

Ch 15: Analysis Case Studies

Paper: Algebraic

Tamara Munzner

Department of Computer Science
University of British Columbia

CPSC 547, Information Visualization

Lecture 15: 10 November 2015

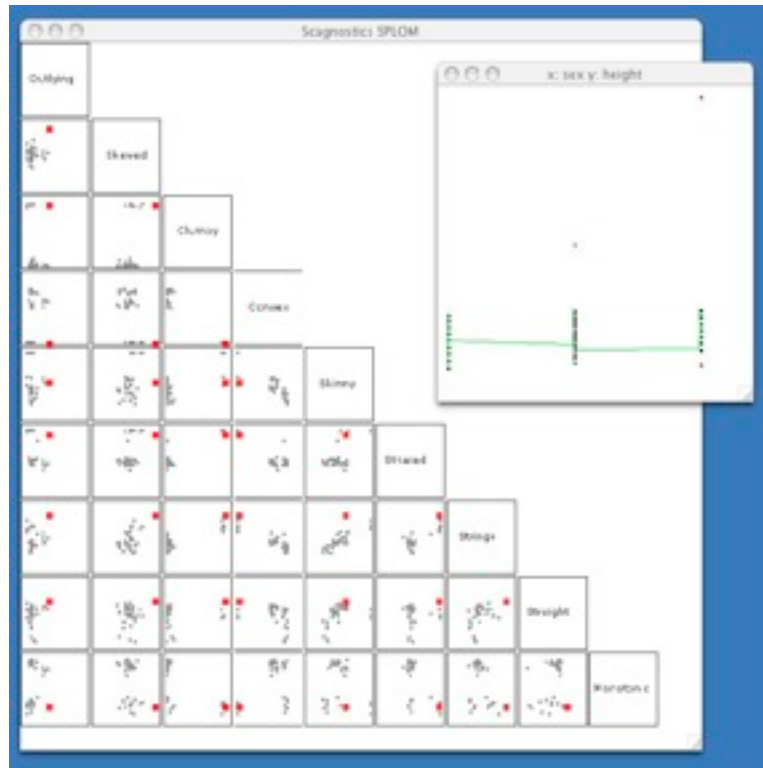
<http://www.cs.ubc.ca/~tmm/courses/547-15>

News

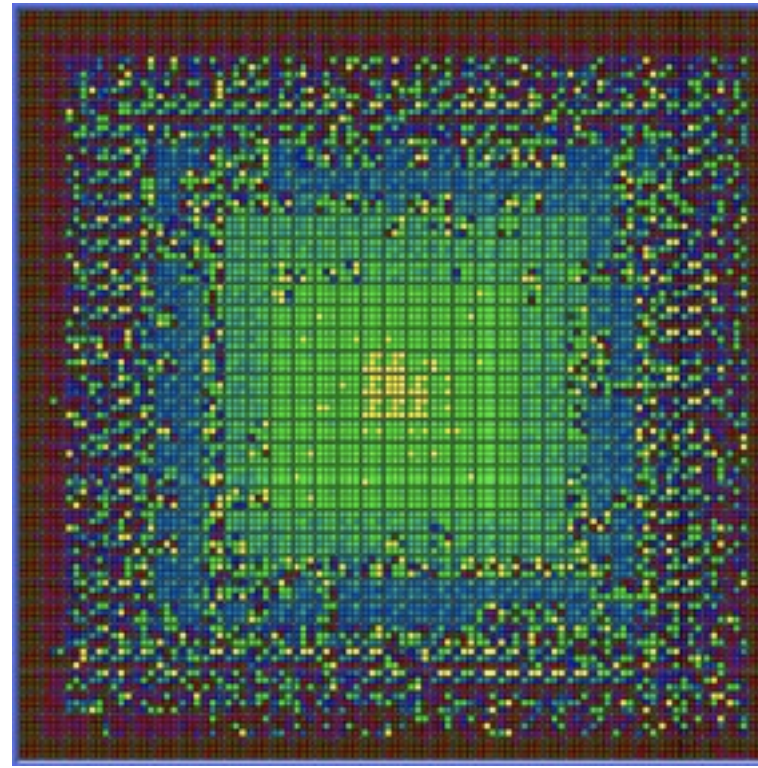
- presentation topics/papers/dates posted

Analysis Case Studies

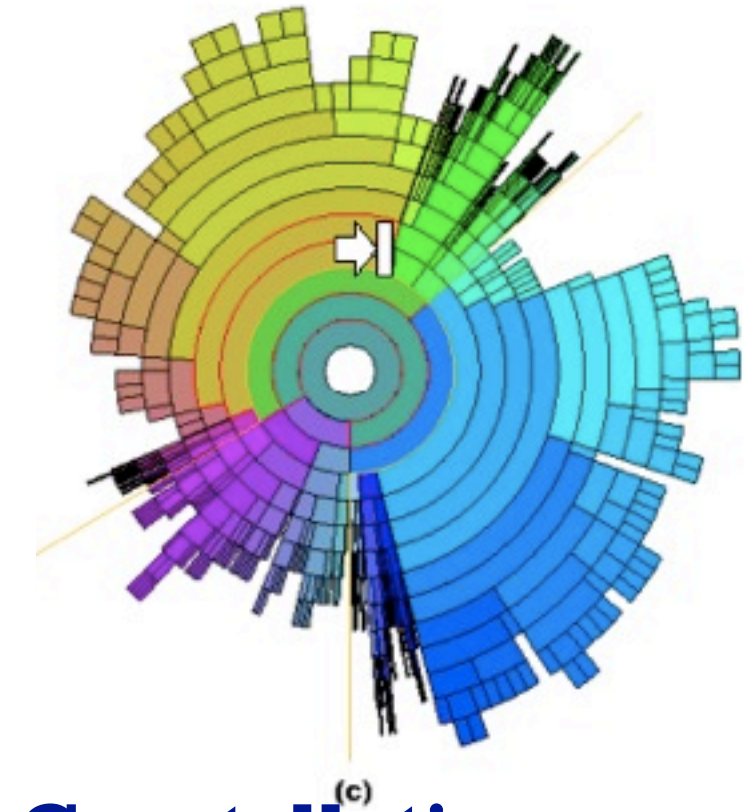
Scagnostics



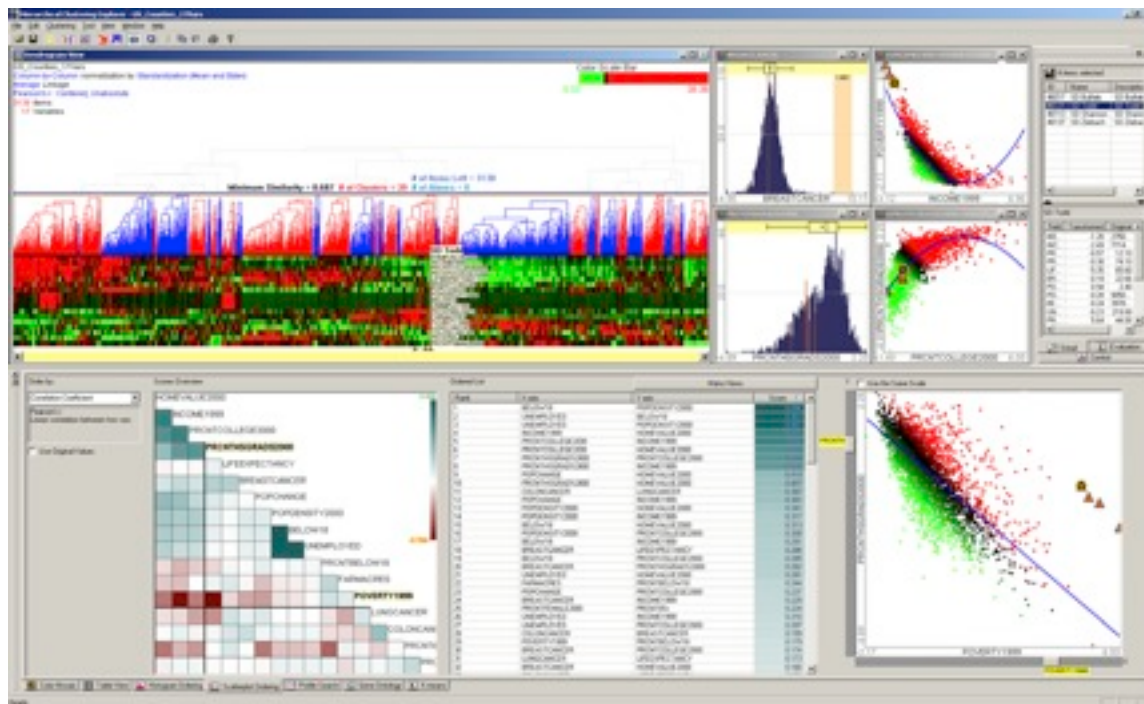
VisDB



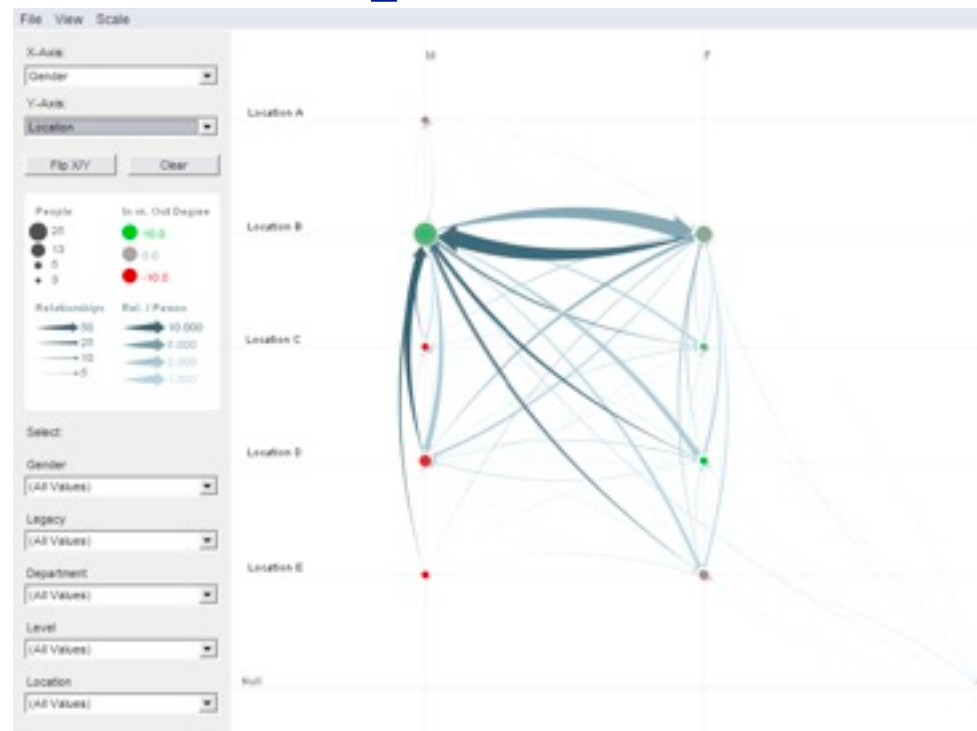
InterRing



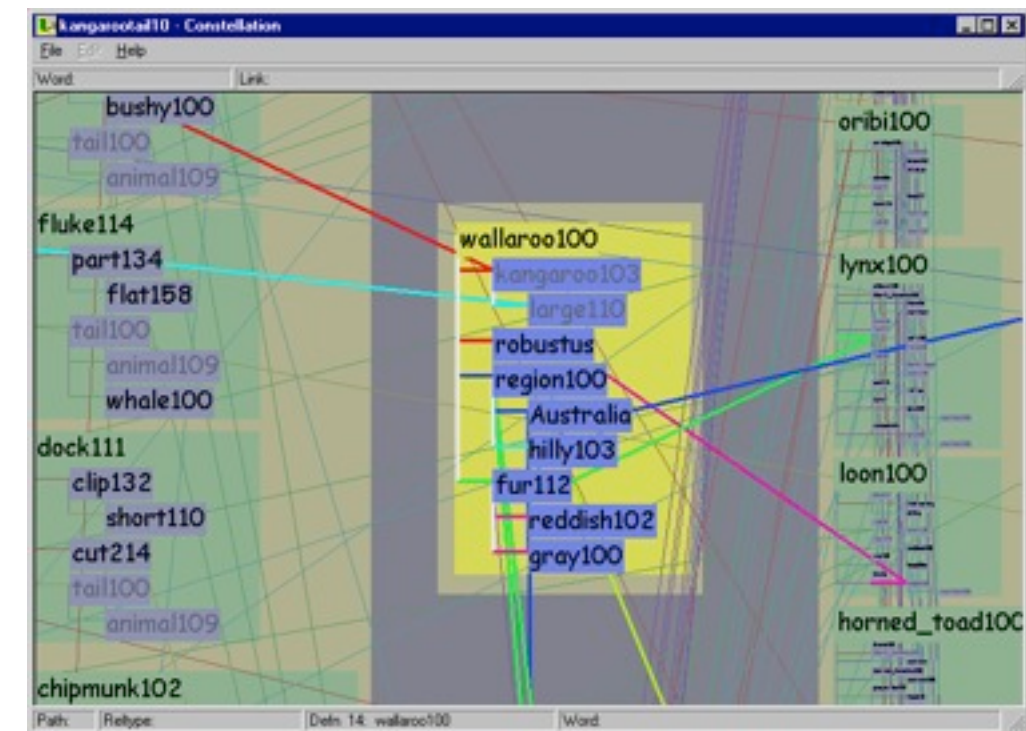
HCE



PivotGraph

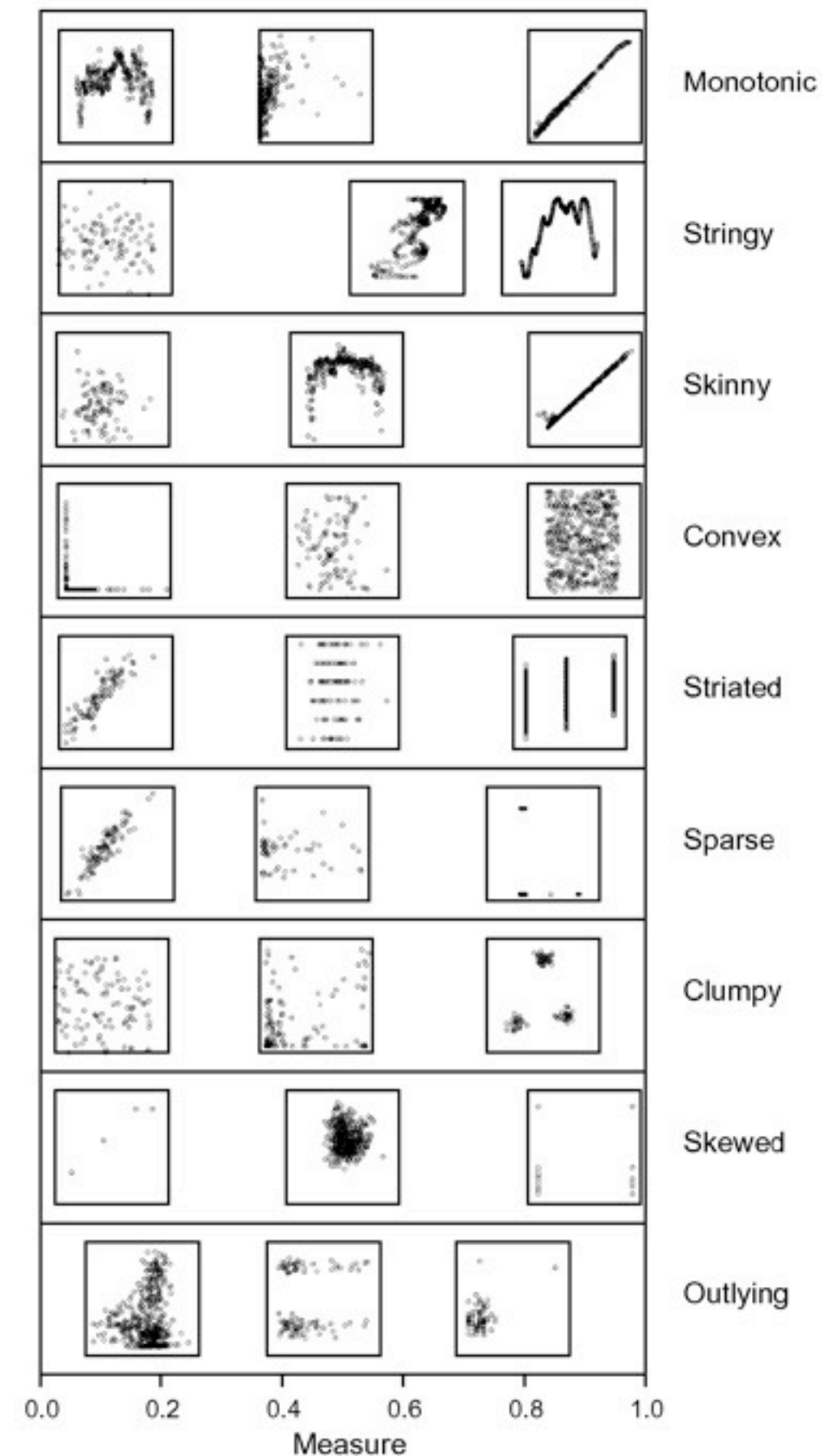
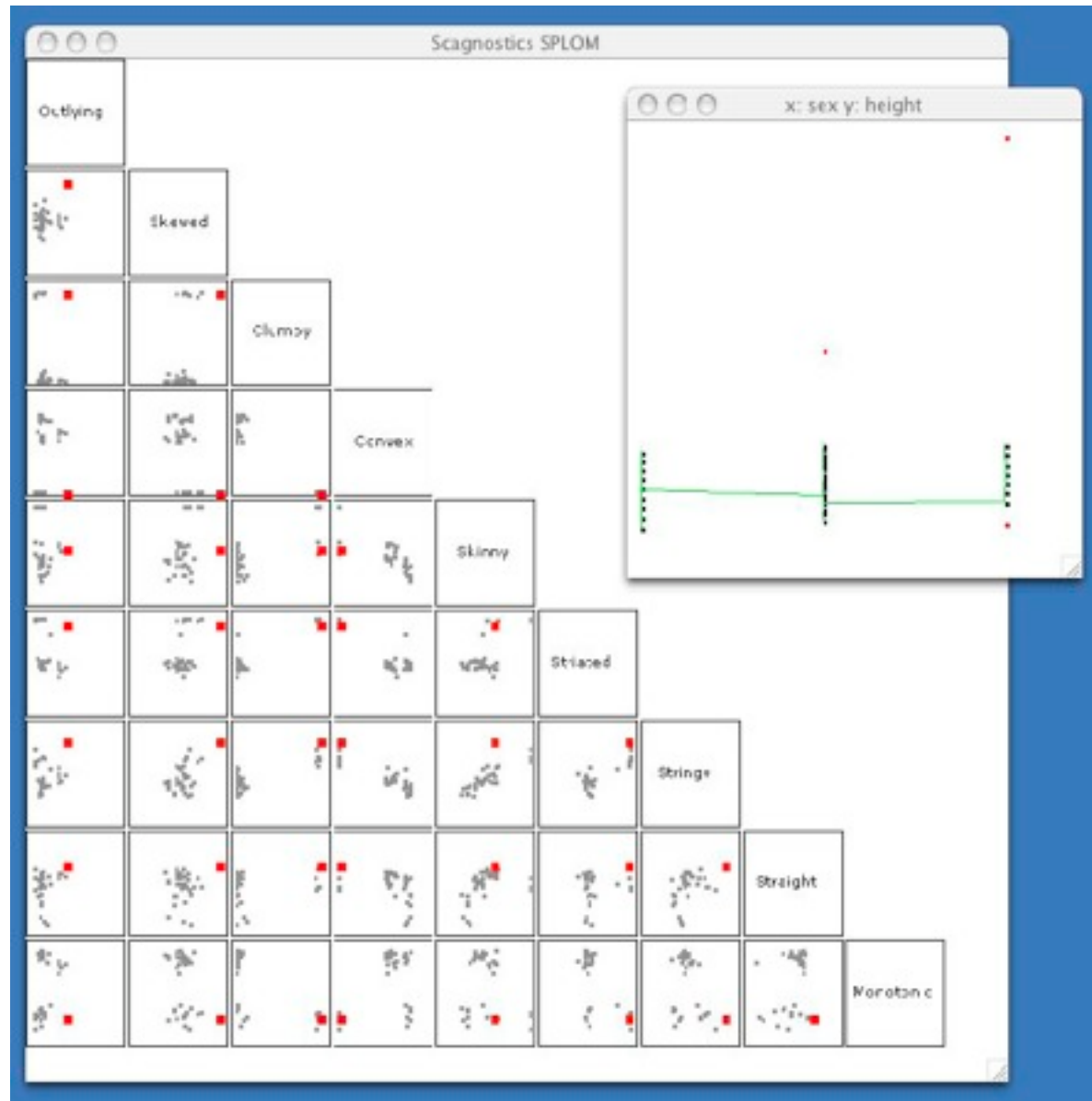


Constellation



Graph-Theoretic Scagnostics

- scatterplot diagnostics
 - scagnostics SPLOM: each point is one original scatterplot



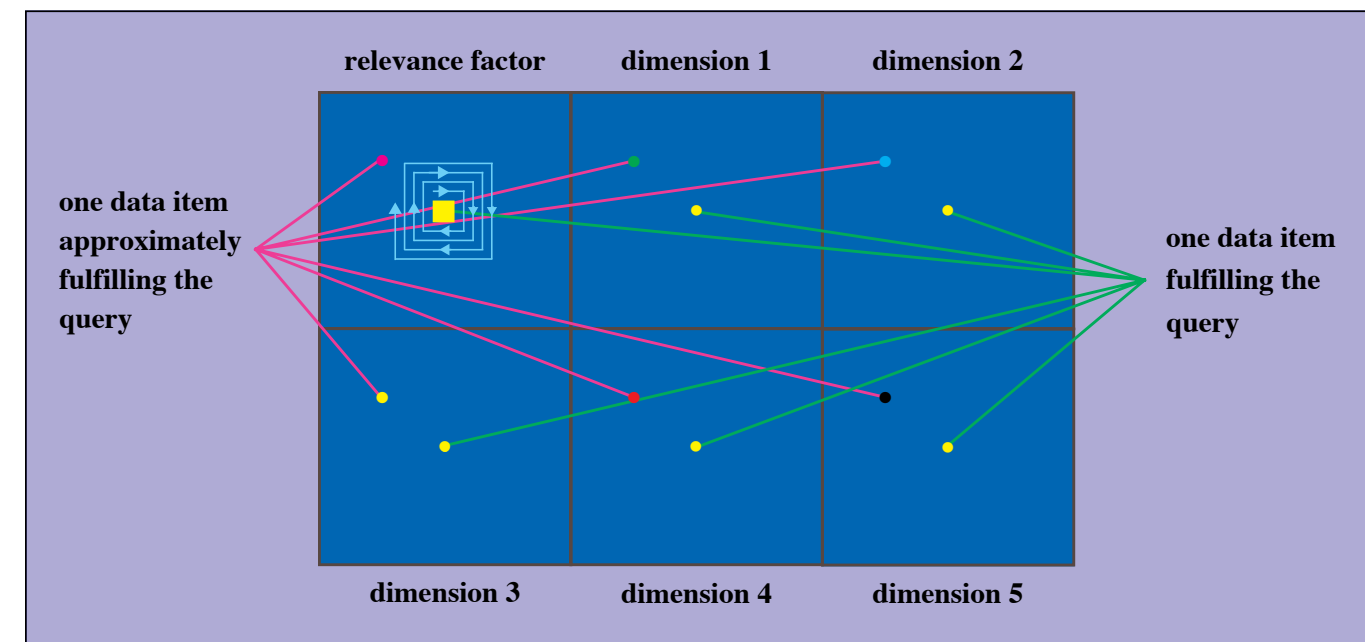
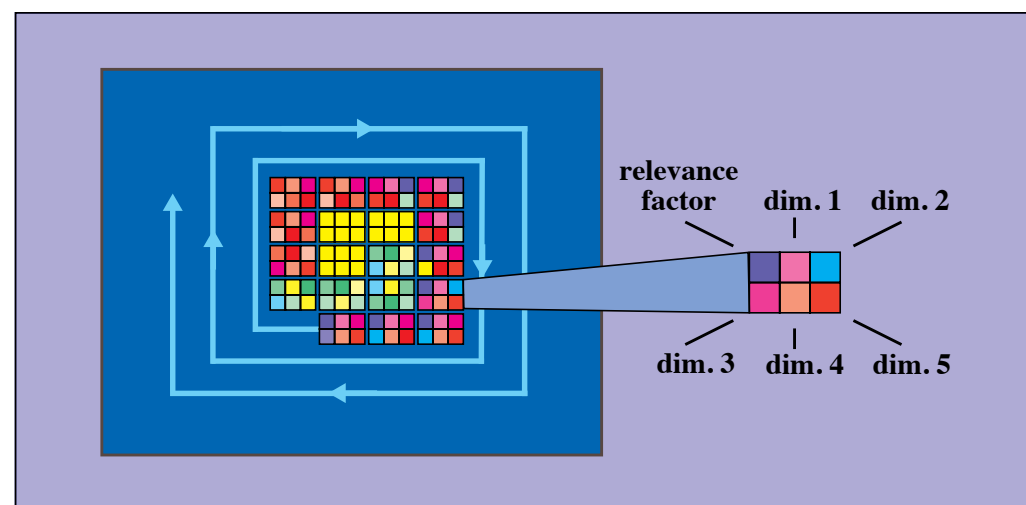
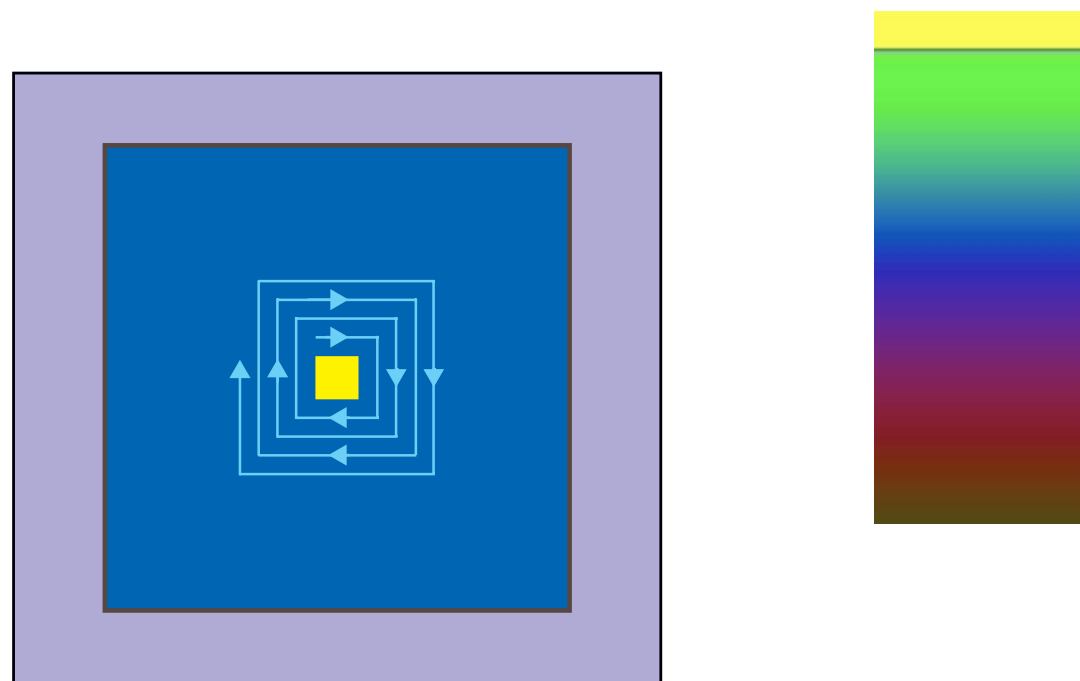
[Graph-Theoretic Scagnostics Wilkinson, Anand, and Grossman. Proc InfoVis 05.]

Scagnostics analysis

System	Scagnostics
What: Data	Table.
What: Derived	Nine quantitative attributes per scatterplot (pairwise combination of original attributes).
Why: Tasks	Identify, compare, and summarize; distributions and correlation.
How: Encode	Scatterplot, scatterplot matrix.
How: Manipulate	Select.
How: Facet	Juxtaposed small-multiple views coordinated with linked highlighting, popup detail view.
Scale	Original attributes: dozens.

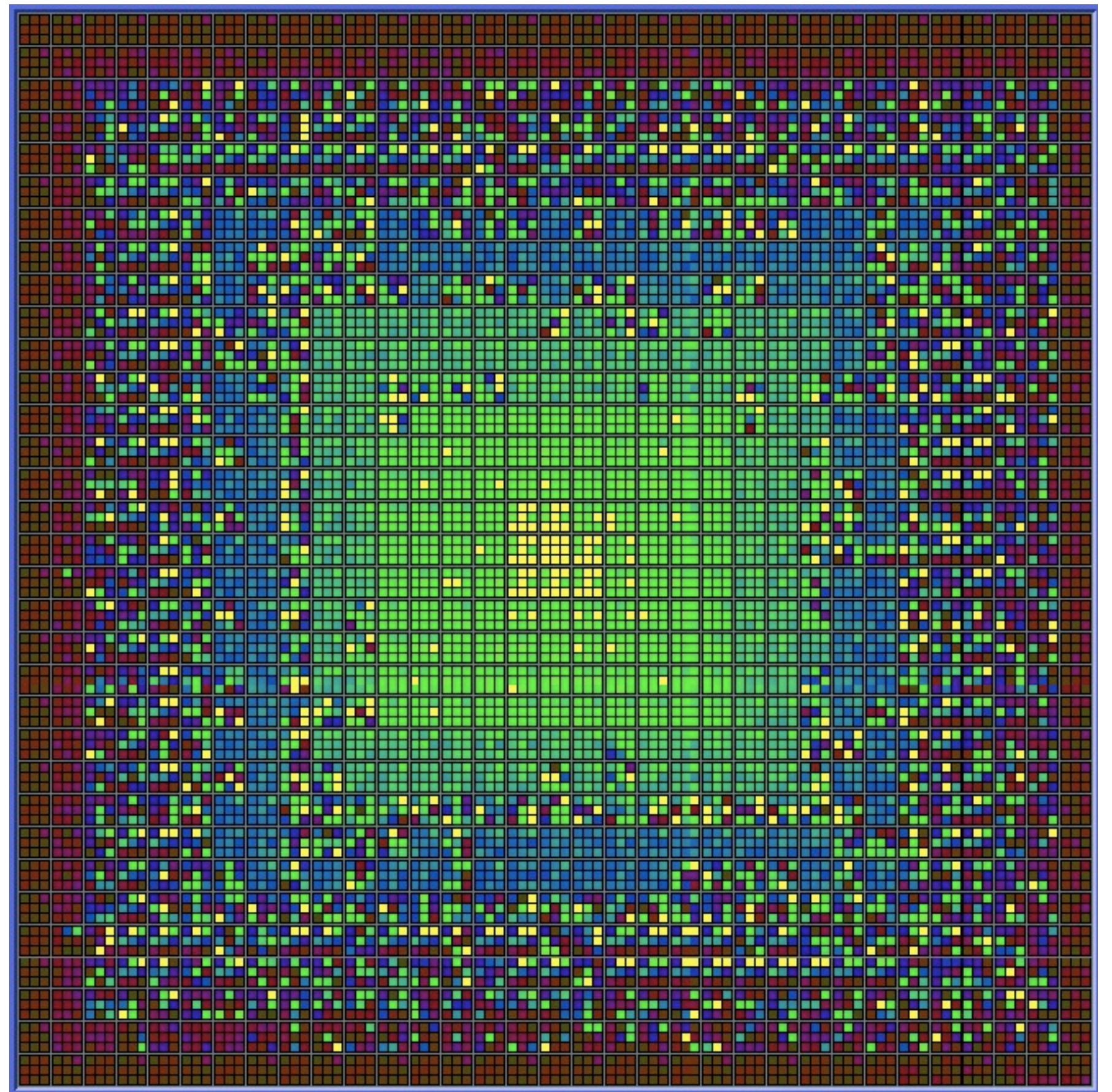
VisDB

- table: draw pixels sorted, colored by relevance
- group by attribute or partition by attribute into multiple views



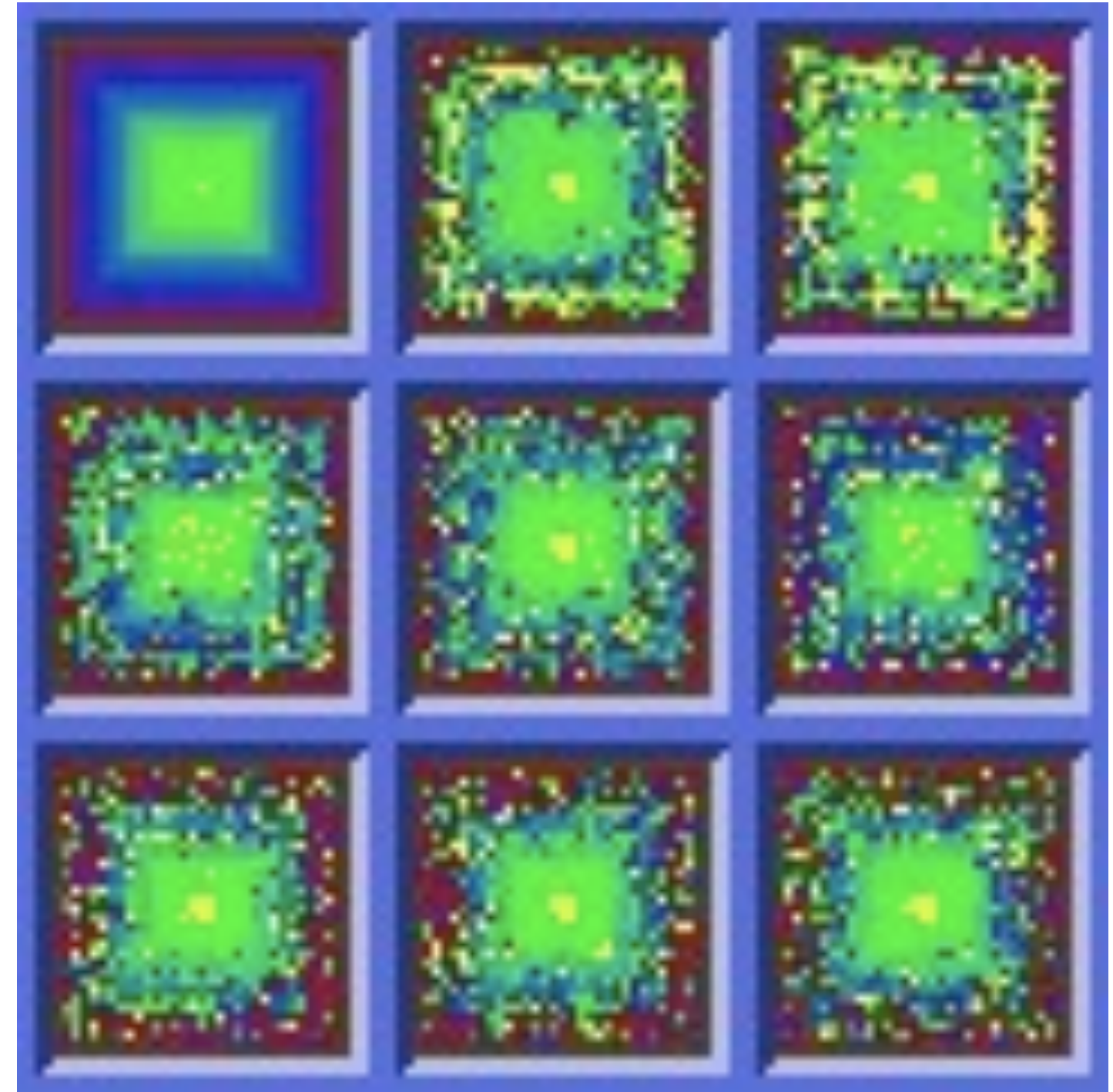
VisDB Results

- partition into many small regions: dimensions grouped together



VisDB Results

- partition into small number of views
 - inspect each attribute

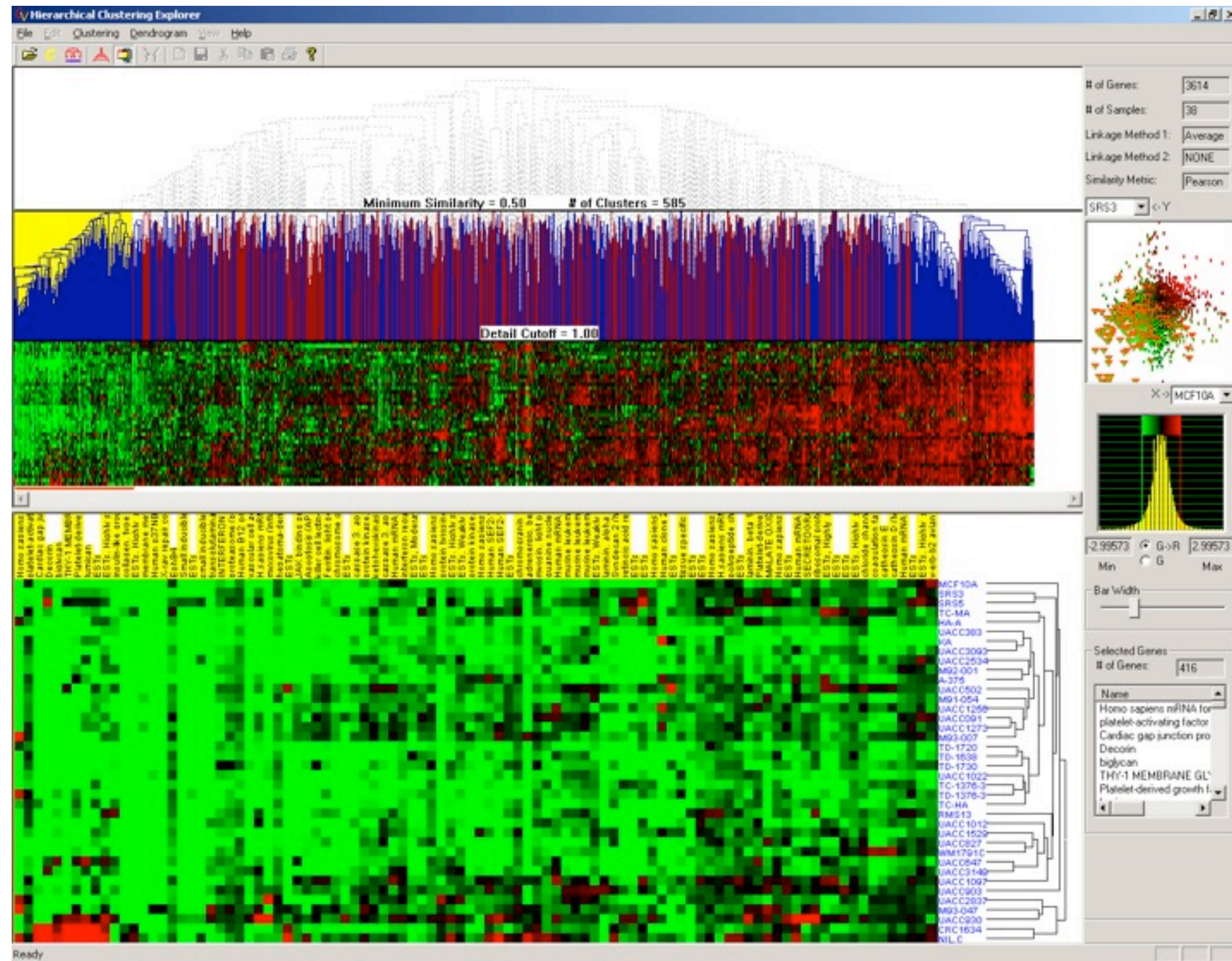


VisDB Analysis

System	VisDB
What: Data	Table (database) with k attributes; query returning table subset (database query).
What: Derived	$k + 1$ quantitative attributes per original item: query relevance for the k original attributes plus overall relevance.
Why: Tasks	Characterize distribution within attribute, find groups of similar values within attribute, find outliers within attribute, find correlation between attributes, find similar items.
How: Encode	Dense, space-filling; area marks in spiral layout; colormap: categorical hues and ordered luminance.
How: Facet	Layout 1: partition by attribute into per-attribute views, small multiples. Layout 2: partition by items into per-item glyphs.
How: Reduce	Filtering
Scale	Attributes: one dozen. Total items: several million. Visible items (using multiple views, in total): one million. Visible items (using glyphs): 100,000

Hierarchical Clustering Explorer

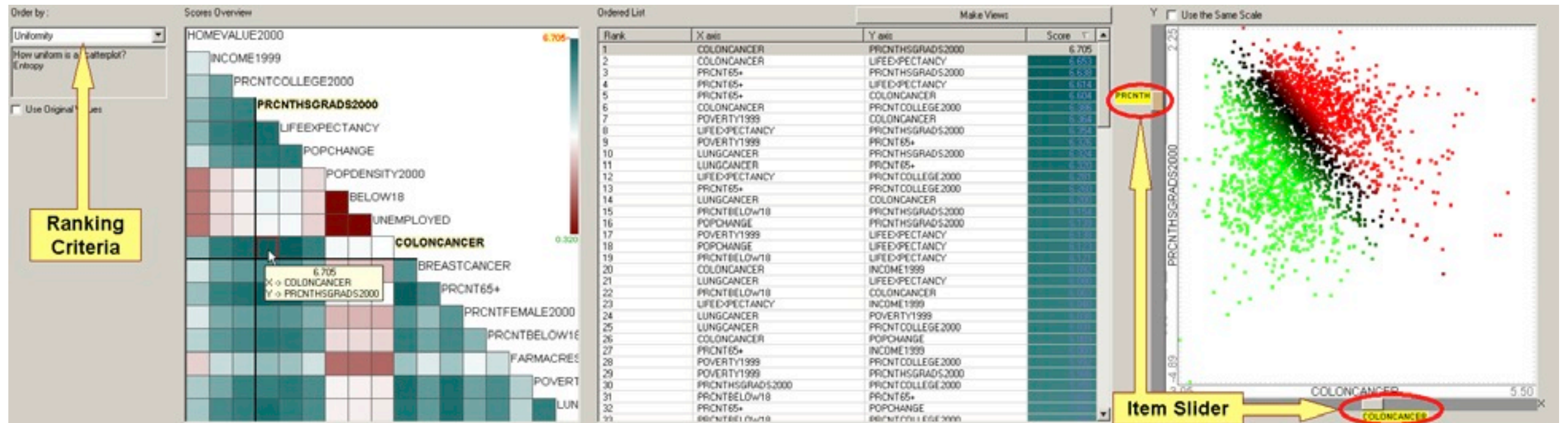
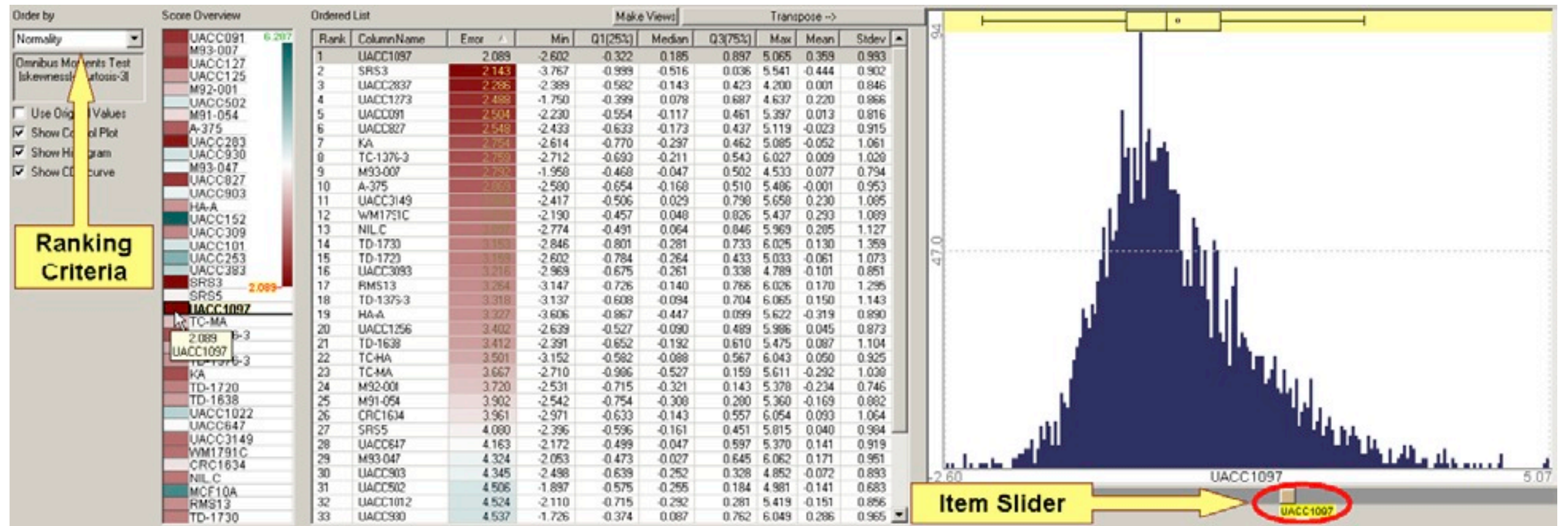
- heatmap, dendrogram
- multiple views



[Interactively Exploring Hierarchical Clustering Results. Seo and Shneiderman, *IEEE Computer* 35(7): 80-86 (2002)]

HCE

- rank by feature idiom
 - 1D list
 - 2D matrix



A rank-by-feature framework for interactive exploration of multidimensional data. Seo and Shneiderman. Information Visualization 4(2): 96-113 (2005)

HCE

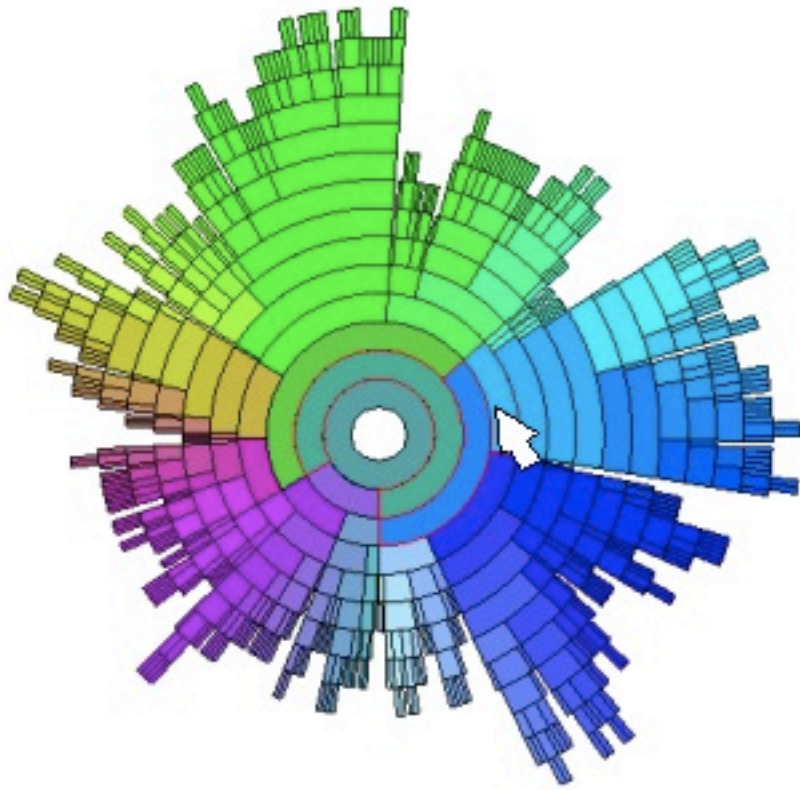


A rank-by-feature framework for interactive exploration of multidimensional data. Seo and Shneiderman. *Information Visualization* 4(2): 96-113 (2005)

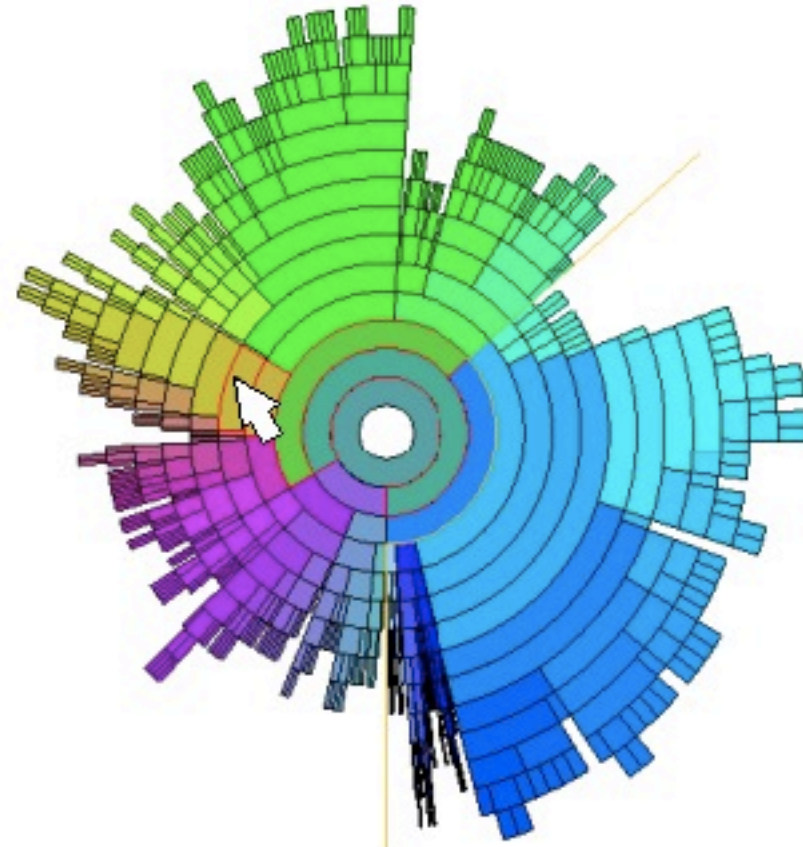
HCE Analysis

System	Hierarchical Clustering Explorer (HCE)
What: Data	Multidimensional table: two categorical key attributes (genes, conditions); one quantitative value attribute (gene activity level in condition).
What: Derived	Hierarchical clustering of table rows and columns (for cluster heatmap); quantitative derived attributes for each attribute and pairwise attribute combination; quantitative derived attribute for each ranking criterion and original attribute combination.
Why: Tasks	Find correlation between attributes; find clusters, gaps, outliers, trends within items.
How: Encode	Cluster heatmap, scatterplots, histograms, boxplots. Rank-by-feature overviews: continuous diverging colormaps on area marks in reorderable 2D matrix or 1D list alignment.
How: Reduce	Dynamic filtering; dynamic aggregation.
How: Manipulate	Navigate with pan/scroll.
How: Facet	Multiform with linked highlighting and shared spatial position; overview–detail with selection in overview populating detail view.
Scale	Genes (key attribute): 20,000. Conditions (key attribute): 80. Gene activity in condition (quantitative value attribute): $20,000 \times 80 = 1,600,000$.

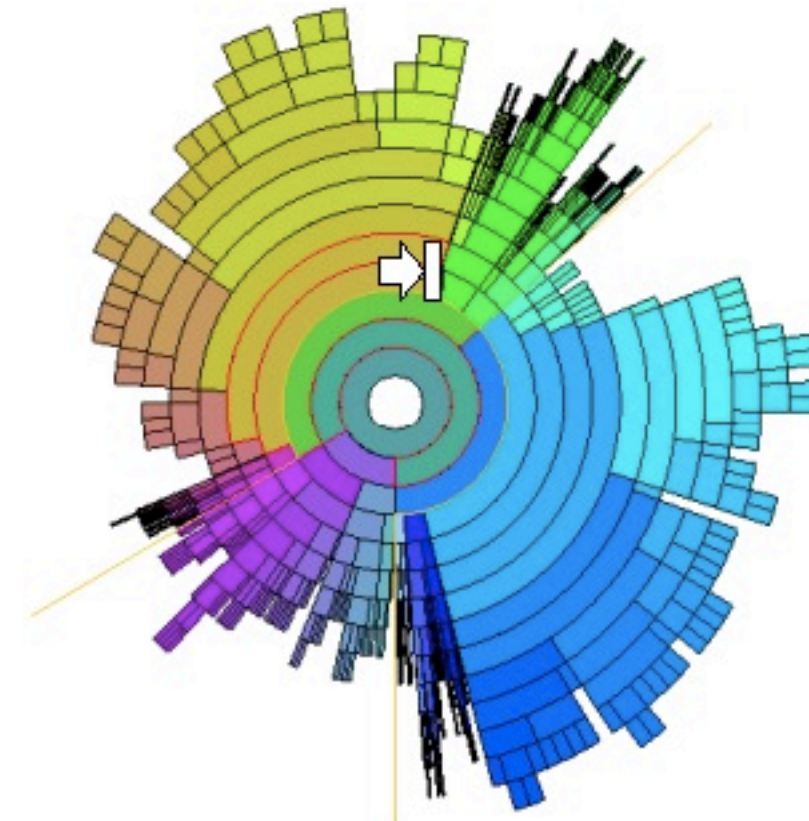
InterRing



original hierarchy



blue subtree expanded



tan subtree expanded

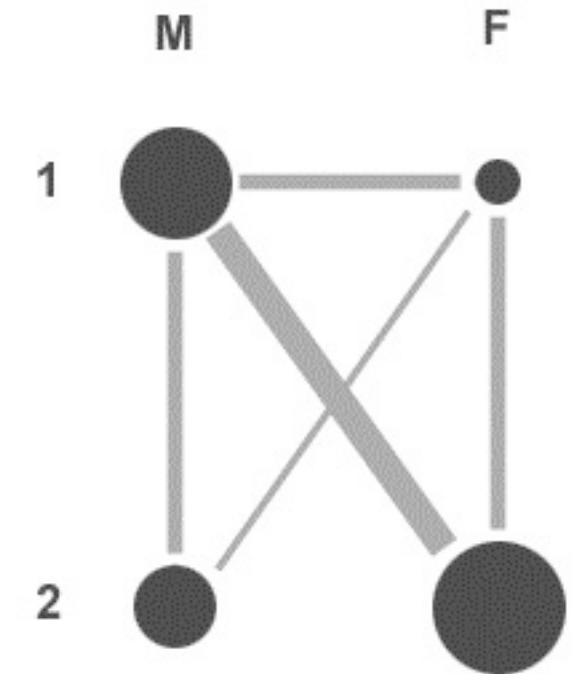
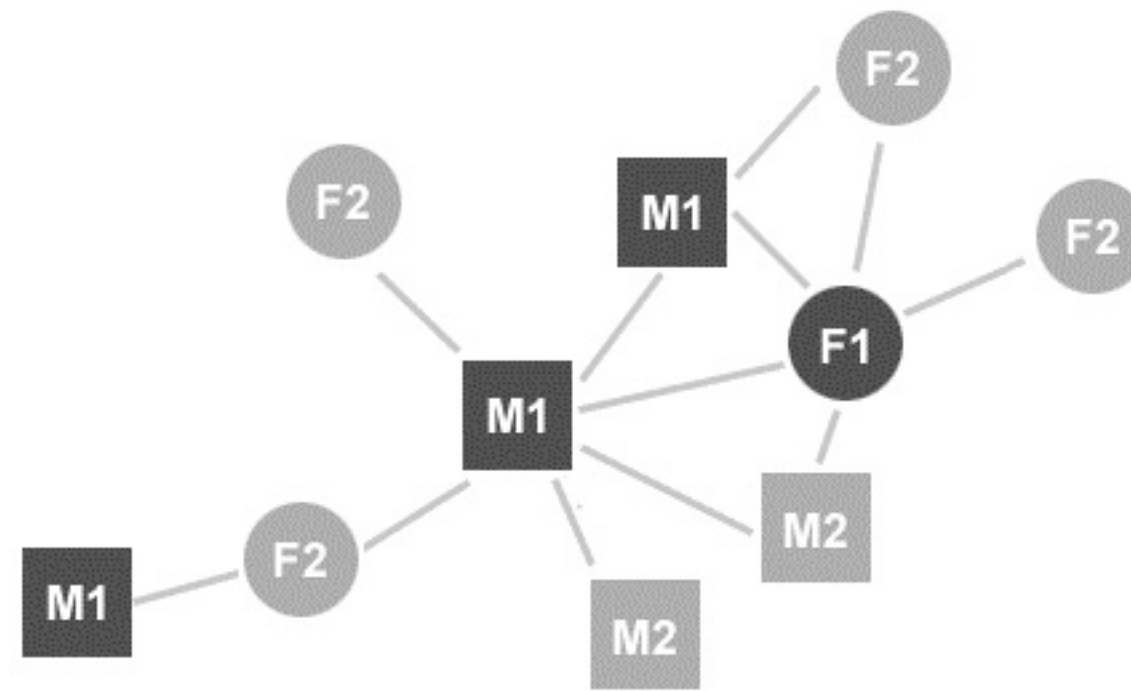
*[InterRing: An Interactive Tool for Visually Navigating and Manipulating Hierarchical Structures.
Yang, Ward, Rundensteiner. Proc. InfoVis 2002, p 77-84.]*

InterRing Analysis

System	InterRing
What: Data	Tree.
Why: Tasks	Selection, rollup/drilldown, hierarchy editing.
How: Encode	Radial, space-filling layout. Color by tree structure.
How: Facet	Linked coloring and highlighting.
How: Reduce	Embed: distort; multiple foci.
Scale	Nodes: hundreds if labeled, thousands if dense. Levels in tree: dozens.

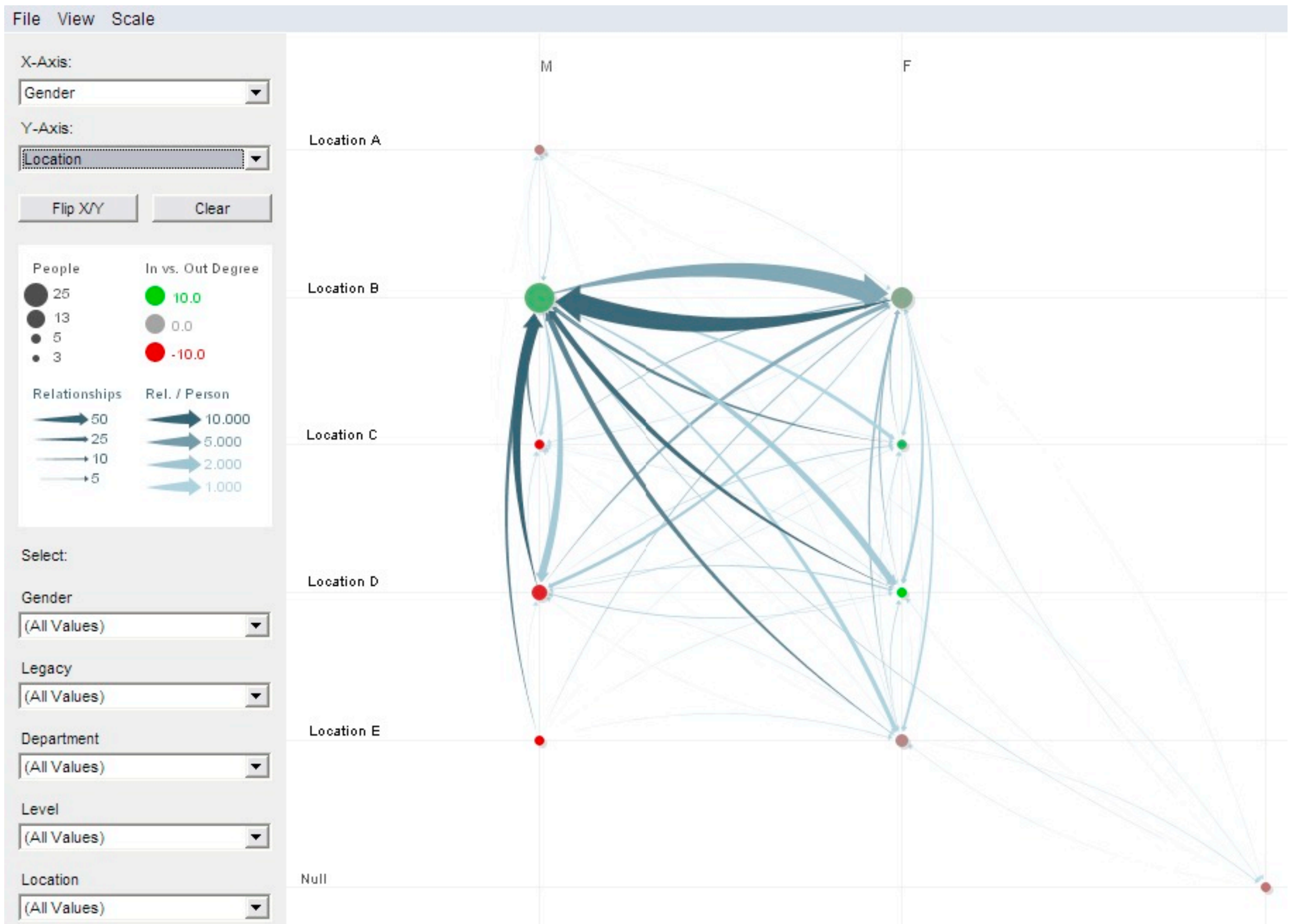
PivotGraph

- derived rollup network



[Visual Exploration of Multivariate Graphs, Martin Wattenberg, CHI 2006.]

PivotGraph



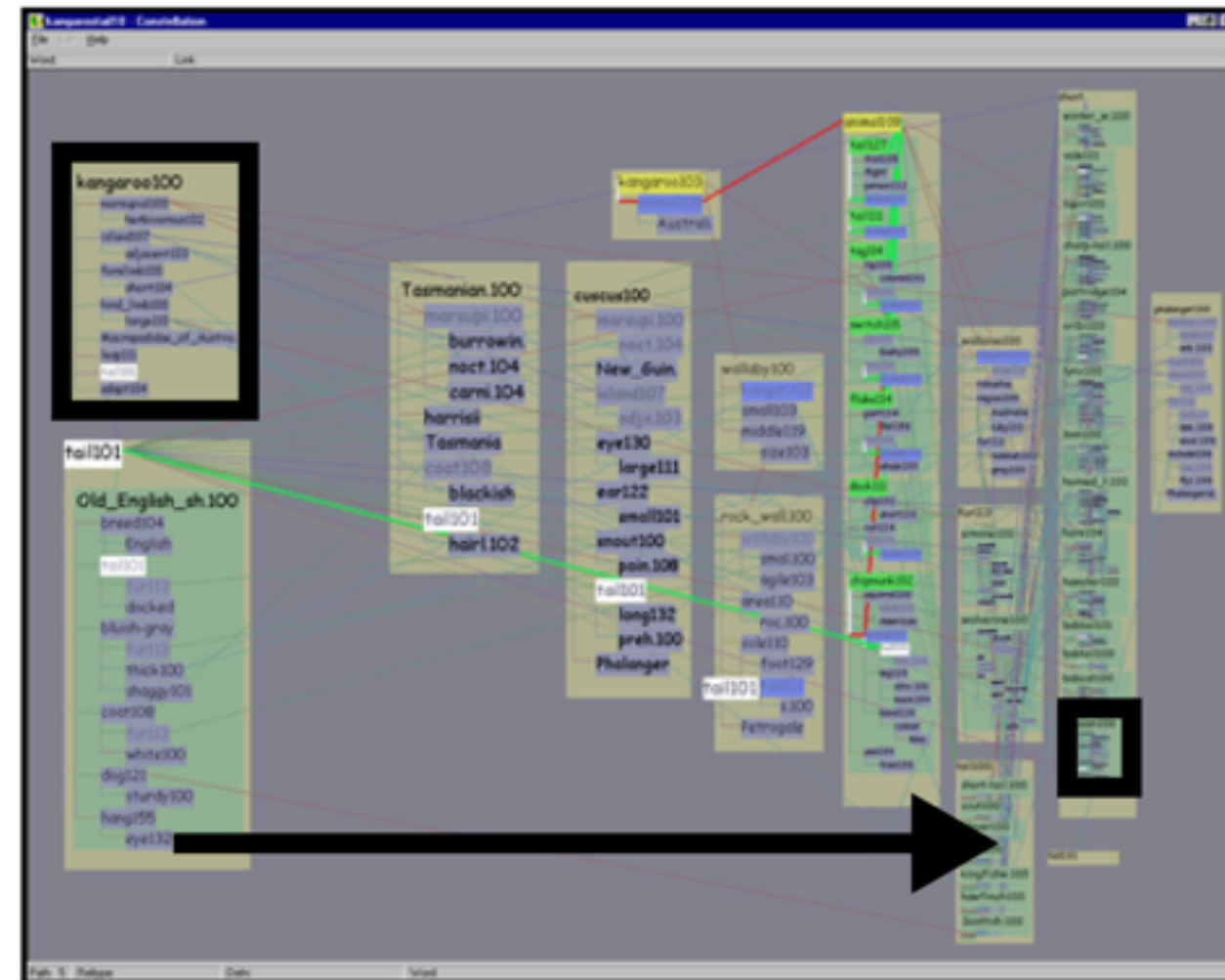
[Visual Exploration of Multivariate Graphs, Martin Wattenberg, CHI 2006.]

PivotGraph Analysis

Idiom	PivotGraph
What: Data	Network.
What: Derived	Derived network of aggregate nodes and links by roll-up into two chosen attributes.
Why: Tasks	Cross-attribute comparison of node groups.
How: Encode	Nodes linked with connection marks, size.
How: Manipulate	Change: animated transitions.
How: Reduce	Aggregation, filtering.
Scale	Nodes/links in original network: unlimited. Roll-up attributes: 2. Levels per roll-up attribute: several, up to one dozen.

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

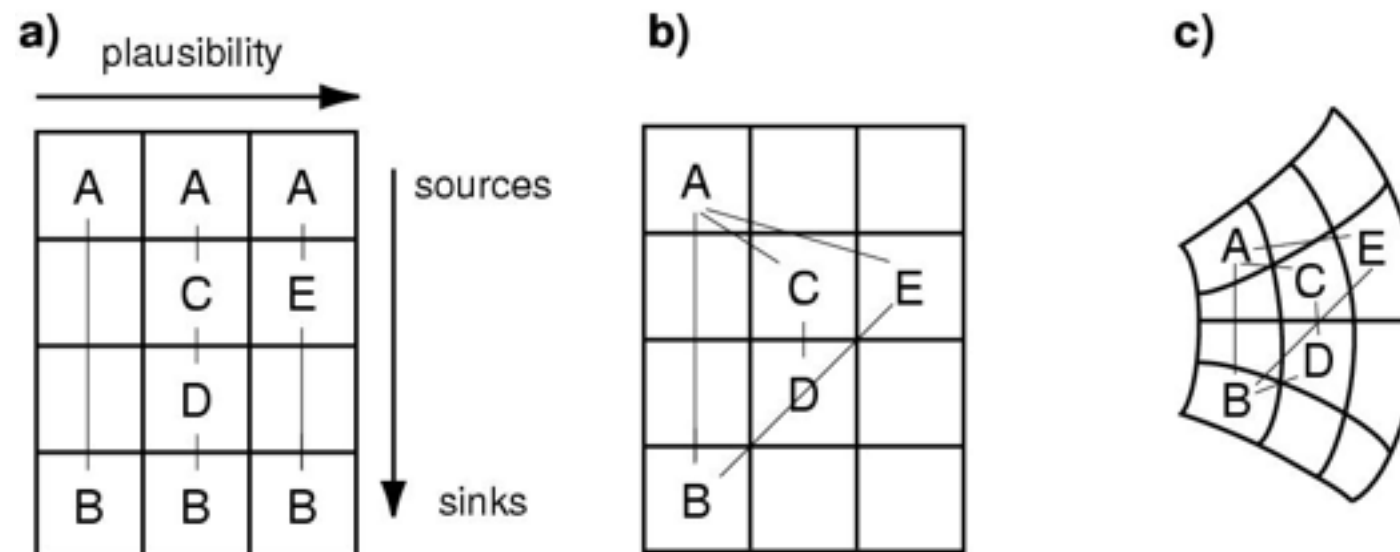
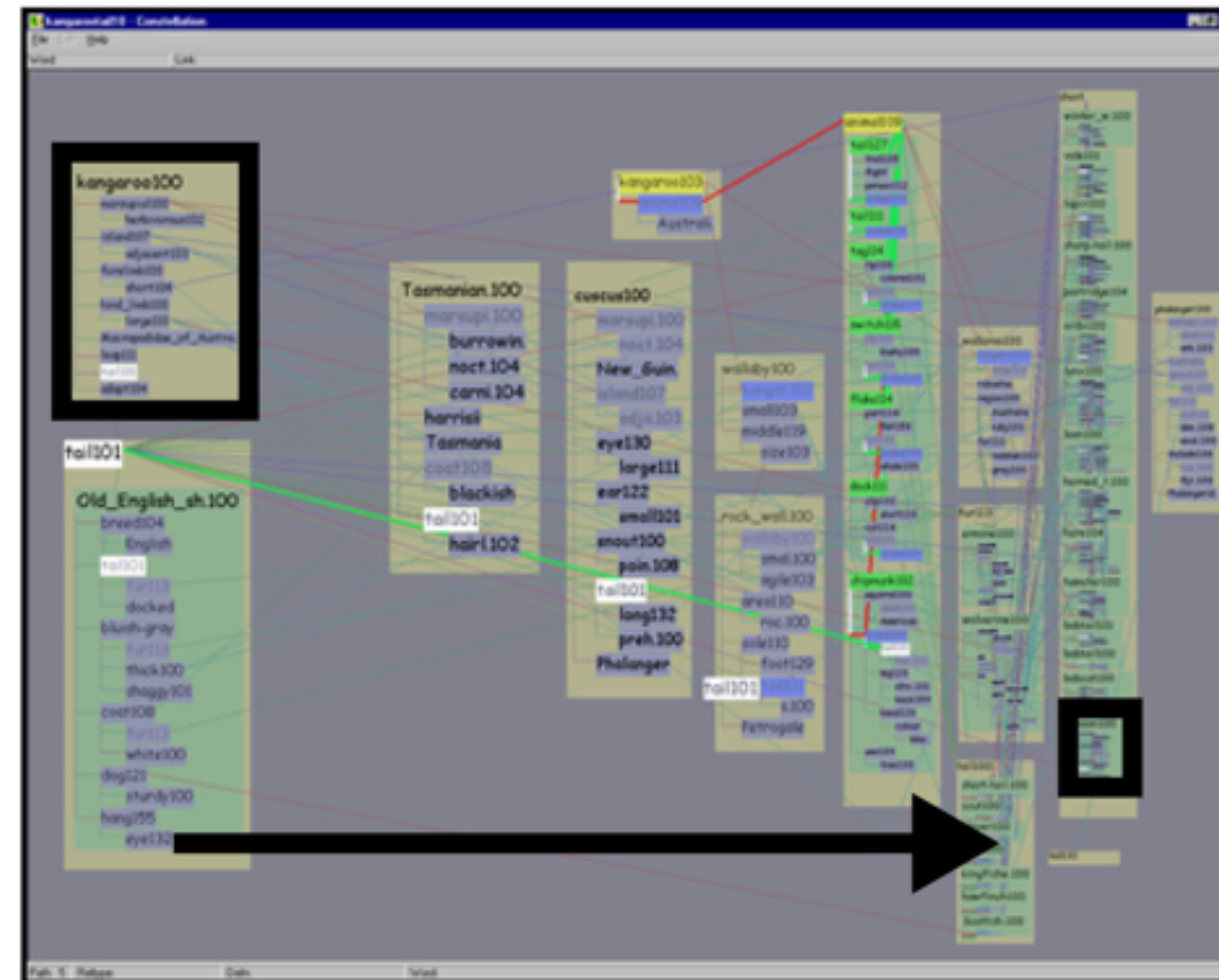


[Interactive Visualization of Large Graphs and Networks. Munzner. Ph.D. Dissertation, Stanford University, June 2000.]

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

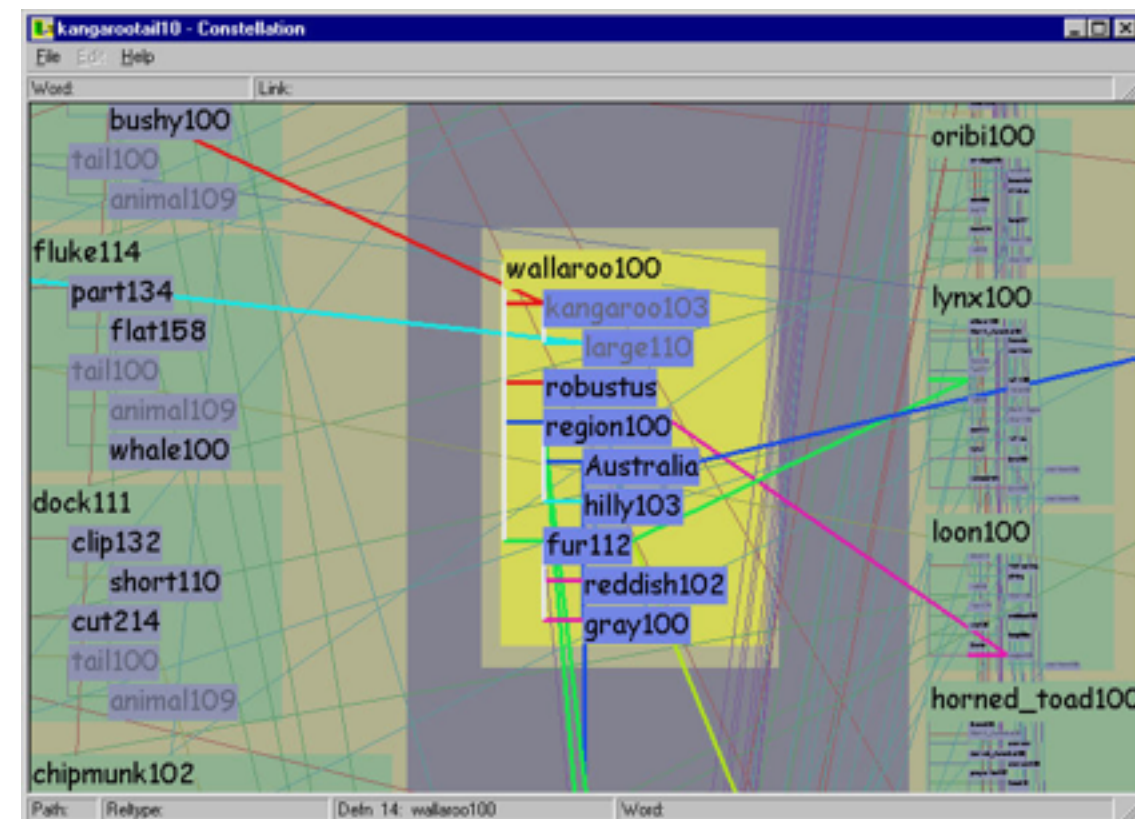
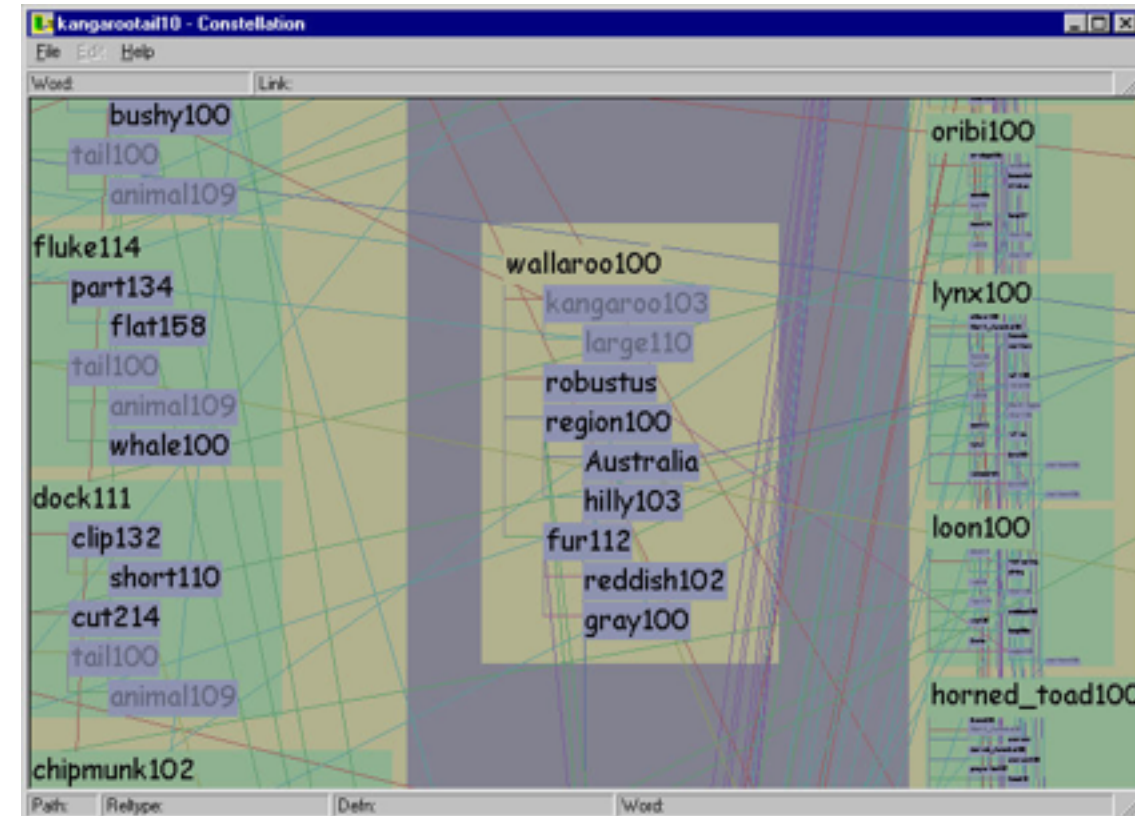
Using space: Constellation

- 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



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.]

Constellation Analysis

System	Constellation
What: Data	Three-level network of paths, subgraphs (definitions), and nodes (word senses).
Why: Tasks	Discover/verify: browse and locate types of paths, identify and compare.
How: Encode	Containment and connection link marks, horizontal spatial position for plausibility attribute, vertical spatial position for order within path, color links by type.
How: Manipulate	Navigate: semantic zooming. Change: Animated transitions.
How: Reduce	Superimpose dynamic layers.
Scale	Paths: 10–50. Subgraphs: 1–30 per path. Nodes: several thousand.

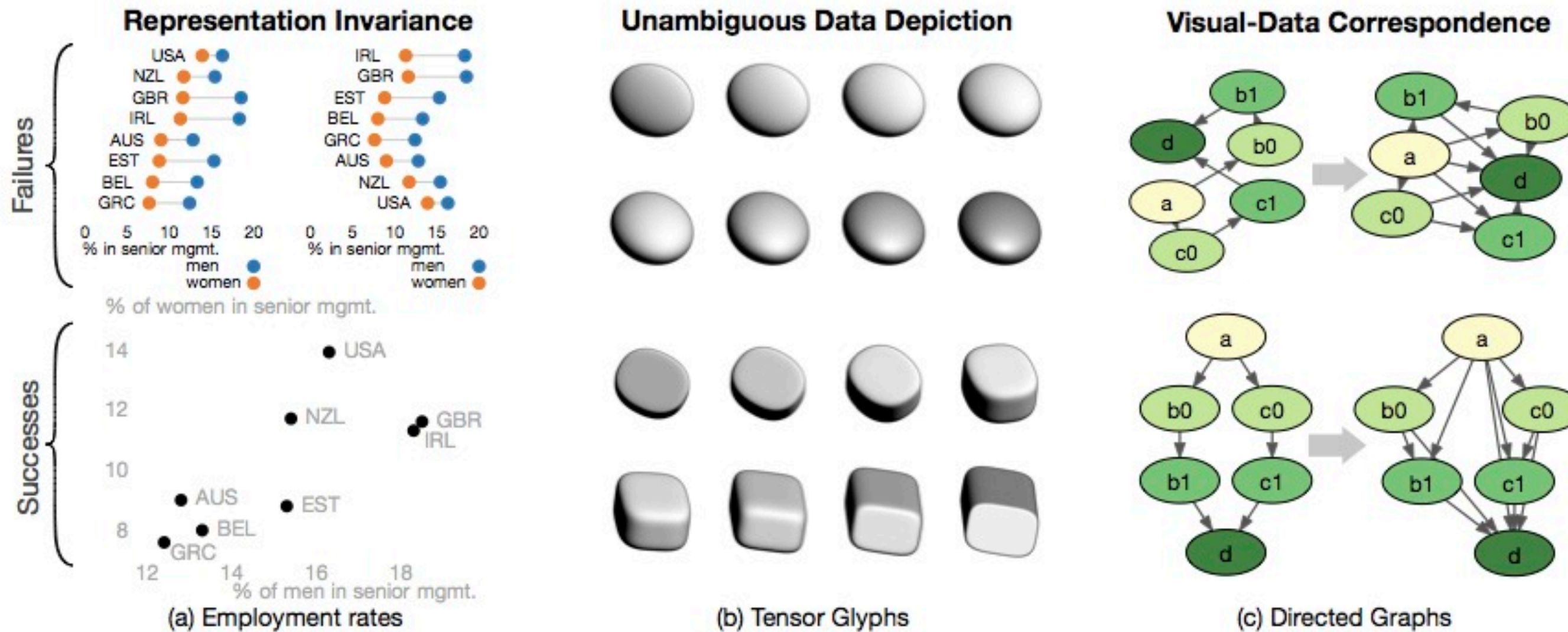
What-Why-How Analysis

- expected in your paper/topic presentations
 - in addition to content summarization and general reflection
- expected in your final projects

- this approach is not the only way to analyze visualizations!
 - one specific framework intended to help you think
 - other frameworks support different ways of thinking
 - today's paper is interesting example!

Algebraic Process for Visualization Design

- which mathematical structures in data are preserved and reflected in vis
 - negation, permutation, symmetry, invariance



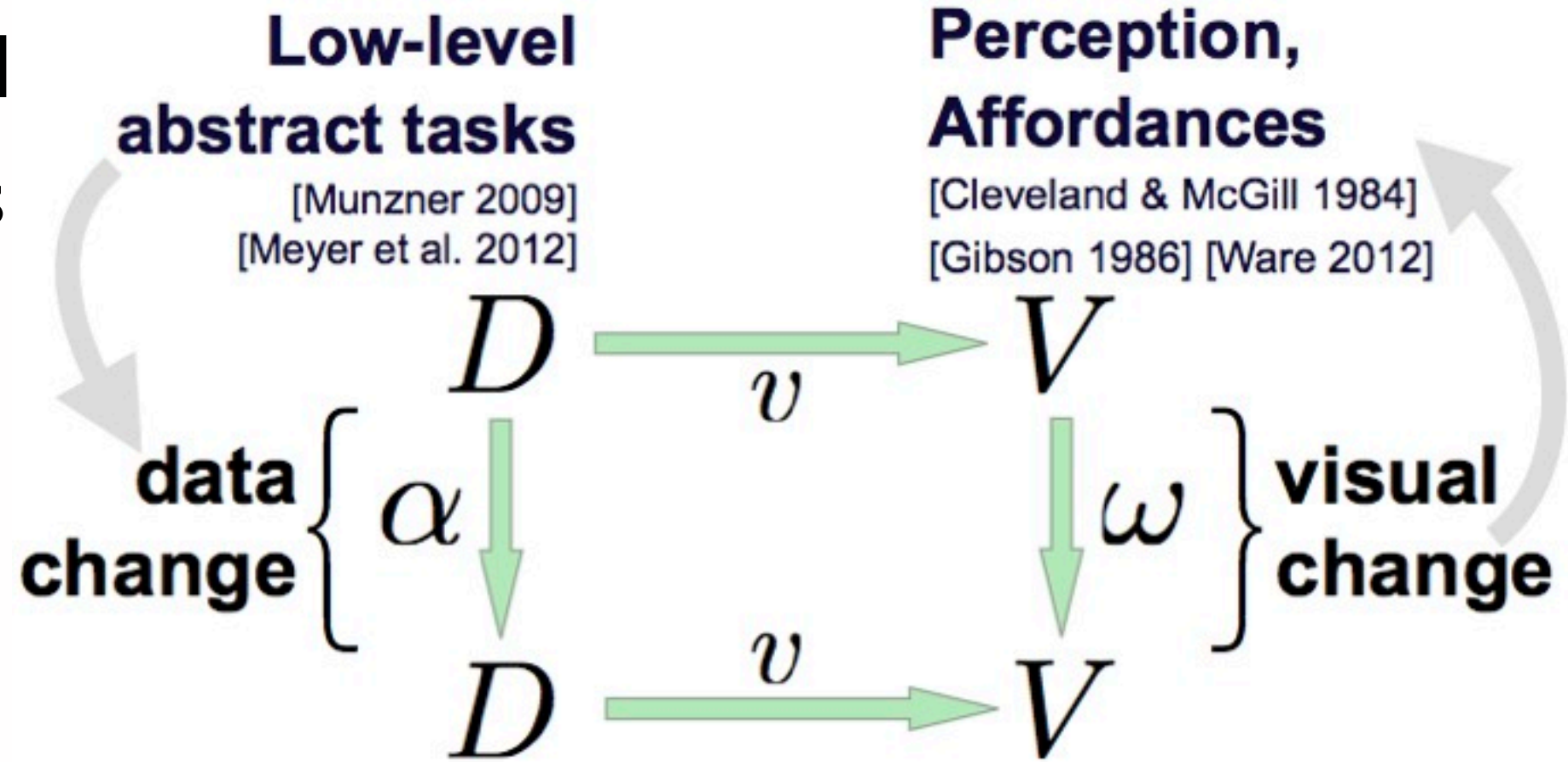
[Fig 1. An Algebraic Process for Visualization Design. Carlos Scheidegger and Gordon Kindlmann. *IEEE TVCG (Proc. InfoVis 2014)*, 20(12):2181-2190.]

Algebraic process: Vocabulary

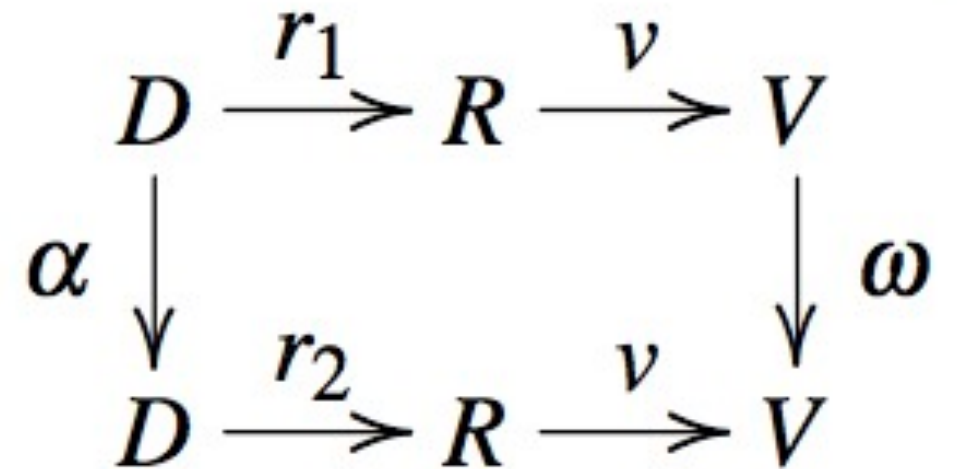
- **invariance** violation: single dataset, many visualizations
 - **hallucinator**
- **unambiguity** violation: many datasets, same vis
 - data change invisible to viewer
 - **confuser**
- **correspondence** violation:
 - can't see change of data in vis
 - **jumbler**
 - salient change in vis not due to significant change in data
 - **misleader**
 - match mathematical structure in data with visual perception
- we can X the data; can we Y the image?
 - are important data changes well-matched with obvious visual changes?

Algebraic process: Model

- D : space of data to be visualized
- R : space of data representations
 - r : mapping from D to R
- V : space of visualizations
 - v : mapping from R to V
- α : data symmetries
- ω : visualization symmetries
- commutative diagram
 - equality between paths

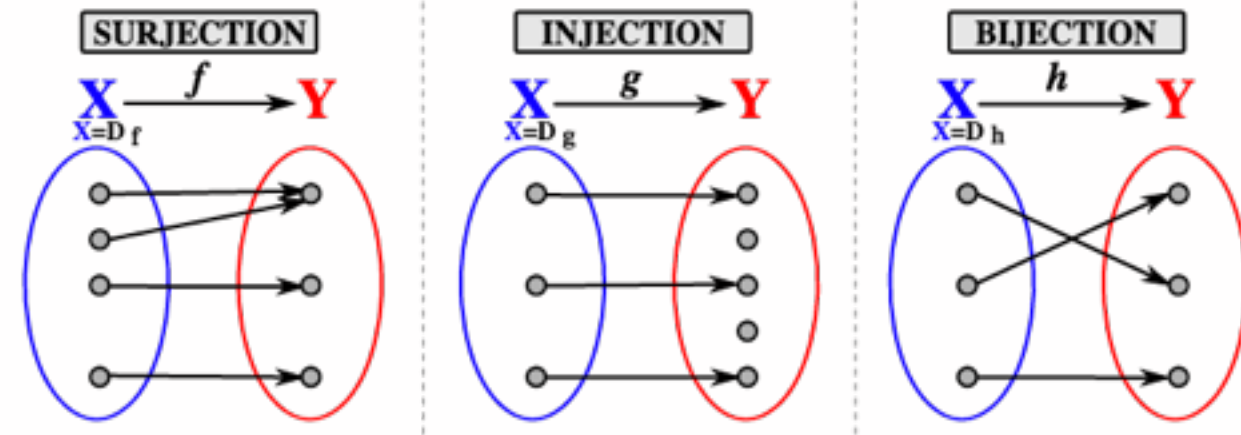


$$v \circ r_2 \circ \alpha = \omega \circ v \circ r_1$$



Algebraic process: Previous work tie-in

- Stevens data types: categorical, ordinal, quant (interval & ratio)
 - defined by symmetry groups and invariances
- Ziemziewicz & Kosara surjective/injective/bijective
 - injectivity: unambiguity
- Mackinlay's Expressiveness Principle
 - convey all and only properties of data
 - invariance/hallucinator, correspondence/misleader
- Mackinlay's Effectiveness Principle
 - match important data attributes to salient visual channels
 - correspondence/jumbler, unambiguity/confuser
- Gibson/Ware affordances
 - perceivable structures show possibility of action
 - correspondence



Algebraic process: Previous work tie-in, cont.

- Tversky Congruence Principle & Apprehension Principle
 - congruence: visual external structure of graphic should correspond to mental internal representation of viewer
 - apprehension: graphics should be readily and easily perceived and comprehended
 - unambiguity and correspondence
- nested model
 - reason about mappings from abstraction to idiom
 - mathematical guidelines for abstraction layer

Next Time

- presentations continue
 - no further assigned readings for everybody
- presentations
 - 4 per class, 20 minutes each total
 - plan on 15-17 min present, 3-5 minute questions
 - note typo in mail!
- update presentations due Mon Nov 23
 - typo on web page - not Mon Nov 14!
 - new this year: full draft of previous work section of final report
 - bulk of your mark will be on what's in the update
 - goal: do this up front not at the end!