Geographic Information System (GIS)
IS 454

Lecture 2: Representing the real world

Professor Dr. Safa A. Najim
Computer IS department
College of CS and IT

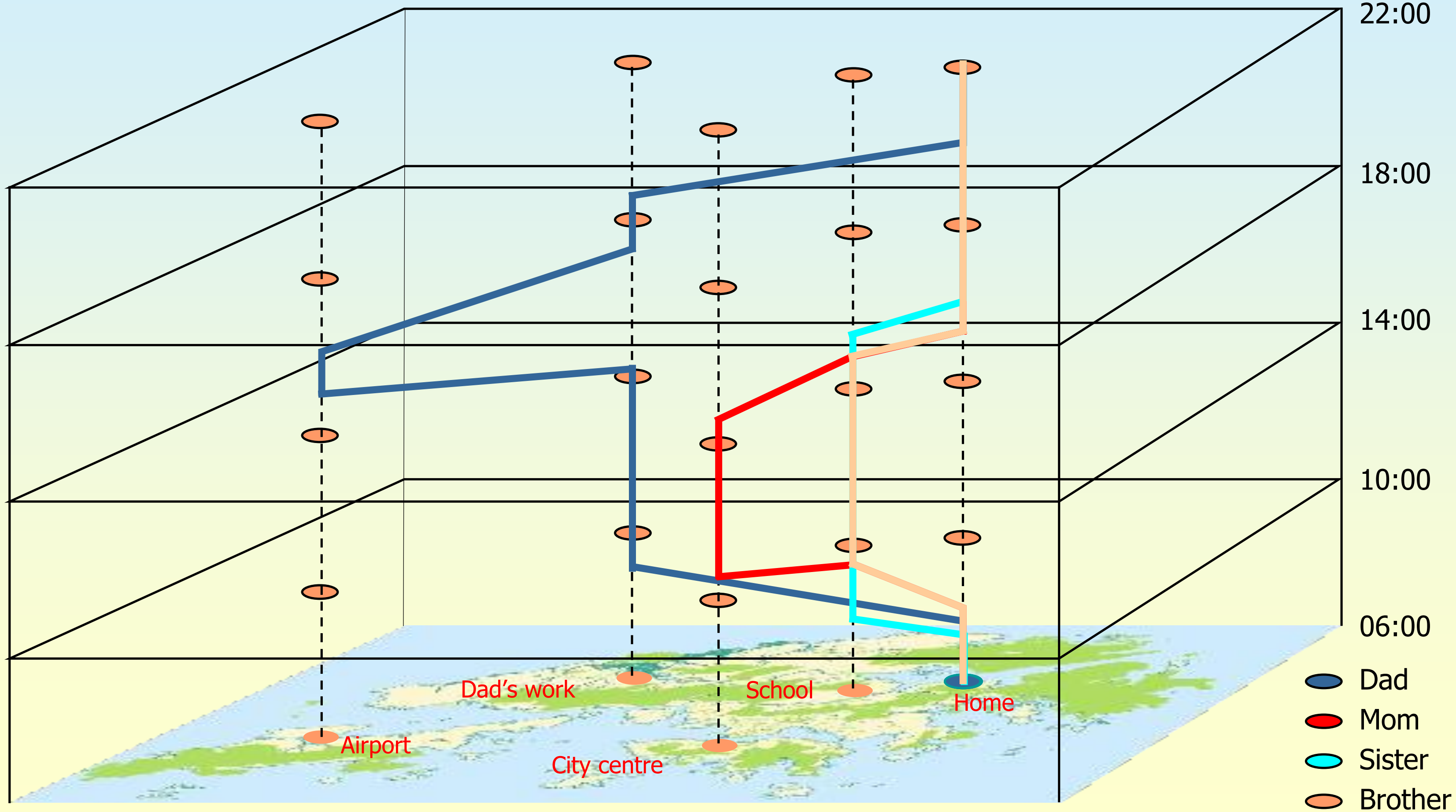
Representing the Real World

- Representation in GIS
- Objects and fields
- Rasters and vectors
- GIS and maps

Law of geography

- All human activities requires knowledge about the Earth.
  - Past, present and future
- Representation occurs
  - in human mind
  - in photographs
  - in spoken and written languages
  - in numbers (i.e. measurements)
- Law of geography: everything relates to everything, but nearby things are more related than those far apart.

Example: representation in space and time



Representation

- Representation help us assemble more knowledge about the Earth than is possible on our own.
  - e.g. maps
- Digital representation
  - Every item of useful information is ultimately reduced by a GIS to combinations of 0s and 1s.
  - Digital representation is simple with low costs.

Geographical representation

- Geographical representation is concerned with the Earth’s surface or near surface.
- The key representation issues are what to represent and how to represent it.
- Many plans for the real world can be tested on models or representations.

Place, time and attributes

- Geographical data link place, time and attributes.
- Place is represented directly by geographical coordinates, or indirectly by other means, e.g. place names, street numbers.
- Time is represented as a relative manner, such as dates, past, present and future.
- Geographical attributes are classified as nominal, ordinal, interval and ratio.
- When enough atoms of geographical information are collected, a complete representation of the world can be built.
  - Digital earth: to explore the world by interacting with its digital representation

A geographical place

Graphical representation

Yellow River (80×1260cm): graphical representation of the geography along mid- to lower-reach of Yellow River, including general shape of the river band, positions of cities along the river, place names and distances between places. It was initially made as an artistic masterpiece as a gift for Emperor Kangxi (康熙).

– A collection in the National Taiwan Central Library

Digital representation

Attributes + time

But …

- The world is infinitely complex.
- Computer systems are finite.
- Representation is, therefore, all about the choices that are made in capturing knowledge about the world.
  - Creating spatial models

The concept of spatial models

The essential function of the spatial data we store and manipulate is to subdivide the Earth’s surface into meaningful entities or objects that can be characterised. In this way, the contents of a spatial database is a model of the Earth.

— Star and Estes, 1990, pp 32

Objects and fields

- The fundamental representation in geography:
  - Discrete objects
  - Fields
- Discrete object view represents the world as objects with well-defined boundaries in empty space.
- The field view represents the real world as a finite number of variables, each one defined at every possible position.

Example of objects

Houses in a countryside are conceived as discrete objects with their clear boundaries and surrounding empty space.

Database entries of discrete objects

House ID	Owner	Built area (m²)	Location	Year built
12	J. Smith	210	42:30:12 E 43:40:10 N	1974
23	K. Jones	155	42:35:40 E 43:38:12 N	1932
51	M. Robert	346	42:41:06 E 43:40:45 N	1992
98	Skyline Co.	622	42:37:38 E 43:36:55 N	1997

Object representation in 2D

Fields

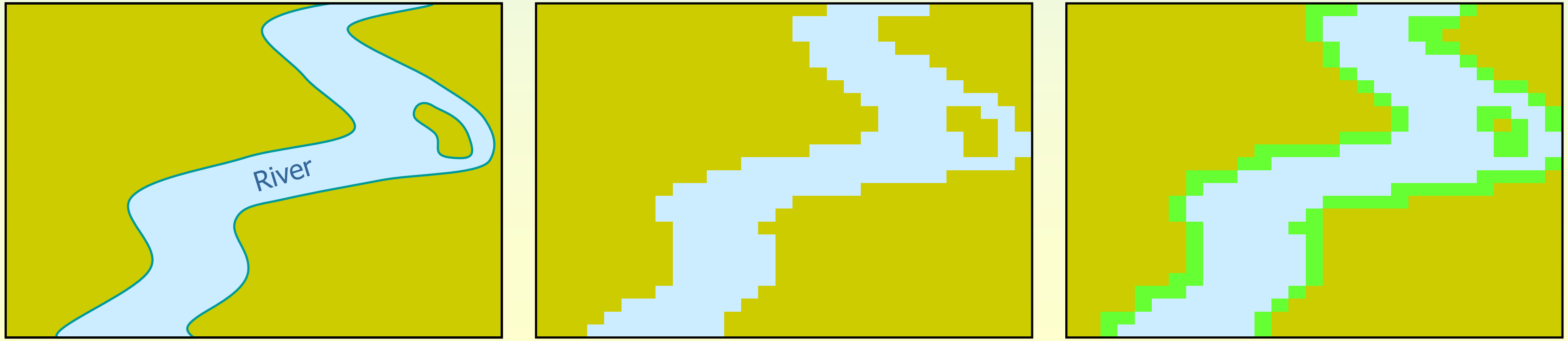
- Fields are used to represent continuous phenomena.
- This form the basis of continuous field view.
- Geography is described by a number of variables, each measurable at any point on the Earth’s surface and changing in value across the surface.
- The continuous field view represents the real world as a finite number of variables, each one defined at every possible position.

Field representation

Object or field?

Definition of river

- Object view: two-dimensional, discrete object in an empty geographic landscape.
- Field view: all points are either river or non-river.
- Fuzziness: the “membership” of river



Membership of "river"

Membership grade (%)	Definition
0	Always dry
25	Sometimes flooded
50	Wetland vegetation river band sediment
75	Most time water flow with river bed sediment
100	Always water flow

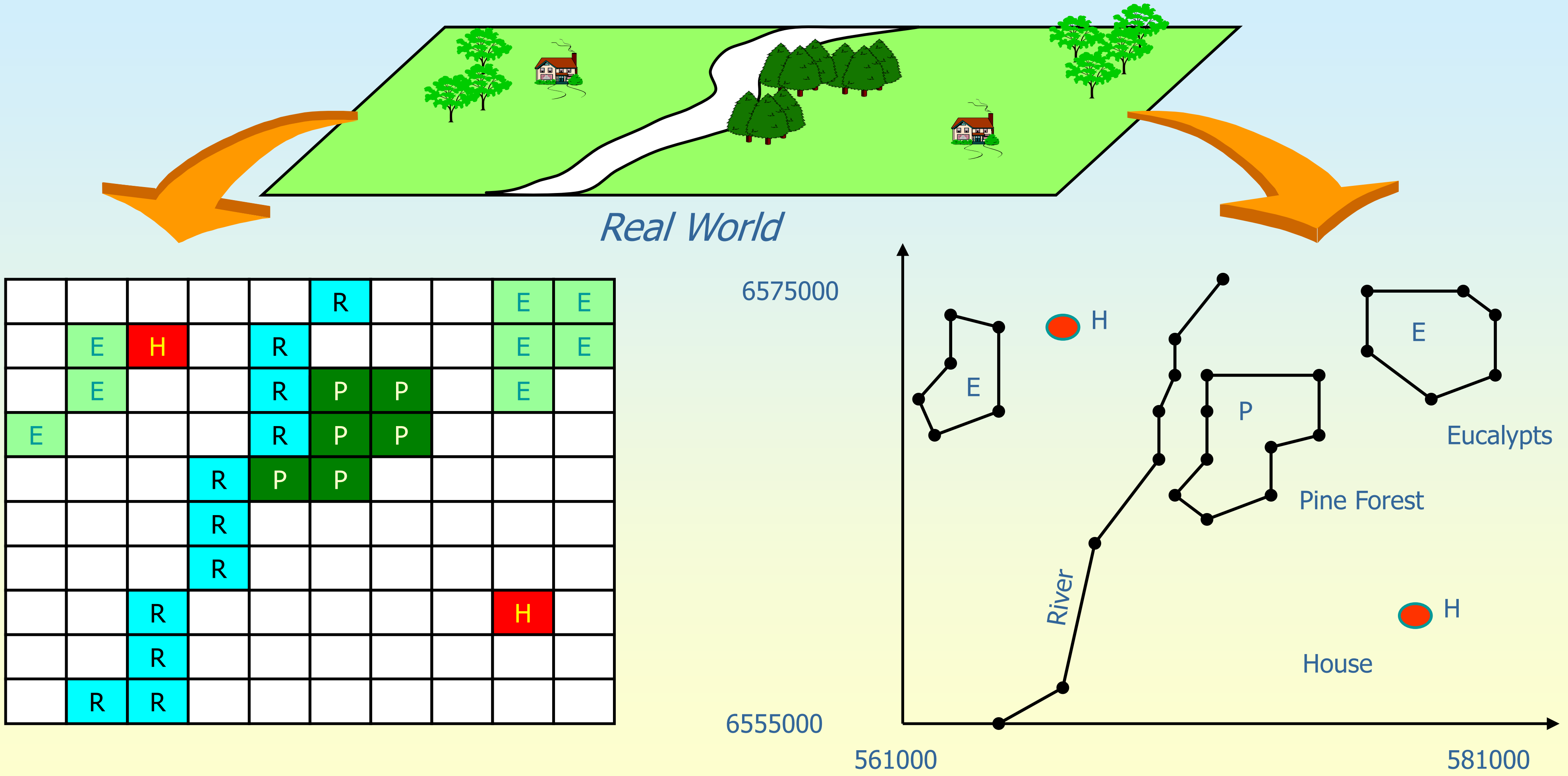
Abstraction of geographical entities

- **Points:** *locations* of, e.g., oil and water wells, bus stops, cities on a small scale map
- **Lines:** *centre lines* of, e.g., railways, highways, natural streams
- **Polygons:** *enclosed regions* such as reservoirs, lakes, local government areas
- **Volumes:** *solid bodies* such as building blocks, mine bodies, capacity of a reservoir
- **Processes:** *changing bodies* such as landslides, moving objects

Rasters and vectors

- There are two fundamental approaches towards the digital representation of the spatial component of geographical information — the rasters and vectors.
- In both methods, the spatial information is represented using finite, discrete homogeneous units.
- Raster representations divide the world into arrays of cells and assign attributes to the cells (or *pixels*).
- Vector representations use *points, lines* and *polygons*, in a way very similar as paper maps.

Raster and vector representations



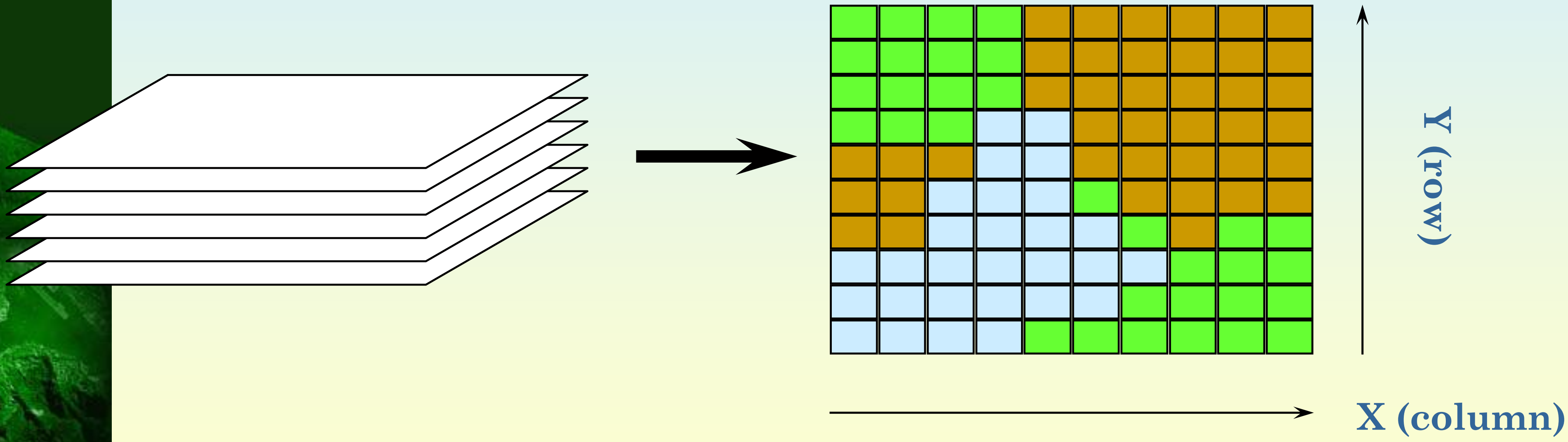
Raster Representation

Vector Representation

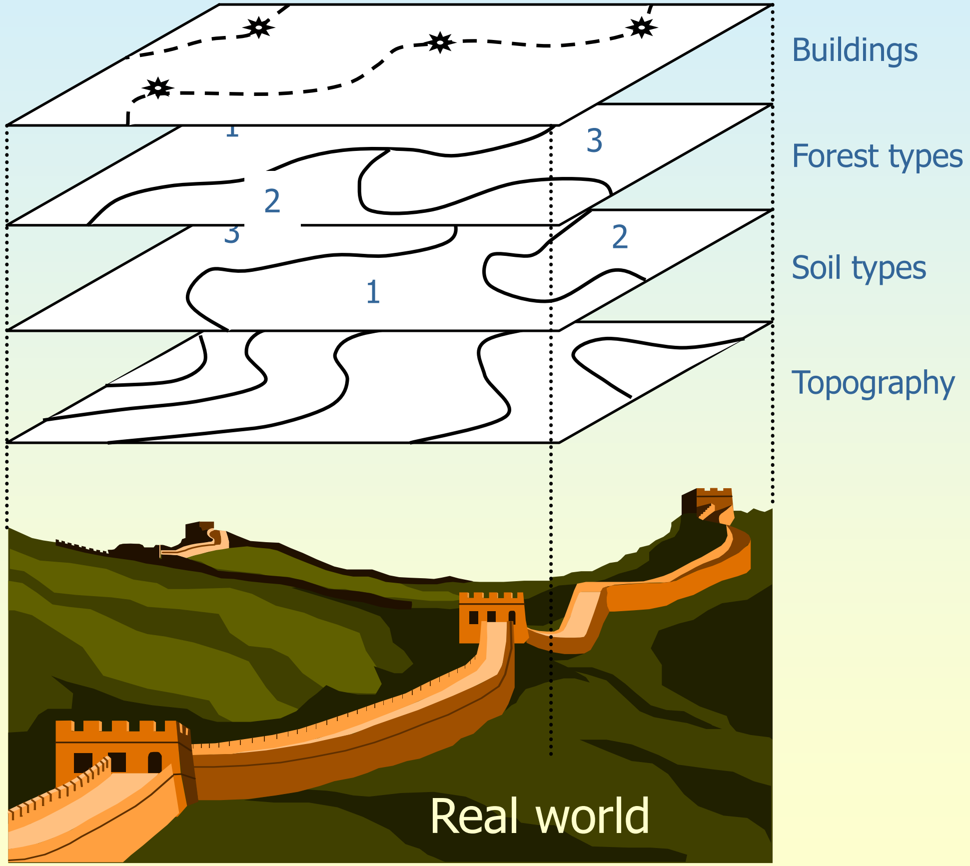
The raster representation

- Divides the entire study area into a regular grid of cells
- Each cell contains a single value
- Is space-filling since every location in the study area corresponds to a cell in the raster

The raster coordinates



Raster geographical database

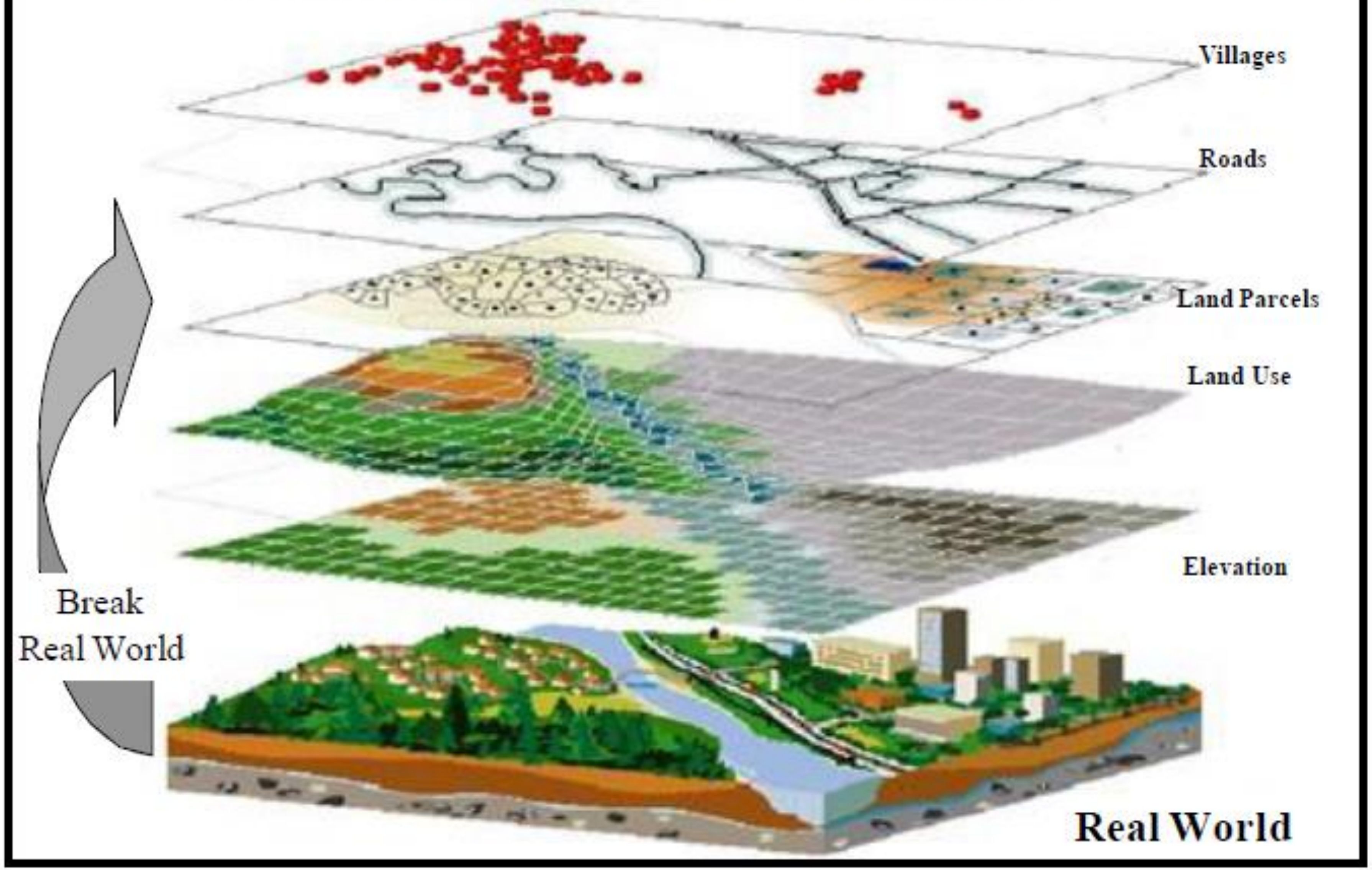


3	3	3	3	1	2	2
3	3	3	1	1	2	2
3	3	3	1	1	2	2
3	3	1	1	2	2	2
3	1	1	1	2	2	2
3	1	1	1	1	2	1
3	3	1	1	1	1	1

Soils

Interpretation of Real World in GIS Terms

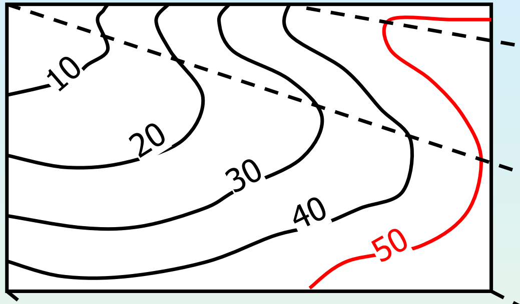
Ex:Locate a New Bus Terminal



Resolution

- The minimum linear dimension of the smallest unit of geographical space for which data are recorded.
- In the raster model, the smallest units are rectangular (for most systems).
- The smallest units are known as cells or pixels.
- The array of cells is known as lattice, grid or matrix.
- High resolution refers to raster with small cells.

3-D representation

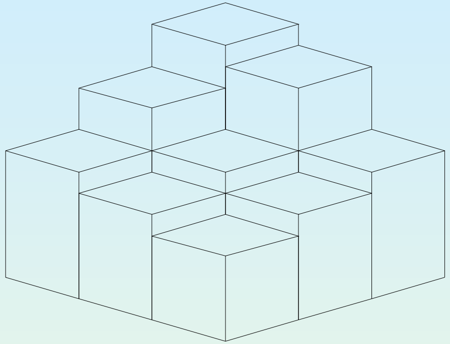


Column number

01	02	03	04	05	06	07	08	09	10	11	
08	09	10	15	20	28	35	41	48	50	51	01
08	09	11	14	19	25	34	40	47	52	55	02
10	12	15	18	19	25	33	38	48	51	52	03
12	12	15	17	20	15	33	40	47	50	51	04
20	20	20	25	30	32	38	40	45	50	51	05
25	24	26	30	32	37	41	44	47	50	53	06
31	30	32	35	38	40	41	45	50	52	53	07
38	37	39	39	40	41	43	49	51	53	56	08
42	41	42	44	45	48	49	50	52	54	57	09

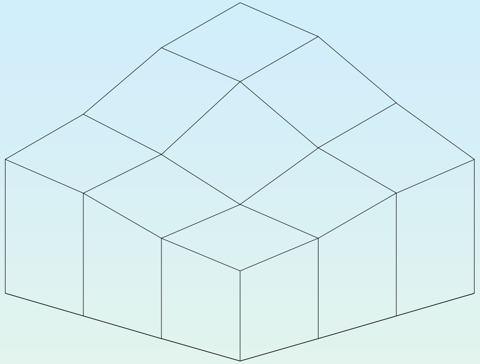
Row number

3-D surfaces



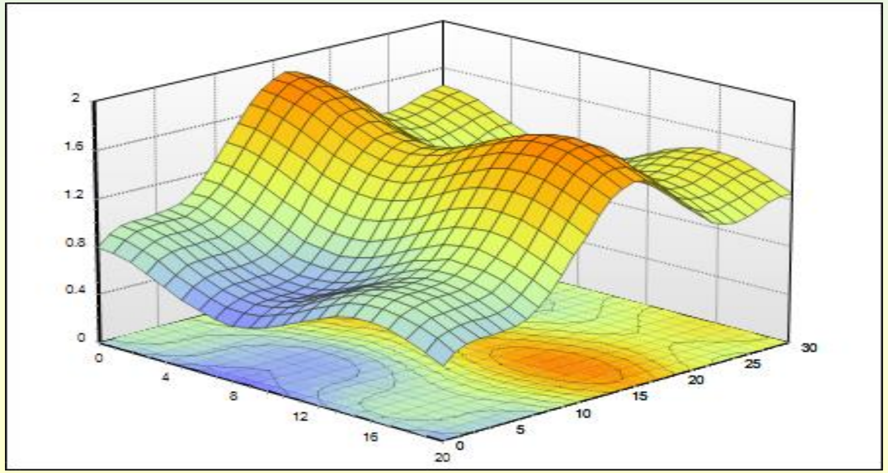
Discrete surface

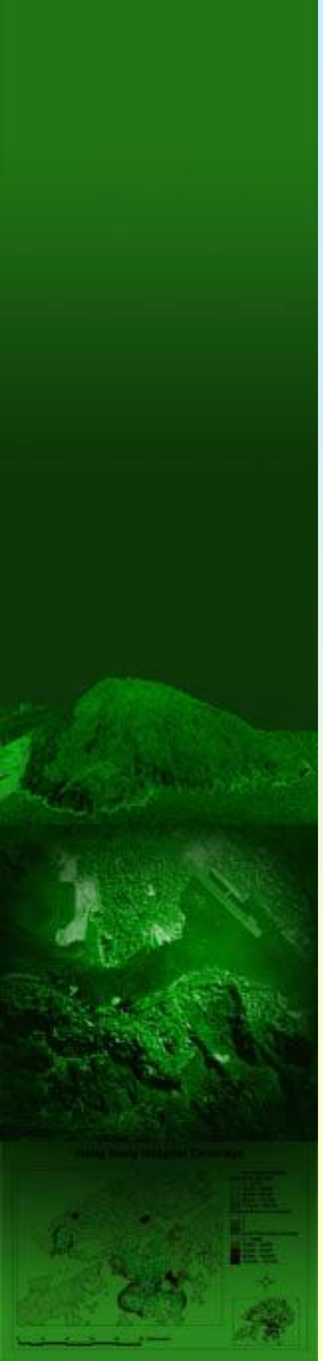
(No slope between Grid Cells)



Continuous surface

(Continuous slope between Grid Cells)

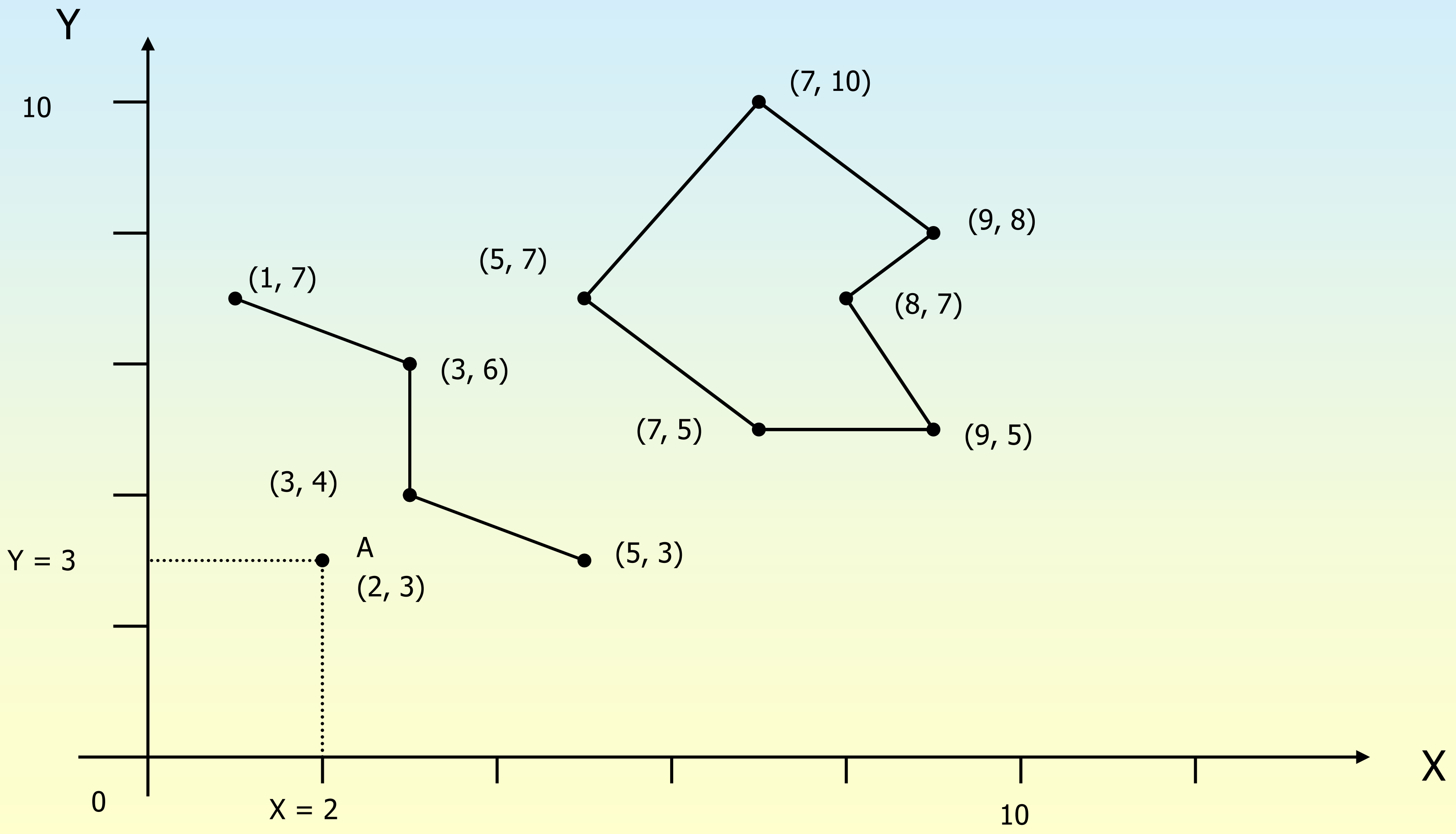




The vector representation

- Based on vectors
- The fundamental primitive is points.
- Objects are created by connecting points with straight lines (or arcs).
- Areas are defined by sets of lines.

The Vector Coordinates

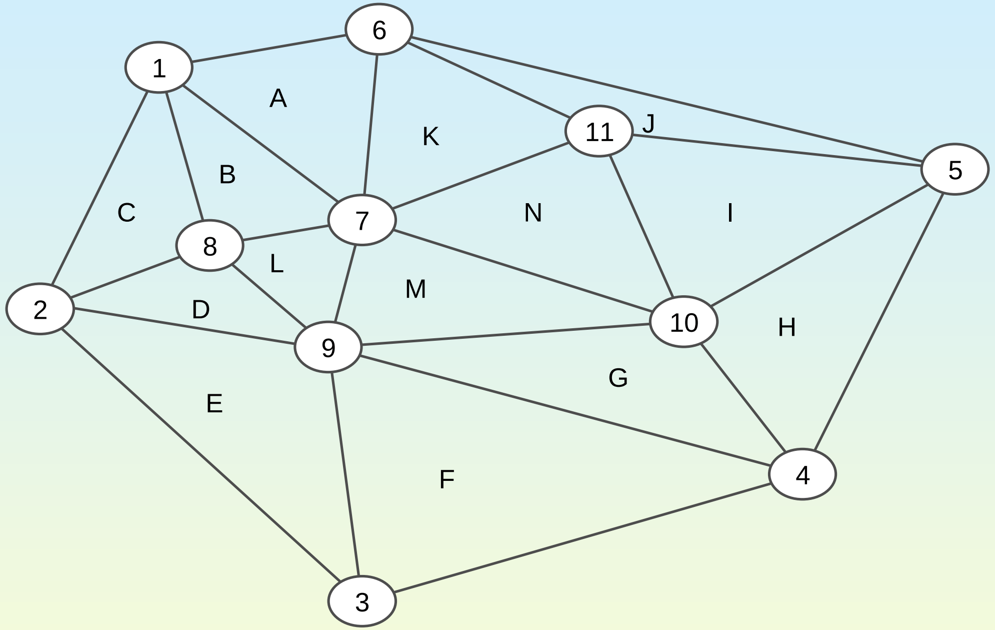


A vector data layer

Triangulated irregular network (TIN)

Triangle Table

Id#	node#	area	slope	… …
A	1, 6, 7	… …	… …	… …
B	1, 7, 8			
C	1, 2, 8			
D	2, 8, 9			
E	2, 3, 9			
F	3, 4, 9			
G	4, 9, 10			
H	4, 5, 10			
I	5, 10, 11			
J	5, 6, 11			
K	6, 7, 11			
L	7, 8, 9			
M	7, 9, 10			
N	7, 10, 11			



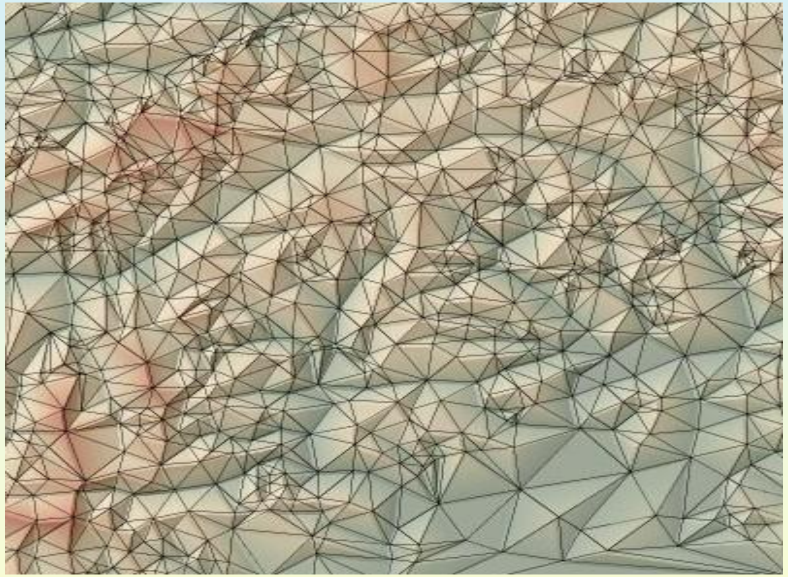
X-Y Coordinates

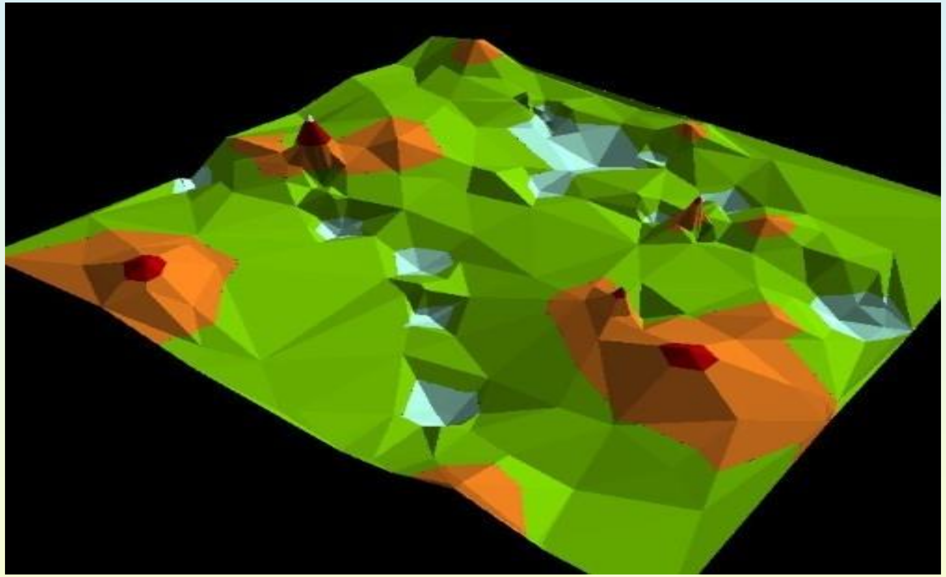
node#	coordinates
1	x1, y1
2	x2, y2
3	x3, y3
. . .	. . .
11	x11, y11

Z Coordinates

node#	z_value
1	z1
2	z2
3	z3
. . .	. . .
11	z11

3-D surfaces represented by TIN

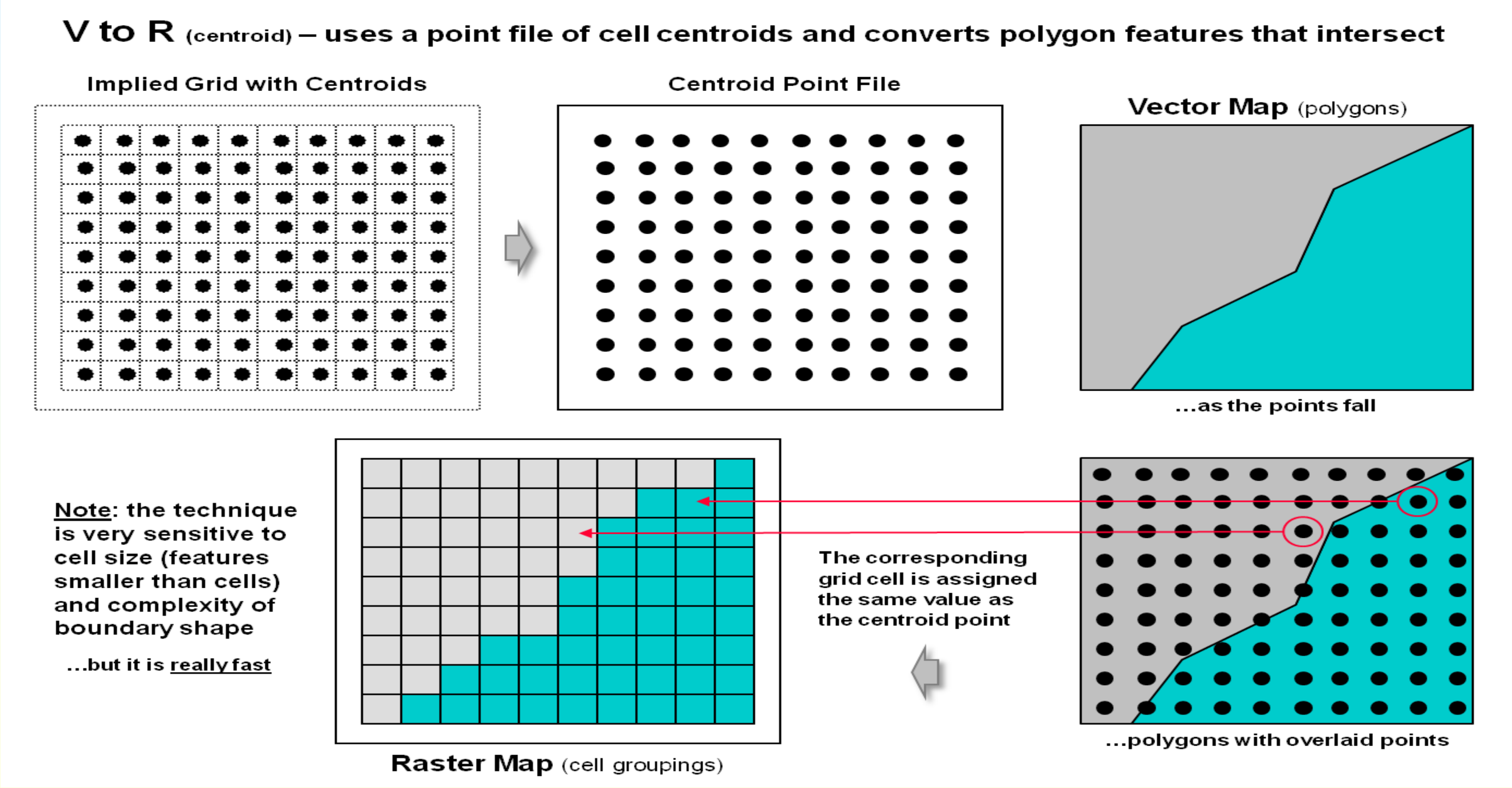




Raster/vector conversion

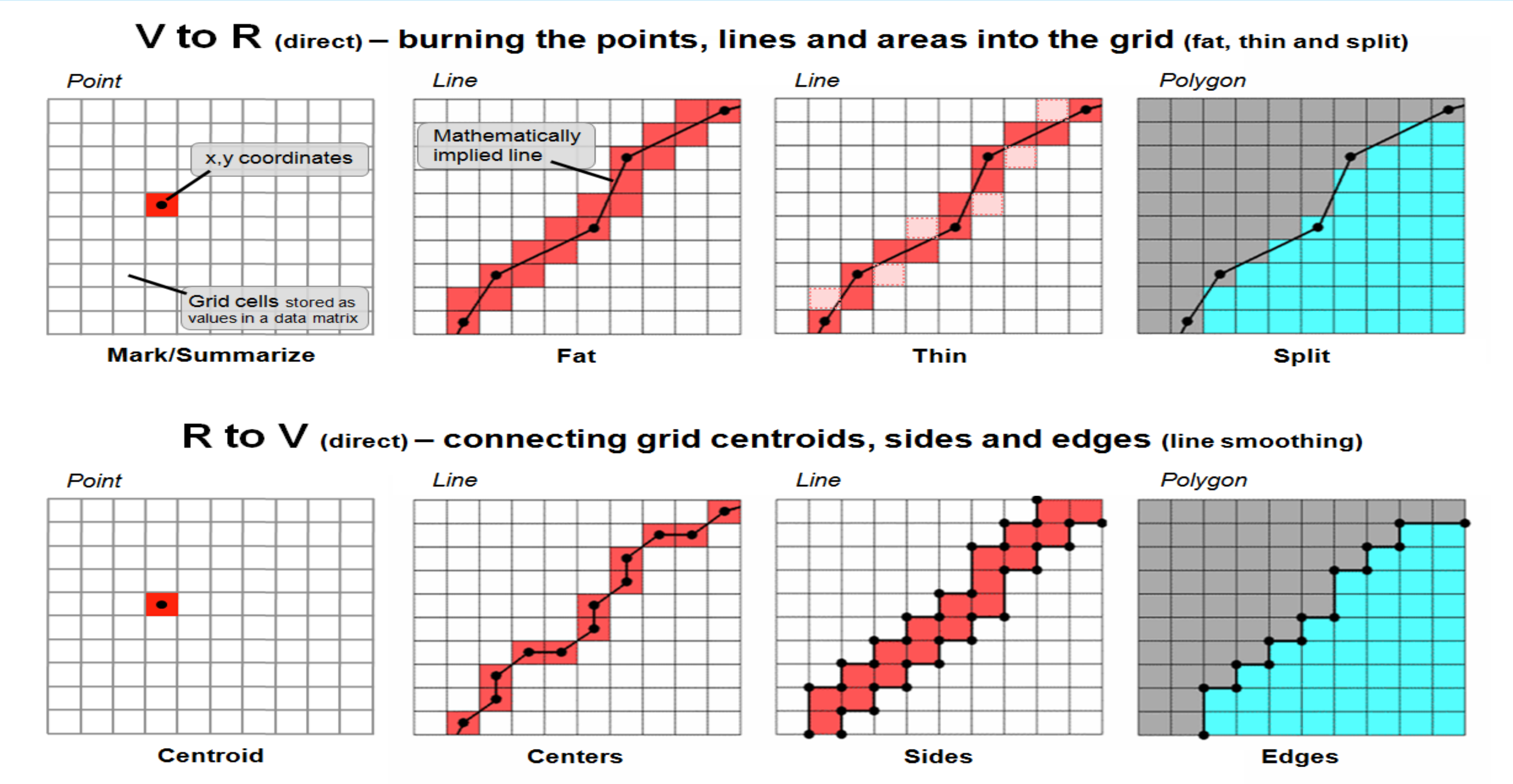
- Raster and vector data can be converted forward and backward
  - Because they represent the same geographical variable
- However, this conversion can be technically complex due to variation in methods and interpretation of the data.

V to R – centroid

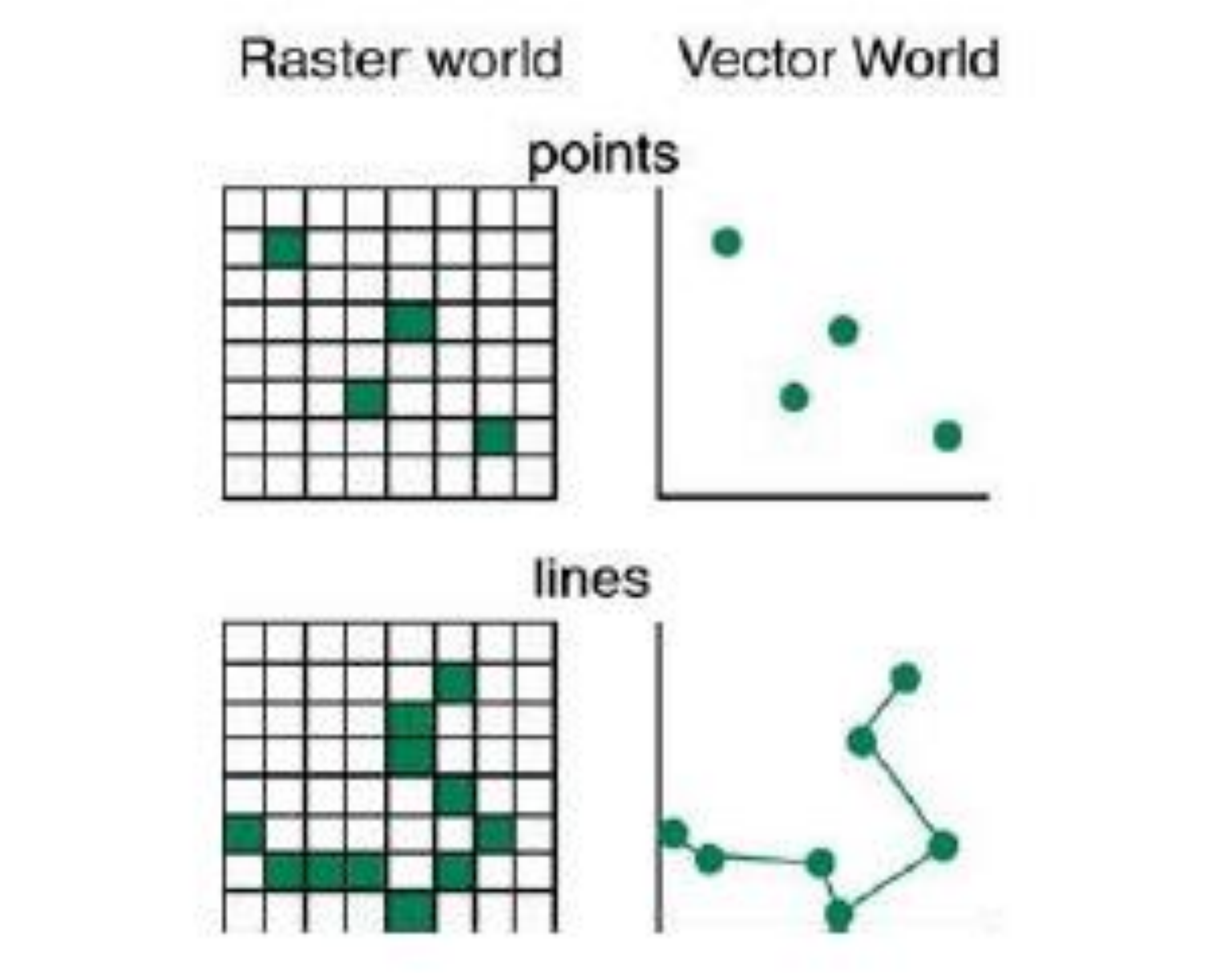


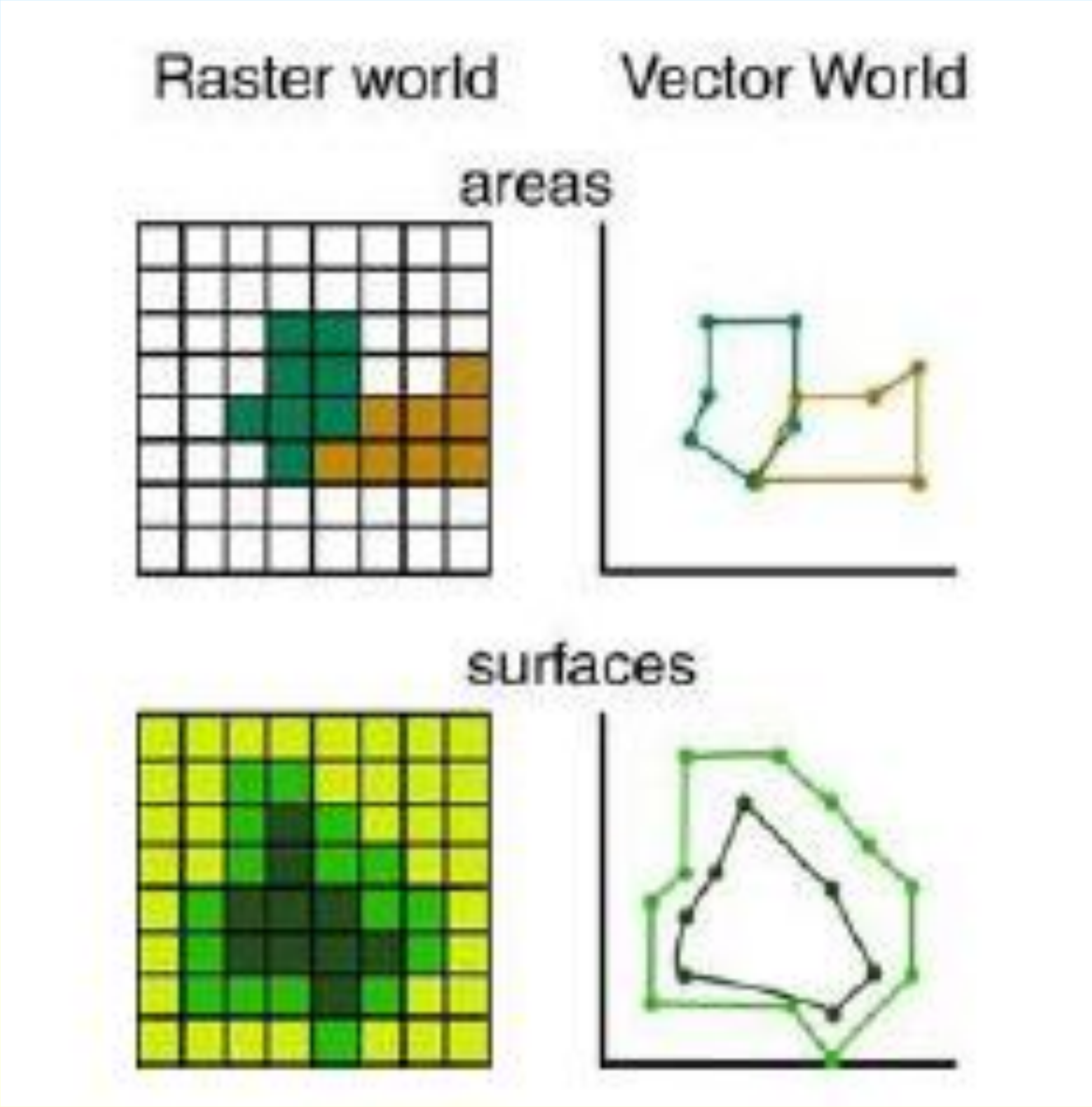
Source: http://www.innovativegis.com/basis/mapanalysis/Topic18/Topic18.htm

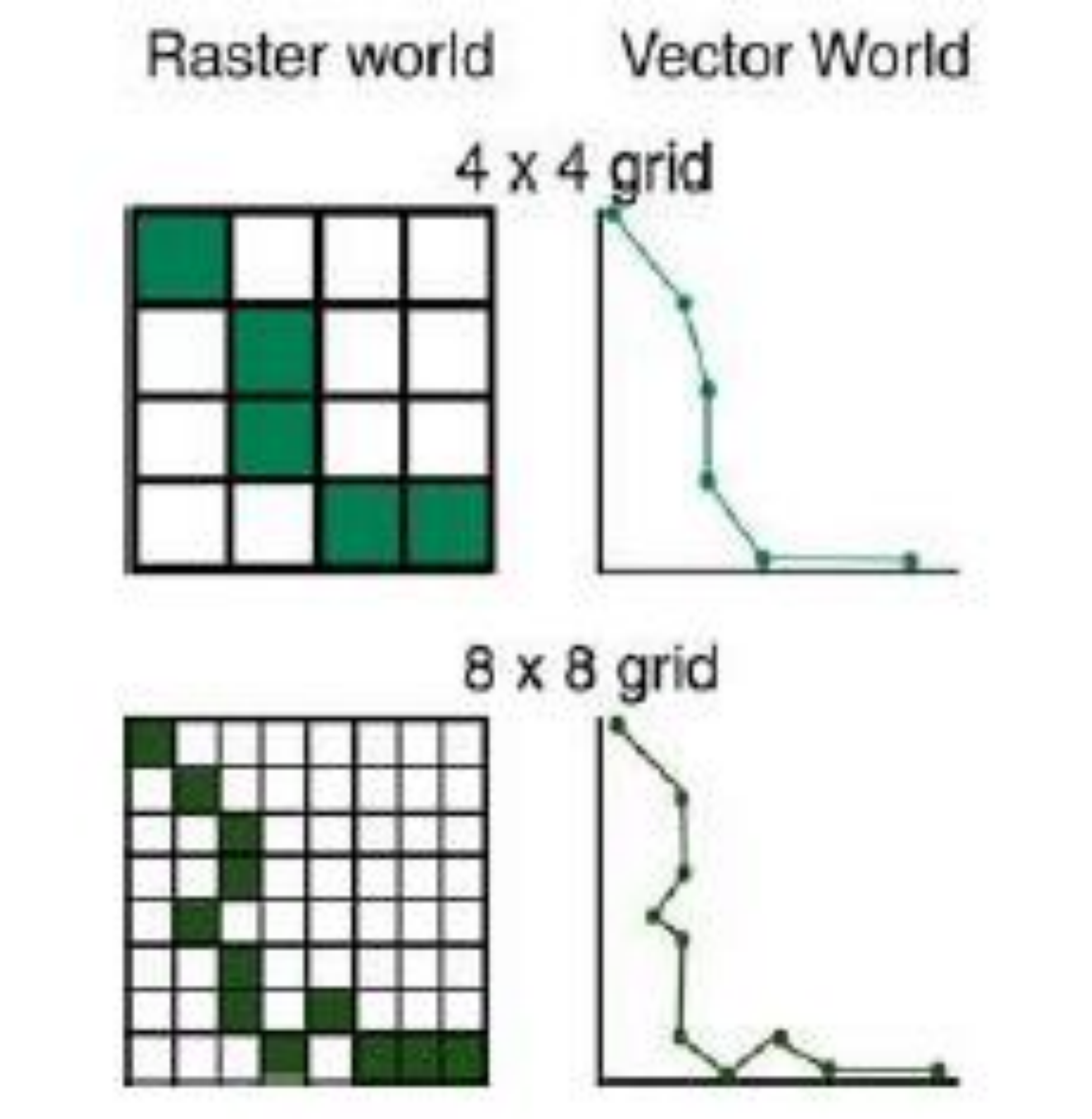
R/V conversion – direct



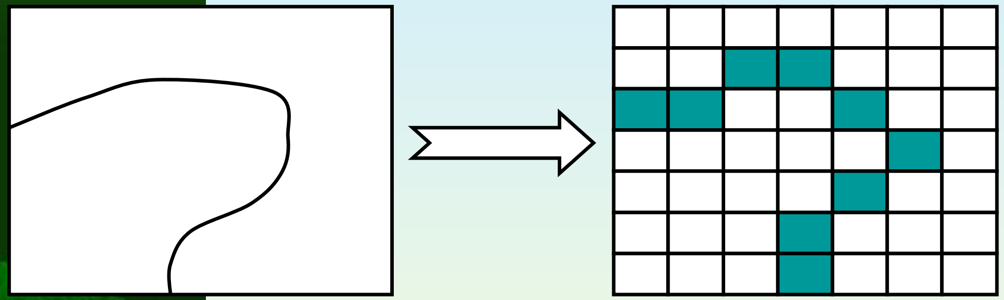
Source: http://www.innovativegis.com/basis/mapanalysis/Topic18/Topic18.htm


Raster world
Vector World
points
lines

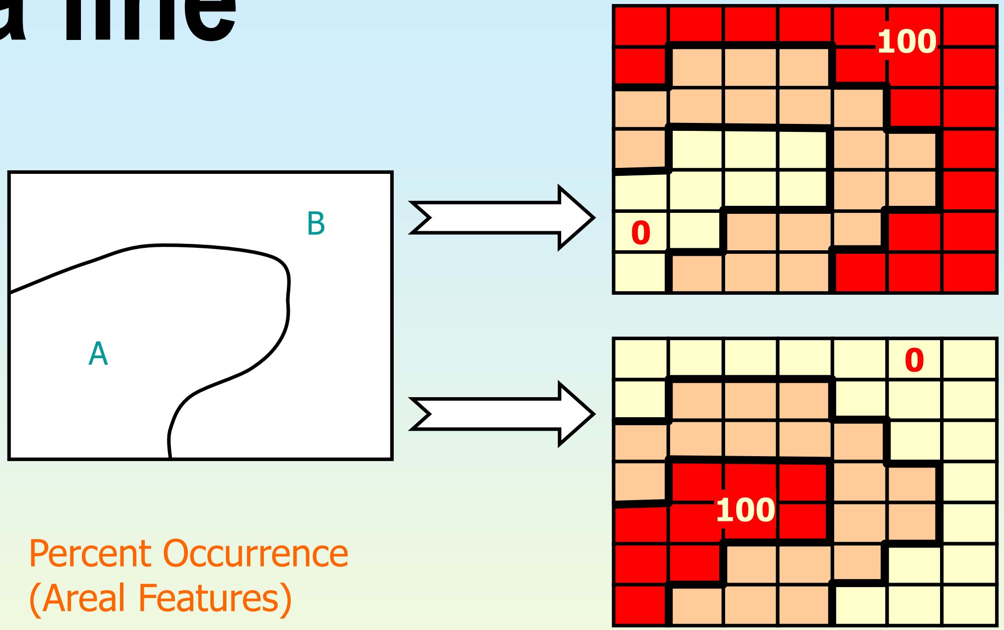

Raster world
Vector World
areas
surfaces


Raster world
Vector World
4 x 4 grid
8 x 8 grid

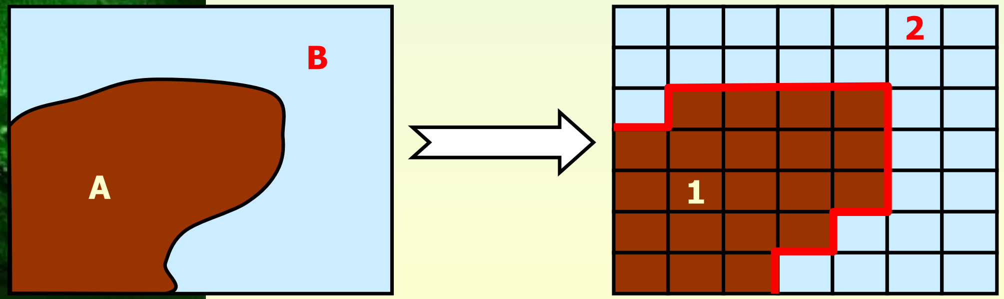
The meaning of a line



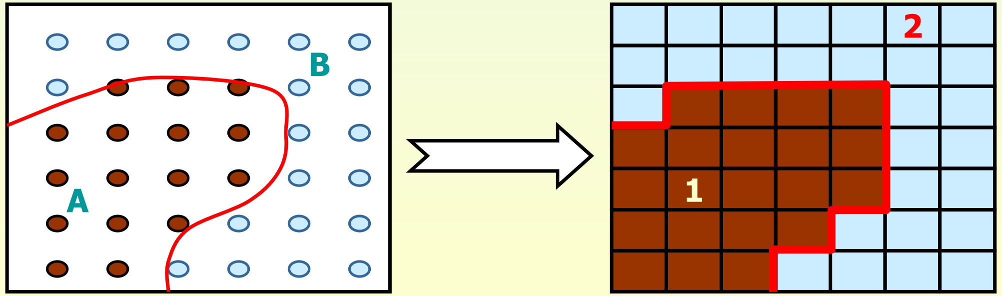
Presence/Absence (Points and Linear Features)



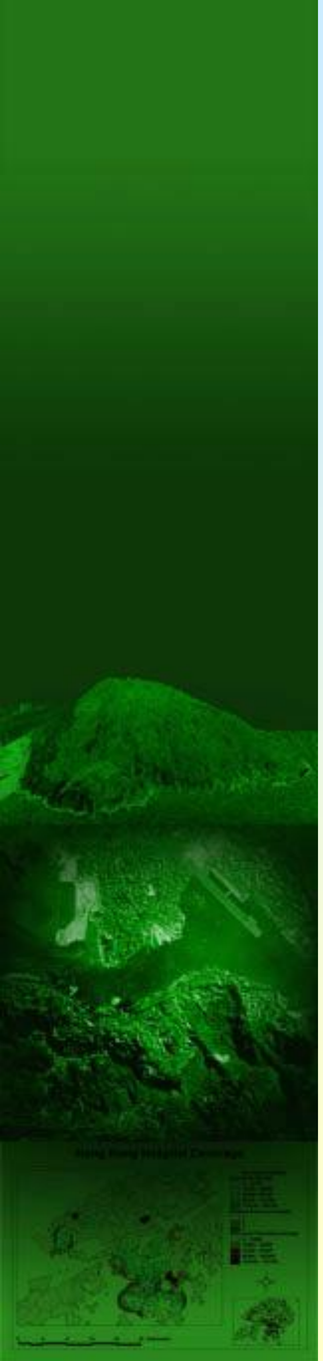
Percent Occurrence
(Areal Features)



Dominant Type (Areal Features)



Centroid of Cell (Areal Features)



Relative advantages

Issue	Raster	Vector
Volume of data	Depends on cell size	Depends on density of vertices
Sources of data	Remote sensing imagery	Social and environmental data
Applications	Resources environmental	Social, economic, administrative
Software	Raster GIS, image processing	Vector GIS, automated cartography
Resolution	Fixed	Variable

Source: Longley et al., 2016, pp68.

Comparison of Raster and Vector Data Models

Raster Model

Advantage:

1. It is a simple data structure.

2. Overlay operations are easily and efficiently implemented.

3. High spatial variability is efficiently represented in raster format.

4. The raster format is more or less required for efficient manipulation and enhancement of digital images.

Vector Model

Advantage:

1. It provides a more compact data structure than the raster model.

2. It provides efficiently encoding of topology and as result more efficiently implementation of operations that require topological information, such as network analysis.

3. The vector model is better suited to supporting graphics that closely approximate Hand-drawn maps.

Comparison of Raster and Vector Data Models

Raster Model

Disadvantage:

1. It is less compact therefore data compression techniques can often overcome this problem.
2. Topological relationships are more difficult to represent.
3. The output of graphics is less aesthetically pleasing because boundaries tend to have a blocky appearance rather than the smooth lines of hand-drawn maps.

Vector Model

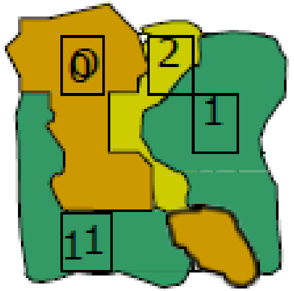
Disadvantage:

1. It is a more complex data structure.
2. Overlay operations are more difficult to Implement.
3. The representation of high spatial variability is inefficient.
4. Manipulation and enhancement of digital images cannot be effectively done in vector domain.

Layer Types

A layer type refers to the way spatial and attribute information are connected. There are two major layer types, vector and raster.

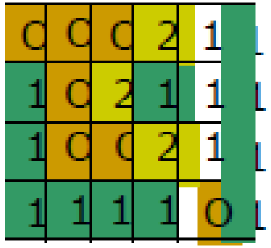
Vector: Points, lines and polygons (spatial data) associated with databases of attributes (attribute data) are considered vector layer types.



Attributes of Theme2.shp

Shape	ID	LANDUSE
Polygon	0	WATER
Polygon	1	HIGHLAND
Polygon	2	WETLAND

Raster: A row and column matrix (pixels) of X & Y space with attribute information associated with each pixel is considered a raster layer type.



1 : WATER
2 : HIGHLAND
3 : WETLAND

GIS and maps

- Paper maps is a source of data for geographical databases.
- A paper map is an analog product from GIS.
- A paper map is an effective communication tool.
  - portable
  - easy to use
  - cheap to make mass production

Summary

- All geographical phenomena need to be represented in some way.
- There are two fundamental representation in geography – objects and fields.
- Geographical entities can be abstracted as points, lines, polygons, volumes and processes.
- Two approaches for representing the geographical entities are vector and raster representations.