Graph Drawing Through the Lens of a Framework for Analyzing Visualization Methods

Tamara Munzner

Department of Computer Science
University of British Columbia

Graph Drawing 2013, Invited Talk
23 Sep 2013

http://www.cs.ubc.ca/~tmm/talks.html#gd13

Why?...

Graph Drawing Through the Lens of a Framework for Analyzing Visualization Methods

Why?...

Graph Drawing Through the Lens of a Framework for Analyzing Visualization Methods

Why analyze vis methods?

- think systematically about space of possibilities
  - methods: design space of techniques
- find gaps in previous work
  - develop new techniques, algorithms
- characterize existing/new work
  - match up algorithms and techniques to real-world problems
  - facilitate broader adoption by establishing suitability

Why?...

Graph Drawing Through the Lens of a Framework for Analyzing Visualization Methods

Why connect graph drawing and visualization?

- vis draws on GD community's work
  - especially algorithms, systems
- GD motivated by vis
  - great connection to application domains

- network data: special case of general principles

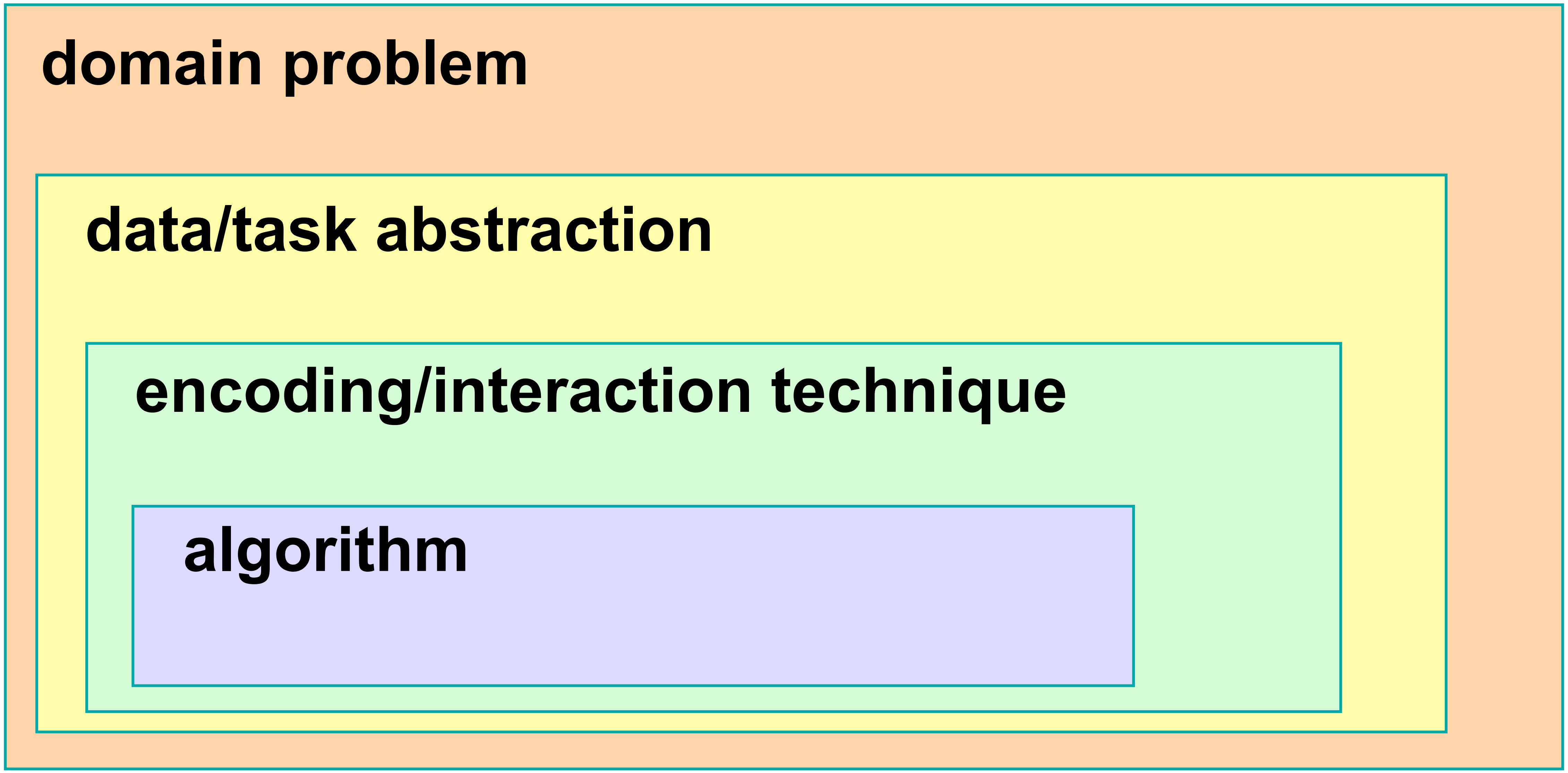
Outline

- Levels of visualization design
- Abstraction for data
- Principles of marks and channels
- Using space
- Further analysis examples
- Conclusions

Levels of visualization design

Separating vis design into four levels

- connecting all the way from real-world problems of target users to algorithms

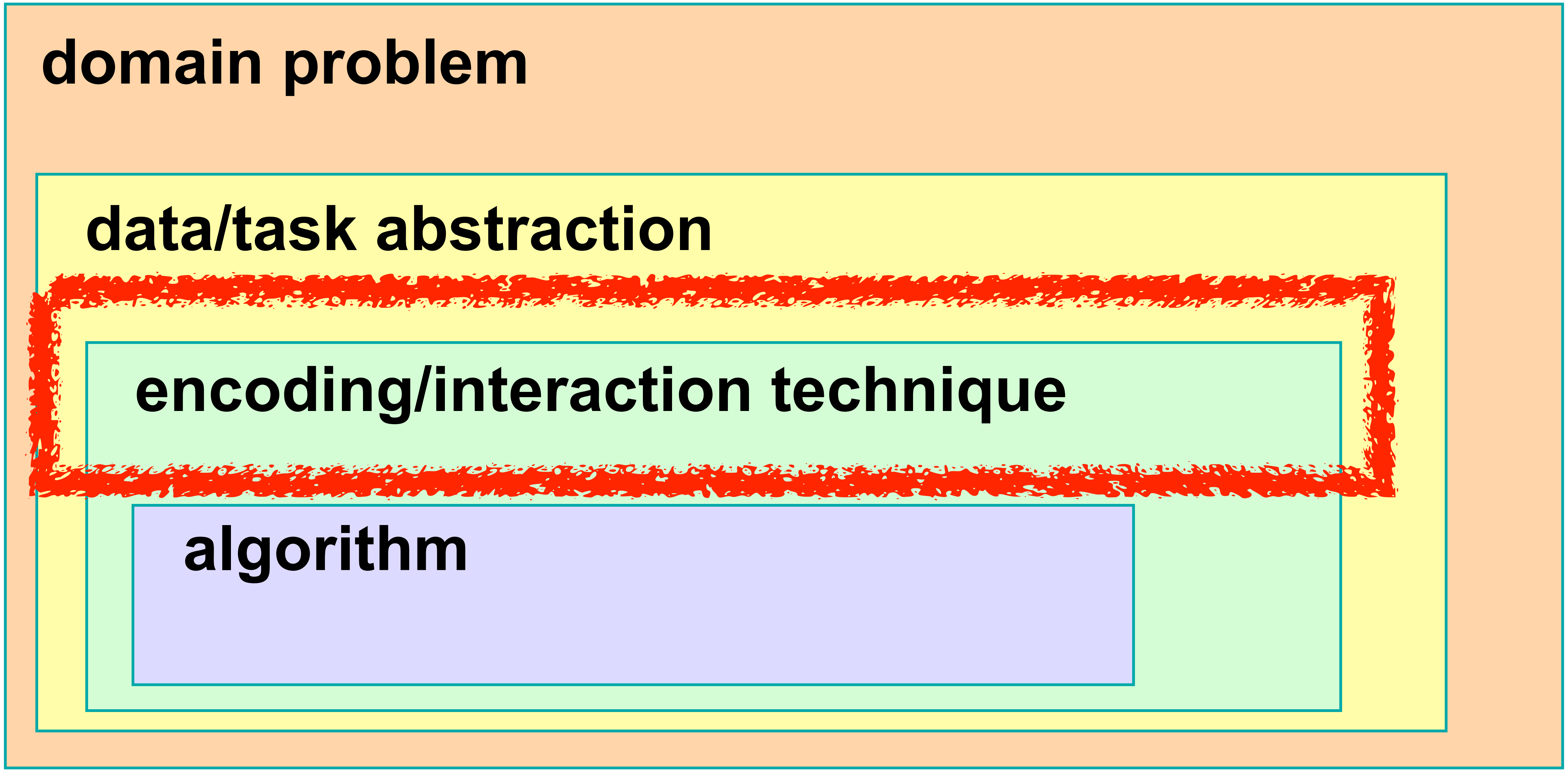


- covered elsewhere: validation

  [A Nested Model for Visualization Design and Validation. Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 09), 15(6):921-928, 2009.]

Emphasis: Technique level

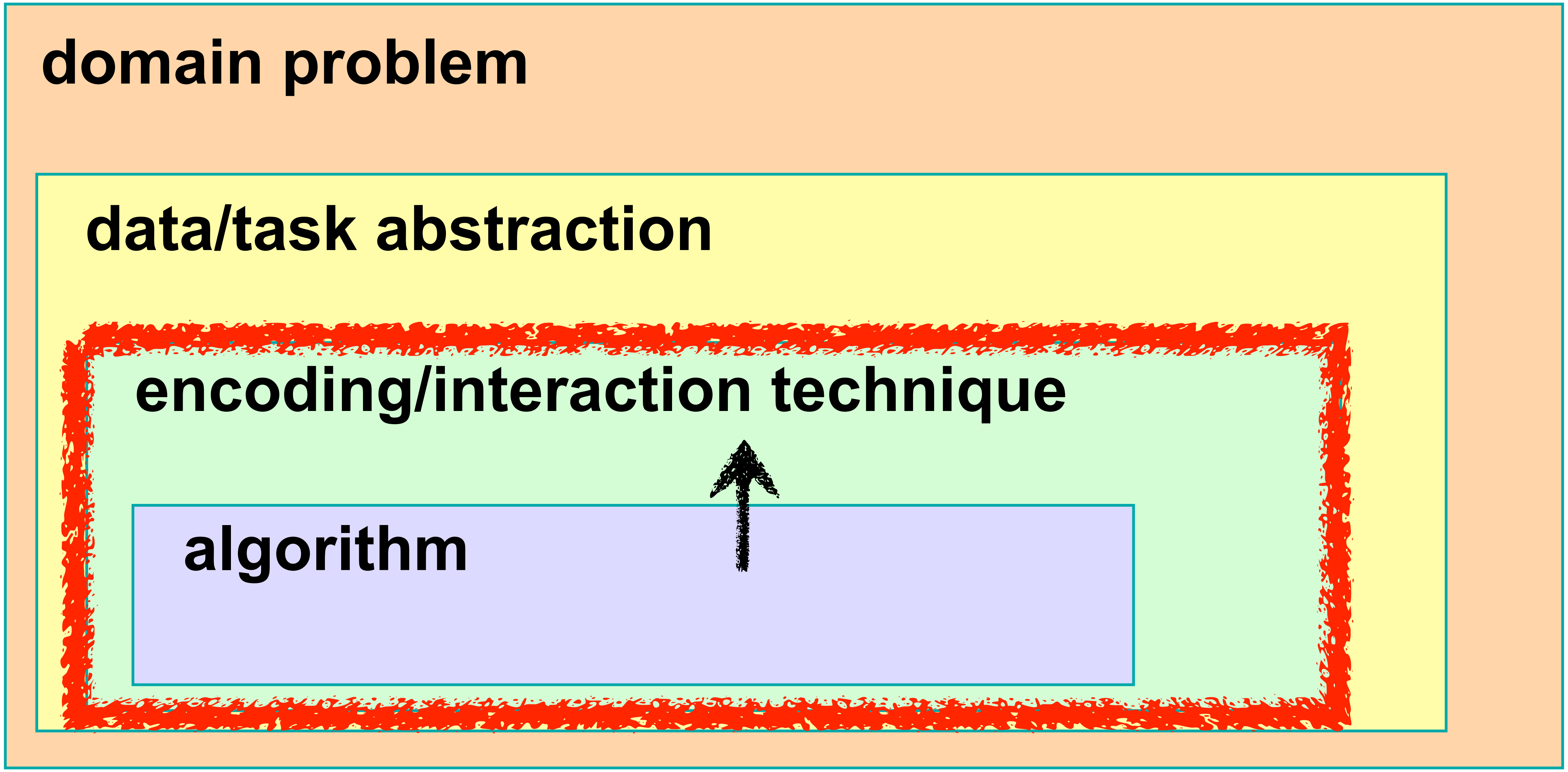
- just above familiar algorithm level, connects directly
- plus a bit of background on abstraction



[A Nested Model for Visualization Design and Validation. Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 09), 15(6):921-928, 2009.]

Goal: More upwards characterization

- map from algorithms up to techniques they support



[A Nested Model for Visualization Design and Validation. Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 09), 15(6):921-928, 2009.]

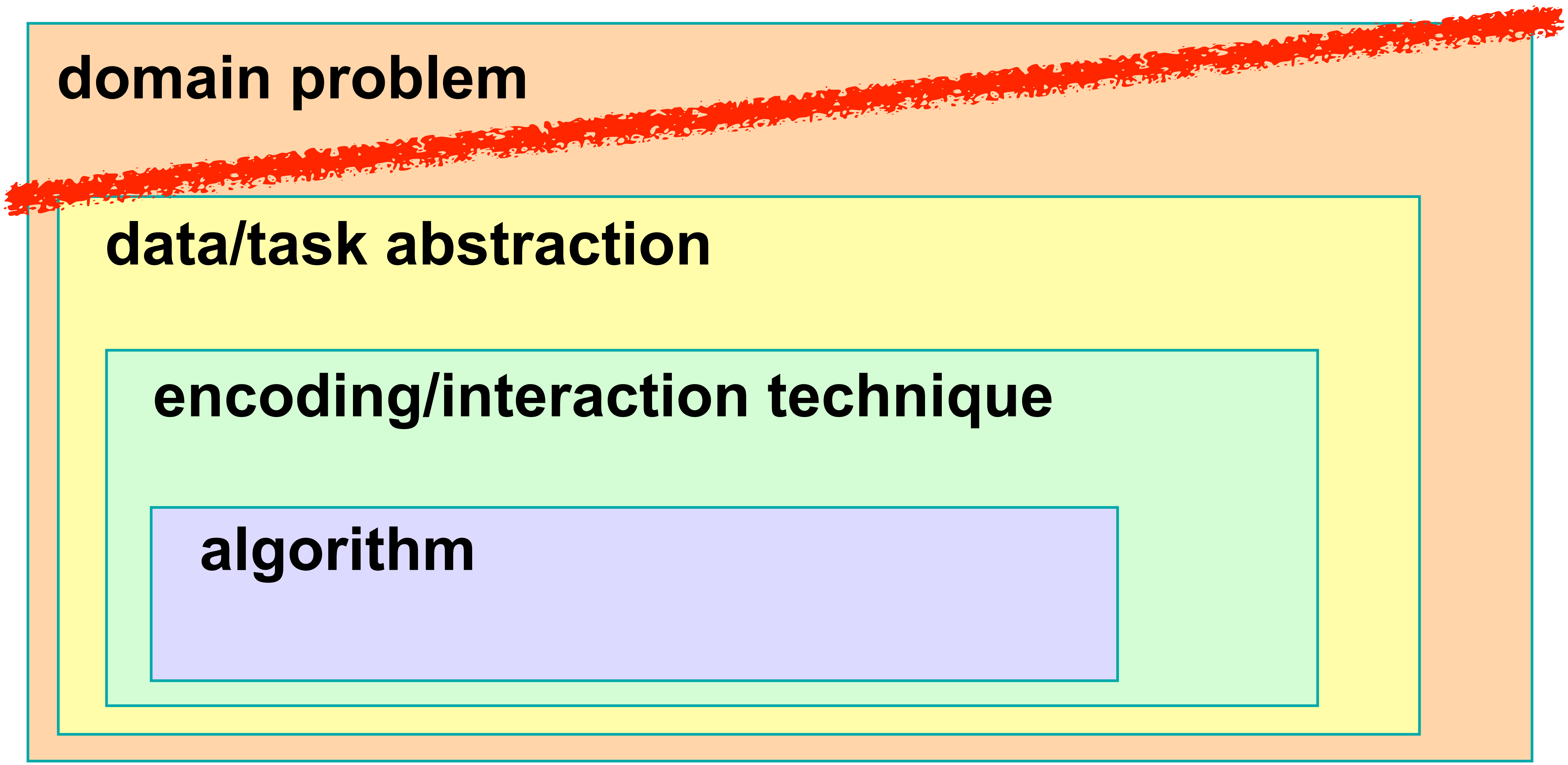
Characterize how?

- focus here on one major issue
  - how is space used?
- explicit consideration in visualization
  - trickier to see from purely graph drawing perspective
    - common cases not trivial to analyze!
      - node-link diagrams, compound graphs

Covered elsewhere: Downwards from real users

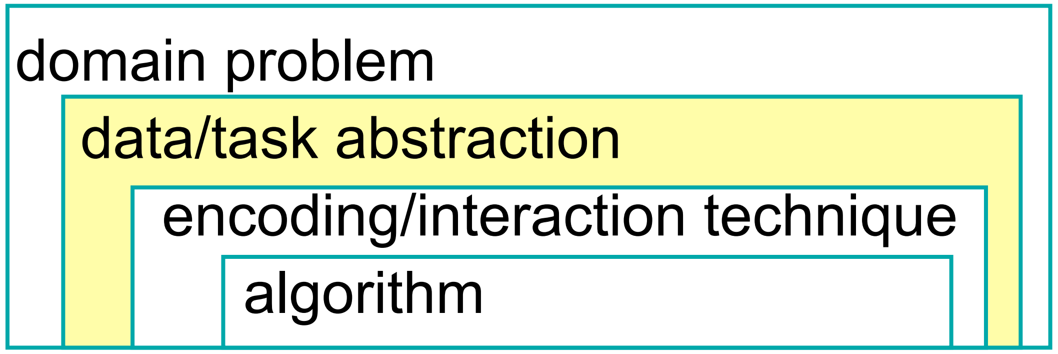
- design study methodology paper
  - problem-driven work: building for specific people to use

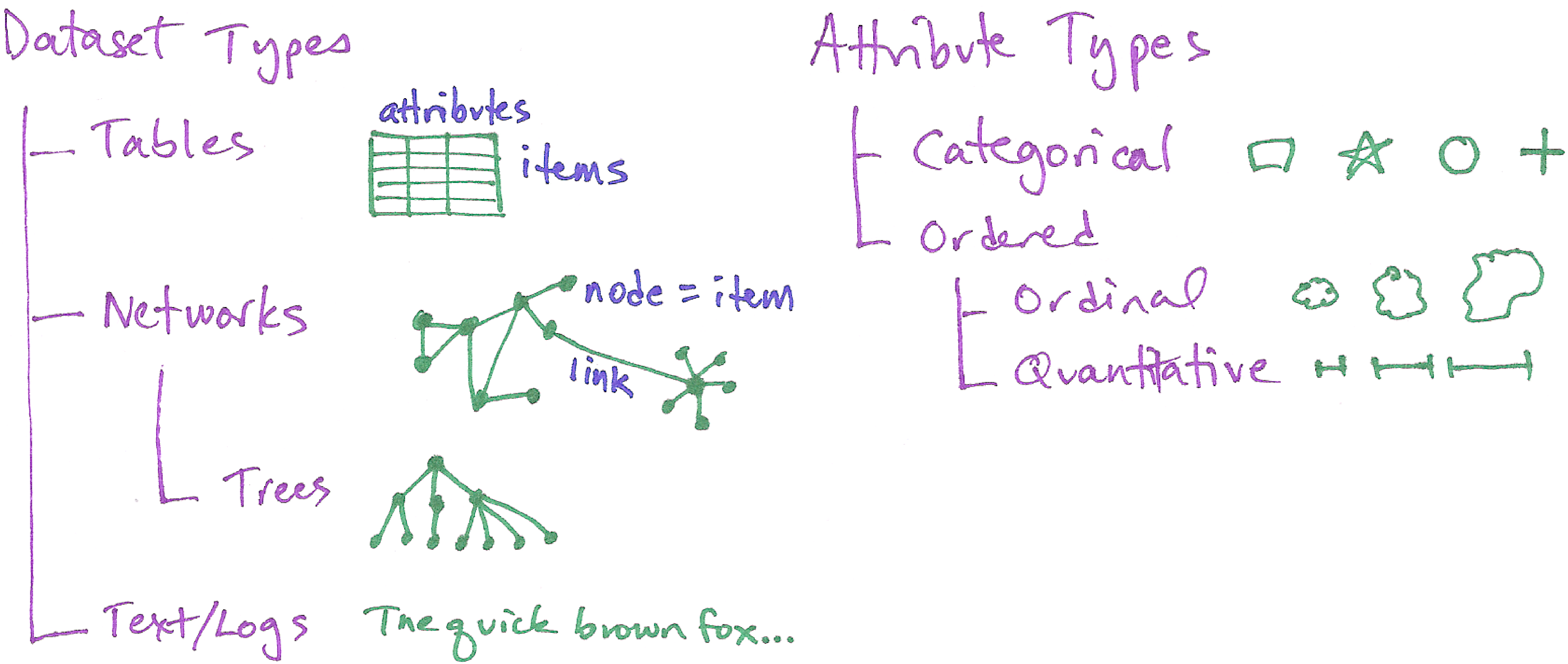
[Design Study Methodology: Reflections from the Trenches and the Stacks. Sedlmair, Meyer, and Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2012), 18(12):2431-2440, 2012.]



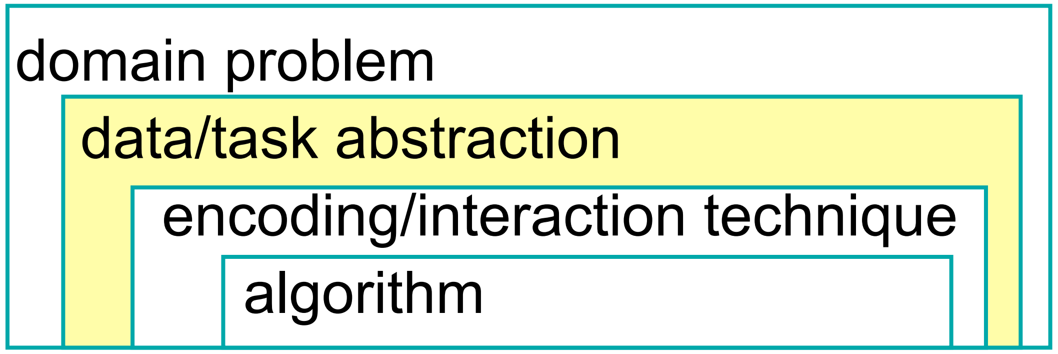
Abstraction for data

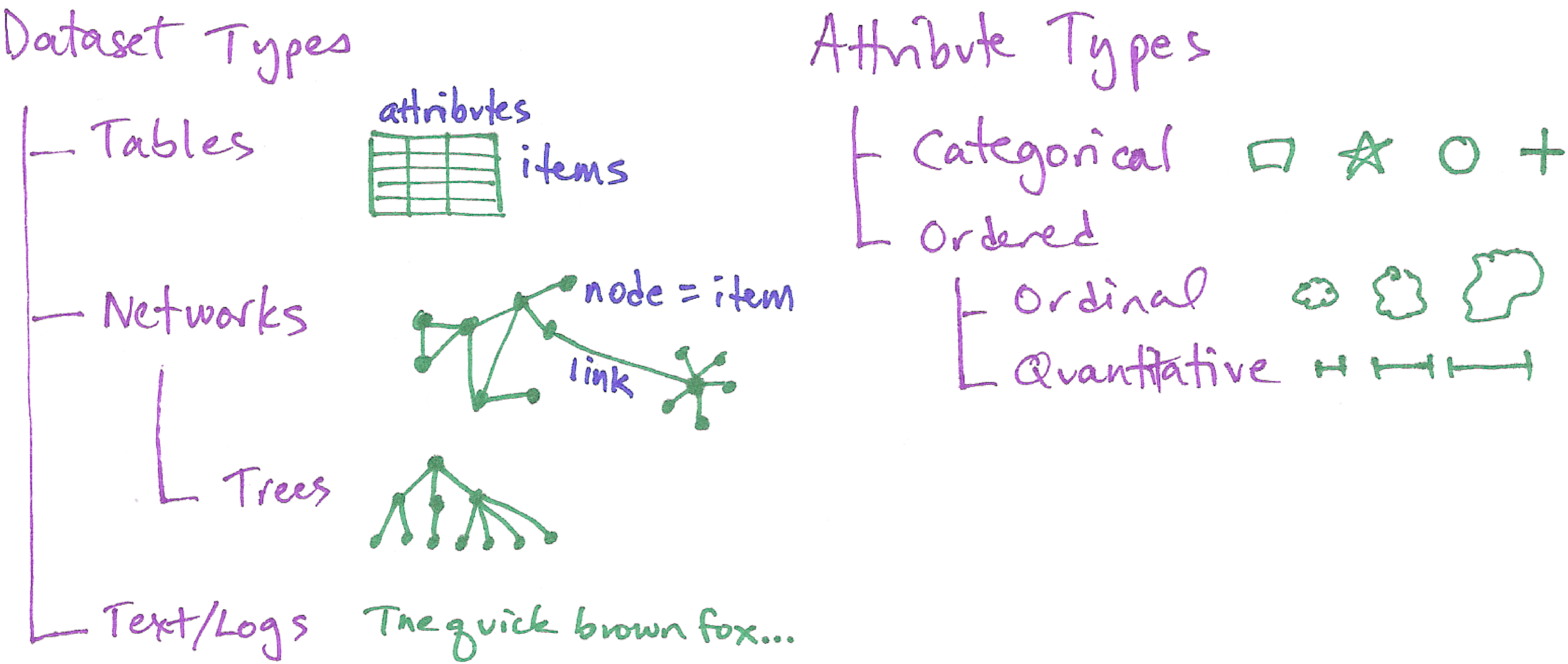
Abstraction: data types





Abstraction: data types



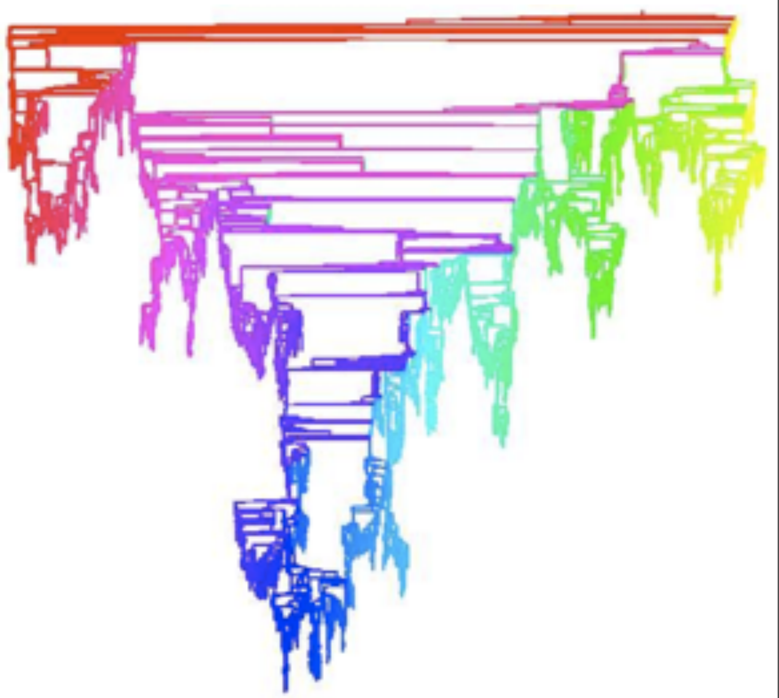


– covered elsewhere: task abstraction

[A Multi-Level Typology of Abstract Visualization Tasks. Brehmer and Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis), to appear 2013.]

Deriving new data: Common case

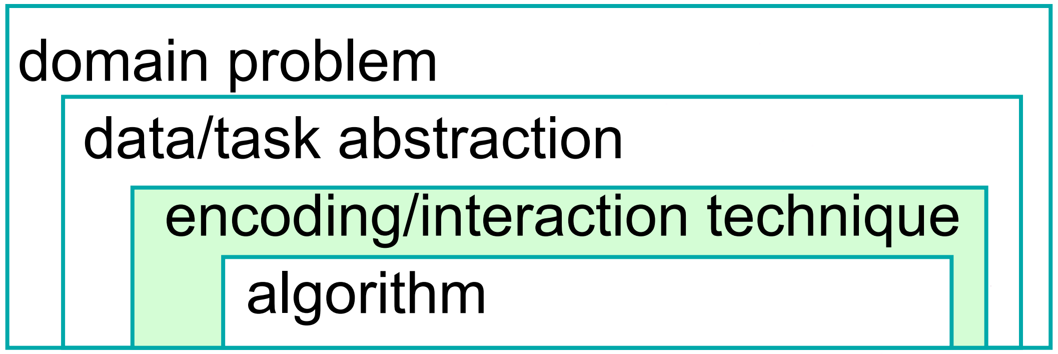
- example: Strahler number for graphs
  - centrality metric: node importance
    - new per-node quantitative attrib
    - result of global calculation
- visualization uses
  - fast interactive rendering: draw nodes in order of importance
  - draw small subset: structure far more understandable than w/ random sampling



  - more detail in Auber02
    [Using Strahler numbers for real time visual exploration of huge graphs. Auber. Intl. Conf. Computer Vision and Graphics, 2002, p. 56-69.]

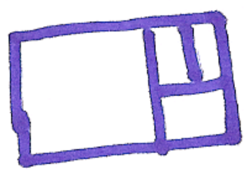
Principles of marks and channels

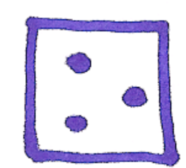
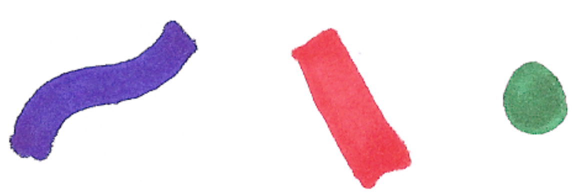
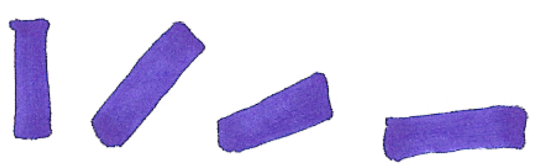
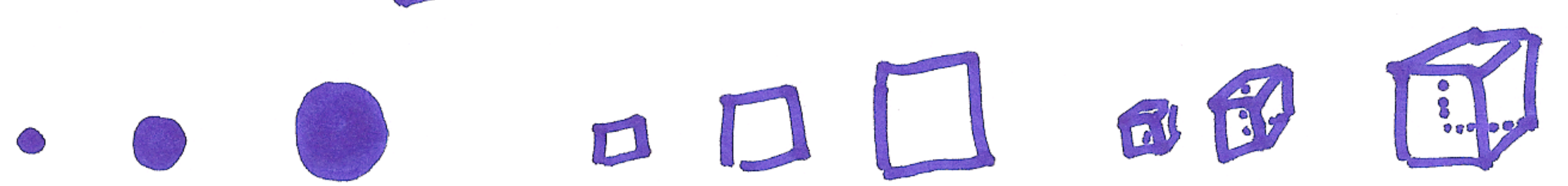
Techniques: Visual encoding



- how to analyze?
  - start with easy cases from statistical graphics

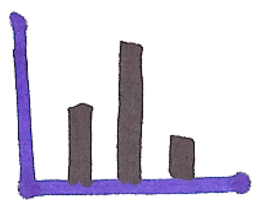
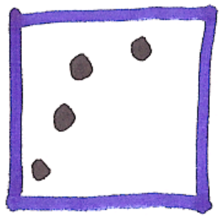
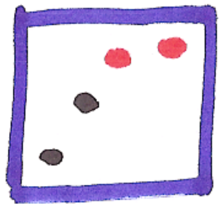
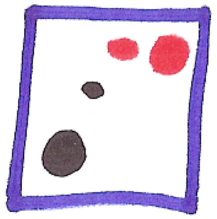
Marks and channels

- marks : geometric primitives
  - points 
  - lines 
  - areas 

- visual channels: control appearance of marks
  - position
    horizontal , vertical , both 
  - color 
  - shape 
  - tilt 
  - size 

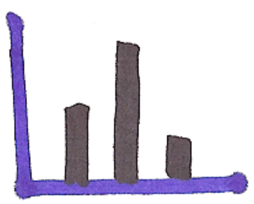
Techniques: Visual encoding analysis principles

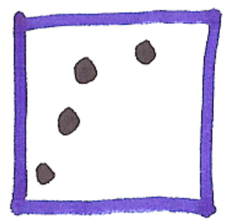
- analyze as combination of marks and channels showing abstract data

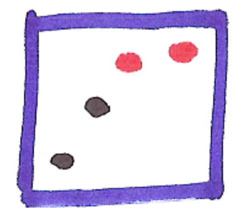
   

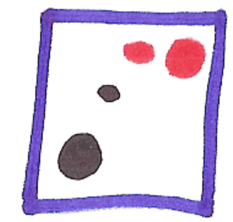
Techniques: Visual encoding analysis principles

- analyze as combination of marks and channels showing abstract data







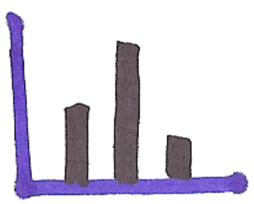


1: vertical position

mark: line

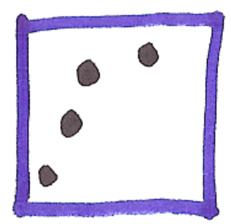
Techniques: Visual encoding analysis principles

- analyze as combination of marks and channels showing abstract data



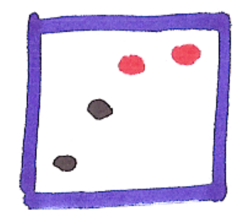
1: vertical position

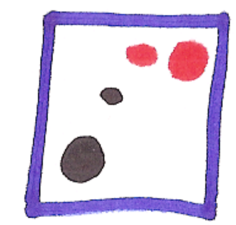
mark: line



2: vertical position, horizontal position

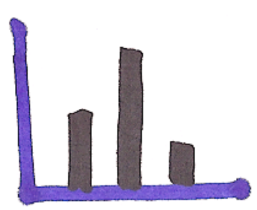
mark: point





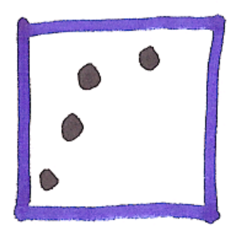
Techniques: Visual encoding analysis principles

- analyze as combination of marks and channels showing abstract data



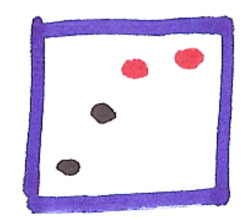
1: vertical position

mark: line



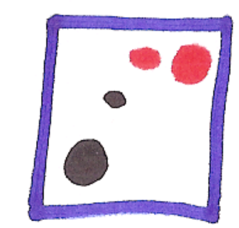
2: vertical position, horizontal position

mark: point



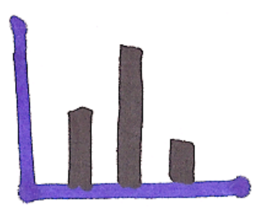
3: vertical position, horizontal position, color

mark: point



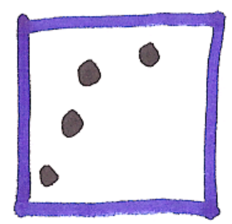
Techniques: Visual encoding analysis principles

- analyze as combination of marks and channels showing abstract data



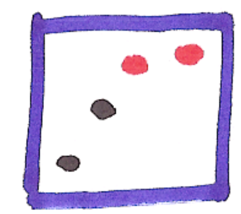
1: vertical position

mark: line



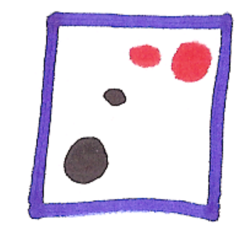
2: vertical position, horizontal position

mark: point



3: vertical position, horizontal position, color

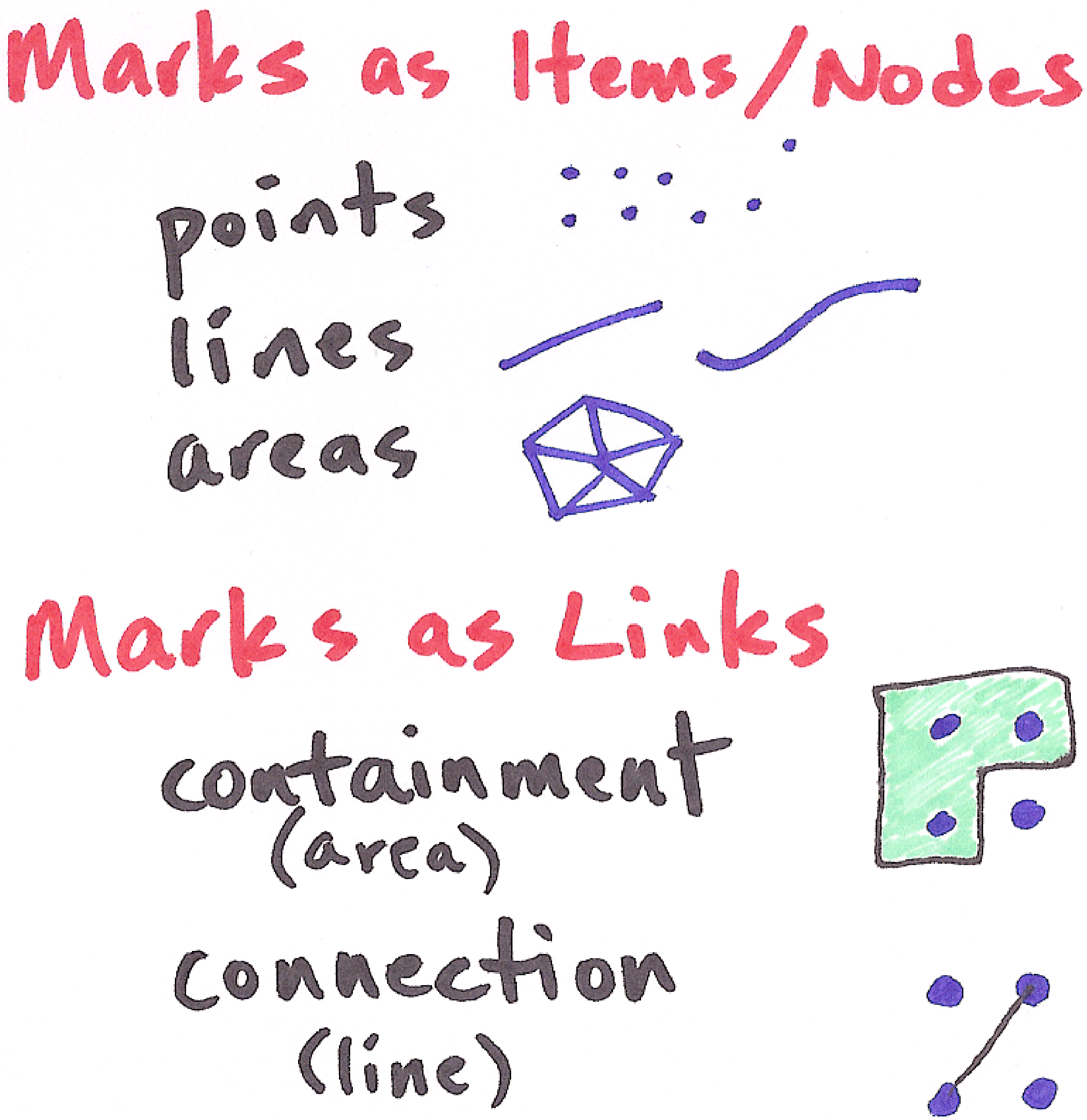
mark: point



4: vertical position, horizontal position, color, size

mark: point

Marks as links

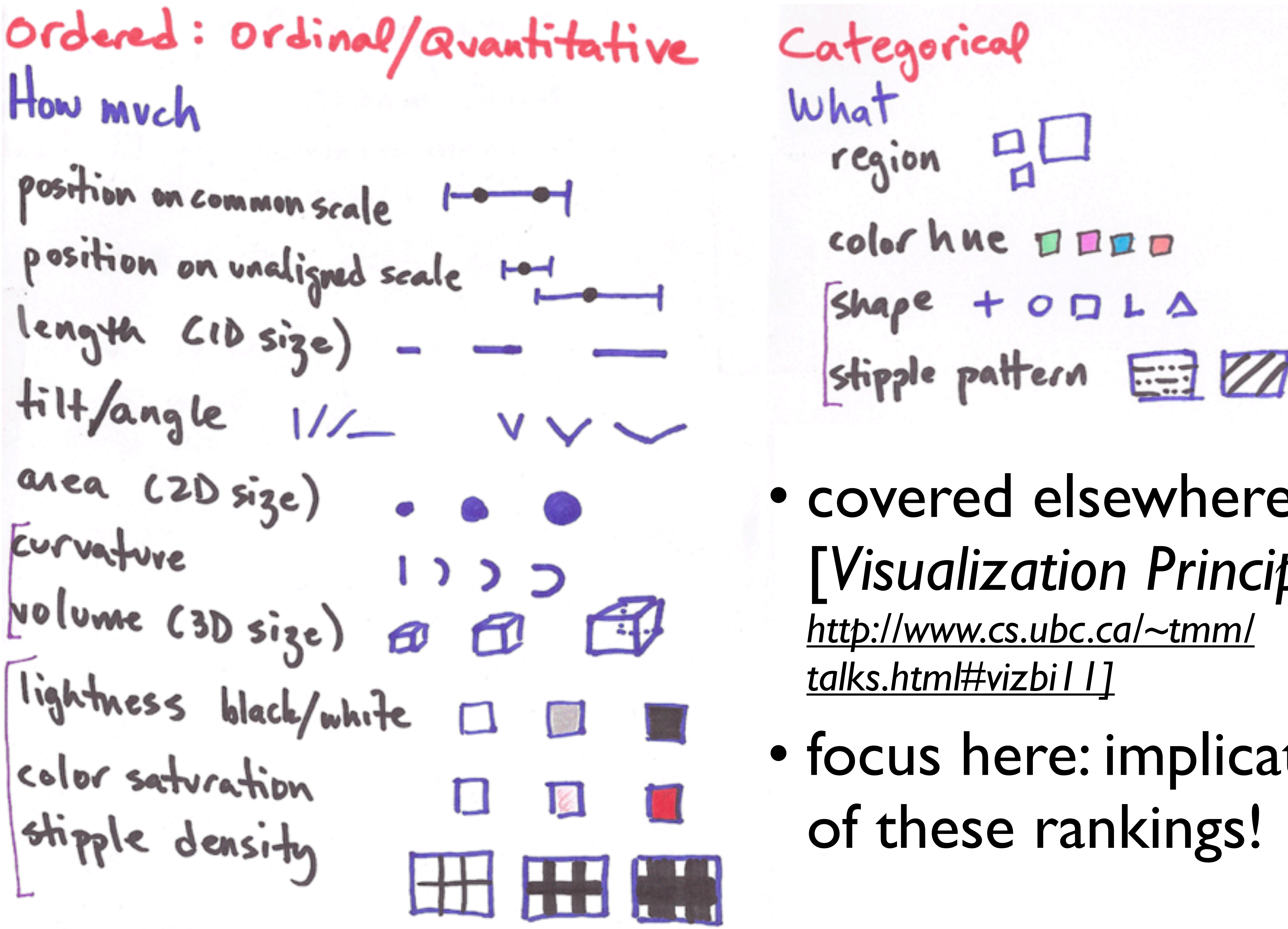


- we implicitly perceive some properties as indicating relationships between items
  - containment
  - connection
  - also, proximity
    - use of space

Channel types

- channels also have implicit perceptual types
  - match them with attribute types
    - avoid losing information or implying incorrect properties
  - *how much:* ordered
    - example: spatial position along a common scale
    - example: length of line mark
  - *what:* categorical
    - example: spatial region
- spatial channels have strongest perceptual impact
  - reason for focus on use of space here
- many other channels: color, size, orientation, ...
  - we know types and ranking in terms of impact (roughly)

Channel rankings



- covered elsewhere: [*Visualization Principles* *http://www.cs.ubc.ca/~tmm/talks.html#vizbi11*]
- focus here: implications of these rankings!

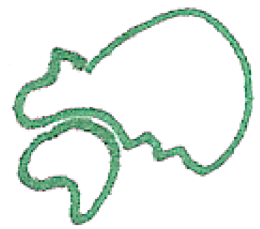
Using space

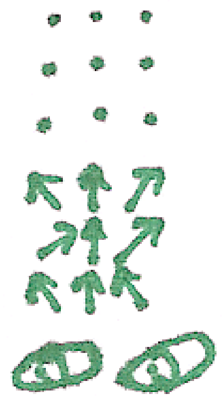
Using space: Channel choices

Given

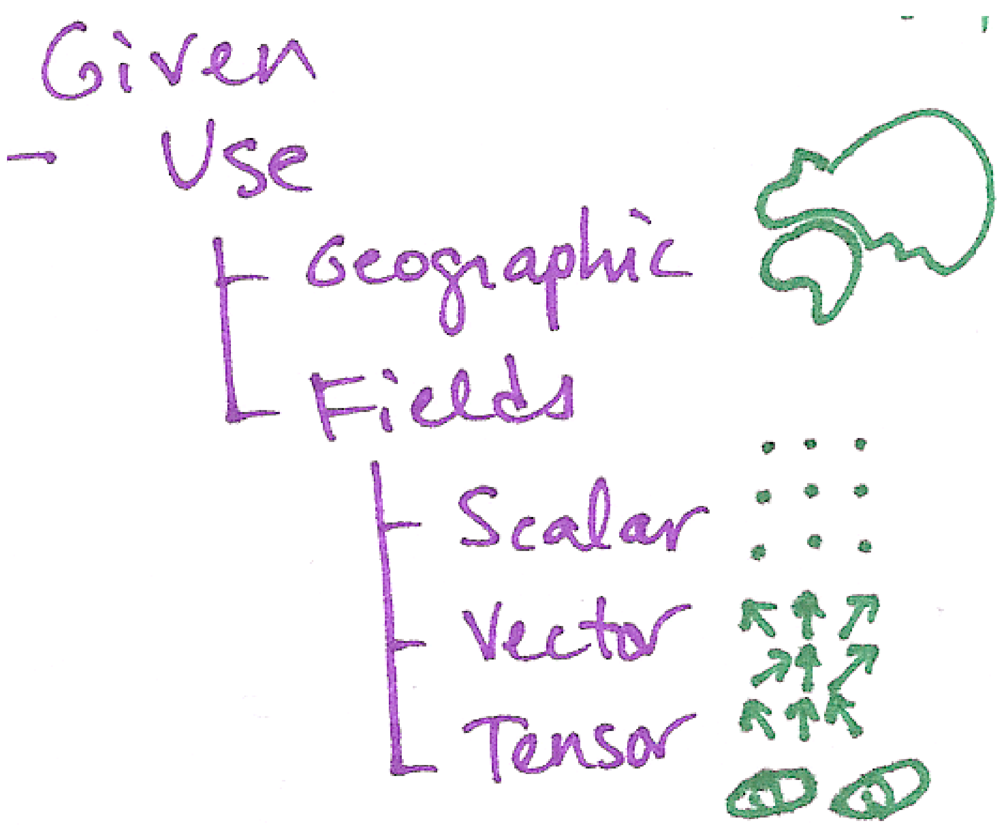
→ Use

- Geographic
- Fields
  - Scalar
  - Vector
  - Tensor



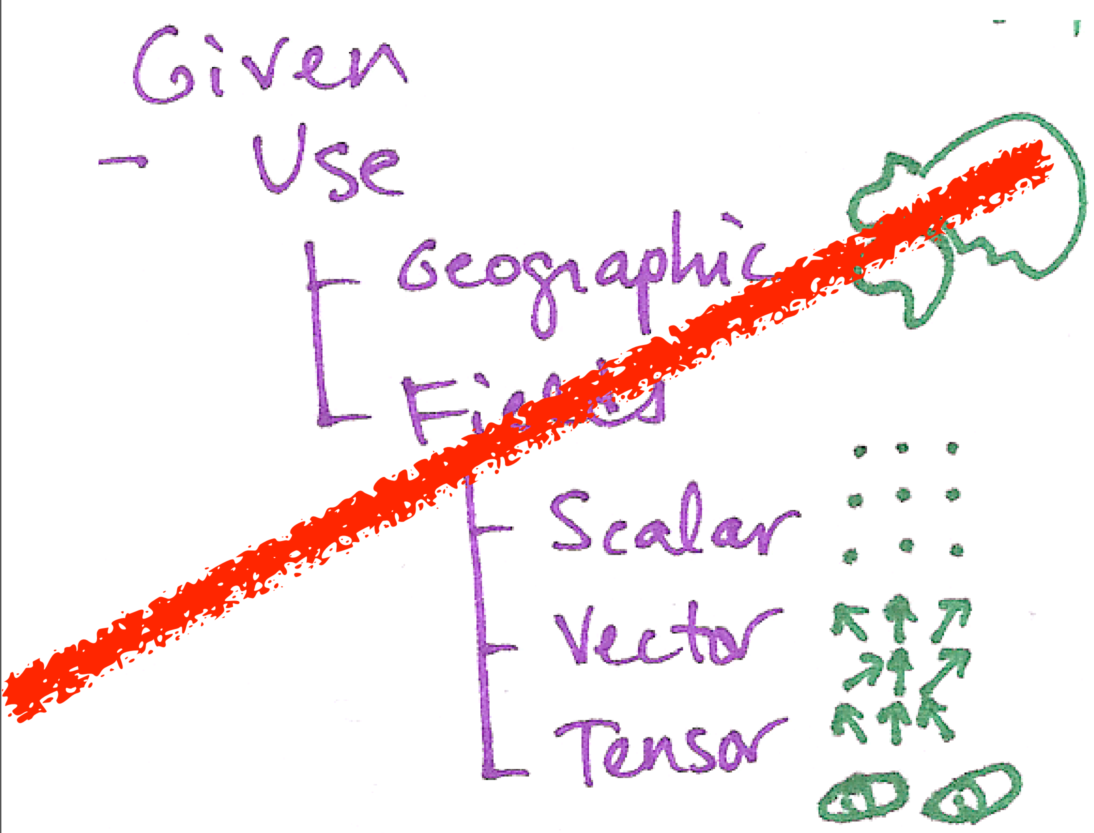


Using space: Channel choices



- could just use data as given
  - cartography
  - volume graphics
  - flow visualization

Using space: Channel choices



- could just use data as given
  - cartography
  - volume graphics
  - flow visualization

- focus: choosing use of space
  - central issue in graph layout

Using space: Channel choices

Spatial Channels

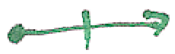
Values

Express 

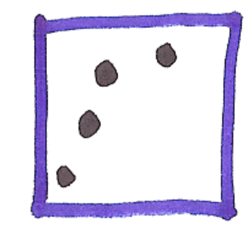
Using space: Channel choices

Spatial Channels

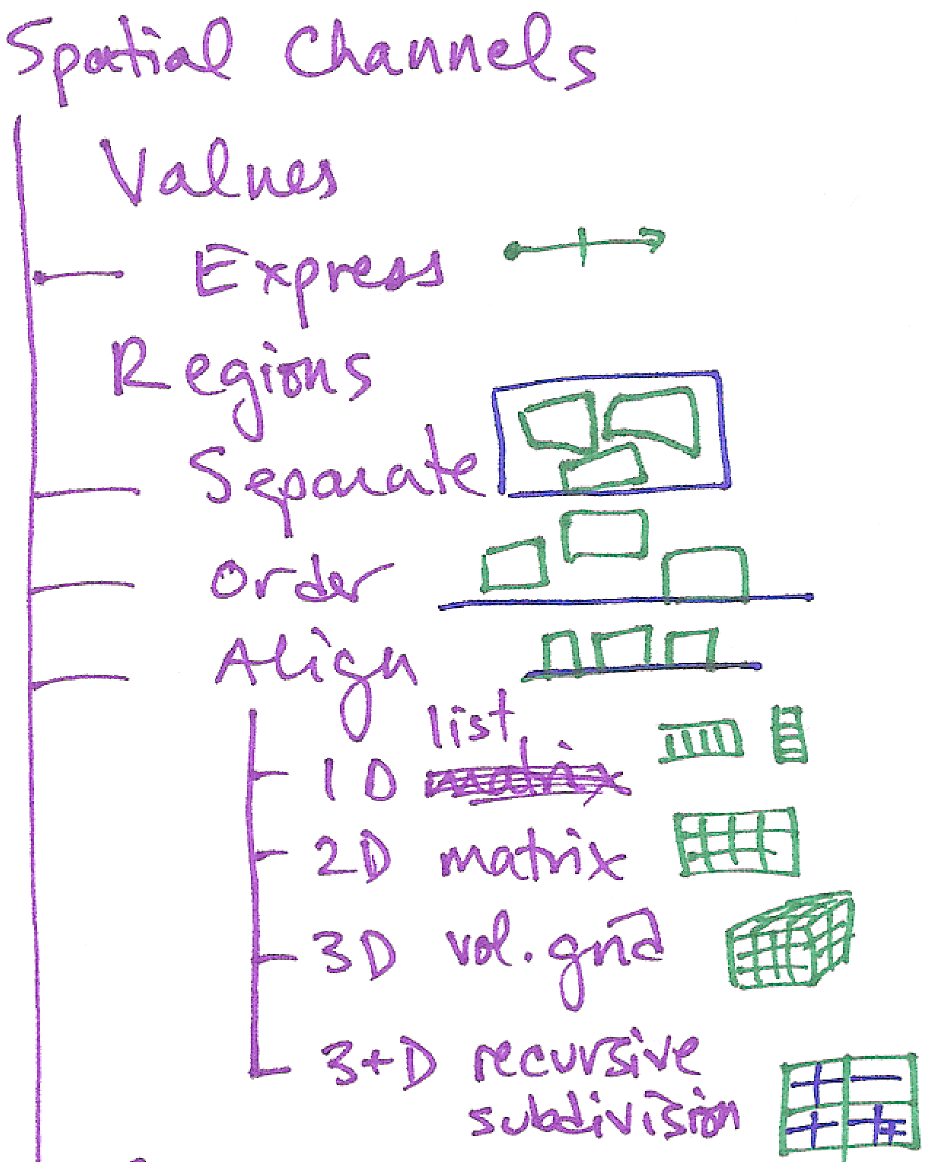
Values

Express 

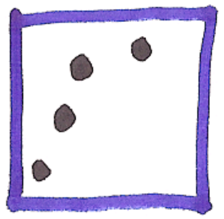
- values expressed spatially
  - encode quantitative attribute using spatial position of mark
    - example: scatterplots



Using space: Channel choices

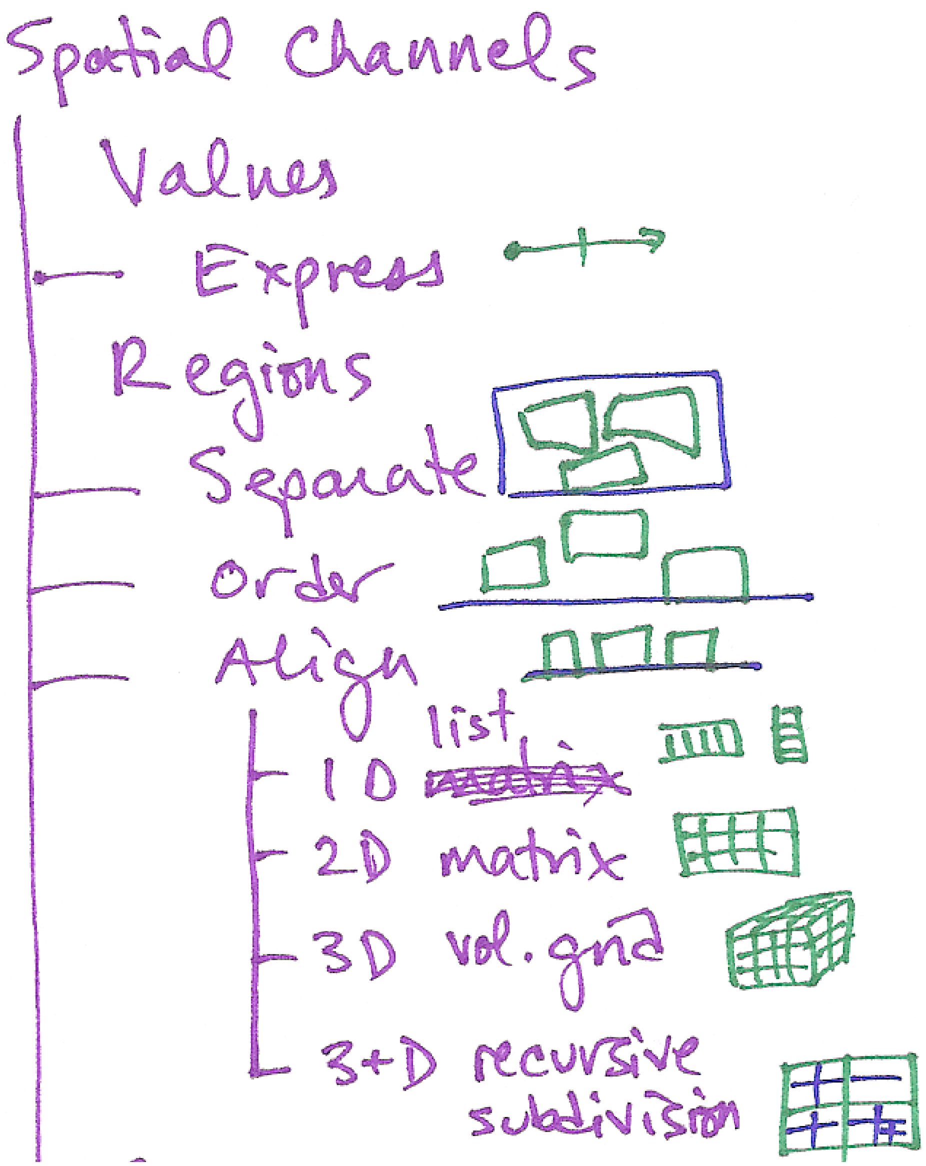


- values expressed spatially
  - encode quantitative attribute using spatial position of mark
    - example: scatterplots

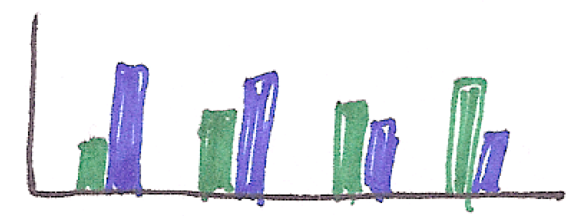


- regions of space
  - separate into regions
    - proximity implies grouping
  - order regions
    - could be data-driven
  - align for more precise judgements
    - can subdivide recursively

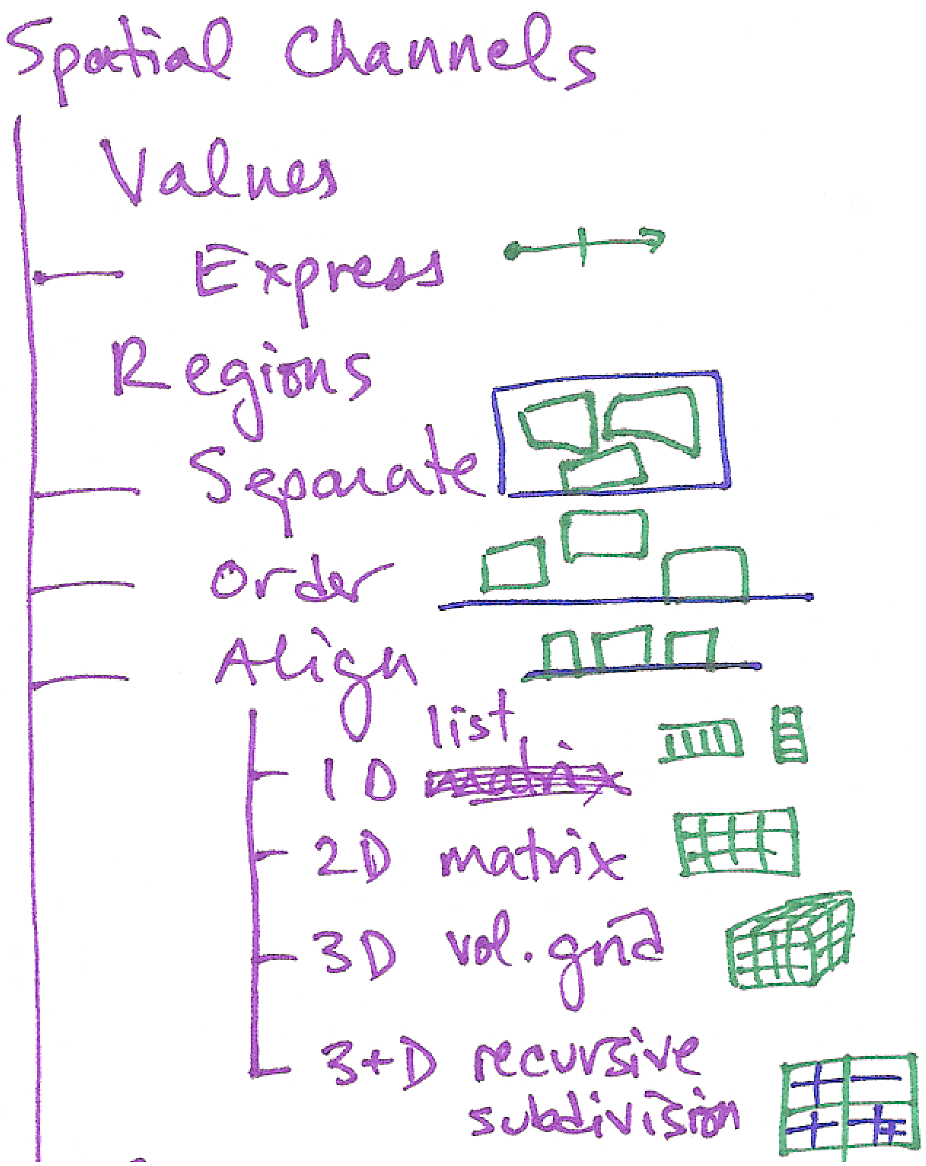
Using space: Examples



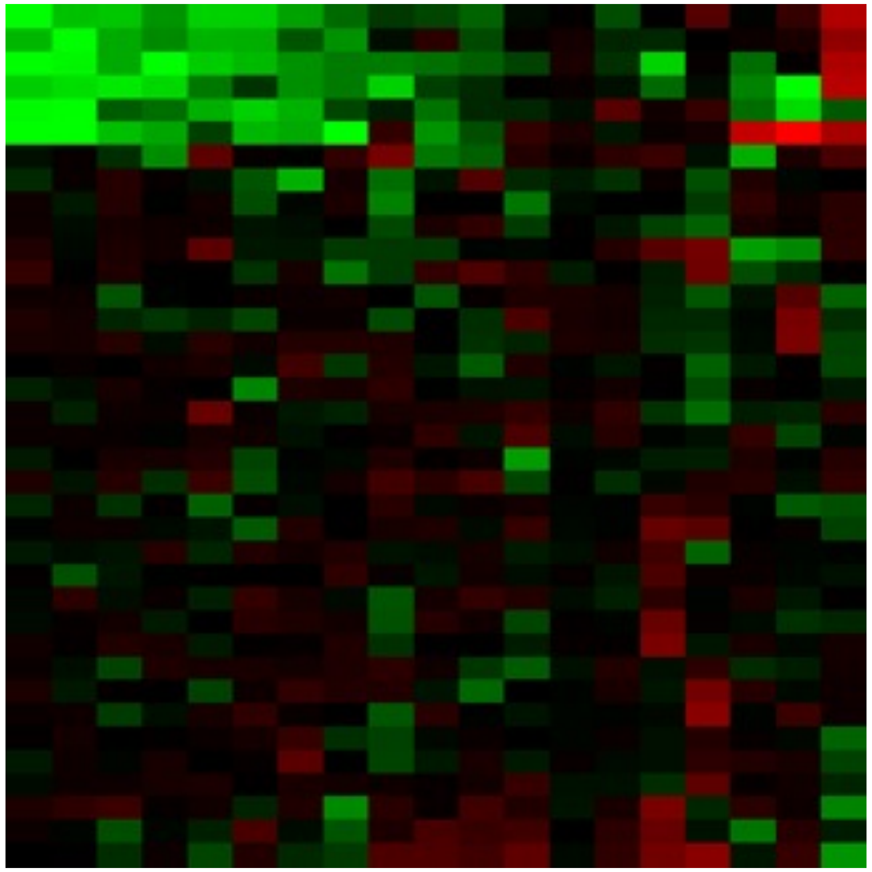
- multiple bar charts
  - data: table, 3 attribs
    - 1 quant, 2 categ
  - marks: line
  - spatial channels
    - within each region
      - express value w/ vert spatial pos
      - align vert
      - order by quant attrib
    - one choice: separate views
      - separate into 2 regions by categ attrib
    - another choice: interleaved view
      - separate into 4 regions, 1 per item
      - draw both attribs within region



Using space: Examples

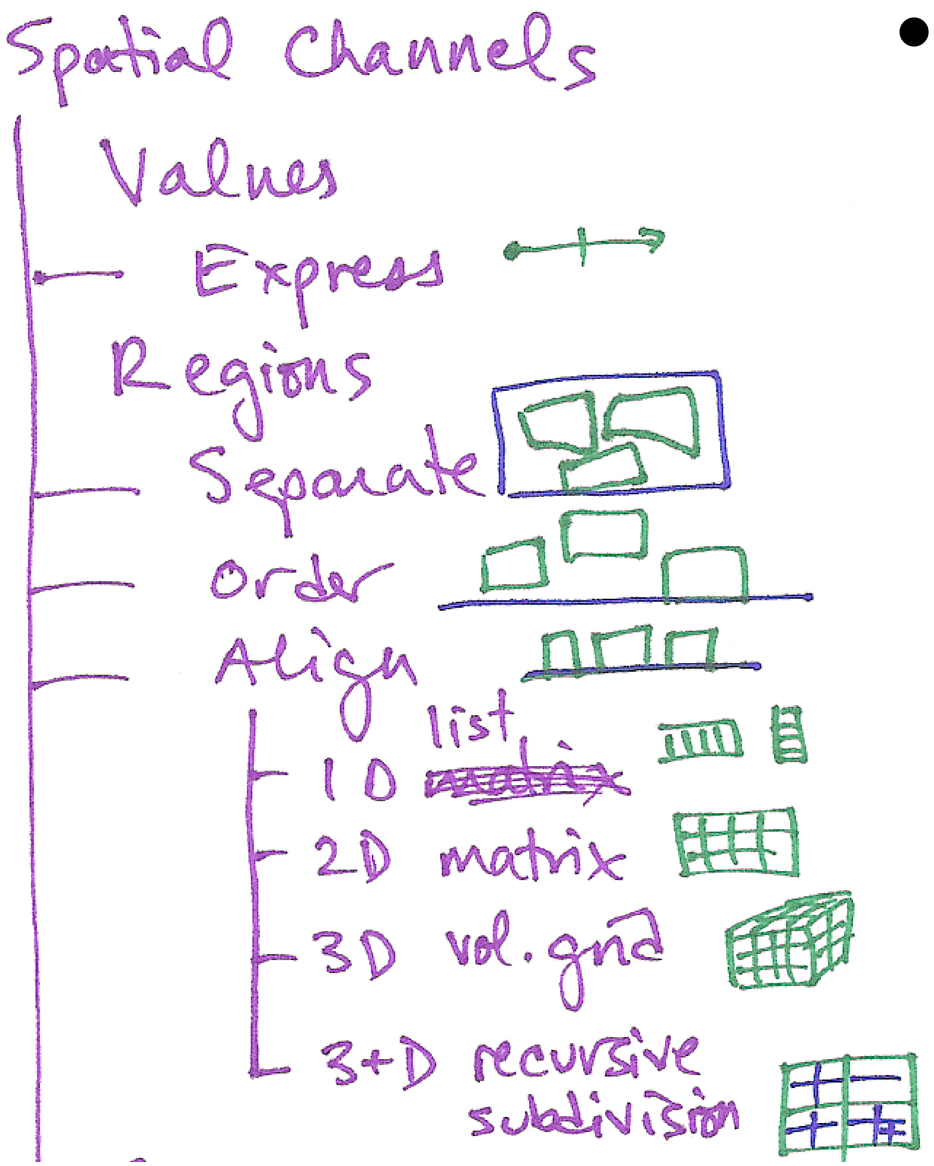


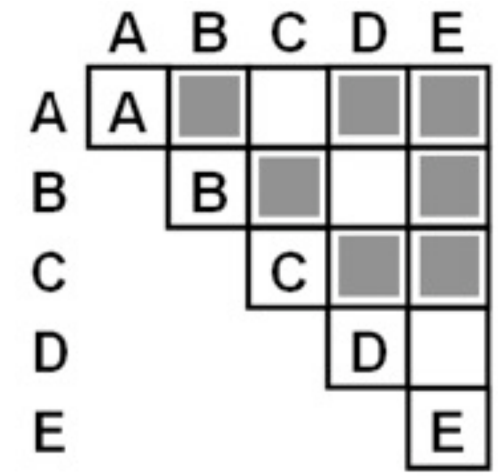
- heatmap
  - data: same!
    - 1 quant, 2 categ
  - marks: area
    - (color by quant attrib)
  - spatial channels
    - separate and align in 2D matrix
      - indexed by 2 categ attribs
    - order: many choices
      - matrix reordering algs



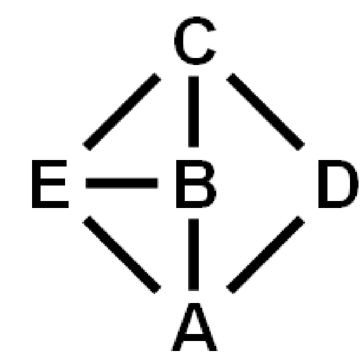
[http://upload.wikimedia.org/wikipedia/commons/4/48/Heatmap.png]

Using space: Examples



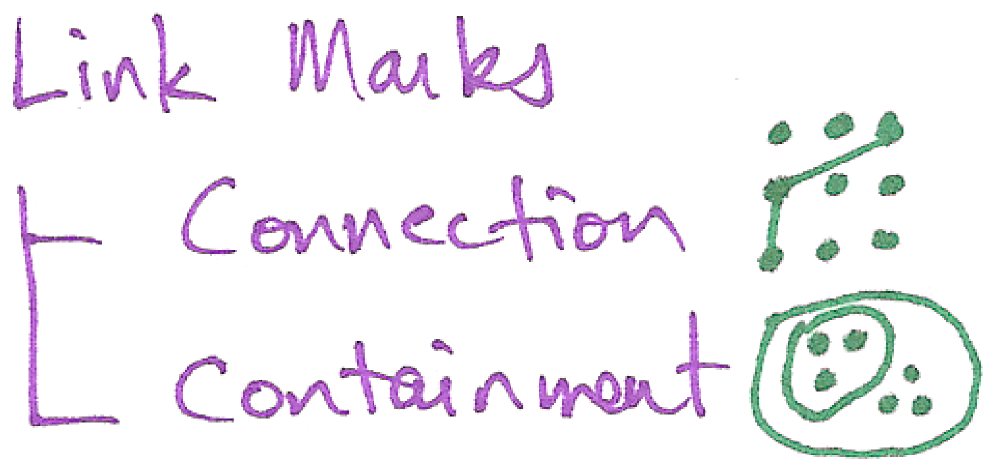


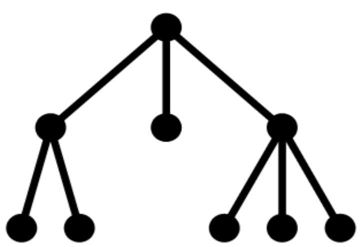
- matrix graph view
  - data, channels: same!
  - derived data: table from network
    - 1 quant attrib
      - weighted edge between nodes
    - 2 categ attribs: node list x 2
  - spatial channels:
    - cell shows presence/absence of edge

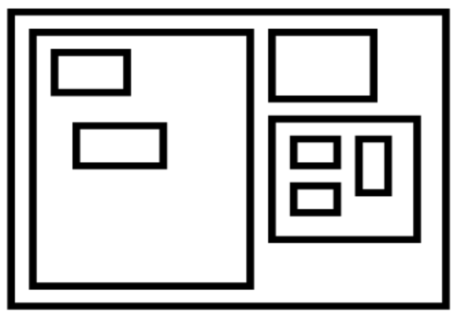


[NodeTrix: a Hybrid Visualization of Social Networks. Henry, Fekete, and McGuffin. IEEE TVCG (Proc. InfoVis) 13(6):1302-1309, 2007.]

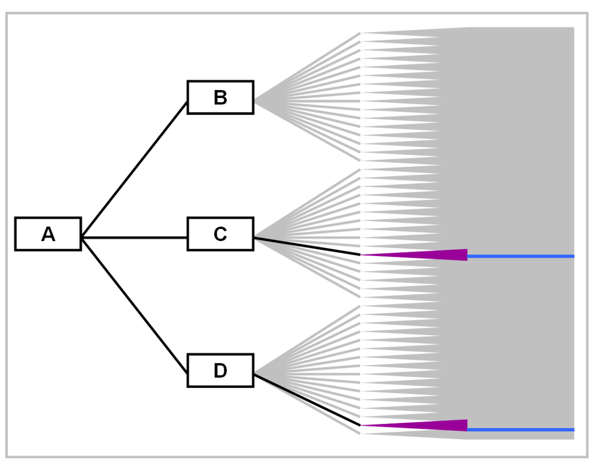
Using space: Links



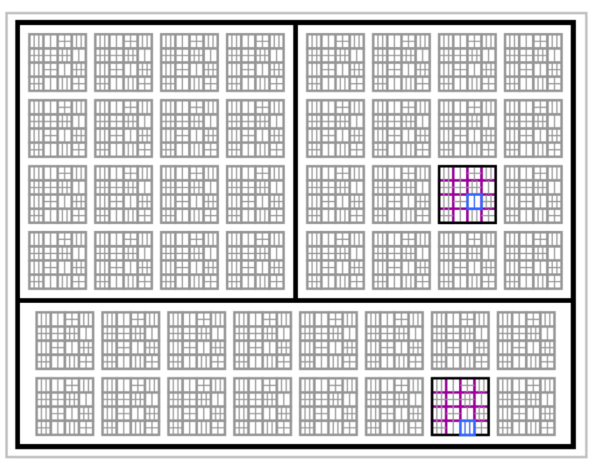




- marks as links (vs. nodes)
  - common case in graph drawing
  - 1D case: connection
    - ex: all node-link diagrams
    - emphasizes topology, path tracing
  - 2D case: containment
    - ex: all treemap variants
    - emphasizes attribute values at leaves (size coding)



Node–Link Diagram



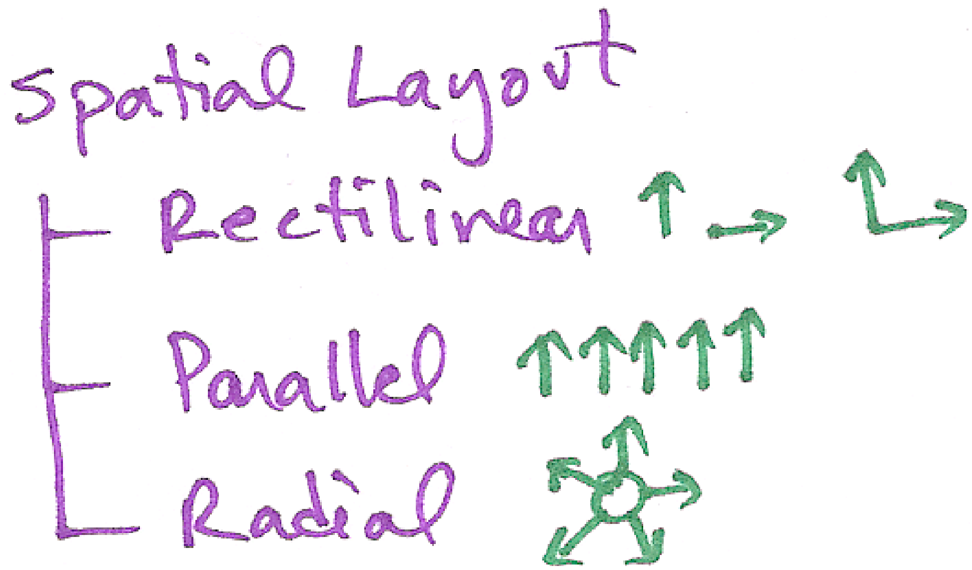
Treemap

[Elastic Hierarchies: Combining Treemaps and Node-Link Diagrams. Dong, McGuffin, and Chignell. Proc. InfoVis 2005, p. 57-64.]

Using space: Layout orientation

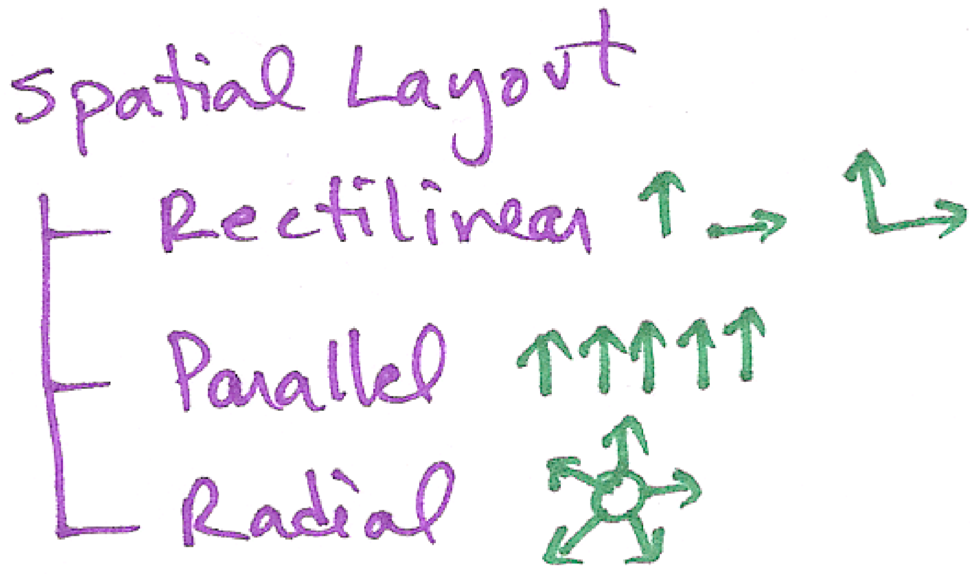
Spatial Layout
- Rectilinear
- Parallel
- Radial

Using space: Layout orientation



- spatial layout
  - orientation of spatial axes

Using space: Layout orientation

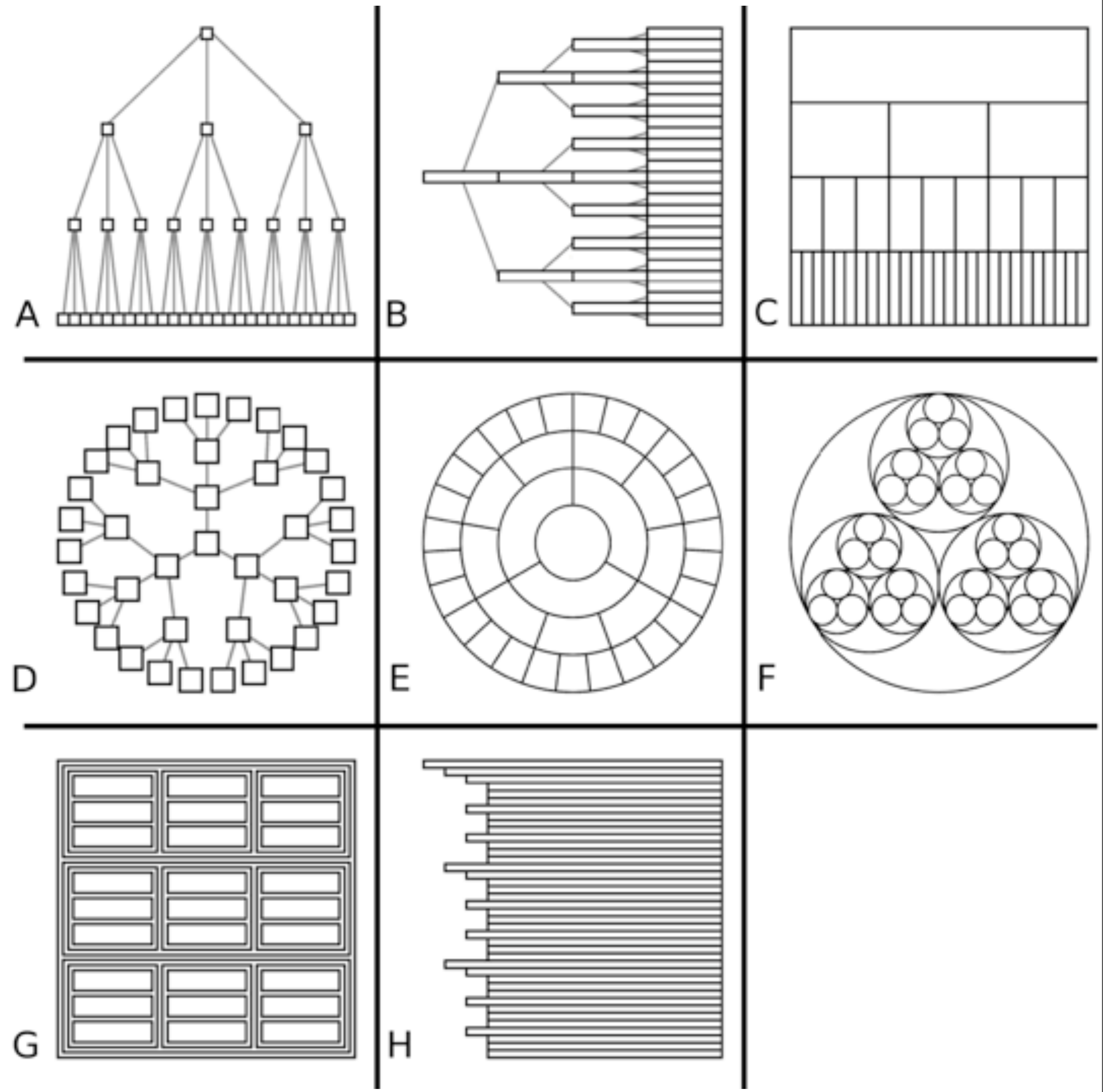


- spatial layout
  - orientation of spatial axes
- limitations studied
  - rectilinear: scalability wrt #axes
    - 2 axes best
    - 3 problematic
      - see Visualization Principles talk
    - 4+ impossible
  - radial: perceptual limits
    - angles lower precision than lengths

[Uncovering Strengths and Weaknesses of Radial Visualizations - an Empirical Approach. Diehl, Beck and Burch. IEEE TVCG (Proc. InfoVis) 16(6):935--942, 2010.]

Analysis examples: Tree drawing

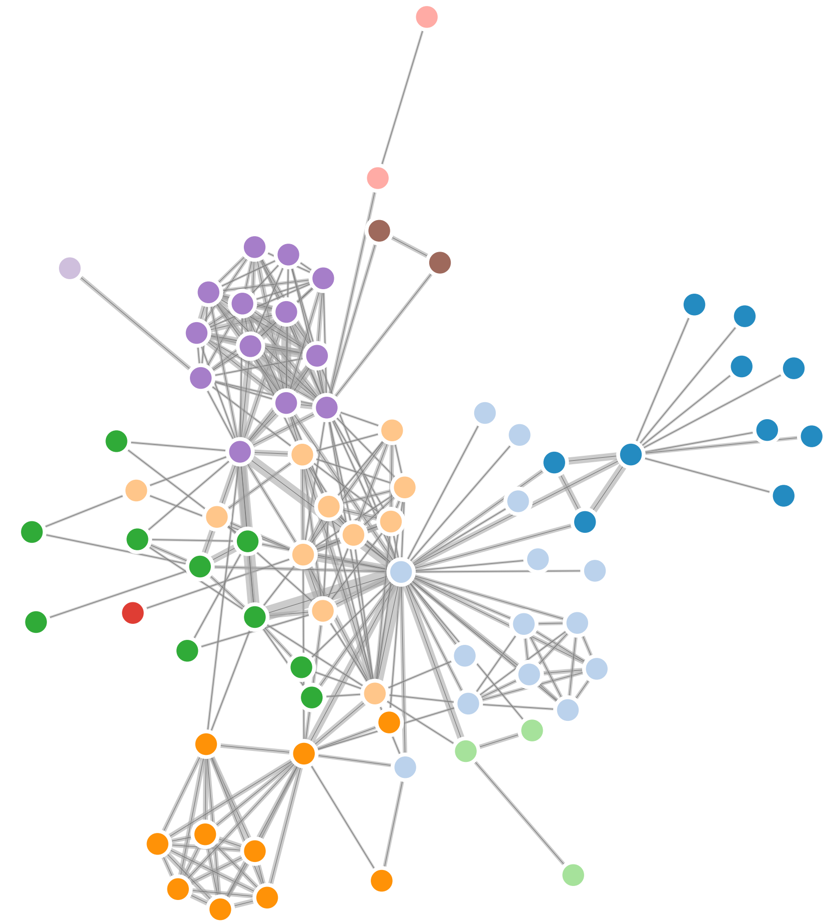
- data shown
  - link relationships
  - tree depth
  - sibling order
- methods
  - connection vs containment link marks
  - rectilinear vs radial layout
  - spatial position channels
- considerations
  - redundant? arbitrary?
  - information density?
    - avoid wasting space



[Quantifying the Space-Efficiency of 2D Graphical Representations of Trees. McGuffin and Robert. Information Visualization 9:2 (2010), 115–140.]

Analysis example: force-directed placement

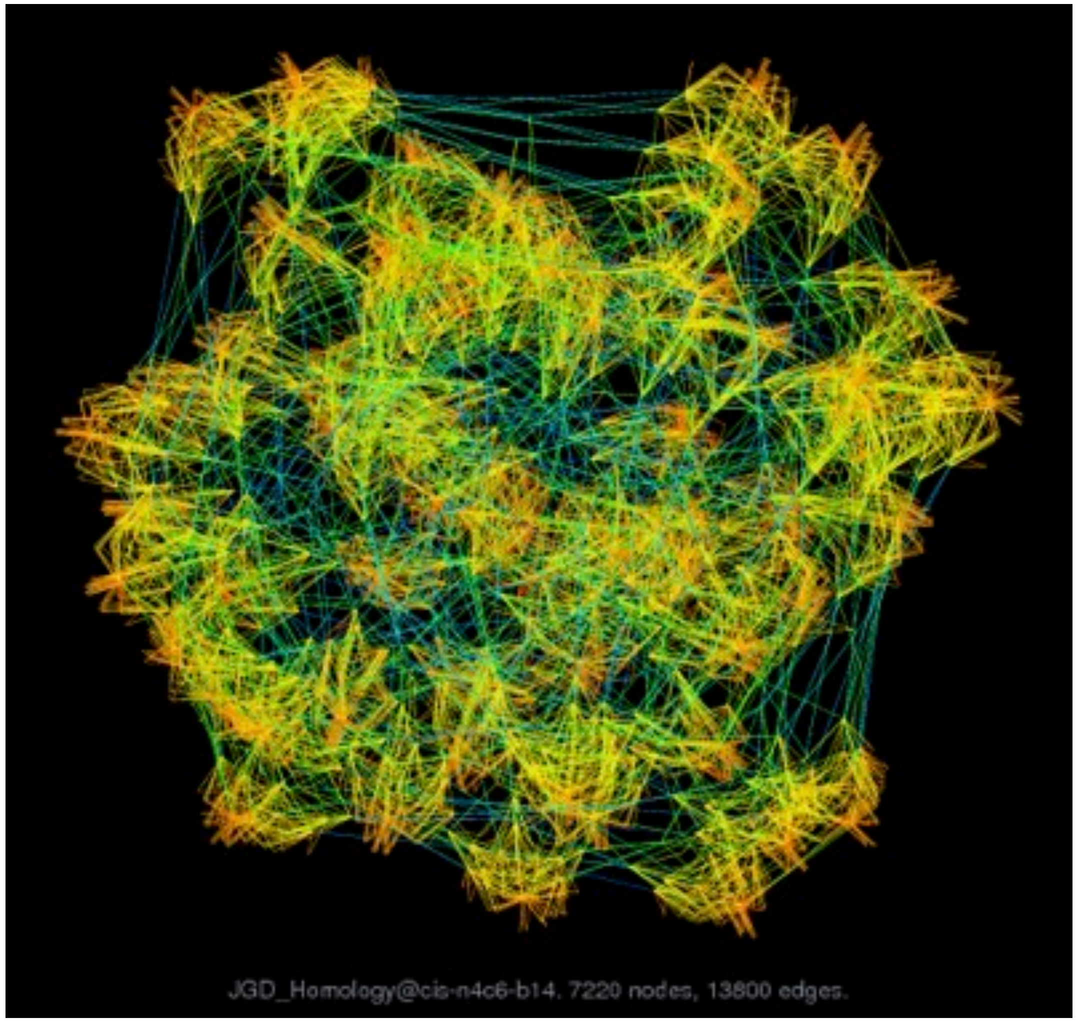
- visual encoding
  - link connection marks
  - node point marks
- considerations
  - spatial position: no meaning directly encoded
    - left free to minimize crossings
  - proximity semantics?
    - sometimes meaningful
    - sometimes arbitrary, artifact of layout algorithm
    - tension with length
      - long edges more visually salient than short



[http://mbostock.github.com/d3/ex/force.html]

Analysis example: multi-level FDP (sfdp)

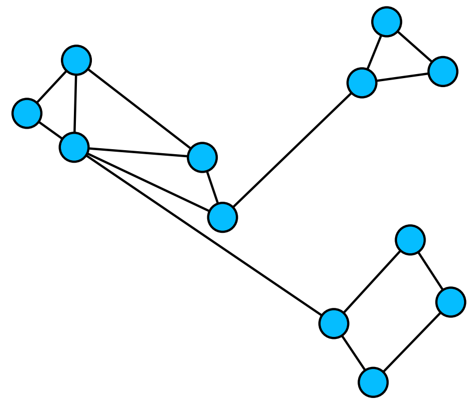
- data
  - original: network
  - derived: cluster hierarchy atop it
- visual encoding
  - same: link connection marks
- considerations
  - better algorithm for same encoding technique
    - same: fundamental use of space
    - hierarchy used in algorithm but not shown explicitly

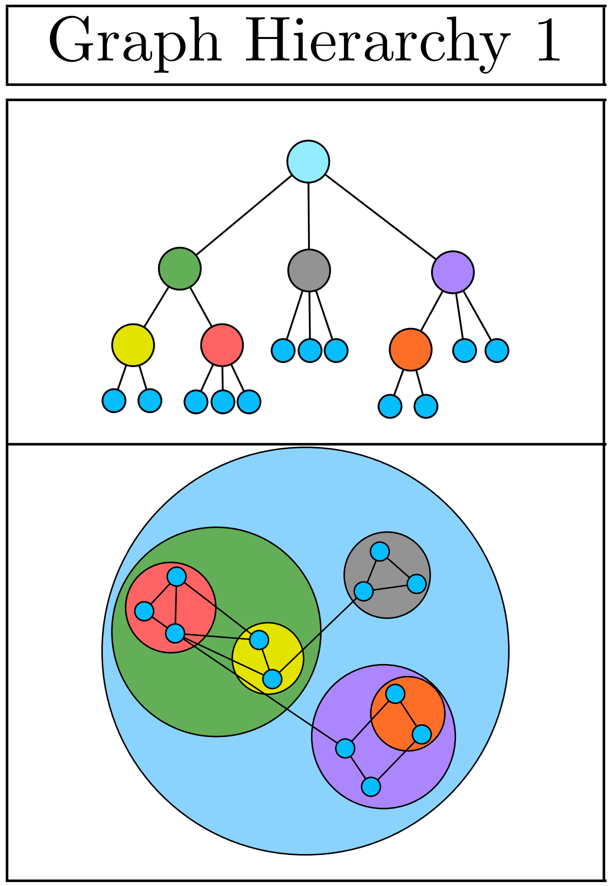


[Efficient and high quality force-directed graph drawing. Hu. The Mathematica Journal 10:37–71, 2005.]

Analysis example: GrouseFlocks

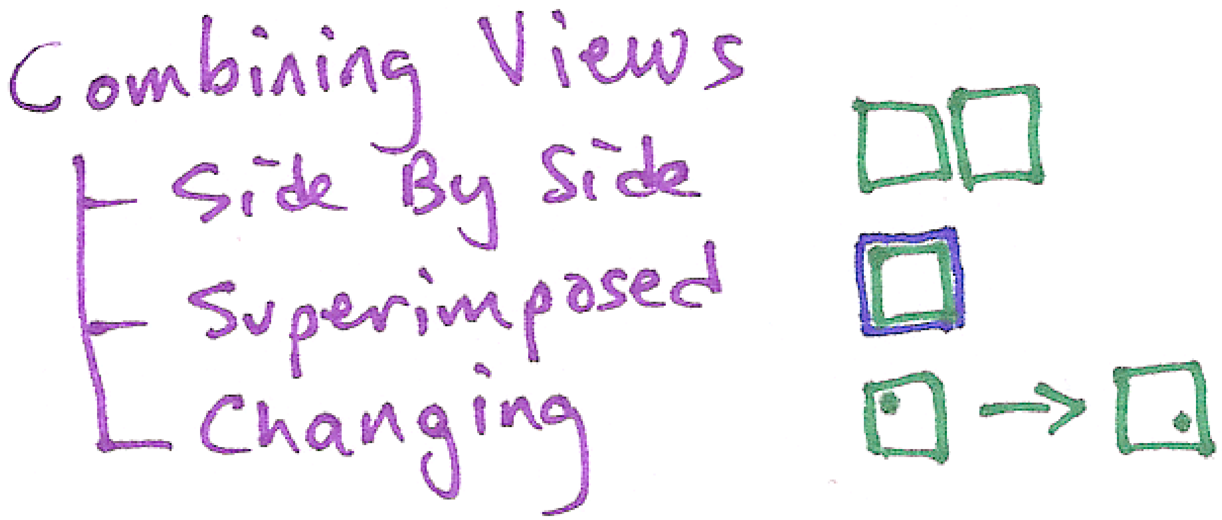
- data: compound graphs
  - network
  - cluster hierarchy atop it
    - derived or interactively chosen
- visual encoding
  - connection marks for network links
  - containment marks for hierarchy
  - point marks for nodes
- dynamic interaction
  - select individual metanodes in hierarchy to expand/contract





[GrouseFlocks: Steerable Exploration of Graph Hierarchy Space. Archambault, Munzner, and Auber. IEEE TVCG 14(4): 900-913, 2008.]

Multiple views vs single views



- powerful method: use multiple views side by side
  - vs. superimposing multiple views as layers atop each other
    - all must have shared spatial layout
  - vs. single view that changes over time
    - as with interactive navigation
- principle: eyes beat memory
  - easy to compare by moving eyes between side-by-side views
    - harder to compare visible item to memory of what you saw
  - external cognition vs. internal working memory limits

Further analysis examples

Analysis example: Cerebral

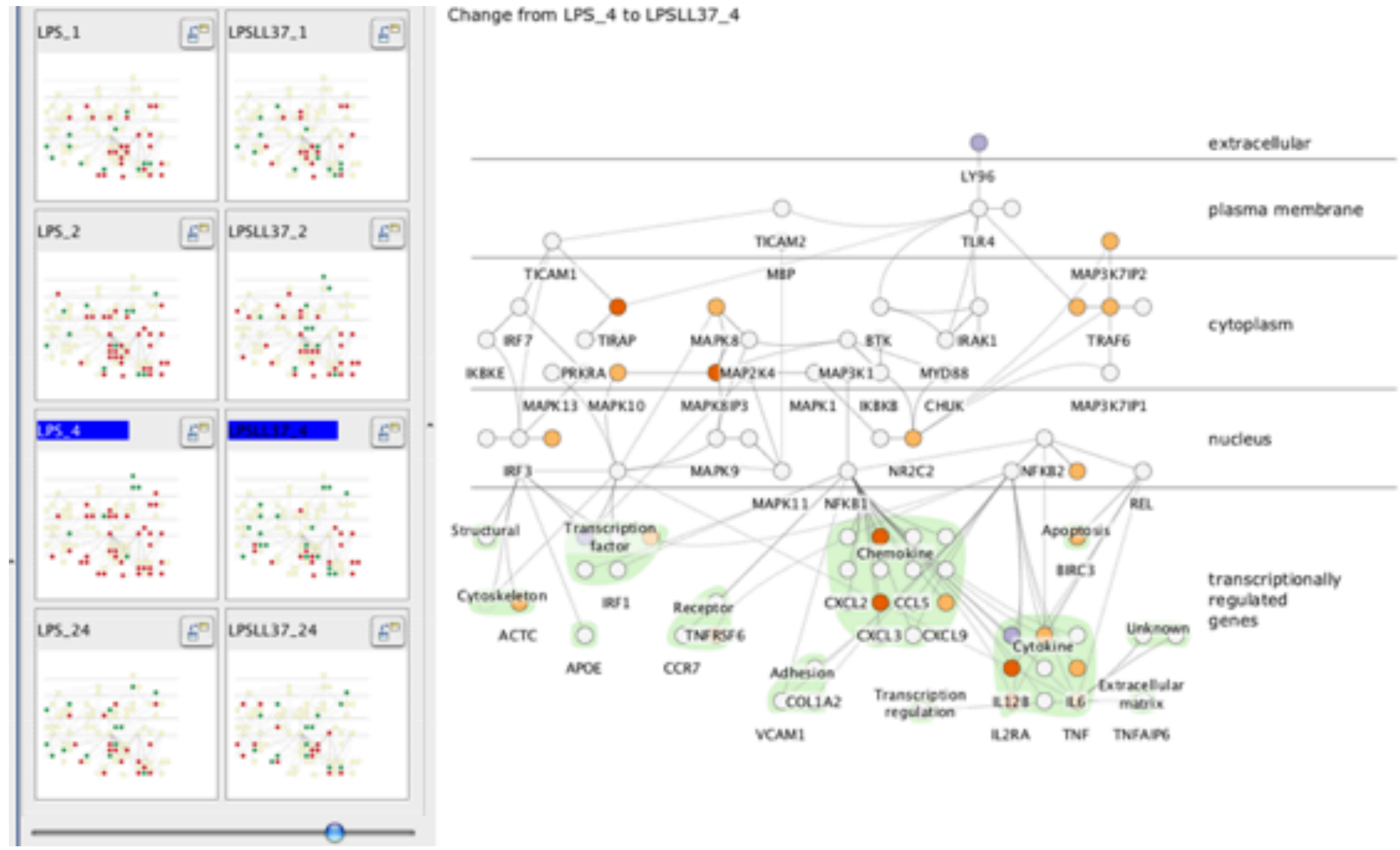
- data
  - network
    - nodes: genes, links: known interaction
    - per-node attribs
      - location within cell where interaction occurs
      - biological function
  - table
    - 1 quant attrib: gene expression level
    - indexed by 2 categ attribs: node/gene, experimental condition

[Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, and Kincaid. IEEE TVCG (Proc. InfoVis) 14(6): 1253-1260, 2008.]

[Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. Barsky, Gardy, Hancock, and Munzner. Bioinformatics 23(8):1040-1042, 2007.]

Use of space: Cerebral

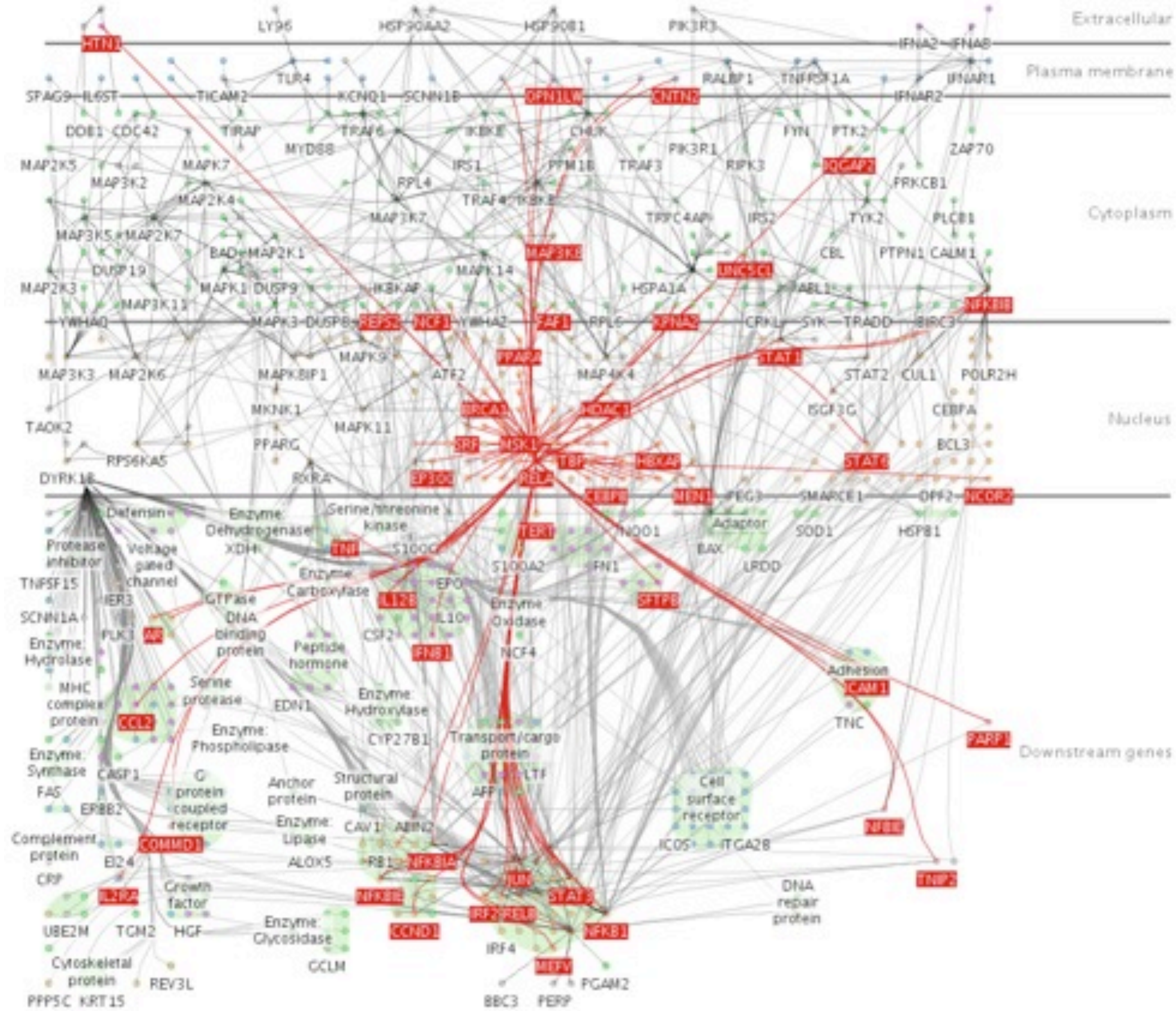
- side by side views
  - *small multiples*
    - same encoding, different data
  - separate into regions
    - each shows entire network
    - color nodes by quant attrib for condition



[Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, and Kincaid. IEEE TVCG (Proc. InfoVis) 14(6):1253-1260, 2008.]

Use of space: Cerebral

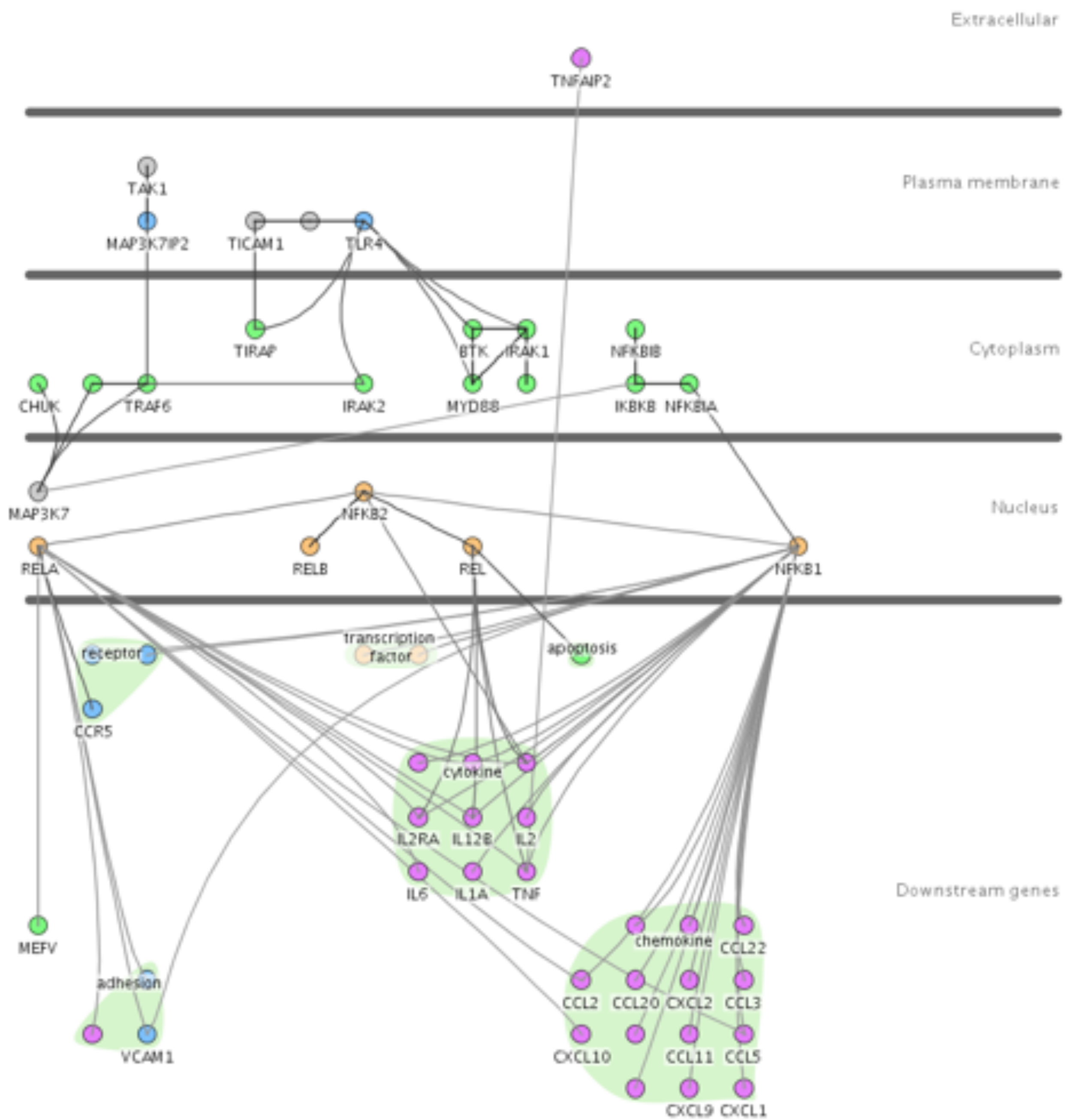
- superimposed layers within each view
  - dynamic interaction technique
- highlight 1-hop neighbors on mouseover
  - foreground layer distinguished by color



Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. Barsky, Gardy, Hancock, and Munzner. Bioinformatics 23(8):1040-1042, 2007.]

Use of space: Cerebral

- network visual encoding
  - consideration
    - mimic stylized spatial semantics of hand-drawn diagrams
  - marks: connection for links
  - spatial channels
    - separate into regions according to subcellular location attrib
    - order regions vert by attrib
    - in bottom region: also separate into subregions by function attrib



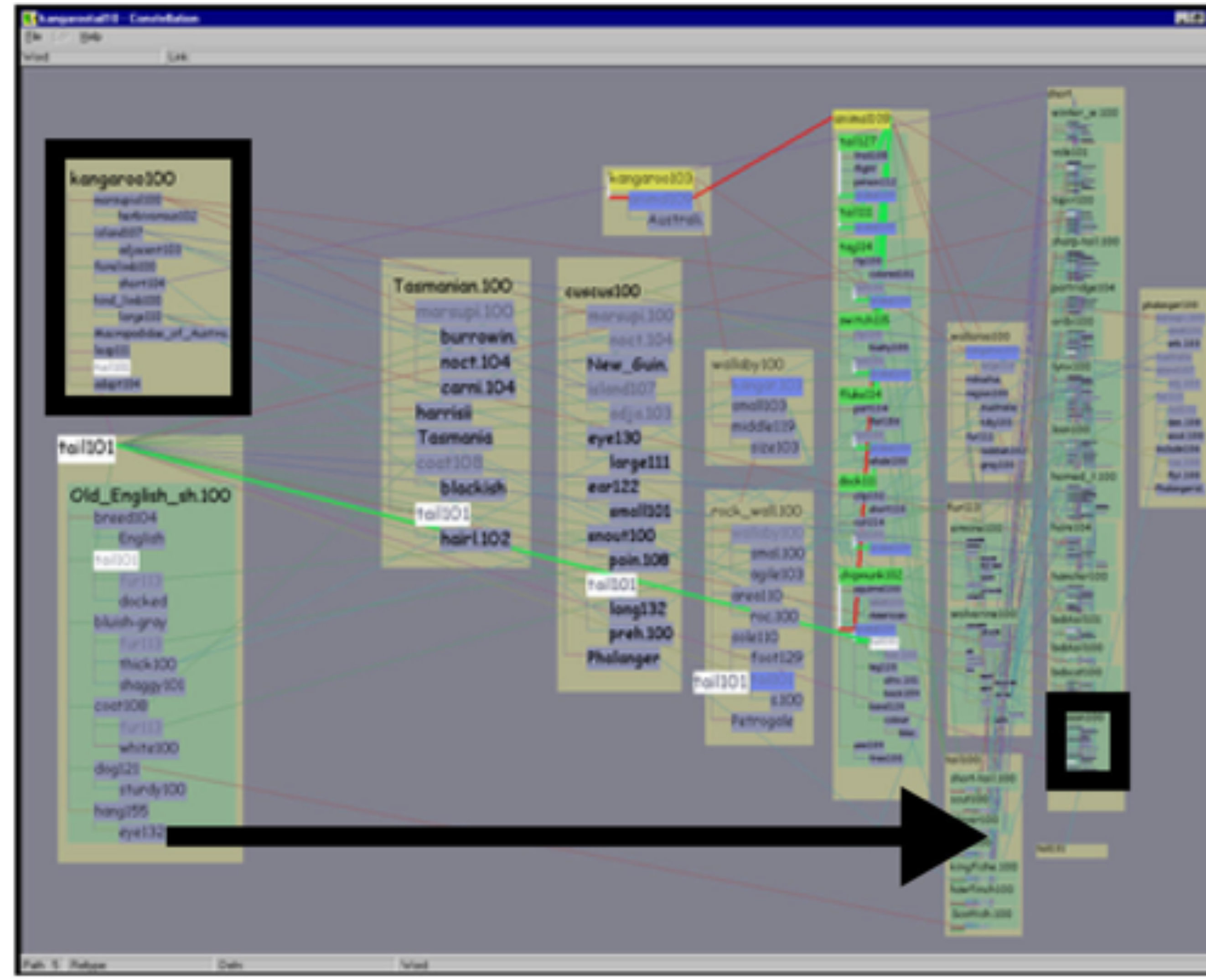
Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. Barsky, Gardy, Hancock, and Munzner. Bioinformatics 23(8):1040-1042, 2007.]

Considerations: Cerebral

- explicit discussion of choices for use of space
  - design motivated by analysis of previous work
  - justified as more suitable than characterized alternatives
    - changing single view with animation: avoided
      - cognitive load
      - hard to track changes across many conditions and many nodes
    - separating into one region per gene: avoided
      - information density
      - not enough space to show multiple attribs within node for big networks
      - enough space to show multiple networks with single mark per node
        - » separating into one region per condition: chosen
  - spatial position: partially constrained

Analysis example: Constellation

- data
  - multi-level network
    - node: word
    - link: words used in same dictionary definition
    - subgraph for each definition
      - not just hierarchical clustering
  - paths through network
    - query for high-weight paths between 2 nodes
      - quant attrib: plausibility

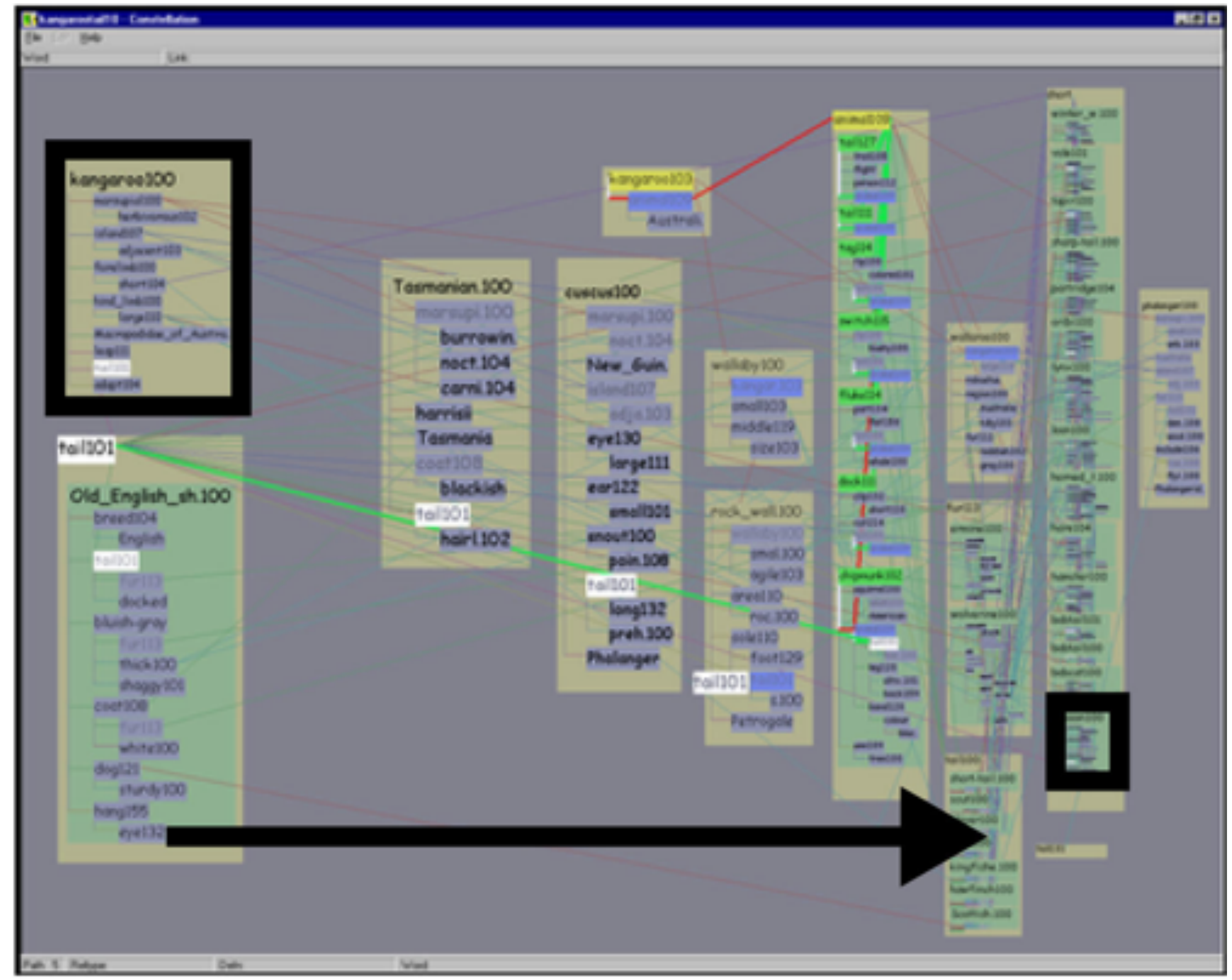


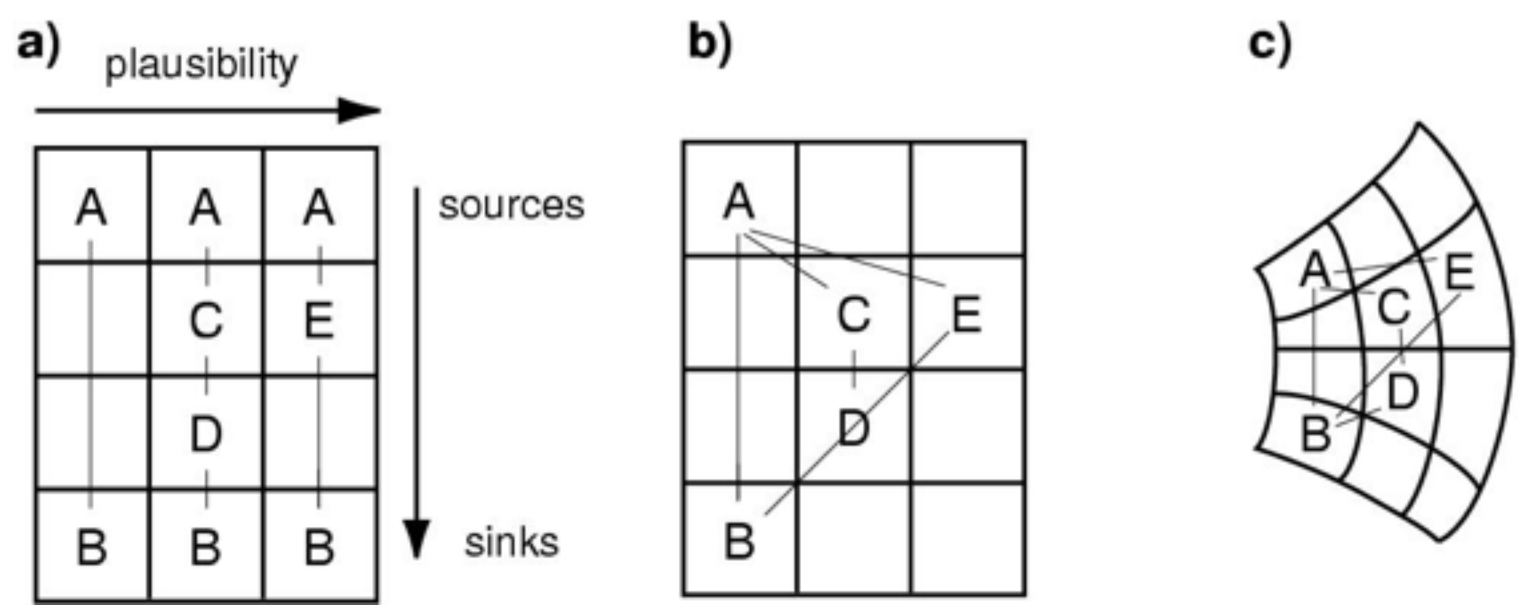
[Interactive Visualization of Large Graphs and Networks. Munzner. Ph.D. Dissertation, Stanford University, June 2000.]

[Constellation: A Visualization Tool For Linguistic Queries from MindNet. Munzner, Guimbretière and Robertson. Proc. IEEE Symp. InfoVis 1999, p.132-135.]

Using space: Constellation

- visual encoding
  - link connection marks between words
  - link containment marks to indicate subgraphs
  - encode plausibility with horiz spatial position
  - encode source/sink for query with vert spatial position
- spatial layout
  - curvilinear grid: more room for longer low-plausibility paths

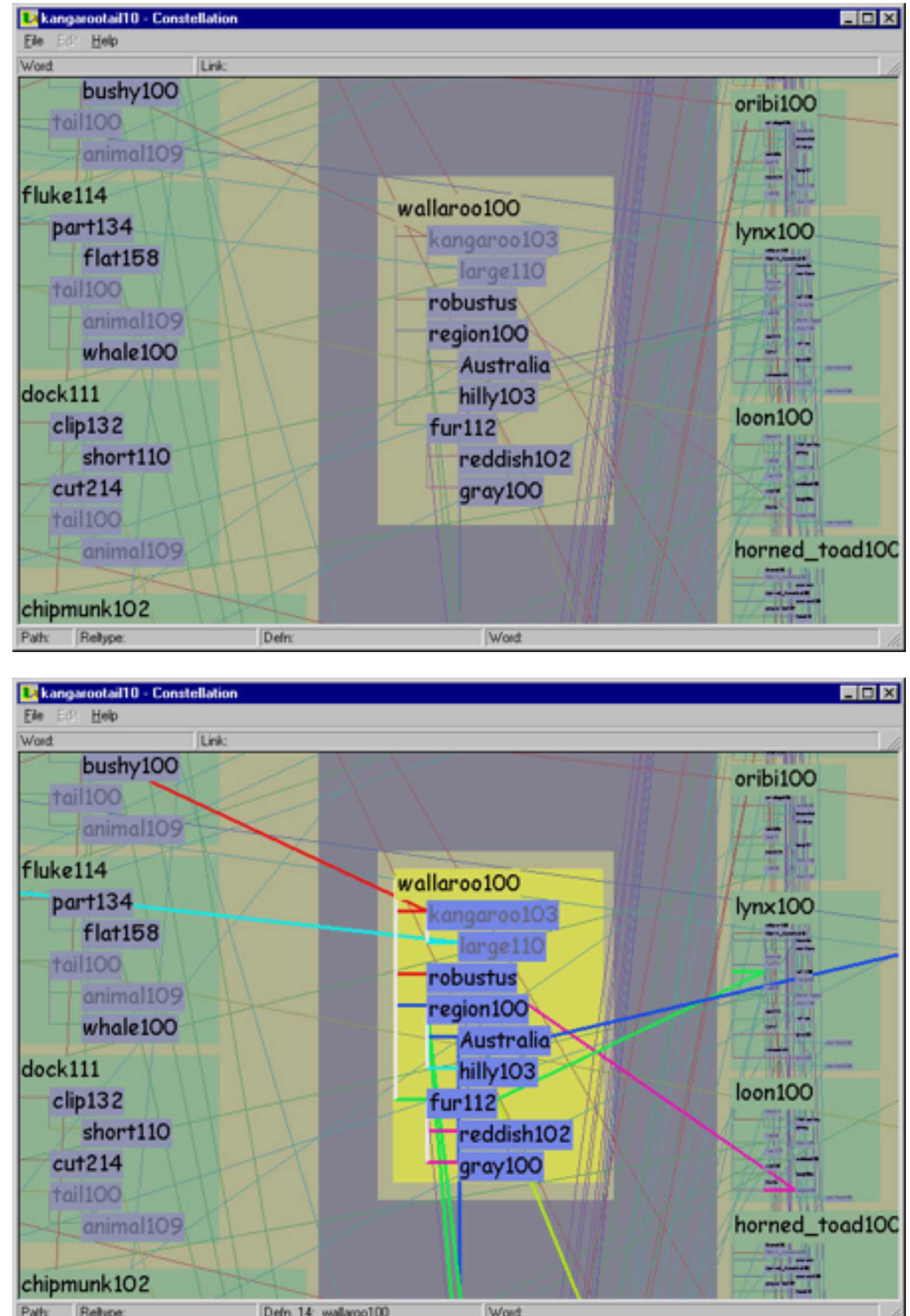




[Interactive Visualization of Large Graphs and Networks. Munzner. Ph.D. Dissertation, Stanford University, June 2000.]

Using space: Constellation

- edge crossings
  - cannot easily minimize instances, since position constrained by spatial encoding
  - instead: minimize perceptual impact
- views: superimposed layers
  - dynamic foreground/background layers on mouseover, using color
  - four kinds of constellations
    - definition, path, link type, word
      - not just 1-hop neighbors



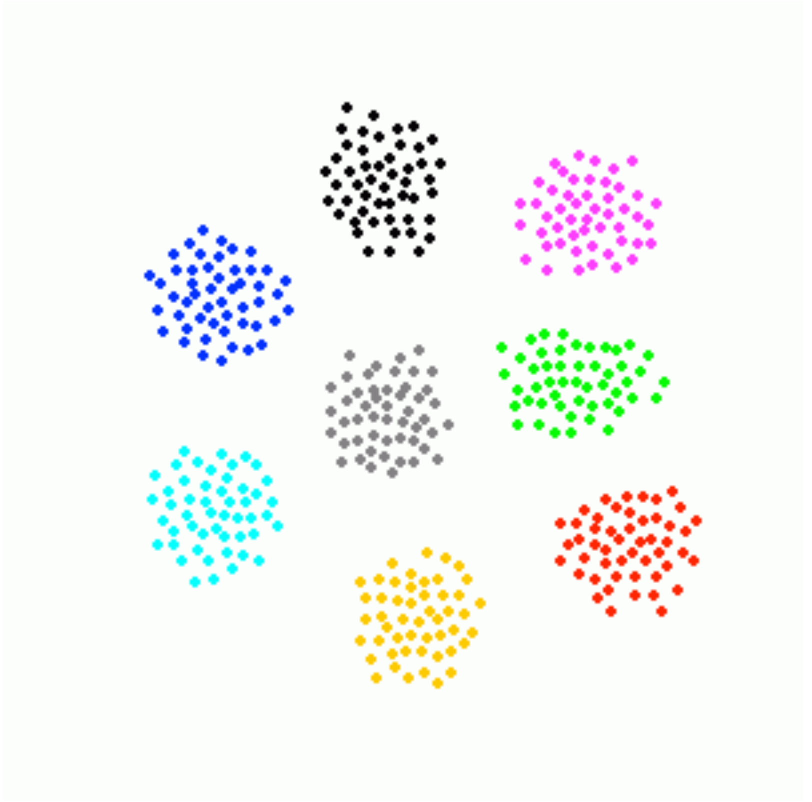
[Interactive Visualization of Large Graphs and Networks. Munzner. Ph.D. Dissertation, Stanford University, June 2000.]

Considerations: Constellation

- another example of design motivated by analysis
  - explicit discussion of choices using space
    - spatial position: highly constrained
    - tradeoffs
      - information density vs spatial encoding semantics
        - » covered elsewhere: iterative refinement of layout
          [Interactive Visualization of Large Graphs and Networks. Munzner. Ph.D. Dissertation, Stanford University, June 2000.]
      - crossings: instances vs salience

Analysis example: Noack LinLog

- energy model designed to reveal clusters in data
  - requires that edges between clusters longer than those within
    - visual encoding technique
  - using same minimization algorithms as previous work
- considerations
  - also design motivated by prior analysis
    - explicit discussion of technique-level issues in GD literature
      - encourage more papers like this!



[An Energy Model for Visual Graph Clustering. Noack. Proc. Graph Drawing 2003, p. 425--436.

Conclusions

Vis methods analysis framework

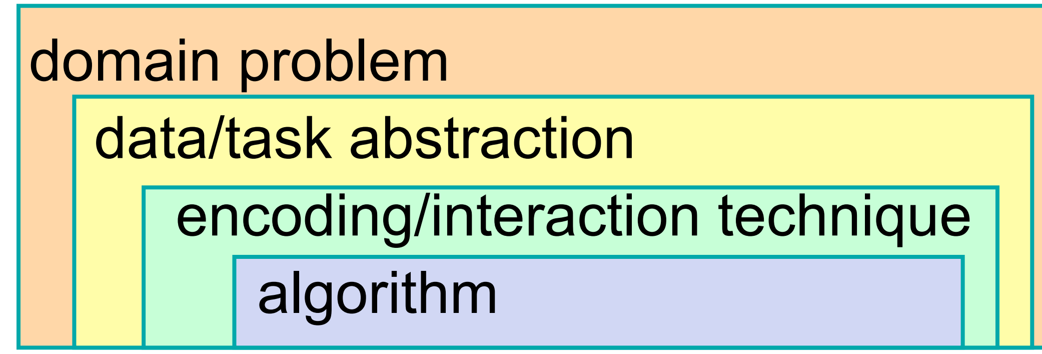
- characterize techniques in terms of methods for using space
    - marks and channels
        - marks for nodes vs marks for links
    - space channel: express, separate, order, align
        - position, proximity, partitioning into groups
- general way to analyze visualizations systematically
    - applied to graph drawing examples in particular

Framework goals

- guide development of new algorithms/techniques
  - in same spirit as examples shown
    - Cerebral, Constellation, LinLog Energy

- characterize existing algorithms/techniques
  - can guide adoption
    - in what context are they suitable?
      - context here: previous design levels

Mapping upwards



- from algorithms to techniques
  - sometimes trivial
    - discussion in paper itself
    - direct citation of previous work for framing context
  - sometimes tricky indeed
    - when algorithm description does not facilitate analysis of resulting visual encoding
      - use for space, or other channels
    - line between algorithm and technique can be blurry
      - does new algorithm support existing technique, or new one?
        - » trivial when speed increase for identical visual results
- from techniques to abstractions to domain problems
  - equally important questions, but beyond scope for today...

Framework goals

- guide development of new algorithms/techniques
  - in same spirit as examples shown
    - Cerebral, Constellation, LinLog Energy

- characterize existing algorithms/techniques
  - can guide adoption
    - in what context are they suitable?
      - context here: previous design levels

- vis methods analysis only one possible route!
  - many others
    - benchmarks, computational complexity, user studies...

More information

- this talk
  http://www.cs.ubc.ca/~tmm/talks.html#gd13

- more on analysis
  - techniques/methods in more depth
  - also, principles and abstractions!

    - single chapter in 2009 Fundamentals of Graphics textbook *Visualization*
      http://www.cs.ubc.ca/~tmm/papers.html#akpchapter

    - full vis textbook: to appear, 2014, AK Peters
      - *Visualization Analysis and Design: Principles, Abstractions, and Methods*