?

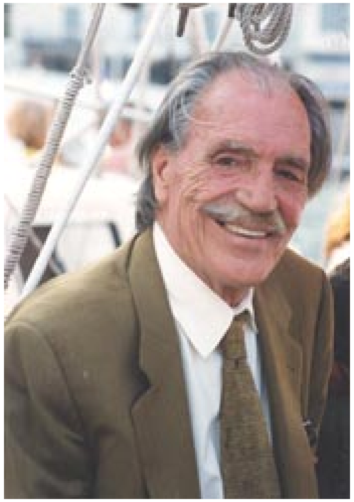
Definitions of GISystems

- GIS is any system designed for capturing, storing, checking, integrating, analyzing and displaying spatially referenced data about Earth.
- A GIS is a system that allows the combination of geographic datasets (or layers) and the creation of new geospatial data to which one can apply standard spatial analysis tools

More definitions

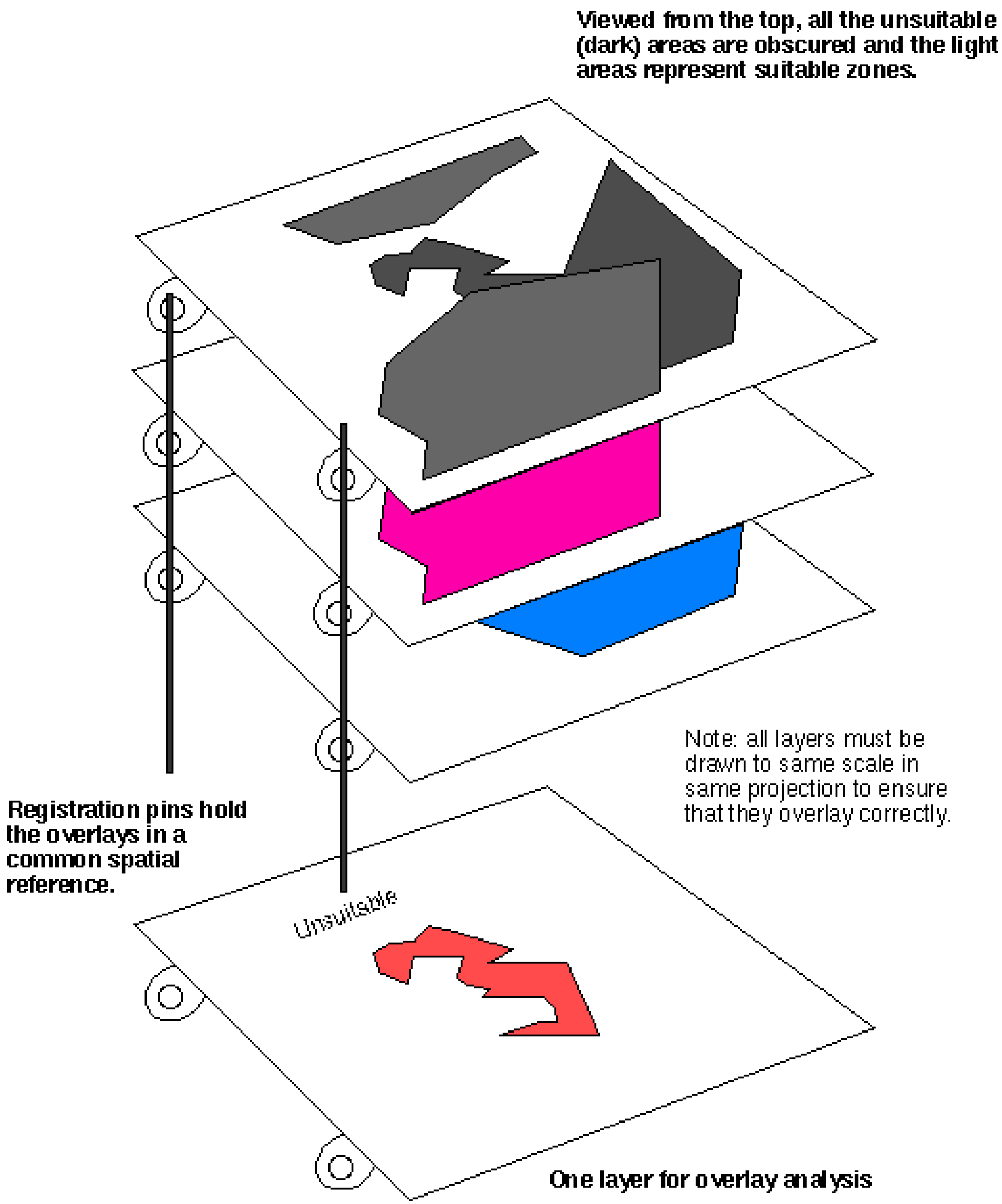
- An organized collection of computer hardware, software and procedures designed to support the capture, editing, management, manipulation, analysis, modelling and display of spatially referenced data for solving complex planning and management problems. A GIS is characterized by its ability to perform topological structuring of data.
- GIS relates a location to an asset or an event. The system may be manual or computerized.
- A relational database with a sense of space

- Best way to start an arguments among GISers: ask them to define GIS.
- Why? Because GIS has evolved so radically since the 1960s.
- Started out as a way of automating cartography.
- Analysis – the defining difference between GIS and cartography – was initially just a bonus.



Ian McHarg and overlay

- Ian McHarg started the program in Landscape Architecture at the University of Pennsylvania.
- Links between GIS and landscape architecture/surveying.
- Remember: GIS evolved as a set of practices linked to hardware and software.


Viewed from the top, all the unsuitable (dark) areas are obscured and the light areas represent suitable zones.
Registration pins hold the overlays in a common spatial reference.
Note: all layers must be drawn to same scale in same projection to ensure that they overlay correctly.
Unsuitable
One layer for overlay analysis

Other highpoints in the history of GIS

- GBF/DIME stands for "Geographic Base File" using "Dual Independent Map Encoding". The DIME system was developed at the US Bureau of the Census in 1967, in preparation for the automation of geocoding of the 1970 census.

Segment name	From node	To node	Block left	Block right
Salem Rd.	75	76	302	306
Elm St.	76	77	304	306
Elm St.	77	80	304	306
Taurus Ave.	80	79		
Belt Dr.	79	75		
Pine Place	77	78		

- DIME was precursor to TIGER, urban areas only
- coded street segments between intersections using
- IDs of right and left blocks
- IDs of from and to nodes (intersections)
- x,y coordinates
- address ranges on each side
- this is essentially the arc structure of CGIS and the internal structure (common denominator format) of POLYVRT

- DIME files were very widely distributed and used as the basis for numerous applications
- topological ideas of DIME were refined into TIGER model
- Topology was further refined in TIGER.
- DIME, TIGER were influential in stimulating development work on products which rely on street network databases
- Harvard Graphics Lab (hothouse of ideas)
- ESRI (implemented CGIS idea of separate

attribute/locational info; RDMS

So what makes it GIScience?

- As the scope of GIS increased and users proliferated, people started to recognize that GIS does much more than could ever be done by manual mapping.
- By 1990, GIS scholars (Michael Goodchild, in particular) had coined the term acronym GISci (Geographic Information Science).

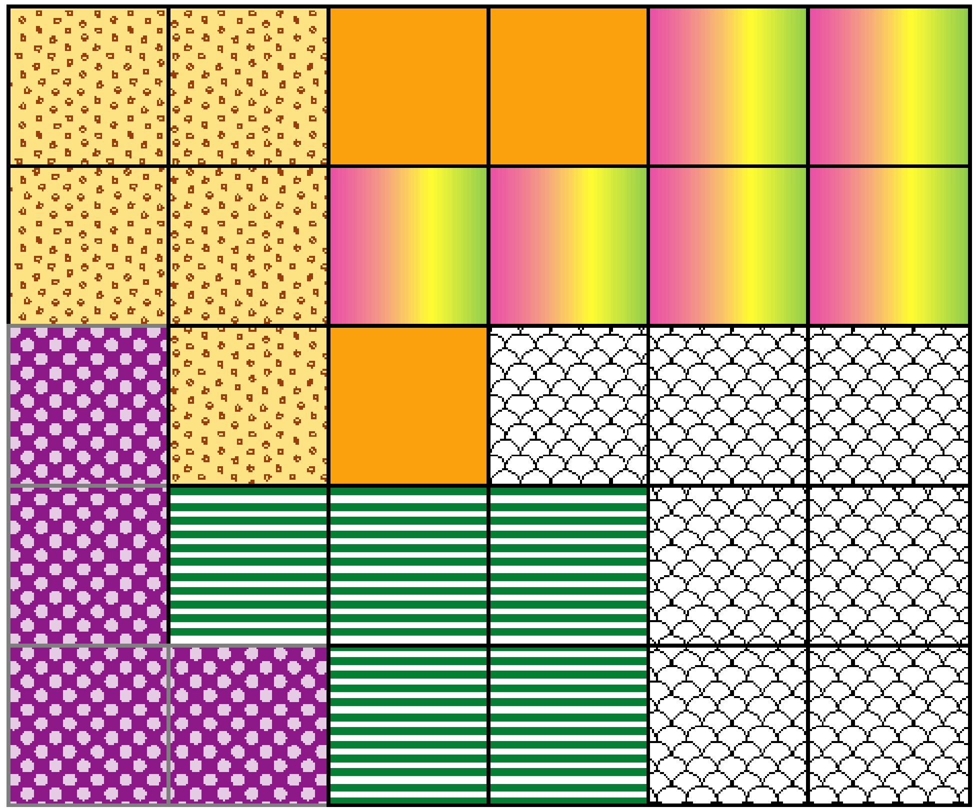
- GIScience is the science behind the technology
- Because GIS can process huge volumes of data and extend the scope (not scale) of geographical analyses, it allows us to ask questions that were never before possible.
- GIS has also changed the nature of the questions. It potentially allows researchers to make queries in spherical and temporal space.

What happened to space?

- The early days of GIS: focus on solving computational problems using paradigms of cartography.
- But RASTER is developed because of computational simplicity and ease of output.
- Results: two data models (raster and vector).

The Raster View of the Earth

The earth is divided by a regular tesselation. Each grid cell represents a location and the resolution of the coverage is determined by the grid size. Attributes are assigned to each cell, one attribute per layer. Location is everything.



A few observations about raster data models

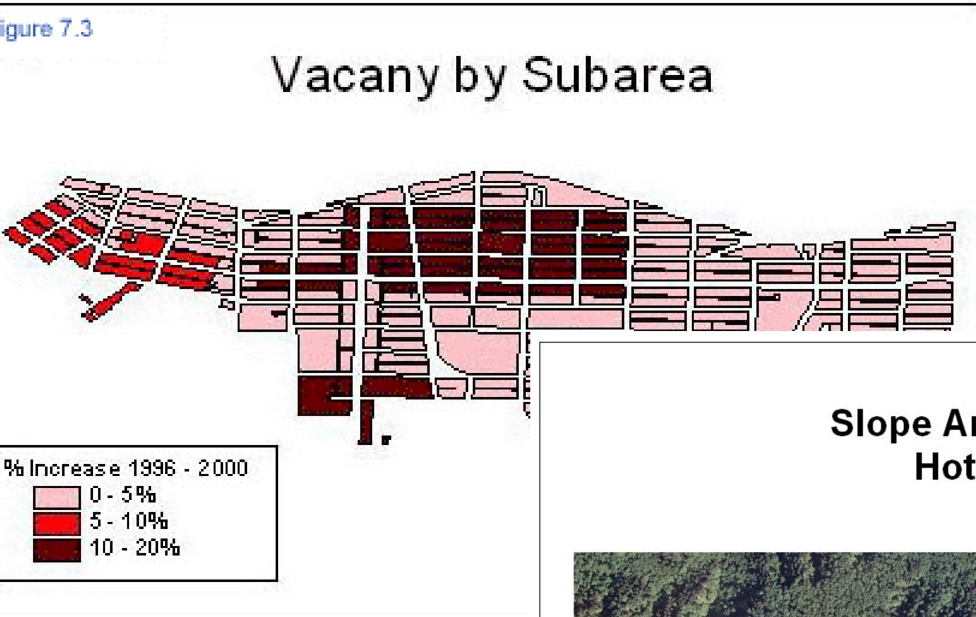
1. Raster data models divide the world into a sequence of identical, discrete entities, by imposing a regular grid. Frequently the grid is square.
2. In each instance, the contents or attributes of each grid cell (or raster) is tied to that location.
3. Each attribute layer includes the same grid cells, identified with a defined geographical area; each cell contains a single value for each attribute.

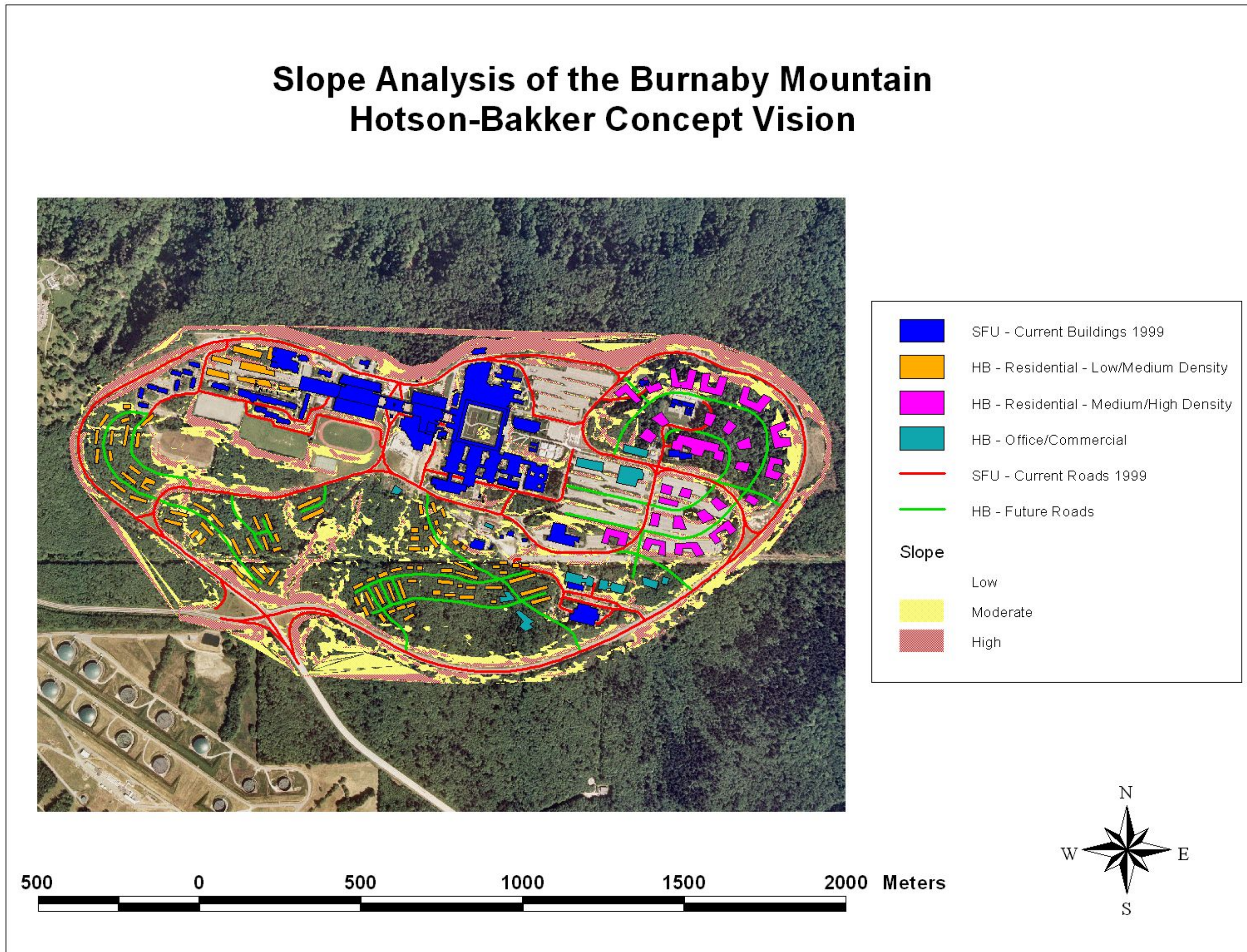
More characteristics of rasters

4. Each attribute layer includes the same grid cells, identified with a defined geographical area and each cell contains a single value for each attribute.
5. The matrix of grid cells is conceptually similar to data structure arrays; each layer is stored as one dimension of an array.
6. Raster systems are widely used in applications which employ remotely sensed images as satellite imagery is itself generated through a regular tesselation of space.

Even more about raster data models…

7. Raster data models are also well-suited to operations which determine friction of movement. (why?)
8. Raster data models have an advantage, in principle, over vector data models as the basic element of geographic data is the mathematically defind tuple $T=(x,y,z_1,z_s,z_3,z_4\ldots,z_n)$ in which x and y are the coordinates and z is an attribute.
9. The infinite number of tuples which populate a raster coverage comprise a *scientific* field. (is that field really continuous?)

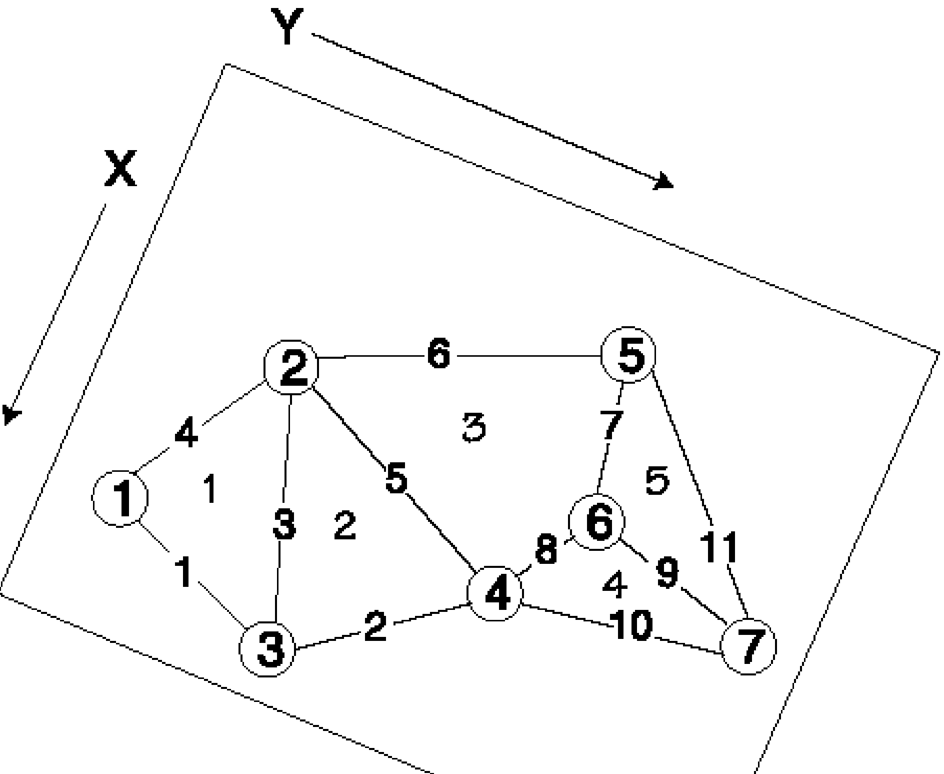

igure 7.3
Vacany by Subarea
% Increase 1996 - 2000
0 - 5%
5 - 10%
10 - 20%


Slope Analysis of the Burnaby Mountain
Hotson-Bakker Concept Vision
SFU - Current Buildings 1999
HB - Residential - Low/Medium Density
HB - Residential - Medium/High Density
HB - Office/Commercial
SFU - Current Roads 1999
HB - Future Roads
Slope
Low
Moderate
High
N
W
E
S
500
0
500
1000
1500
2000
Meters

A few notes about vector data models

1. Vector-based GIS have traditionally been distinguished from cartographic points and lines by their data structures which include topological information, based on adjaceny and connectivity.
2. Topological information has long been a hallmark of vector data models as it allows re-drawing of areas without drawing points or lines twice.

Explicit Vector Topology



Find the mistakes…

Link #	R-H Polygon	L-H Polygon	Node 1	Node 2
1	1	0	3	1
2	2	0	4	3
3	2	1	3	4
4	1	0	1	4
5	3	2	4	4
6	3	0	2	5
7	5	3	5	6
8	4	3	6	4
9	5	4	7	6
10	4	0	7	4
11	0	5	5	7

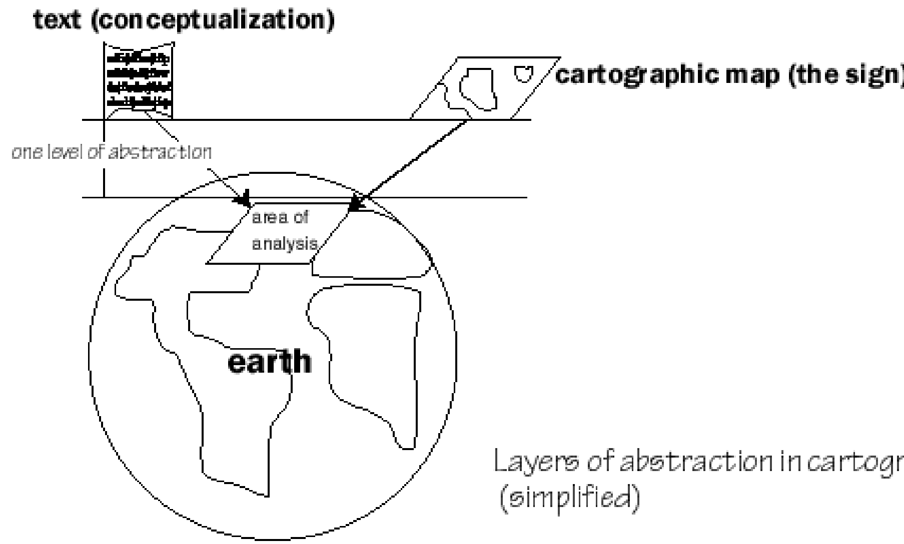
More about vector data models

3. More importantly, topology is used to expedite the computation of spatial queries such as “how many gas stations are within 1000 meters of the intersection of the Trans Canada Highway and route 97?”

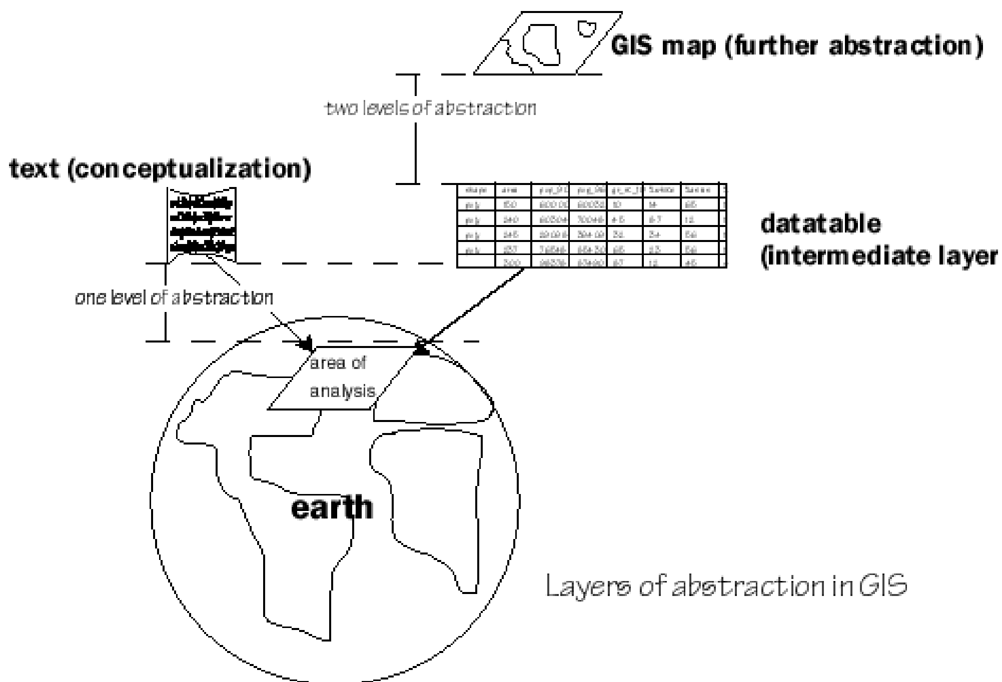
Space in GIS/GIS in space

- Questions about how space and geographic entities are *formalized* – what is lost and gained in moving between semantic and computational descriptions – are being asked.
- Ontologies and methods of formalization have necessarily engendered serious theoretical research in GIS

Tranlsation from the infological to datalogical



Layers of abstraction in cartography (simplified)



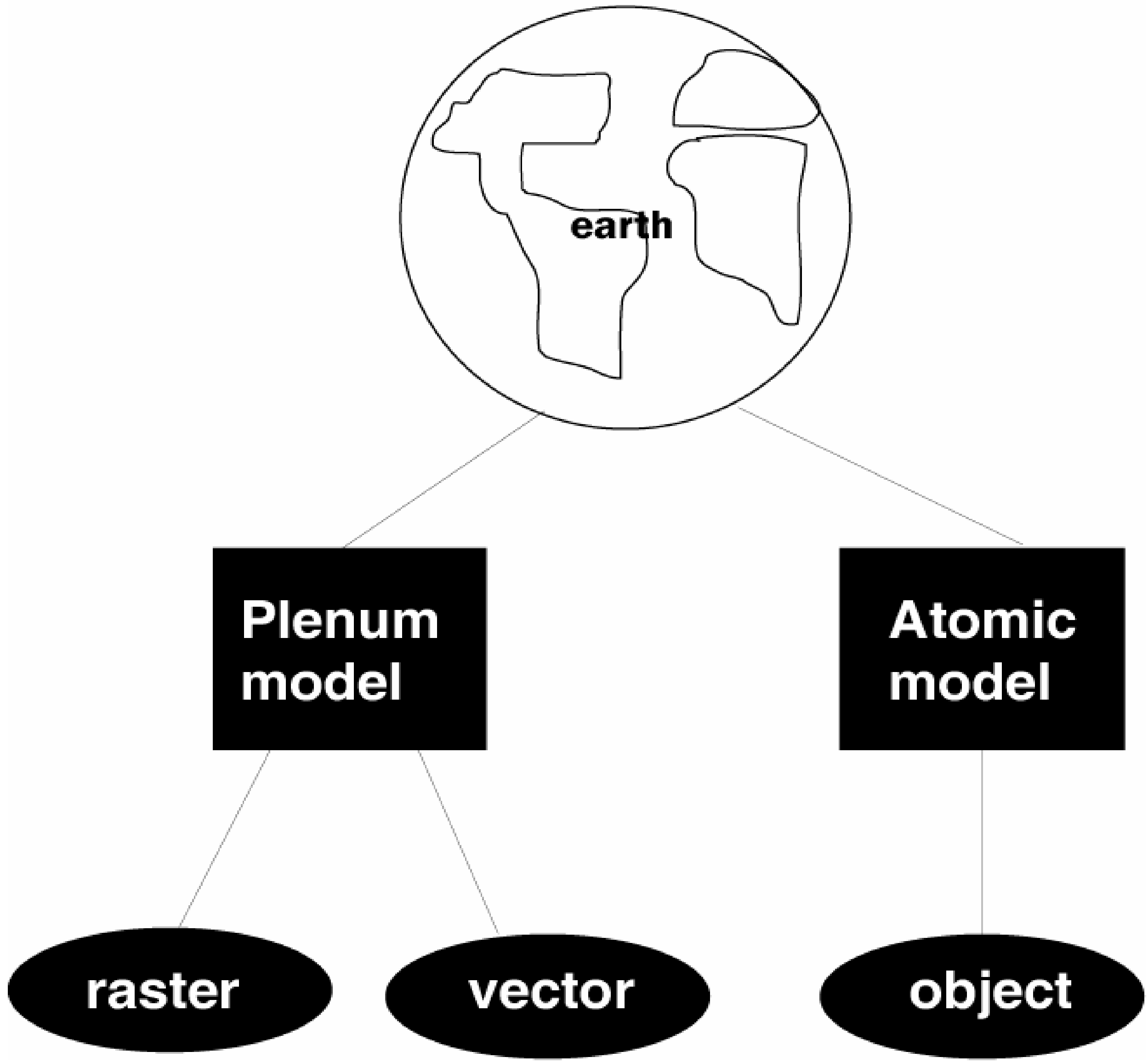
Layers of abstraction in GIS

Atomic vs plenum

- Technical questions about data models (formalized abstraction) are really philosophical questions about how we *conceptualize* space in GIS.
- We use two models to conceptualize space: the atomic (object) and plenum (field).
- The atomic model assumes that entitites are positioned in absolute space and exhibit attributes, including space and time, which we can measure.

- The plenum ontology considers clusters of attributes in space to be the things themselves.
- This is related to the raster or field view. Attributes coagulate to form objects and entities do not have an identity outside of their attributes.

Field and Object Data models



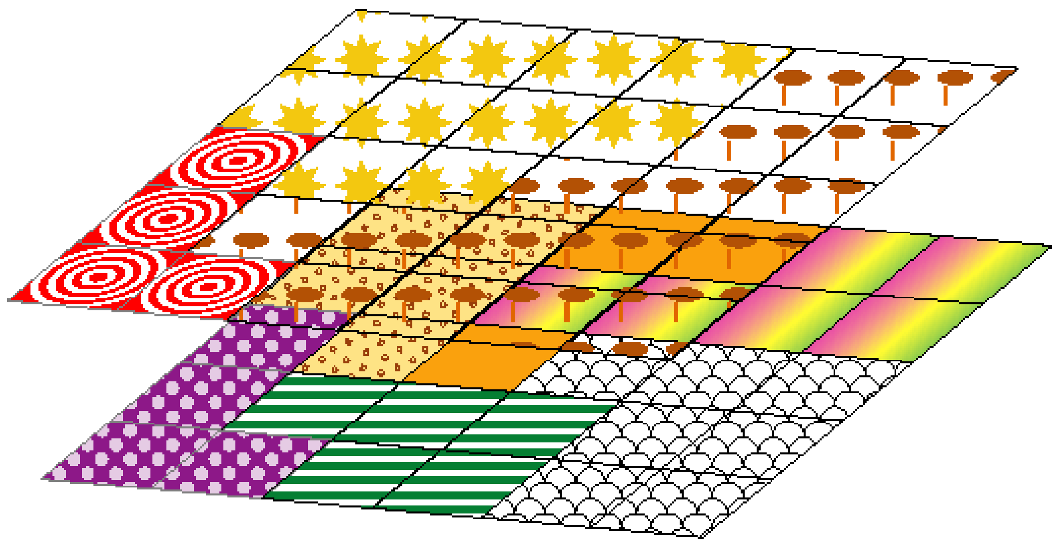
- In the plenum view, different combinations of attributes are different things.
- This is the inverse of the object (atomic) view in which objects exist in their own right and *manifest* or *exhibit* attributes.
- We assume that objects can be moved about, stacked and manipulated like building blocks, as if they had no relation to the space they fill but *geographic* objects are famously fuzzy and can scarcely be considered separately from the space they occupy.
- What to do?

Objects: reworking the data model paradigm in the 90s

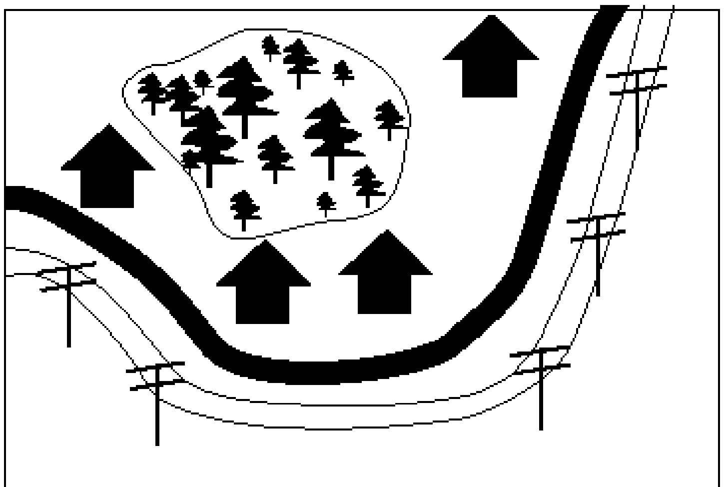
- Objects came into GIS from computer science in the late 1980s and early 1990s.
- *Objects* interrupt the vision of the world as a series of locationally registered layers, each representing a single attribute.
- Object-oriented GIS defines geographical phenomena, such as telephone poles or streets, as objects.
- Location becomes one of many other attributes associated with a particular object.

How can I tell an object from a field?

Field and object data models

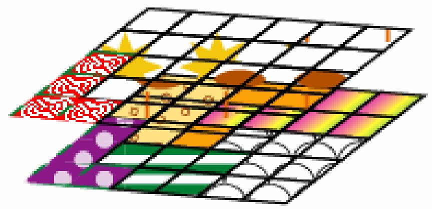


Field data models can be envisaged as layers which register to the same geographical coordinates, each containing information about one attribute or theme.

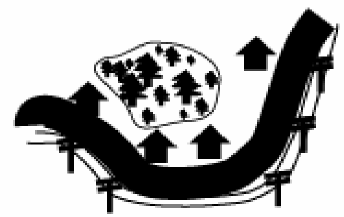


Object data models do not necessarily acount for every point in the map area. Rather they portray individual objects which can range from forests to telephone poles.

Fields and objects, rasters and vectors: a clarification

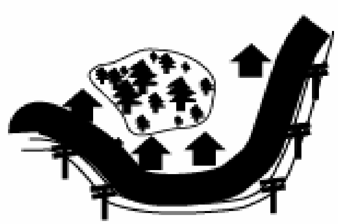


Fields

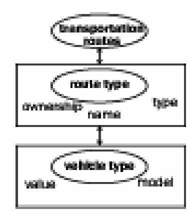


Objects

Both fields and objects can be implemented using either rasters (grid cells) or vectors (points and lines). Fields and objects are conceptualizations (infological), while raster and vector are means of implementation (datalogical) .



Infological objects



Datalogical objects

The meaning of objects at the datalogical and infological levels vary. At a representational level, objects are digital versions of spatial entities. At the level of implementation, object oriented systems allow child classes of object to inherit attributes and methods from parent classes. Datalogical objects can be used (but are not required) to implement infological objects. Raster and vector are used, in either instance, to portray digital objects.

A few questions:

1. Does the object model allow empty space?
2. Does the object model use layers?
3. How do objects handle the fuzzy boundaries that characterize geographical entities?
4. Do object models allow overlapping objects?
5. Are vector GIS based on object models?
6. What is the difference between the two?