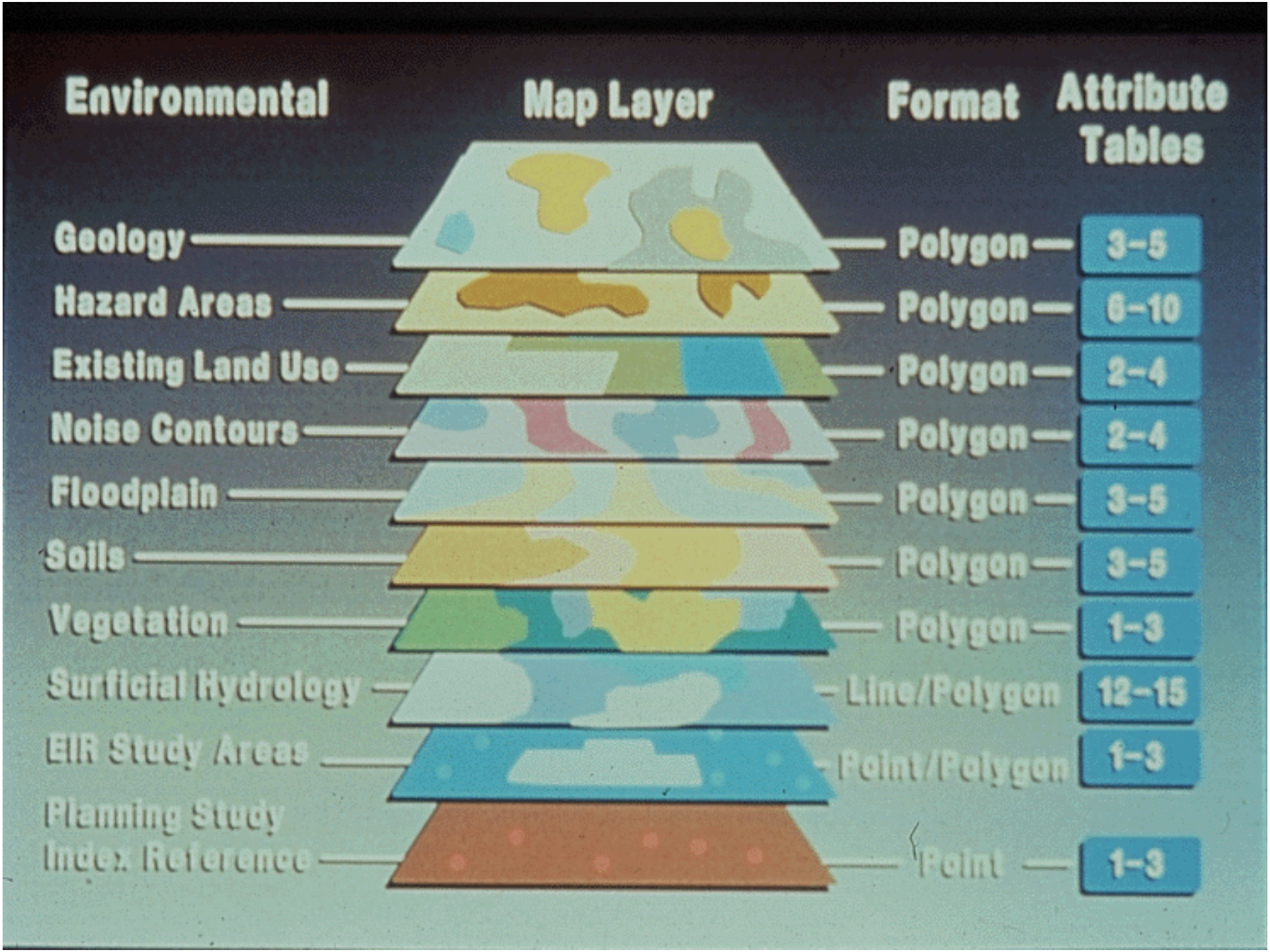
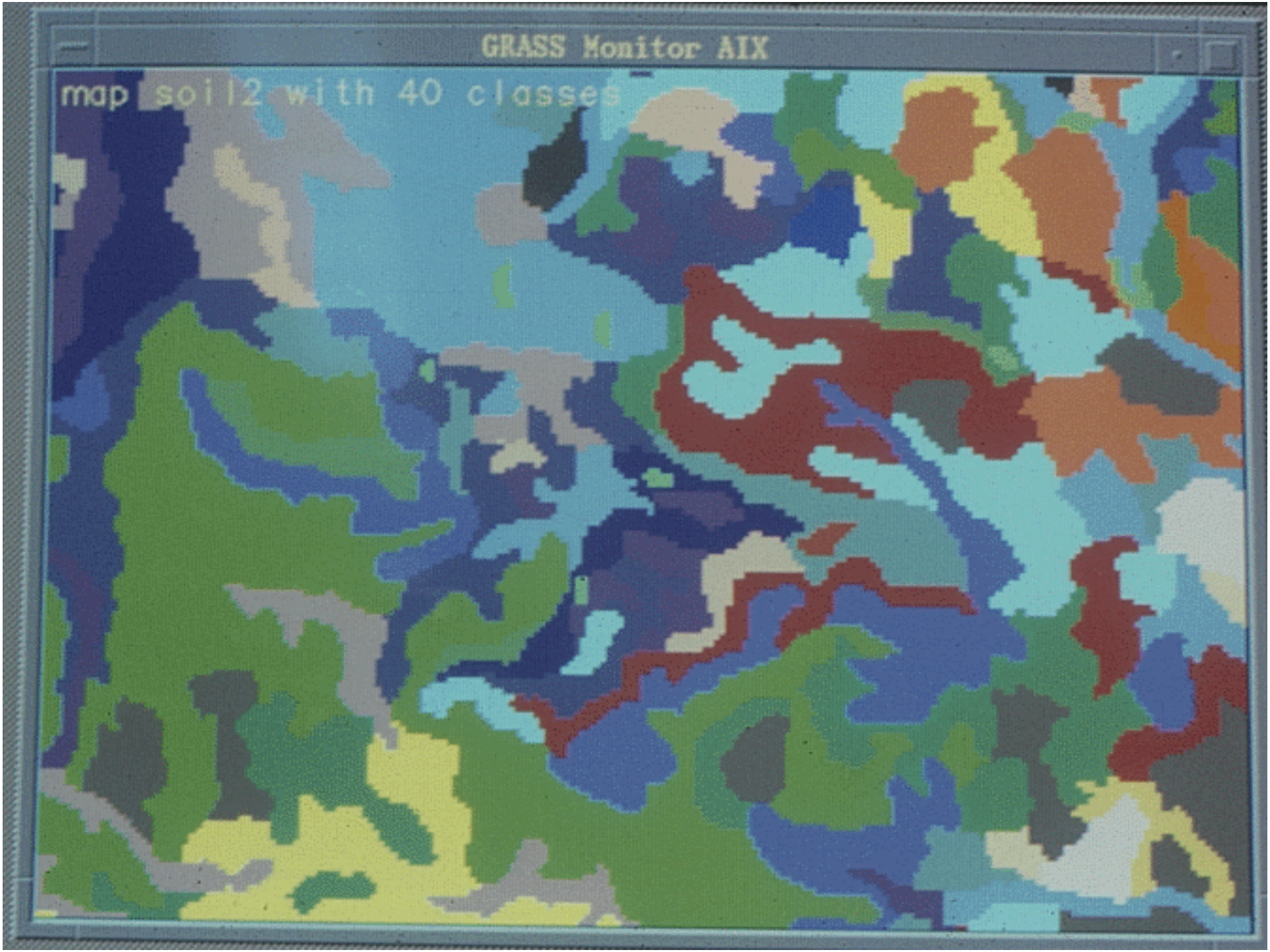
Towards a General Theory of Geographic Representation in GIS

Michael F. Goodchild
University of California
Santa Barbara

The Canada Geographic Information System

- Maps inventorying the Canadian land resource
  - order 10^6 sheets
- Measurement of area
  - reporting statistics on existing and potential land use
- $20 million in 1965 Canadian dollars


Environmental
Map Layer
Format
Attribute Tables
Geology
Hazard Areas
Existing Land Use
Noise Contours
Floodplain
Soils
Vegetation
Surficial Hydrology
EIR Study Areas
Planning Study Index Reference
Polygon
Polygon
Polygon
Polygon
Polygon
Polygon
Polygon
Line/Polygon
Point/Polygon
Point
3–5
6–10
2–4
2–4
3–5
3–5
1–3
12–15
1–3
1–3

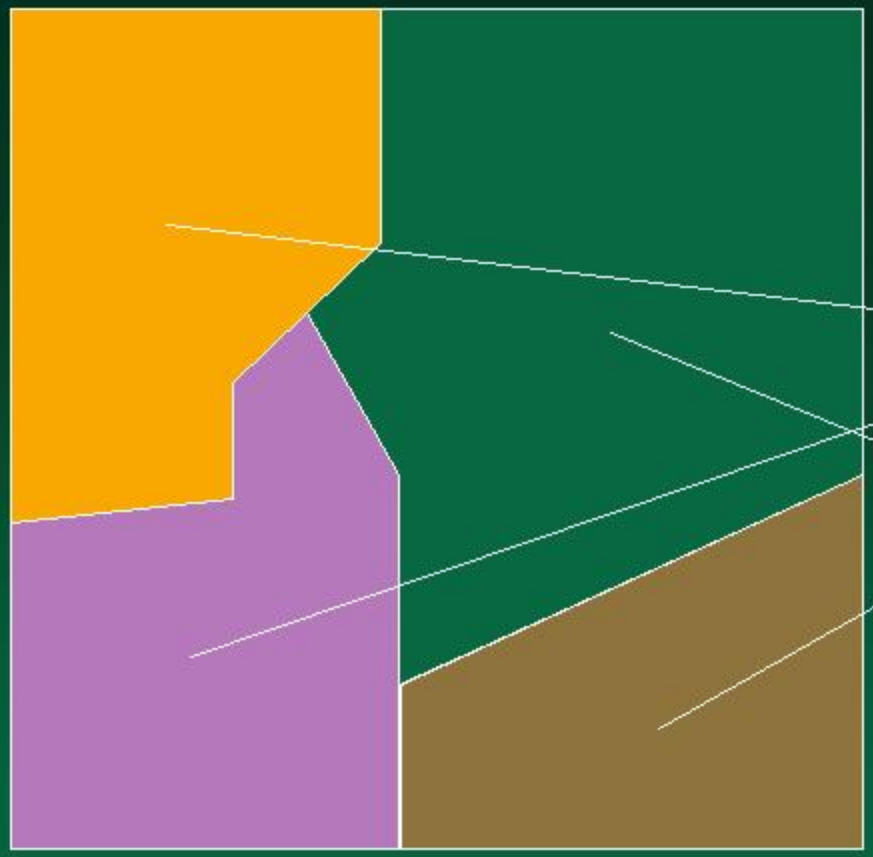

GRASS Monitor AIX
map soil2 with 40 classes

Flat-file options (tape)

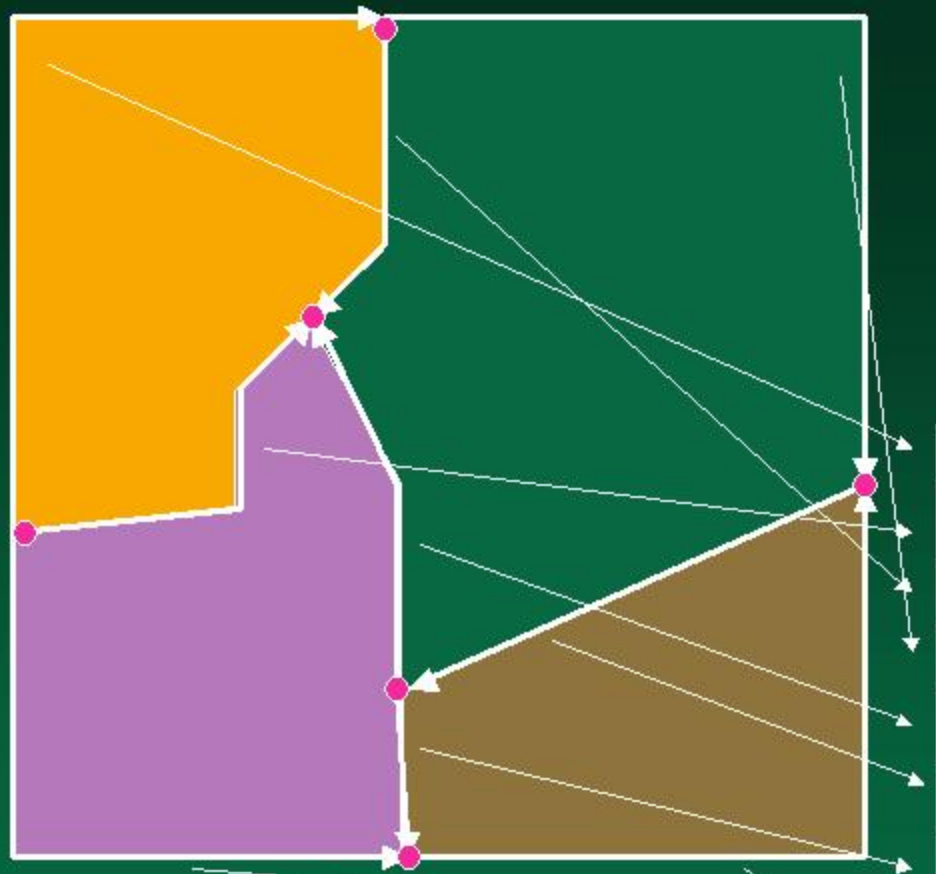
- By face/polygon
  - double recording of internal boundaries
  - spurious differences
- By edge/arc
  - half the data volume
  - compute area in O(vertices)
  - simplify overlay
  - attributes of adjacent polygons
  - no polygon records

The relational model

- The map as a collection of arcs, nodes, and faces
  - F-A+N = 2
- Stored in tables with keys
- GIS built on RDBMS
  - INFO
- Vertices left out
  - a hybrid solution

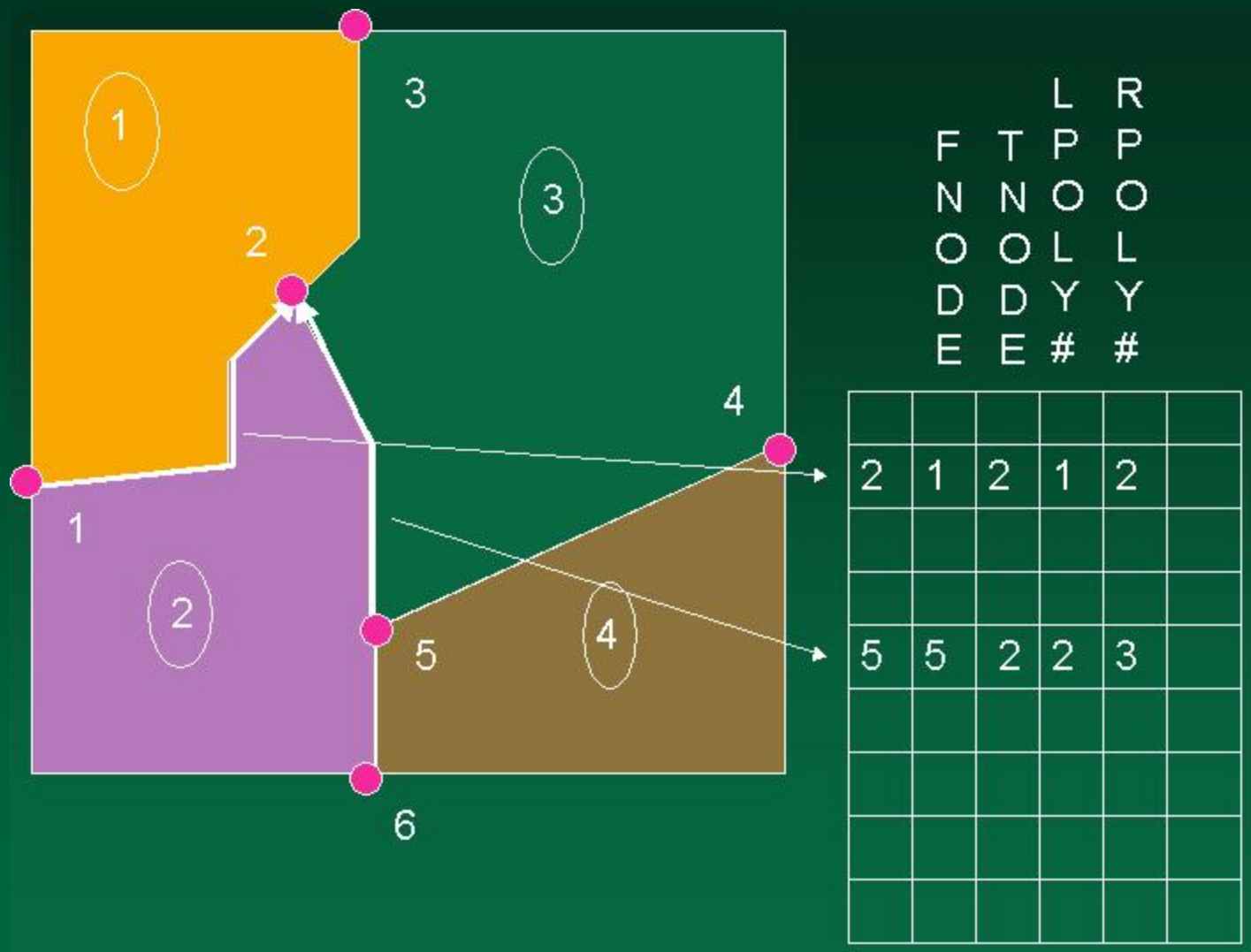


A	874	7	
B	95	8	
C	199	8	
D	767	5	



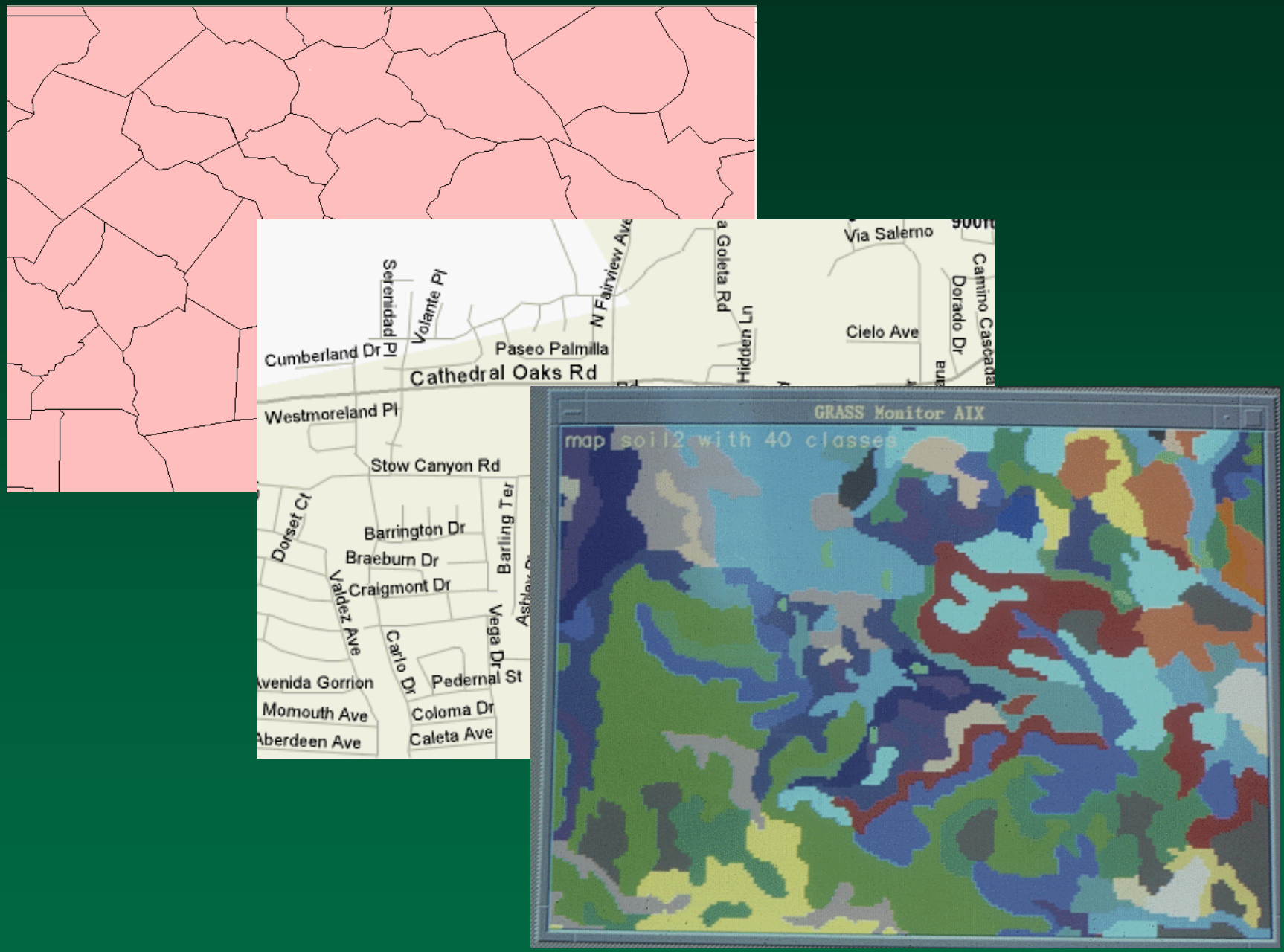
A	874		
B	95		
C	199		
D	767		

1					
2					
3					
4					
5					
6					
7					
8					
9					


1
3
2
3
4
1
2
5
4
6
FNODE
TNODE
LPOLY#
RPOLY#
2 1 2 1 2
5 5 2 2 3

Other types of maps

- Transportation links
  - linear features
  - networks
  - U.S. Bureau of the Census
  - blocks = 2-cells
  - street segments = 1-cells
  - intersections = 0-cells


Via Salerno
Camino Cascada
Dorado Dr
Cielo Ave
Hidden Ln
a Goleta Rd
N Fairview Ave
Volante Pl
Serenidad Pl
Cumberland Dr
Paseo Palmilla
Cathedral Oaks Rd
Westmoreland Pl
Stow Canyon Rd
Dorset Ct
Barrington Dr
Braeburn Dr
Craigmont Dr
Barling Ter
Valdez Ave
Carlo Dr
Vega Dr
Pedernal St
Avenida Gorrion
Momouth Ave
Aberdeen Ave
Coloma Dr
Caleta Ave
GRASS Monitor AIX
map soil2 with 40 classes

Square pegs in round holes

- Cul-de-sacs
  - allow 1-nodes
- Properties of parts of edges
  - dynamic segmentation
  - linear referencing
- Non-planarity
  - overpasses and underpasses
  - turntables

A 1990s house of cards

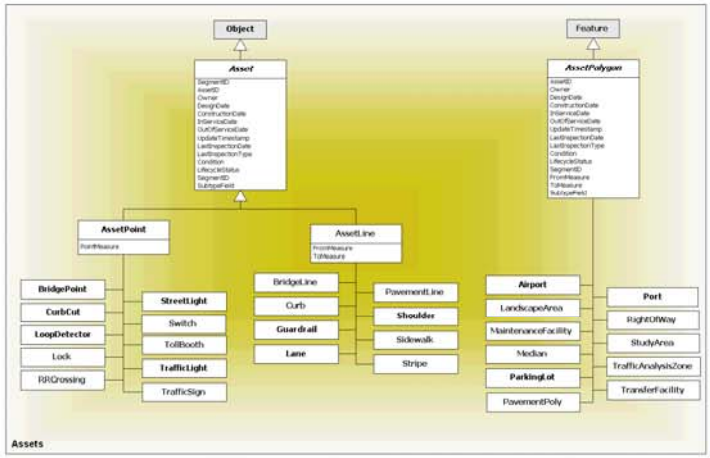
- Still no vertices in the RDBMS
- Points
  - coordinates stored in tables
  - no topological relationships with other features
- Does it have to be this hard?
  - simple CAD data model
  - points, lines, and areas in an empty space
  - potentially overlapping
  - no topological relationships
  - compute on the fly

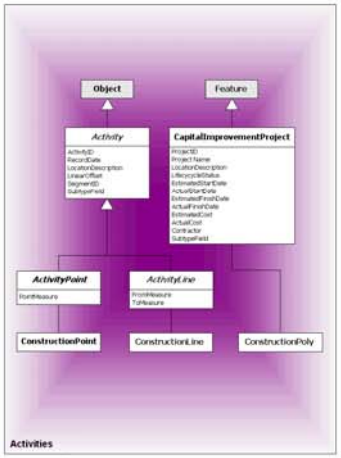
Object-oriented data modeling

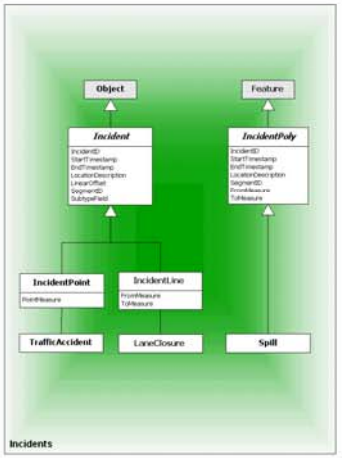
- All features are instances of classes
- Classes inherit properties from more general classes
- Features can be aggregates of other features
- Features can be composed of other features
- Features can be associated

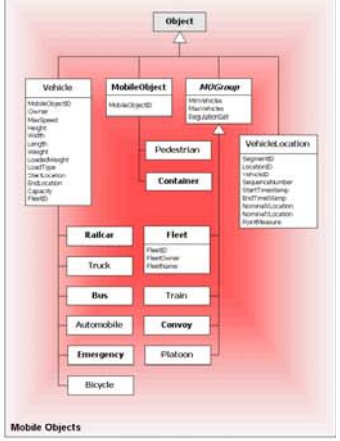
ArcGIS™ Transportation Data Model (Draft)

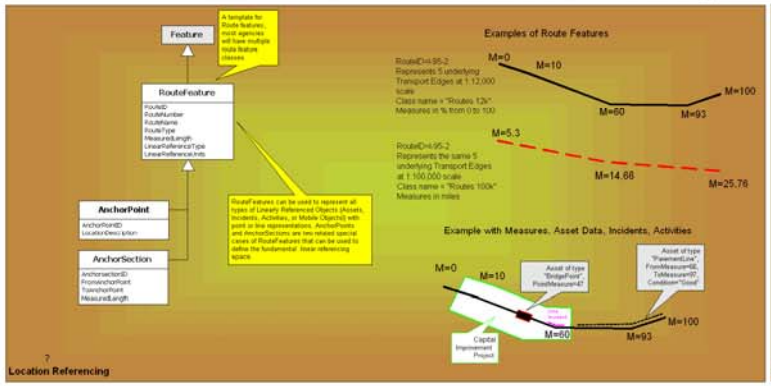
May 11, 2001

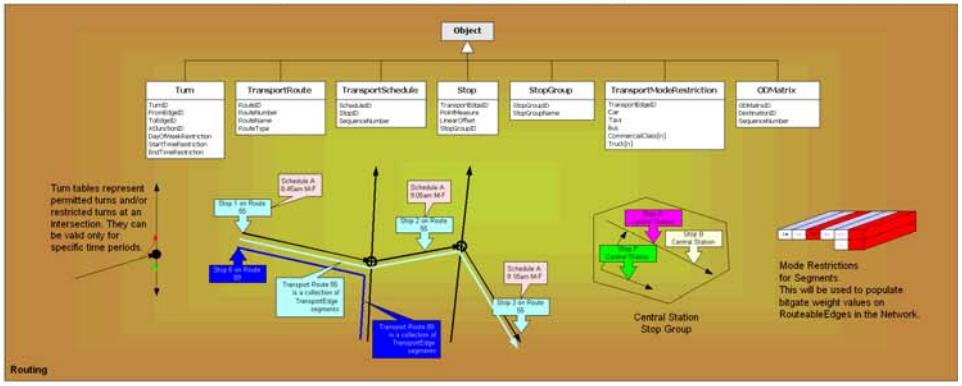


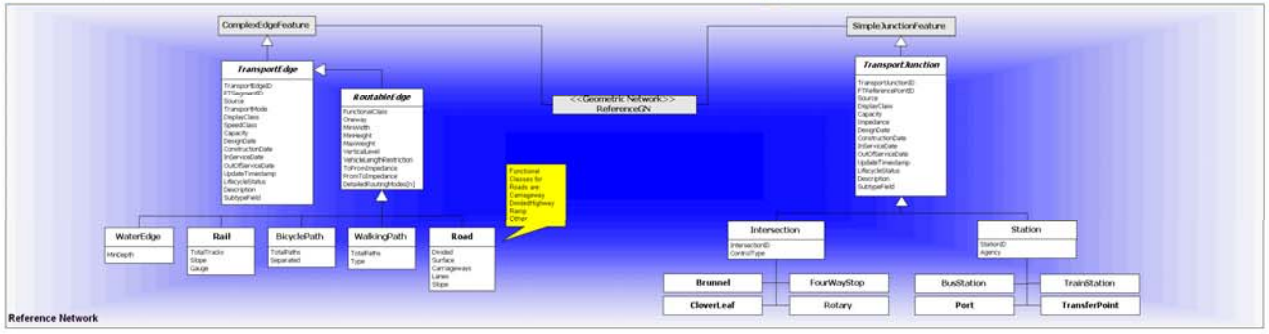


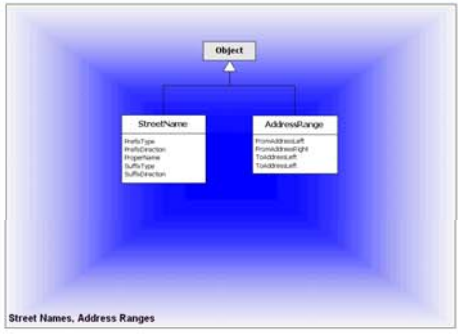


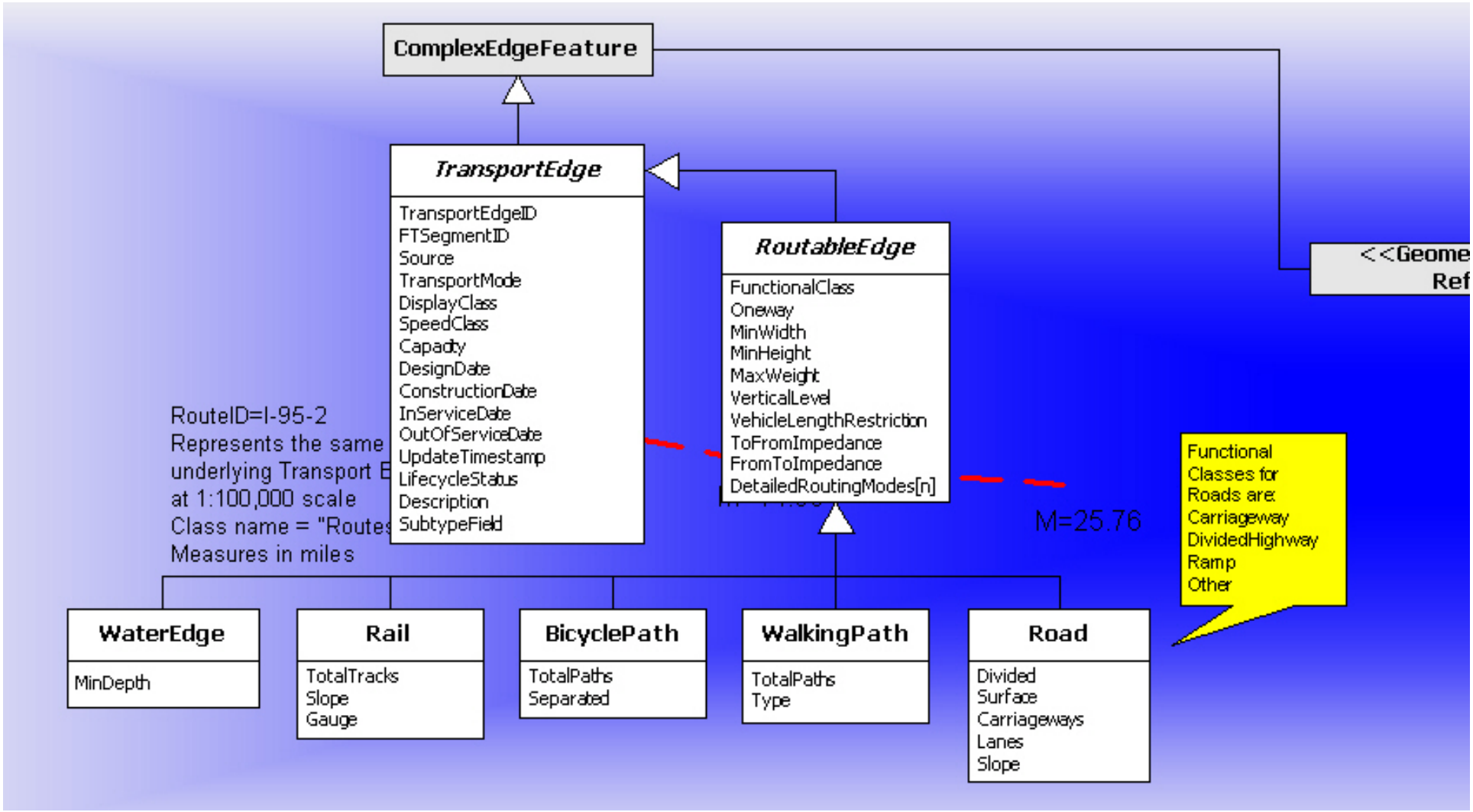




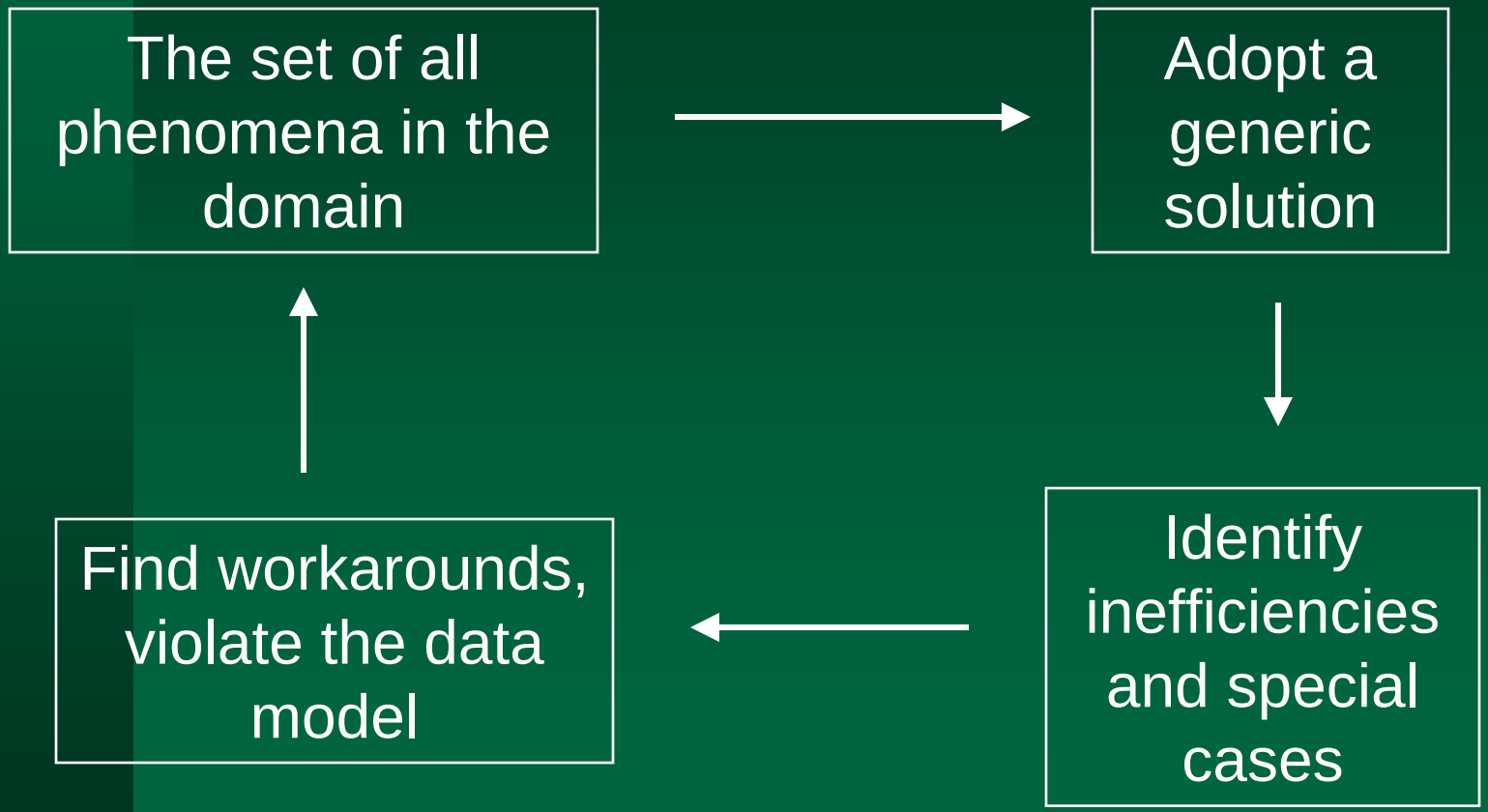







ComplexEdgeFeature
TransportEdge
TransportEdgeID
FTSegmentID
Source
TransportMode
DisplayClass
SpeedClass
Capacity
DesignDate
ConstructionDate
InServiceDate
OutOfServiceDate
UpdateTimestamp
LifecycleStatus
Description
SubtypeField
RoutableEdge
FunctionalClass
Oneway
MinWidth
MinHeight
MaxWeight
VerticalLevel
VehicleLengthRestriction
ToFromImpedance
FromToImpedance
DetailedRoutingModes[n]
<<Geome
Ref
RouteID=I-95-2
Represents the same
underlying Transport E
at 1:100,000 scale
Class name = "Routes
Measures in miles
M=25.76
Functional
Classes for
Roads are:
Carriageway
DividedHighway
Ramp
Other
WaterEdge
MinDepth
Rail
TotalTracks
Slope
Gauge
BicyclePath
TotalPaths
Separated
WalkingPath
TotalPaths
Type
Road
Divided
Surface
Carriageways
Lanes
Slope

The data modeling cycle



Is the process beginning again?

- All features are instances of classes
  - are all phenomena naturally features?
  - is there a pre-feature stage?
- Inherently continuous phenomena
  - roads, rivers
  - topography
  - the pre-patch ecological landscape

Current stresses

- Fields
  - and other continuous phenomena
  - not naturally partitioned into discrete features
  - must be discretized
  - results unsatisfactory
- Other complex, dynamic phenomena
  - moving objects
  - changing shapes
  - dynamic internal structure

An increasing confusion

- Rasters and vectors (1960s)
- Topological data structures (1977)
- Relational databases (1980)
- Object-oriented databases (1990)
- Fields and objects (1990)
- Time and the third spatial dimension
- Object fields, metamaps, ...
- Why does it have to be so complex?

Desiderata of a general theory (Galton, 2003)

- To provide suitable forms of representation and manipulation to do justice to the rich network of interconnections between field-based and object-based views of the world
- To extend the field-based and object-based views, and the forms of representation developed to handle them, into the temporal domain
- To provide a means to develop different views of spatio-temporal extents and the phenomena that inhabit them, especially with reference to those phenomena which seem to represent dual aspects as both object-like and field-like

Major questions

- Can the complexity be reduced?
  - can some of these options be integrated under a single umbrella?
- Can objects and fields be integrated?
  - and are they the only options?

Relevant literature

- Berry's geographical matrix
  - Sinton's three-dimensional schema
  - control one dimension, hold a second constant, observe variation in the third
- Li, Ning, and Sun (1999)
  - UML framework
  - schemata for objects and fields
  - merged into a single schema
  - Parameterized Geographic Object Model

Geographic information

- Associates places with properties
- Reducible to an atomic form
  - <**x**,*Z*,*z*(**x**)>
  - **x** is a location in space–time
  - *Z* is a property
  - *z*(**x**) is the value of that property at **x**
  - the geo-atom
  - point properties are sometimes measured and managed
  - other structures are aggregations of point sets

Geo-fields

- Defined over a domain *D*
- Aggregates all geo-atoms within the domain for some property *Z*
- Scalar
  - a single property
  - nominal, ordinal, interval, ratio
- Vector
  - directional
  - one property per dimension of space–time
- Aliases
  - surface, coverage

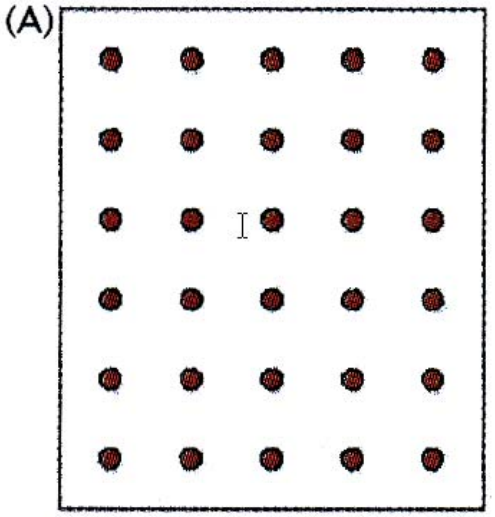
Representation of geo-fields

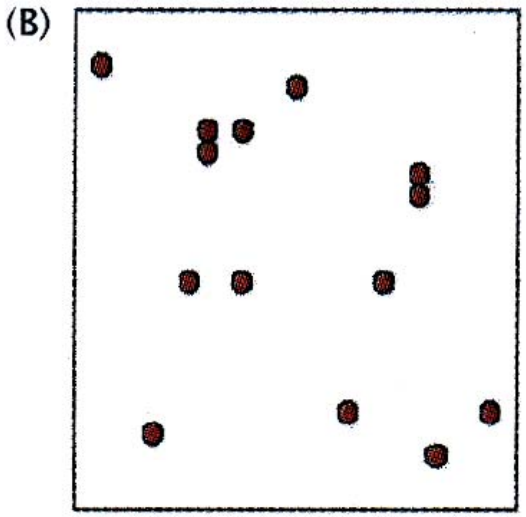
- Domain and variation within domain
- Six familiar methods
  - discretization is a required and inherent property of the representation of any geo-field

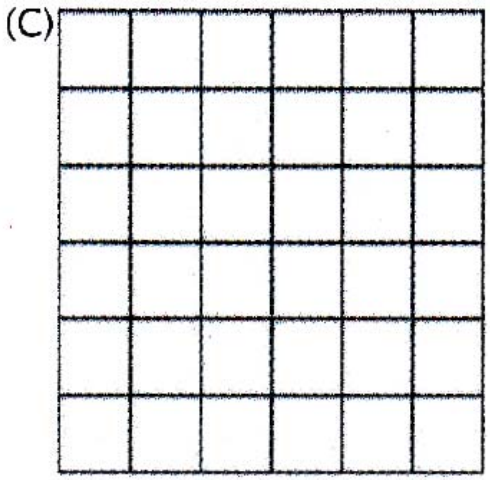
raster

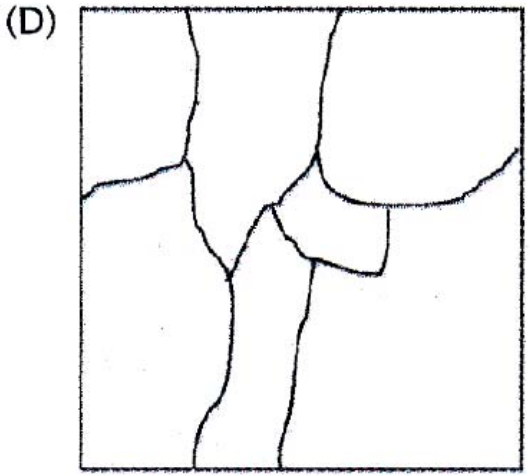
points

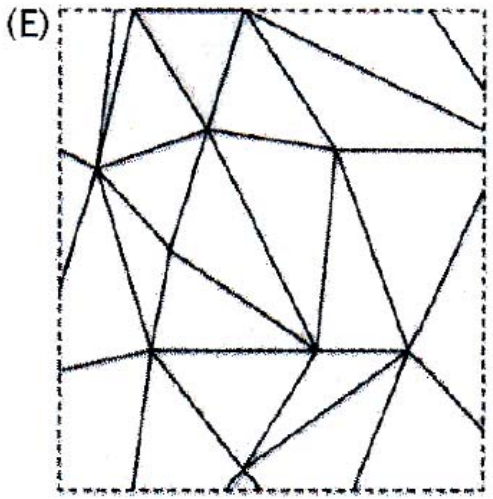
raster

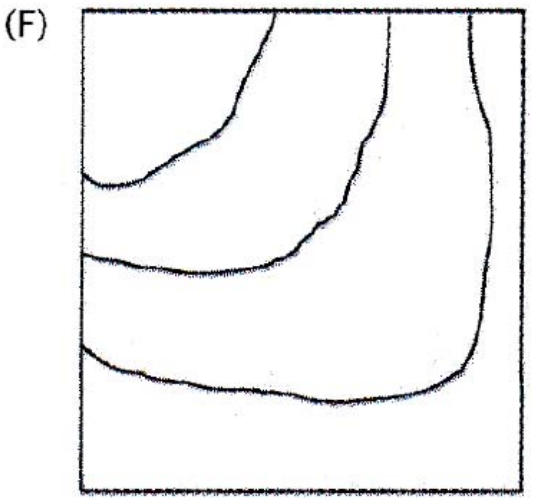












polygons

polygons

polylines

Measurement of geo-field properties

- Point measurements
  - weather data
- Convolutions
  - population density
  - width of kernel

Geo-objects

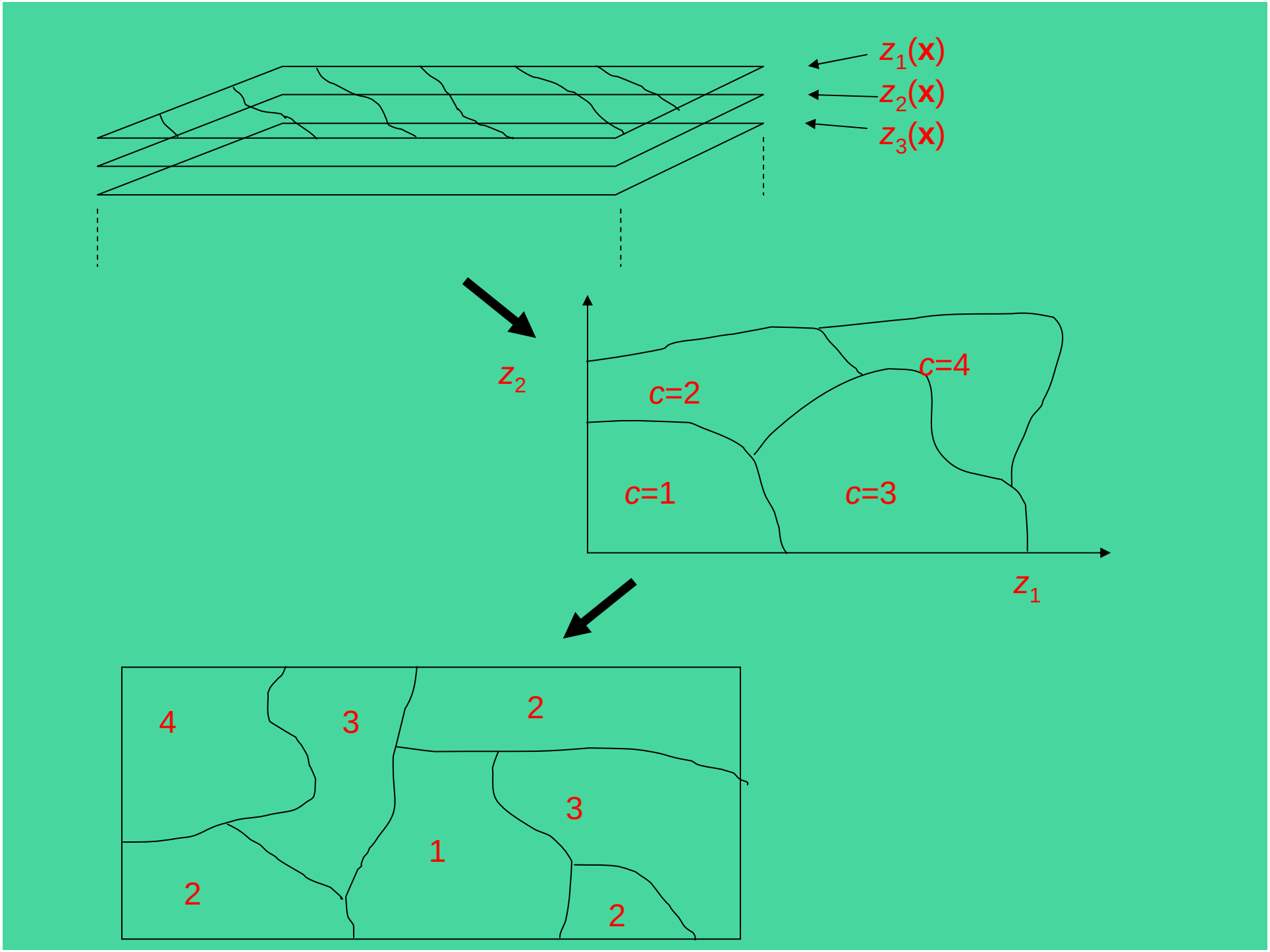
- Aggregations of geo-atoms
  - based on certain rules applied to values
    - all geo-atoms with elevation > 100
    - all geo-atoms with elevation = 100
    - all geo-atoms with county = “jefferson”
  - aliases
    - entity, feature
- Tobler’s First Law
  - ensures geo-objects of non-trivial size
  - multipart polygons
  - multipart polylines

Criteria for geo-objects

- Homogeneity
  - *bona fide* geo-objects
- Complementarity
  - functional regions
- Fiat
  - administrative decisions
- Fuzzy membership
  - geo-objects with indeterminate boundaries

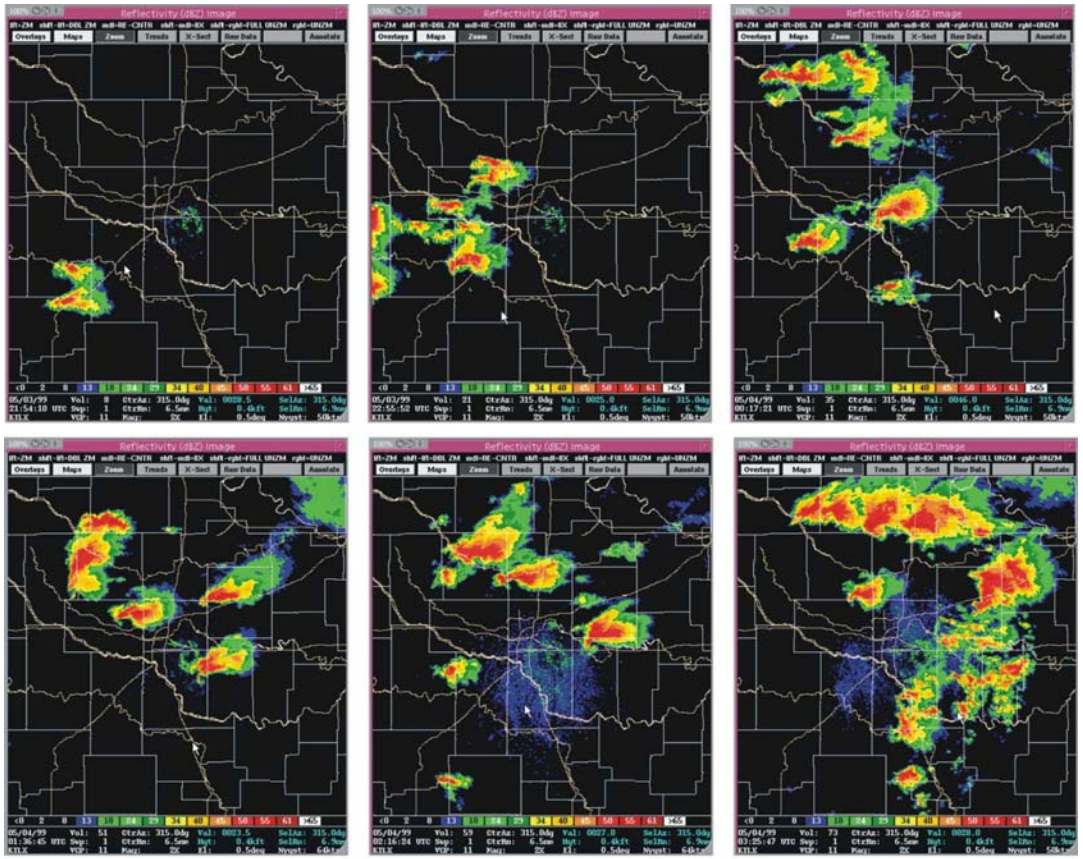
A general theory of bona-fide geo-objects

- *m* dimensional "phase" space defined by field variables
  - partition into *n* regions
- *m* fields locate **x** in phase space
- Assign **x** to one of *n* classes
  - compare classifiers


$z_1(\mathbf{x})$
$z_2(\mathbf{x})$
$z_3(\mathbf{x})$
z_2
c=2
c=4
c=1
c=3
z_1
4
3
2
3
1
2
2

Field objects (Yuan, 1999)

- A geo-object with internal variation
  - characterized by a field
- A geo-field whose domain *D* is *bona fide*
- A moving object that carries its geo-field with it
  - compare moving objects in video games
  - physics may be better understood in a moving frame



Representative radar images to show the evolution of supercell storms that produced F5 tornadoes in Oklahoma City, May 3, 1999. WSR-88D radar TKLX scanned the supercells every 5 minutes, but the images shown here were selected at approximately every two hours.

May Yuan, University of Oklahoma

Hurricane Frances


30.jpg 35.jpg 40.jpg 45.jpg 50.jpg 55.jpg


60.jpg 65.jpg 70.jpg 75.jpg 80.jpg 85.jpg


90.jpg 95.jpg 100.jpg 105.jpg 110.jpg 115.jpg


120.jpg 125.jpg 130.jpg 135.jpg 140.jpg 145.jpg


150.jpg 155.jpg 160.jpg 165.jpg 170.jpg 175.jpg


180.jpg 185.jpg 190.jpg 195.jpg 200.jpg 205.jpg


210.jpg 215.jpg 220.jpg 225.jpg 230.jpg

Tables and classes

- Groups of similar geo-objects
  - same topological dimension
  - same set of attributes
- Relational and object-oriented models
- Two steps removed from geo-atoms
  - aggregated to geo-objects, aggregated to tables and classes
  - aggregated to geo-fields and then discretized

Geo-dipoles

- Properties associated with pairs of points
  - $\langle \mathbf{x}_1, \mathbf{x}_2, Z, z(\mathbf{x}_1, \mathbf{x}_2) \rangle$
- Interaction properties
  - distance, interaction, flow, direction, ...

Visible areas

- *Z*: “is visible from”
  - $z(\mathbf{x}_1,\mathbf{x}_2) = 1$ if $\mathbf{x}_2$ is visible from $\mathbf{x}_1$ else 0
  - aggregate all such geo-atoms to form a geo-object
    - the “area visible from” $\mathbf{x}_1$
    - $O(\mathbf{x}_1)$
- Object field (Cova and Goodchild, 2002)
  - a mapping from location to an object
  - an object for every point
  - can have indeterminate boundary
    - if *Z* denotes membership

Metamaps (Takeyama and Couclelis, 1997)

- A raster of cells
  - $\{O_i, i = 1,n\}$
  - a pair of such cells $\{O_i,O_j\}$
  - a metamap is the set of such pairs for all i,j and their properties z_{ij}
  - the set of all interactions
- Each $<O_i,O_j,z_{ij}>$ is an aggregation of geo-dipoles

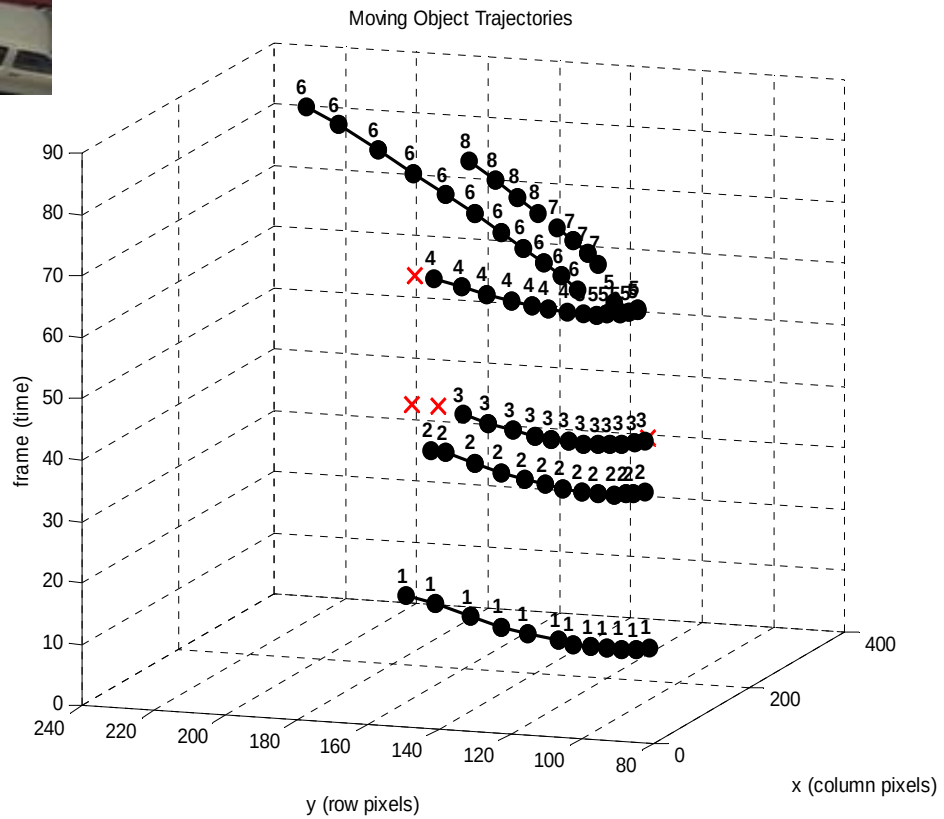
Association classes

- The object pair (Goodchild, 1991)
  - a pair of geo-objects having properties that exist only for the pair
    - distance, interaction, direction, flow...
  - distance matrices, turntables, the **W** matrix
- An association class
- An aggregation of geo-dipoles

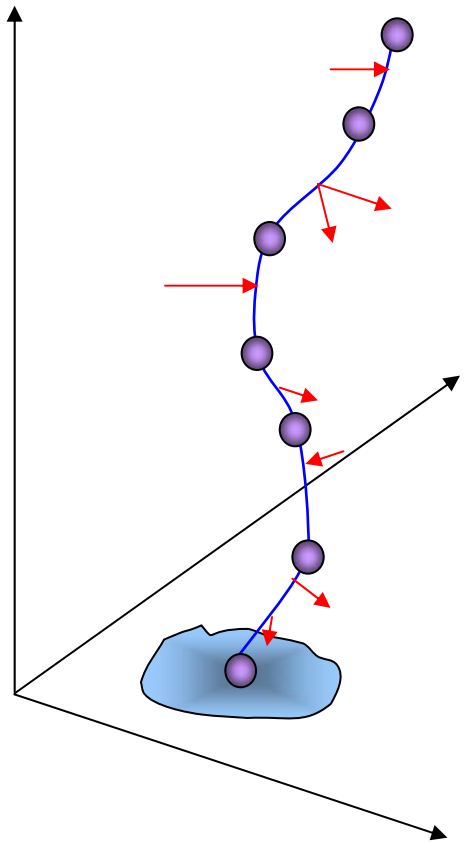
Dynamics and time dependence

- Already included in the definition of a geo-atom
- Dynamic geo-fields
  - a discretization of space–time
  - an ordered sequence of spatial fields
    - with a common discretization?
    - isolines, bona fide polygons, and TINs rediscretized
- Dynamic geo-objects

Peggy Agouris (Maine): Tracking vehicles through feature extraction from imaging camera



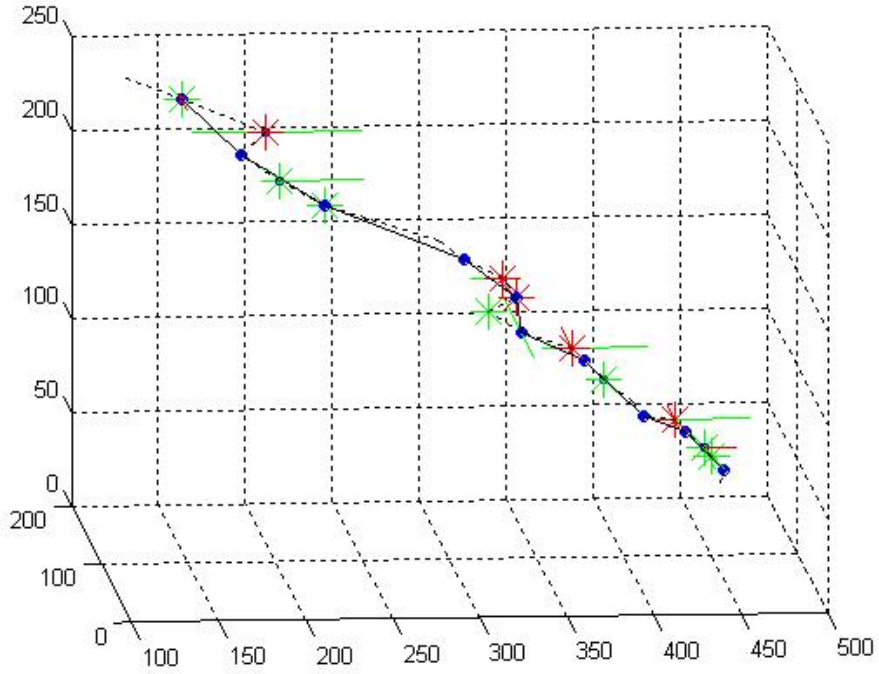
Helix representation

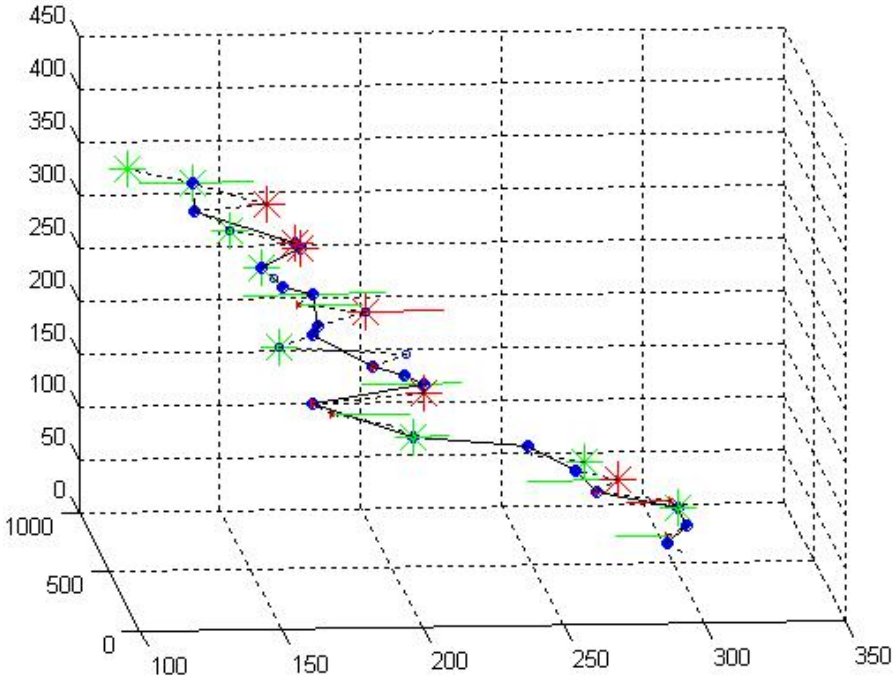


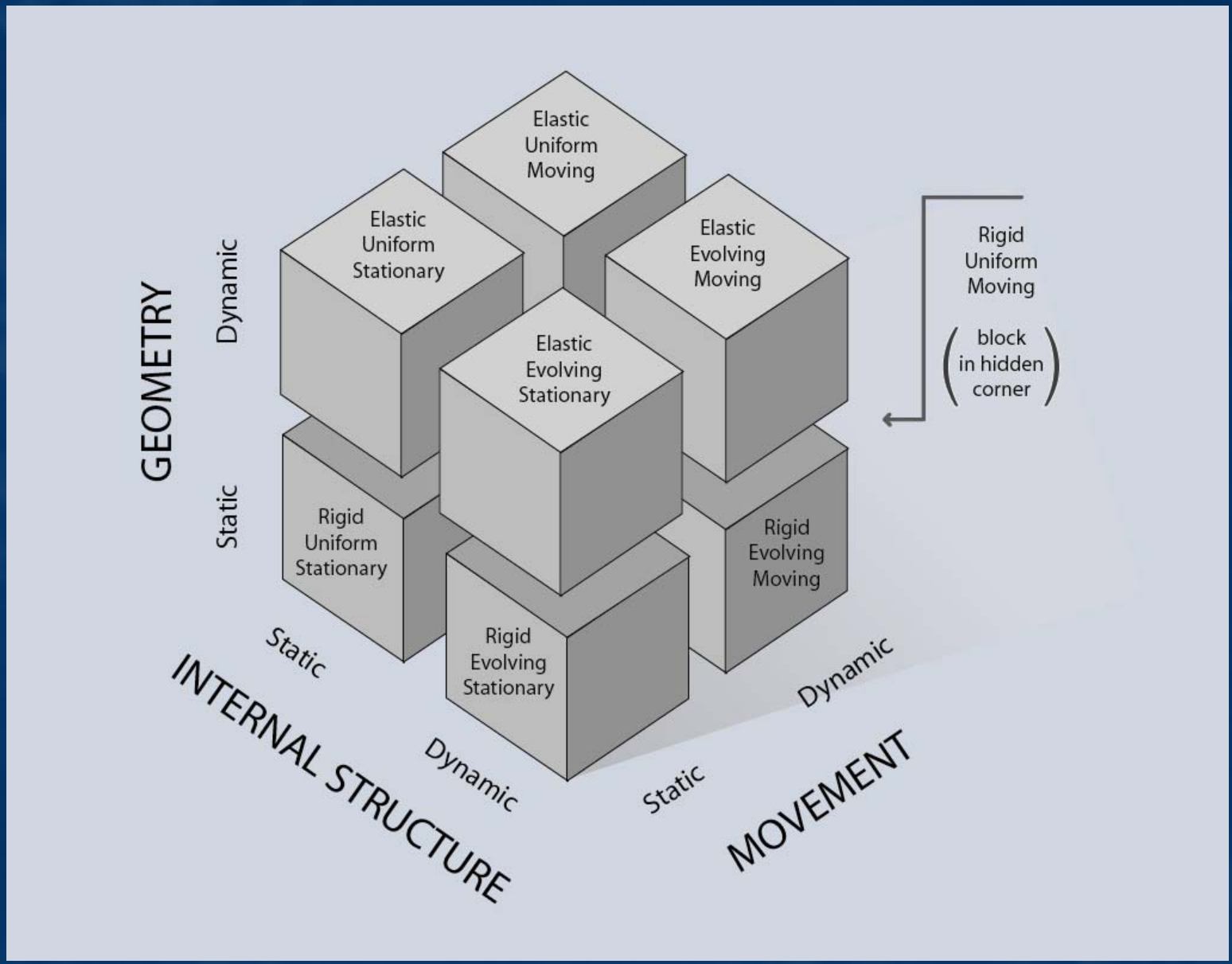
Spine: expresses spatio-temporal 3-D movement of the center of mass.

Prongs: express expansion or collapse of the object's outline

Hurricane helixes






Elastic
Uniform
Moving
Elastic
Uniform
Stationary
Elastic
Evolving
Moving
Elastic
Evolving
Stationary
Rigid
Uniform
Moving
block
in hidden
corner
GEOMETRY
Dynamic
Static
Rigid
Uniform
Stationary
Rigid
Evolving
Moving
Rigid
Evolving
Stationary
Static
Dynamic
INTERNAL STRUCTURE
Dynamic
Static
MOVEMENT

Concluding comments

- Increasing complexity
  - mainstream database concepts
  - increasing stress
- A single, unifying theory
  - integrating objects and fields
  - integrating all major concepts
- Geo-atoms and geo-dipoles
  - aggregation of point sets and point pairs
  - fiat and *bona fide* geo-objects
  - geo-fields