CSE 564
Visualization & Visual Analytics

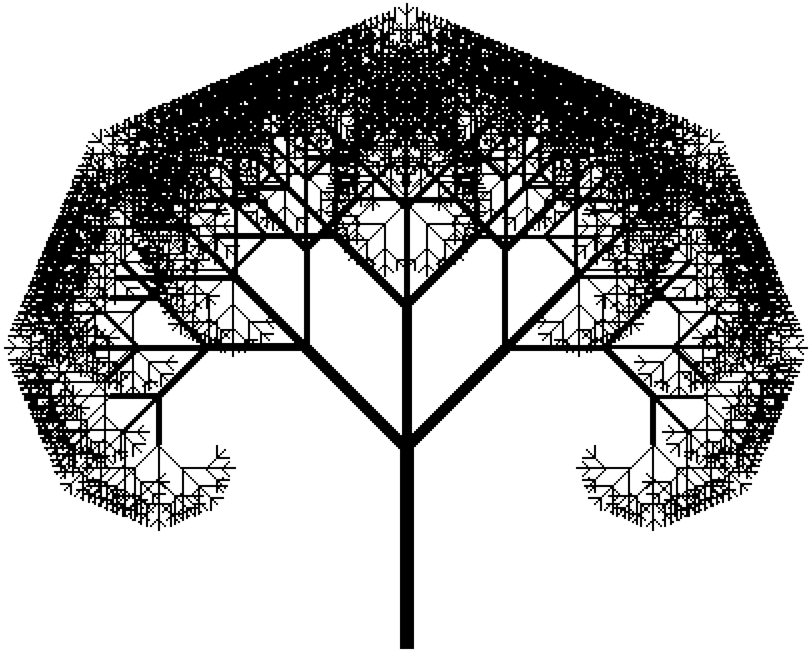
Visualization of Hierarchies

Klaus Mueller

Computer Science Department
Stony Brook University

Lecture	Topic	Projects
1	Intro, schedule, and logistics	
2	Applications of visual analytics, basic tasks, data types	
3	Introduction to D3, basic vis techniques for non-spatial data	
4	Data assimilation and preparation	Project #1 out
5	Bias in visualization	
6	Data reduction and dimension reduction	
7	Visual perception and cognition	
8	Visual design and aesthetics	Project #1 due/Project #2 out
9	High-dimensional data vis. & dimensionality reduction	
10	Visual design and aesthetics	
11	Cluster analysis and dimension reduction	
12	Cluster analysis and dimension reduction	
13	Cluster analysis and dimension reduction	Project #2 due
14	Foundations of scientific and medical visualization	Project #3 out
15	Computer graphics and volume rendering	
16	Scientific and medical visualization	Project #3 due/Final project proposal call
17	Illustrative rendering	
18	Principles of interaction	
19	Visual analytics and the visual sense making process	Final project proposal due
20	Midterm #1	
21	Midterm discussion	
22	Tasks, VA design and evaluation with user studies	
23	Visualization of time-varying, time-series & streaming data	Final project prelim report due
24	Visualization of geospatial data	
25	Visualization of graphs and hierarchies	
26	Memorable vis., visual embellishments, infographics	
27	Midterm #2	
	Final project presentations	Final Project slides, video, report due

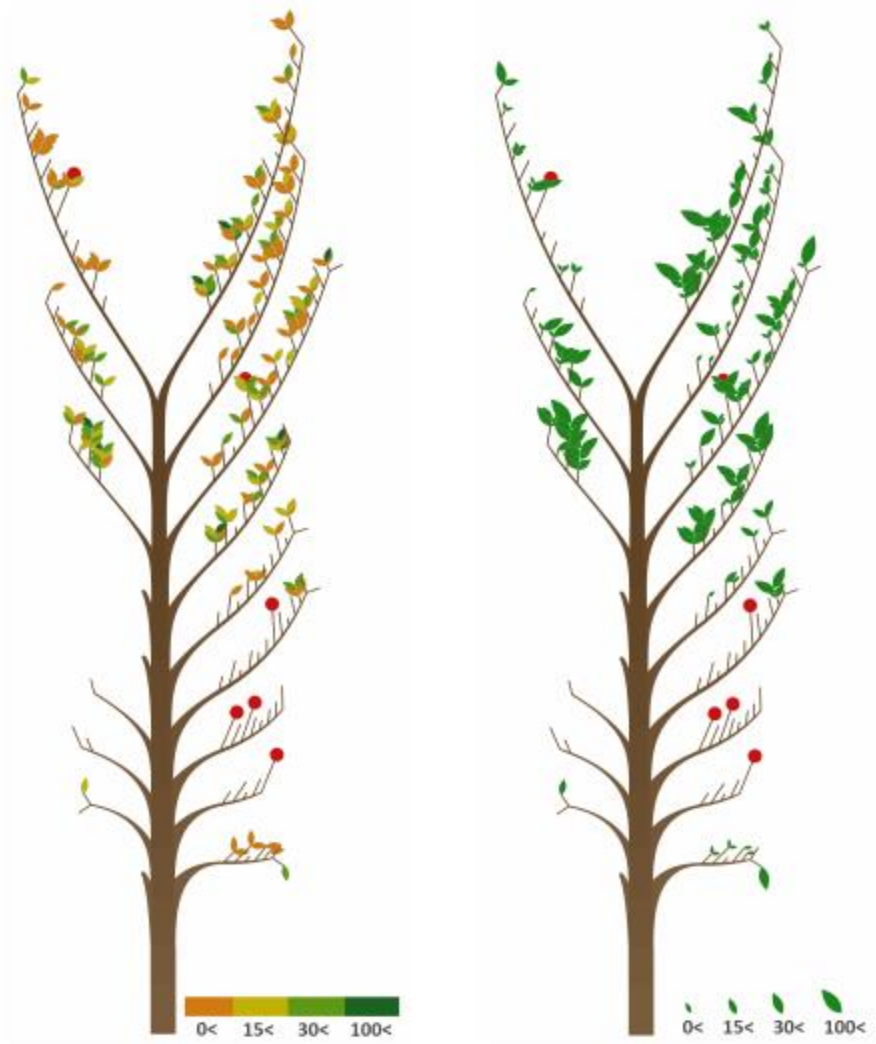
Hierarchies = Trees



Tree – A Natural Metaphor

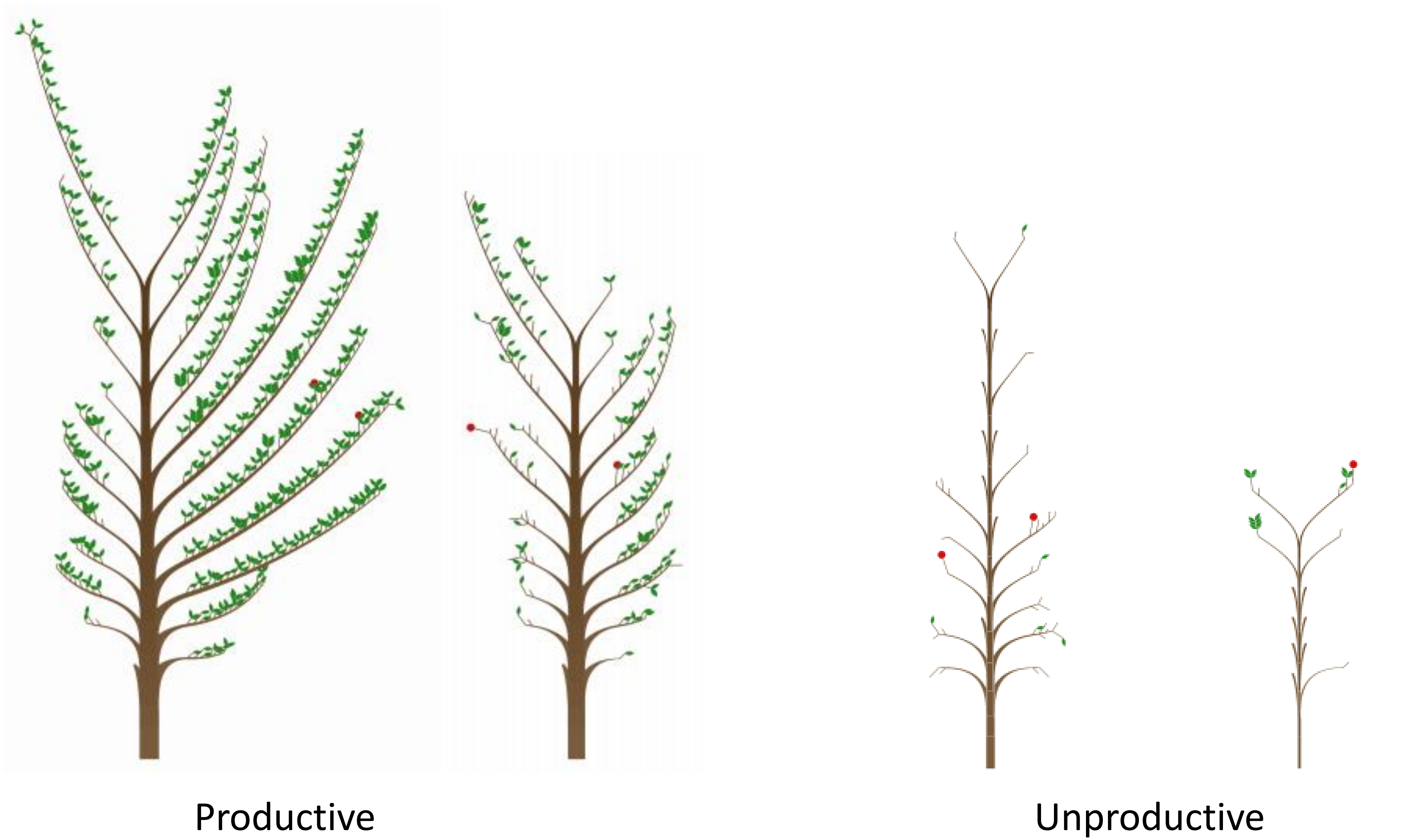
Mapping publications to a tree

- major leaves are papers
- minor leaves are co-authors
- height is time
- fruit are comments
- size or color is number of paper's citations
- journal papers on right side
- conference papers left side



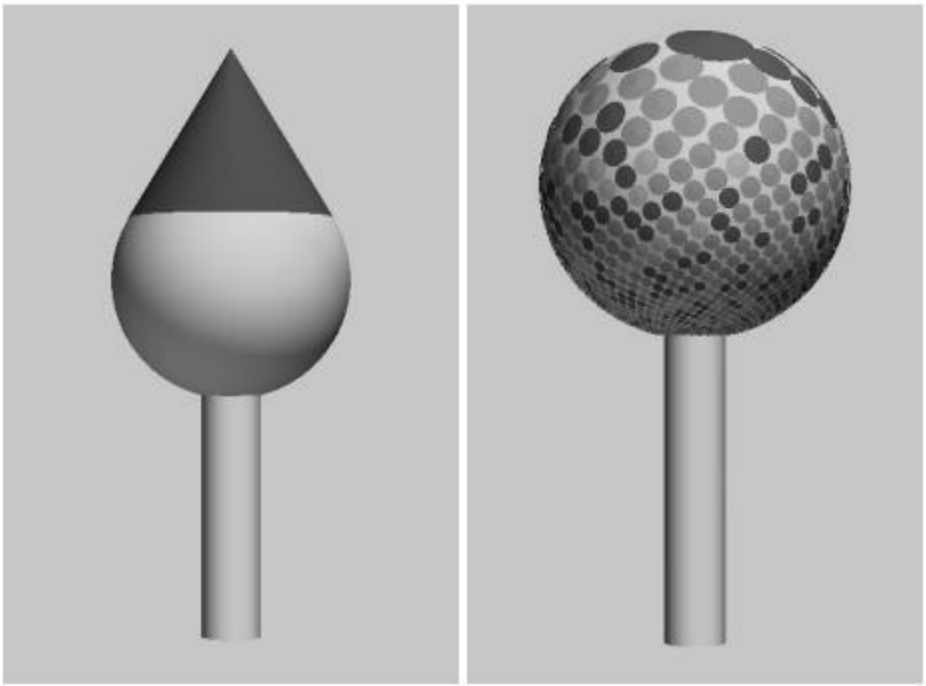
Fung and Ma, Personal VIS 2015

Productive vs. Unproductive Researchers



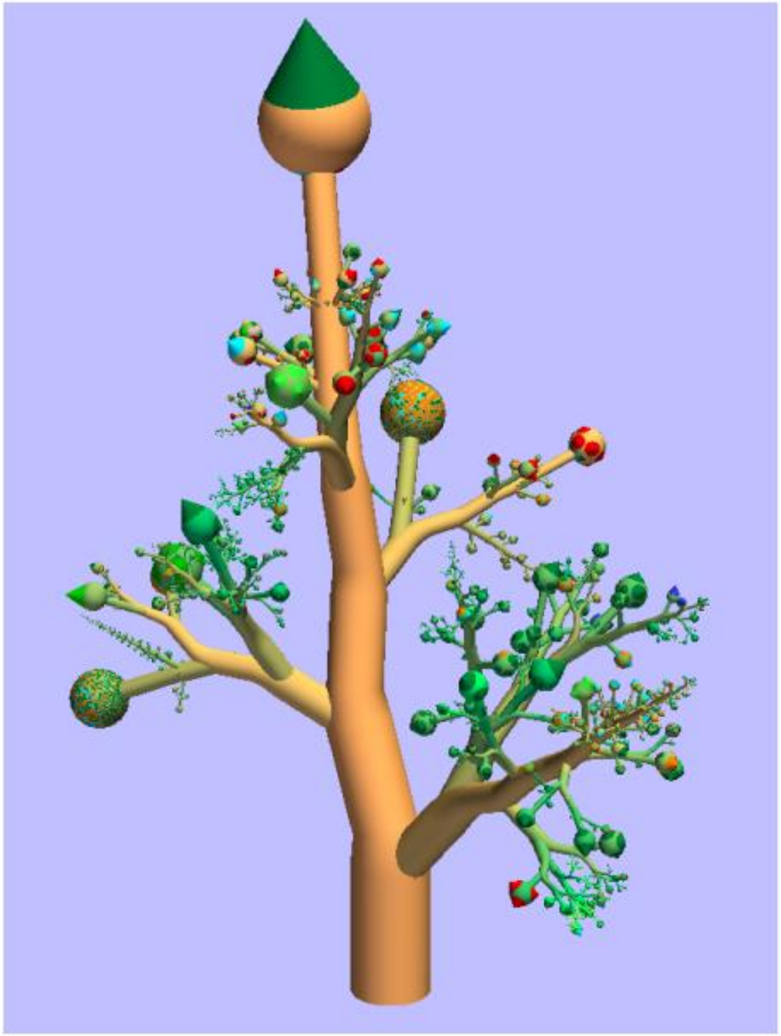
Botanical-Inspired Visualizations

Visualizing hard drives with tree cartoons



one file many files

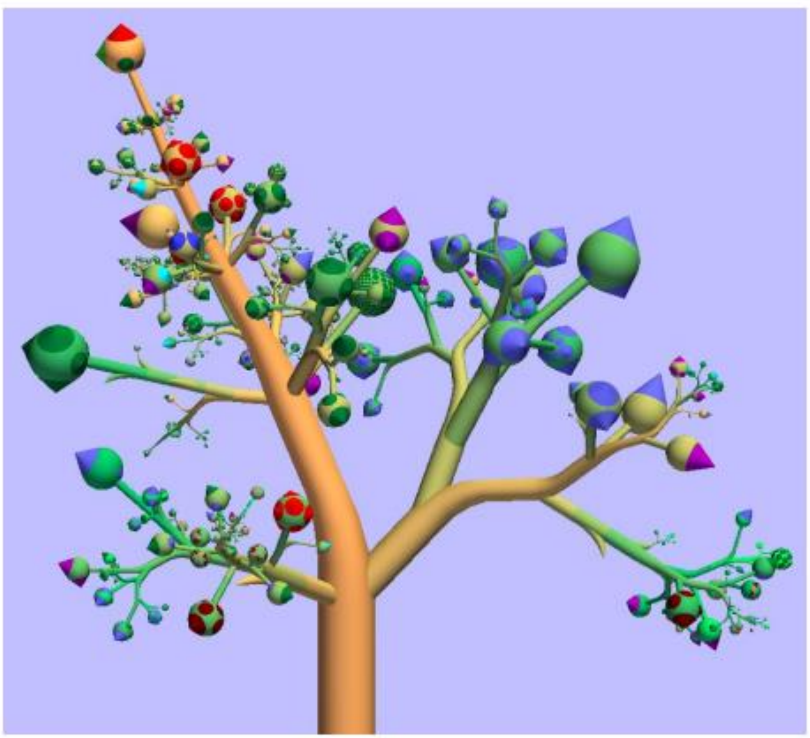
Kleiberg et al., IEEE InfoVis 2001

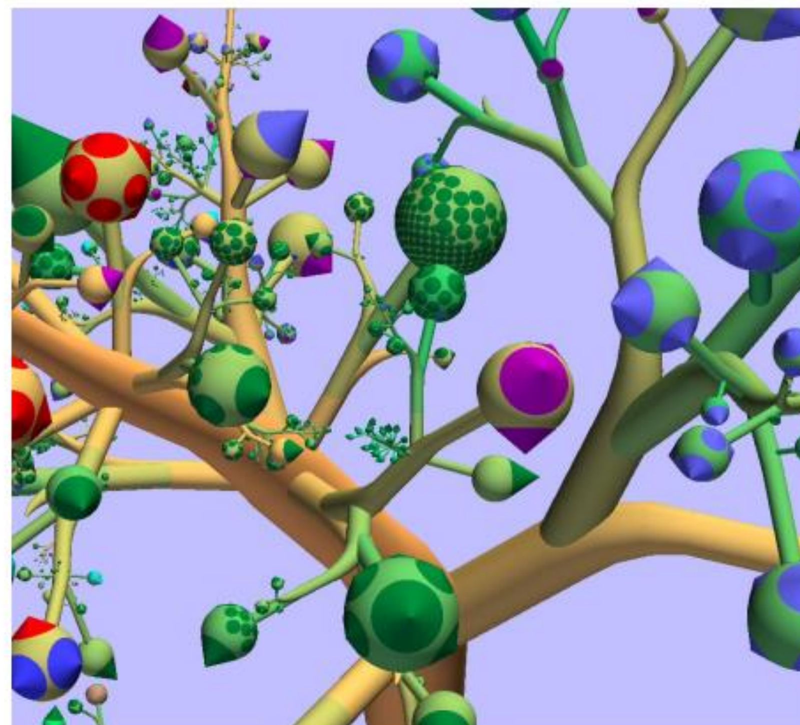


Botanical-Inspired Visualizations

Color maps to file type

- blue are pdf files, red are image files

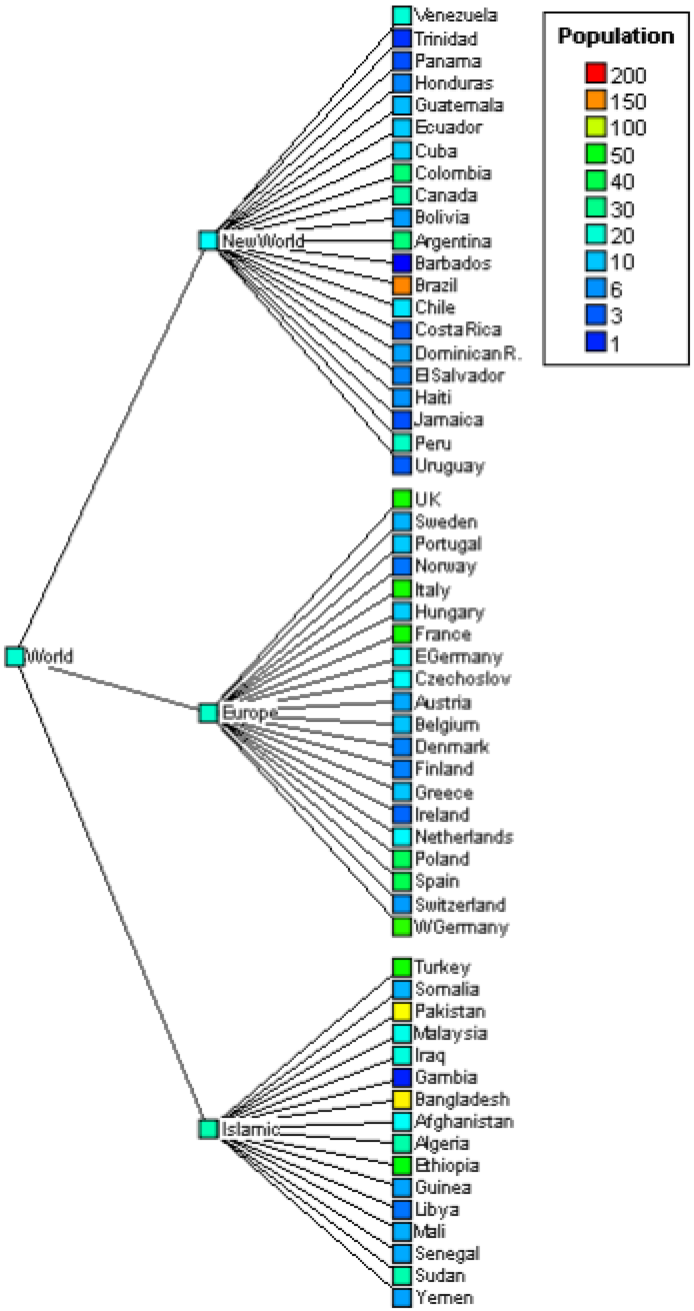




Conventional

Standard Node-Edge layout for a hierarchical network

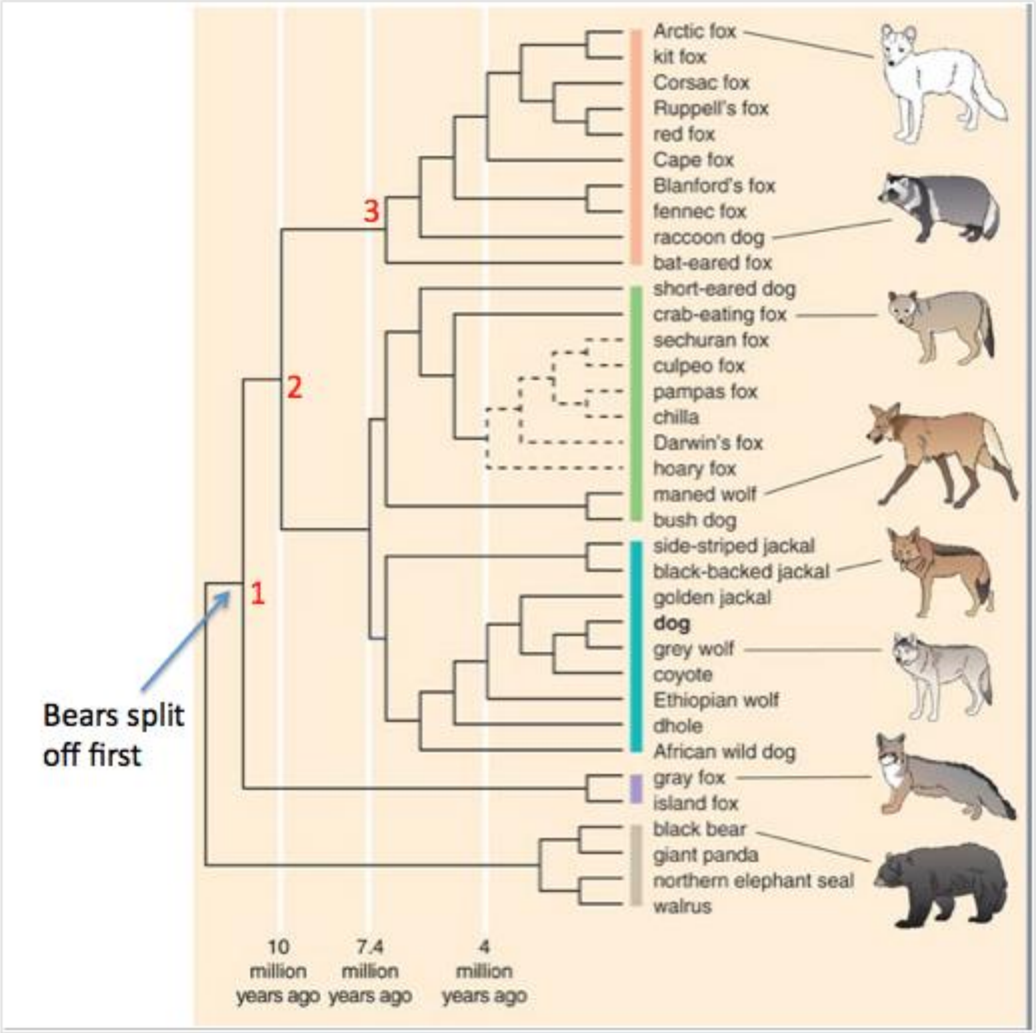
- 3 levels
- color maps to quantitative information (here population)



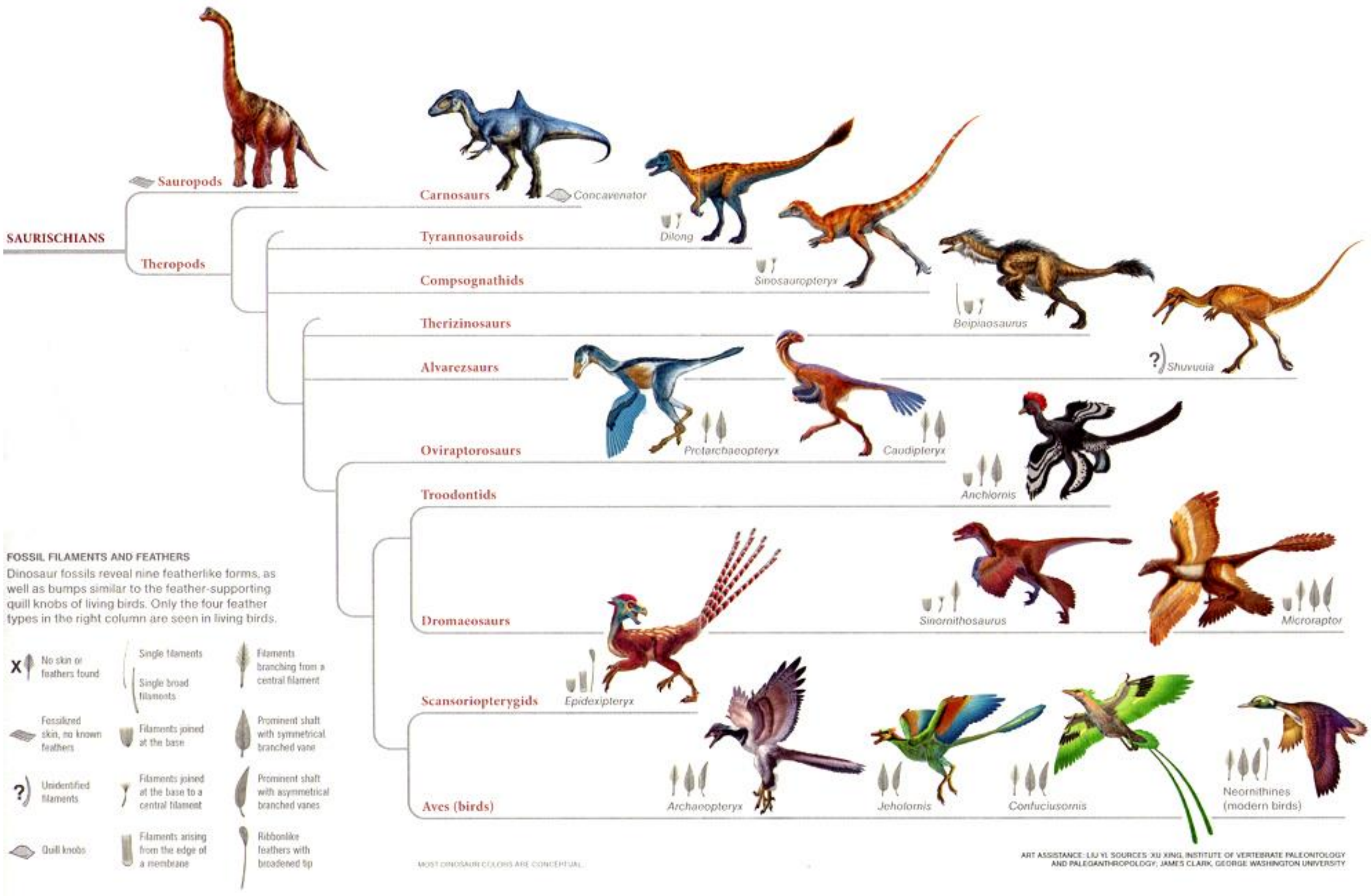
DENDROGRAM

Typically used to depict classification hierarchies

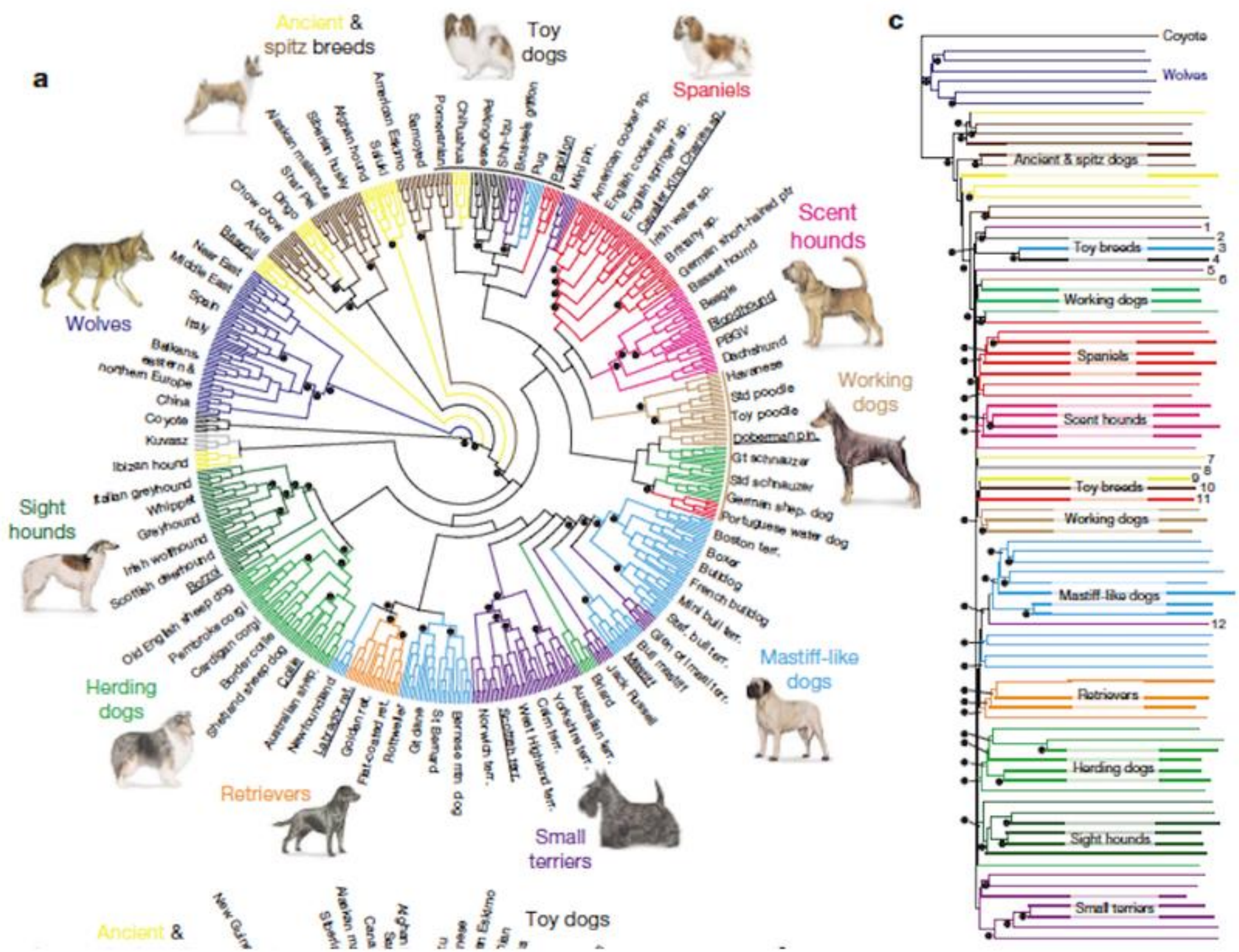
- split-off points visualize proximity



Birds and Dinosaurs



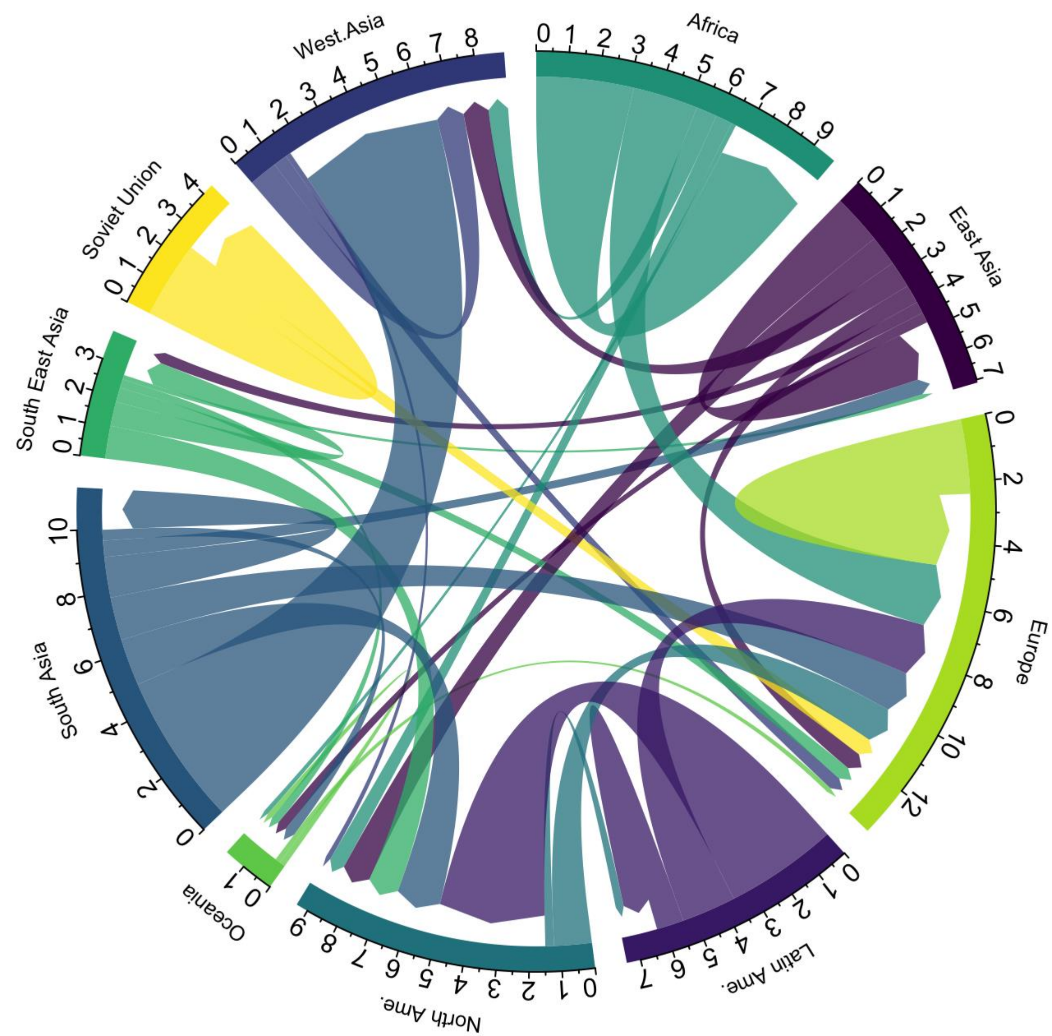
Circles Are More Space-Efficient



Chord Diagrams

Represents flows or connections between several entities

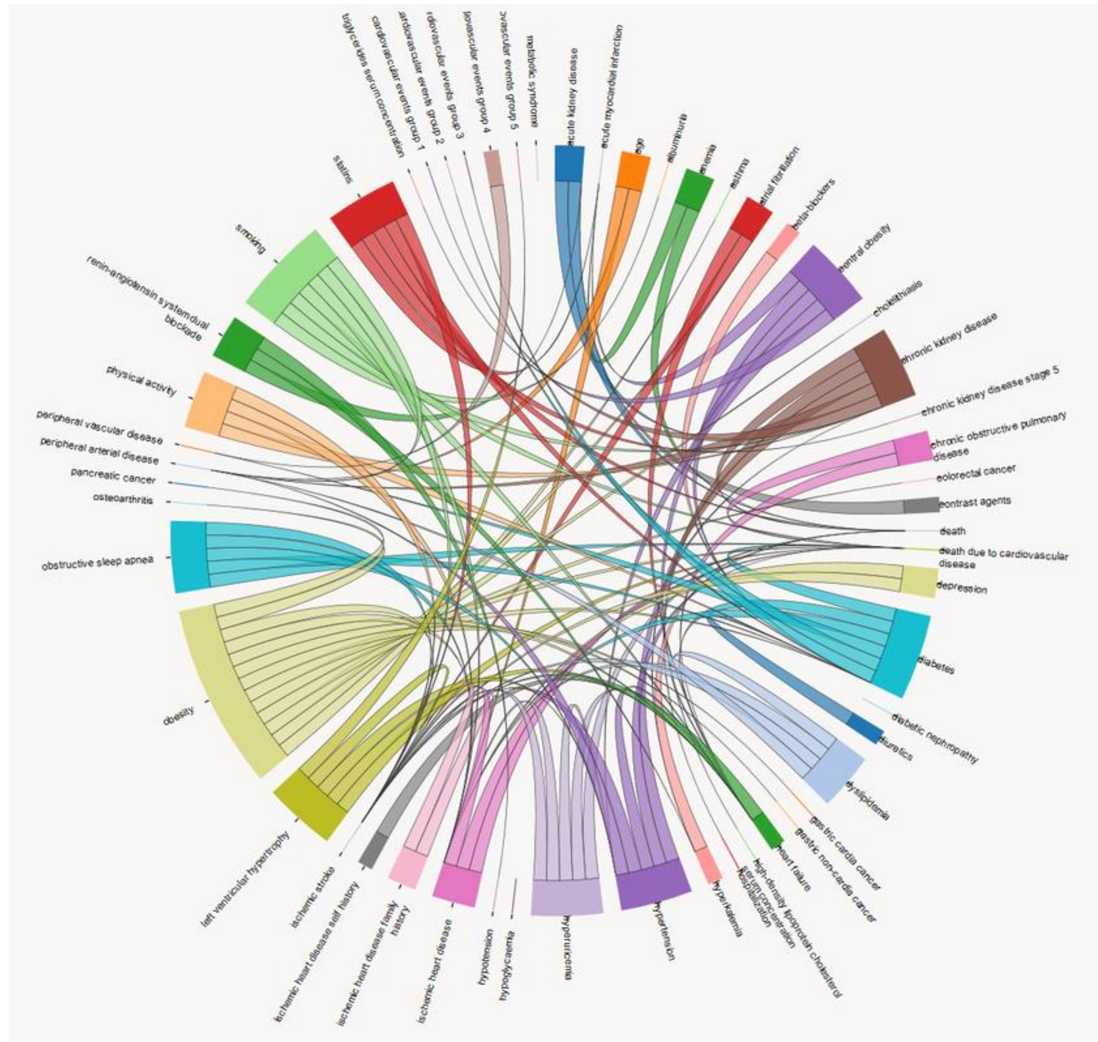
- for example the number of people migrating from one country to another



More Complex Chord Diagram

Can we make it easier to read?

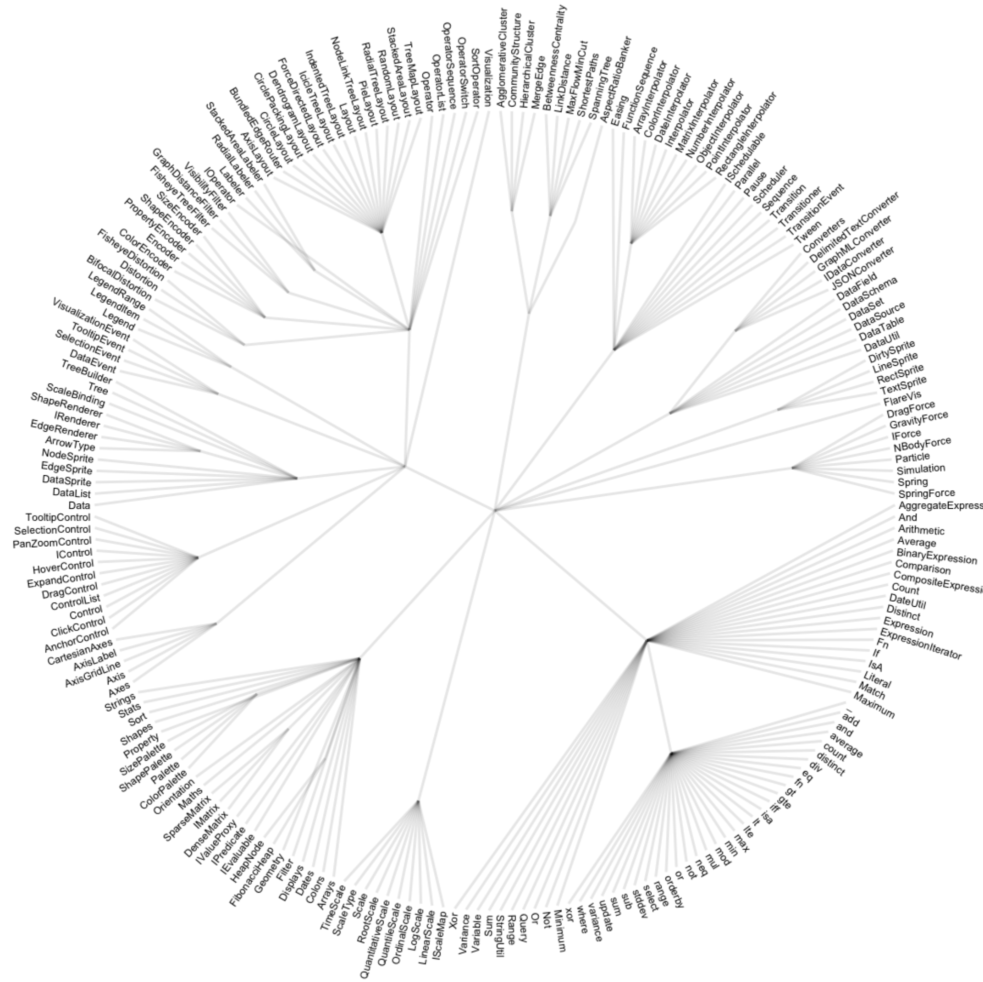
- yes
- via edge bundling



Hierarchical Chord Diagram

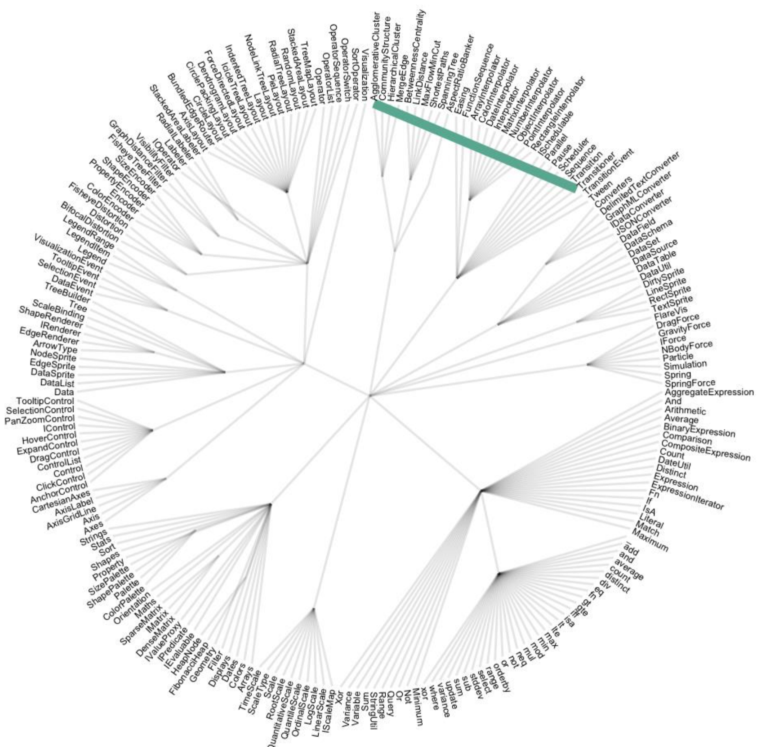
Hierarchy of the Flare ActionScript visualization library

- elements are organized in several folders, such as query, data, scale…
- each folder is then subdivided in subfolders and so on.
- can be visualized as a radial dendrogram

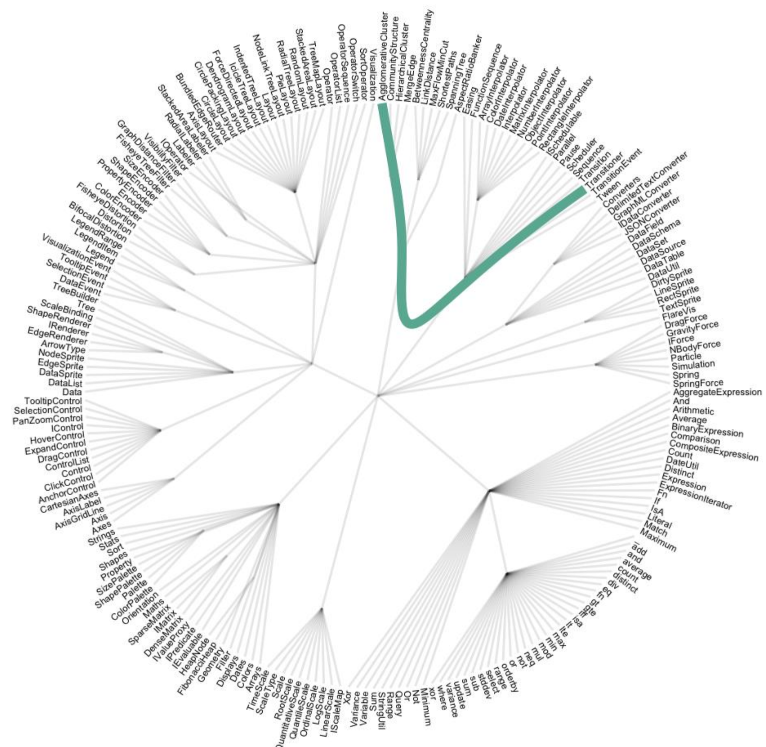


Hierarchical Chord Diagram

Visualize dependencies in the library



bad: straight line

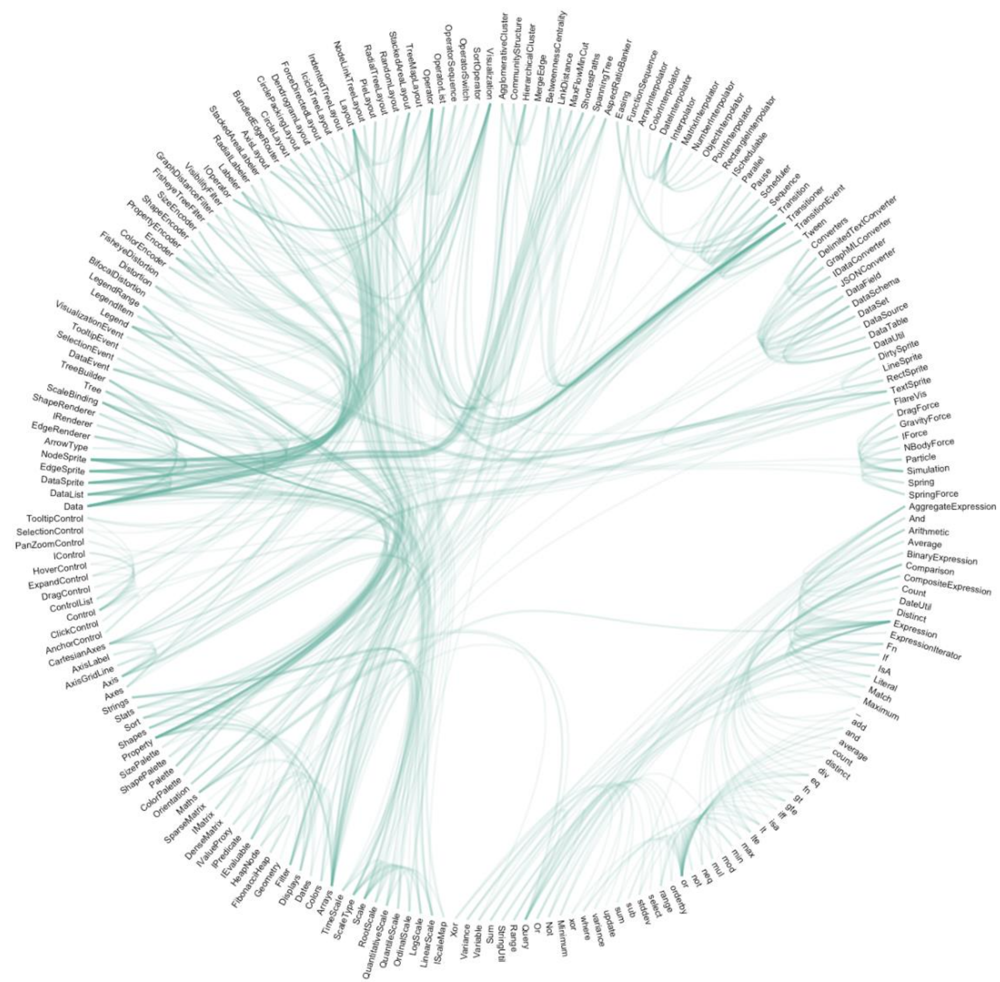


better: follow a hierarchical edge bundling line

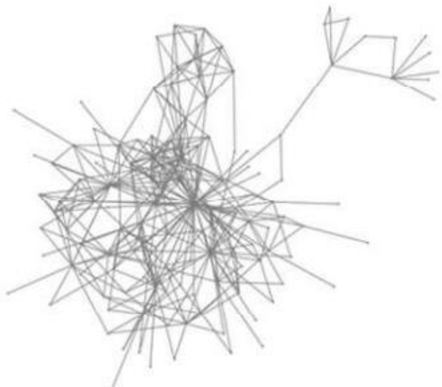
Edge Bundling

Apply the bundling to every adjacency connection of the dataset

- show the hierarchy of the dataset
- decrease the clutter as much as possible
- bundling the electrical wires together in order to reduce clutter
- and fan them out at their terminus in order to connect them to the terminals



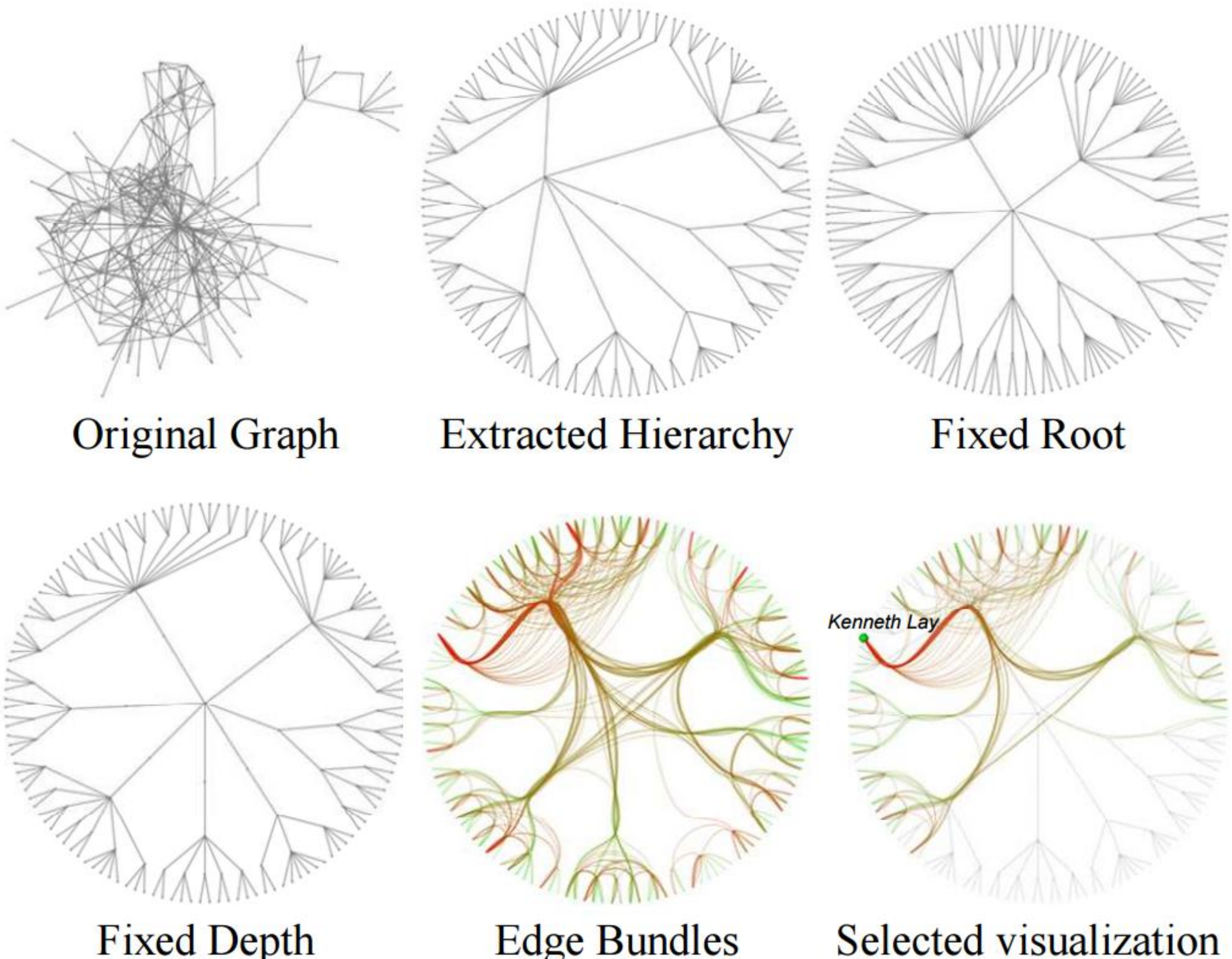
Radial Plots and Edge Bundles



Original Graph

Jia et al., Computer Graphics Forum 2011

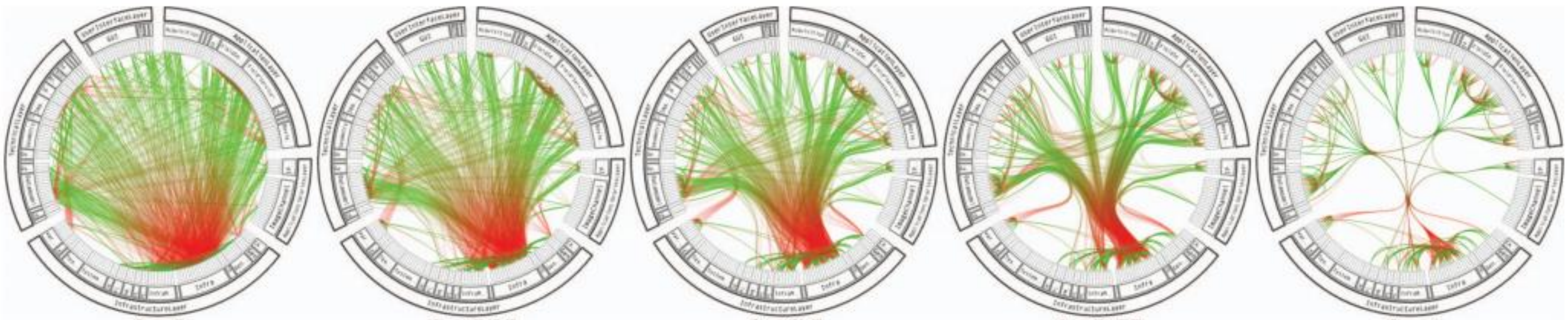
Radial Plots and Edge Bundles



Jia et al., Computer Graphics Forum 2011

Levels of Edge Bundling

Edges are represented by splines with tension β



$\beta = 0$ $\beta = 0.25$ $\beta = 0.5$ $\beta = 0.75$ $\beta = 1$

Setting β

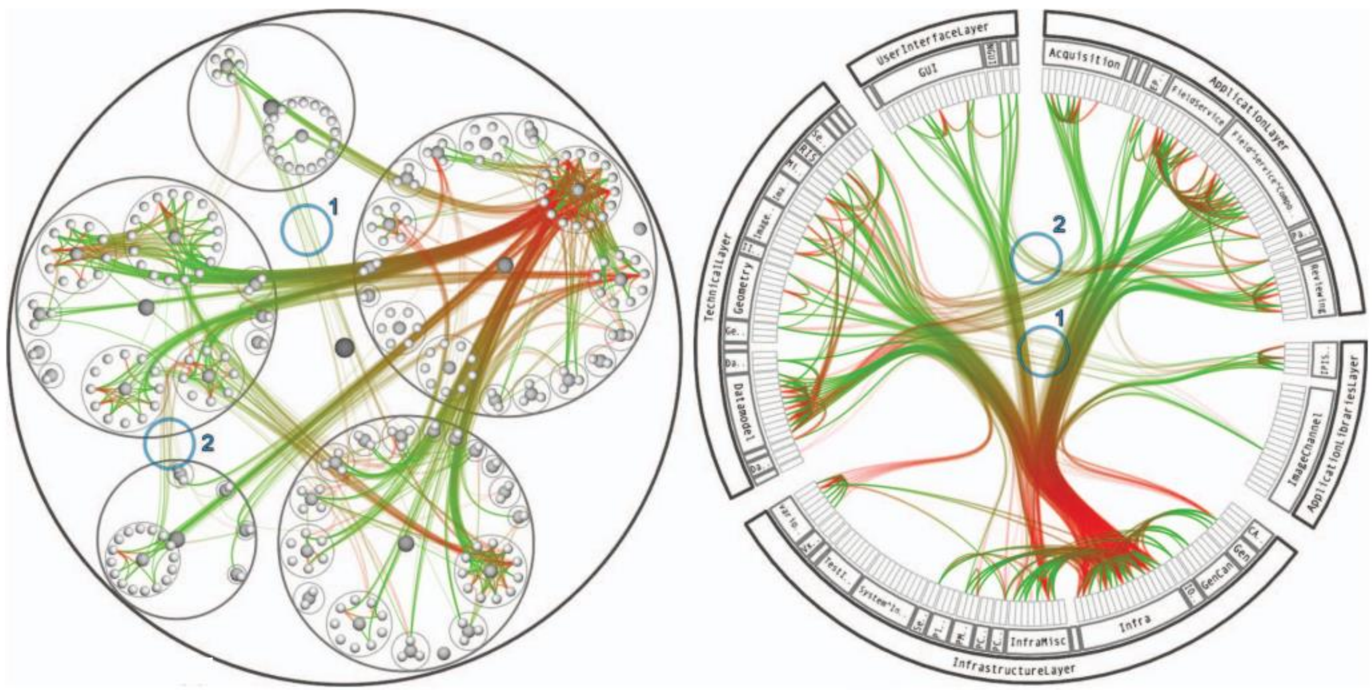
- low values mainly provide low-level, node-to-node connectivity information
- high values provide high-level information

Holten, IEEE TVCG 2006

Edge Bundling Example

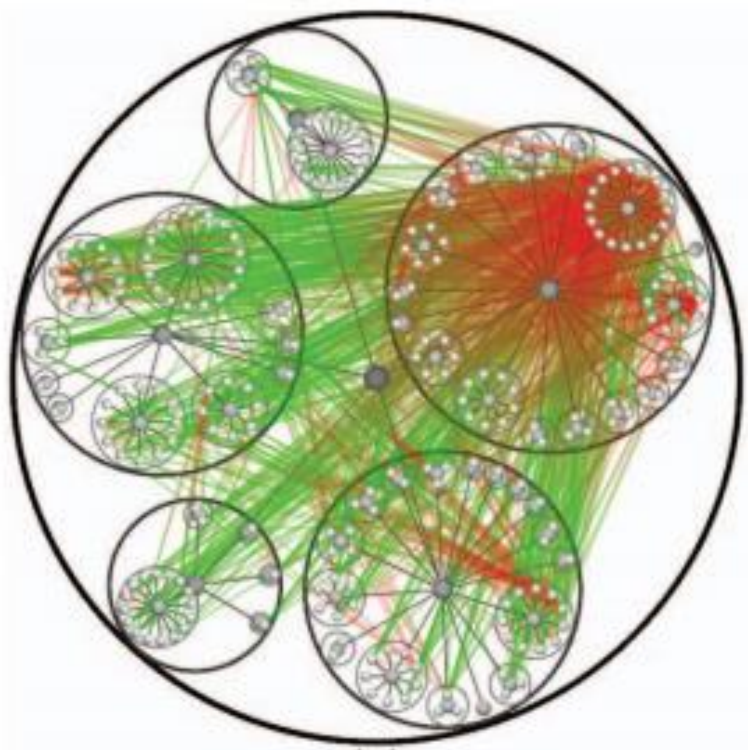
Software system call graph

- green is caller, red is callee

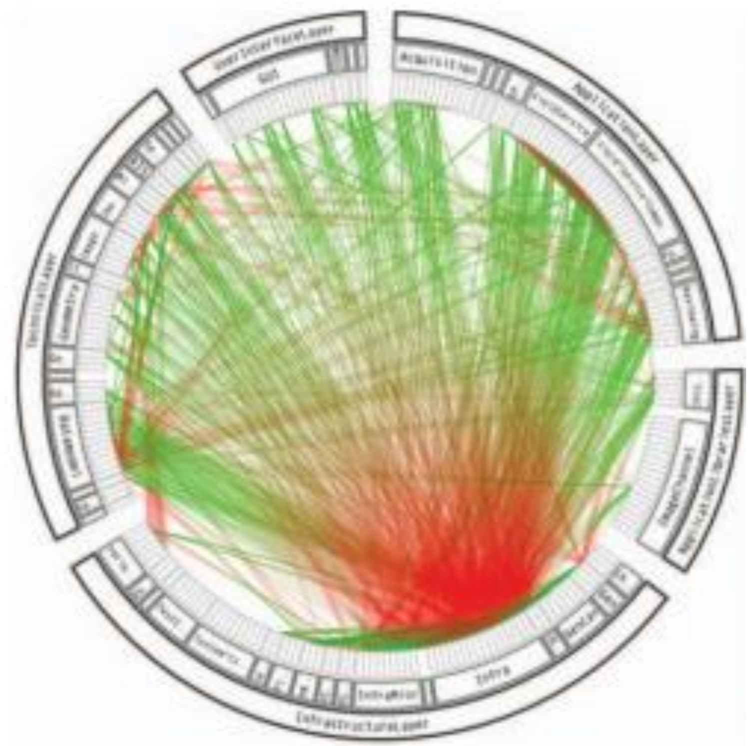


balloon layout (isolated processes) radial layout (more integrated)

Without Edge Bundling



balloon layout

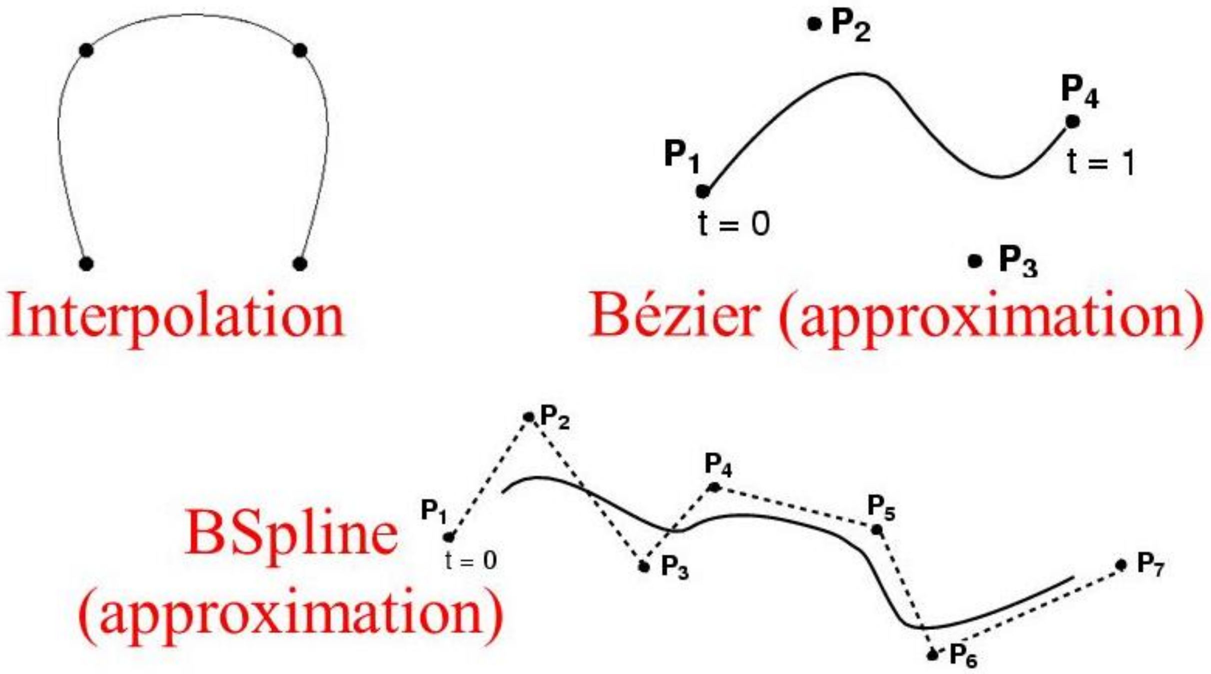


radial layout

Curved Edges Modeled as Splines

Curved edges are represented as *splines*

- a spline is a smooth curve defined by some control points
- moving the control points changes the curve



Primer: Uniform Cubic B-Spline

A B-Spline curve is defined as follows:

$$X(t) = \sum_{k=0}^{n} P_k B_{k,d}(t)$$

- n is the total number of control points
- d is the order of the curves, $2 \le d \le n+1$, d typically 3 or 4
- $B_{k,d}$ are the uniform B-spline blending functions of degree d-1
- P_k are the control points
- Each $B_{k,d}$ is only non-zero for a small range of t values, so the curve has local control

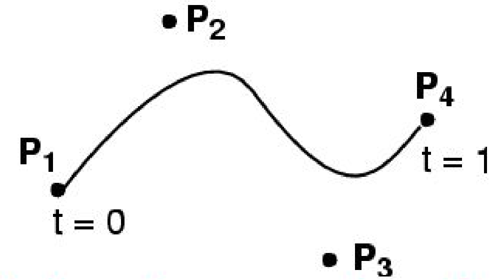
$$x(t) = \frac{1}{6}\begin{bmatrix} P_0 & P_1 & P_2 & P_3 \end{bmatrix} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 0 & 4 \\ -3 & 3 & 3 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} t^3 \\ t^2 \\ t \\ 1 \end{bmatrix}$$

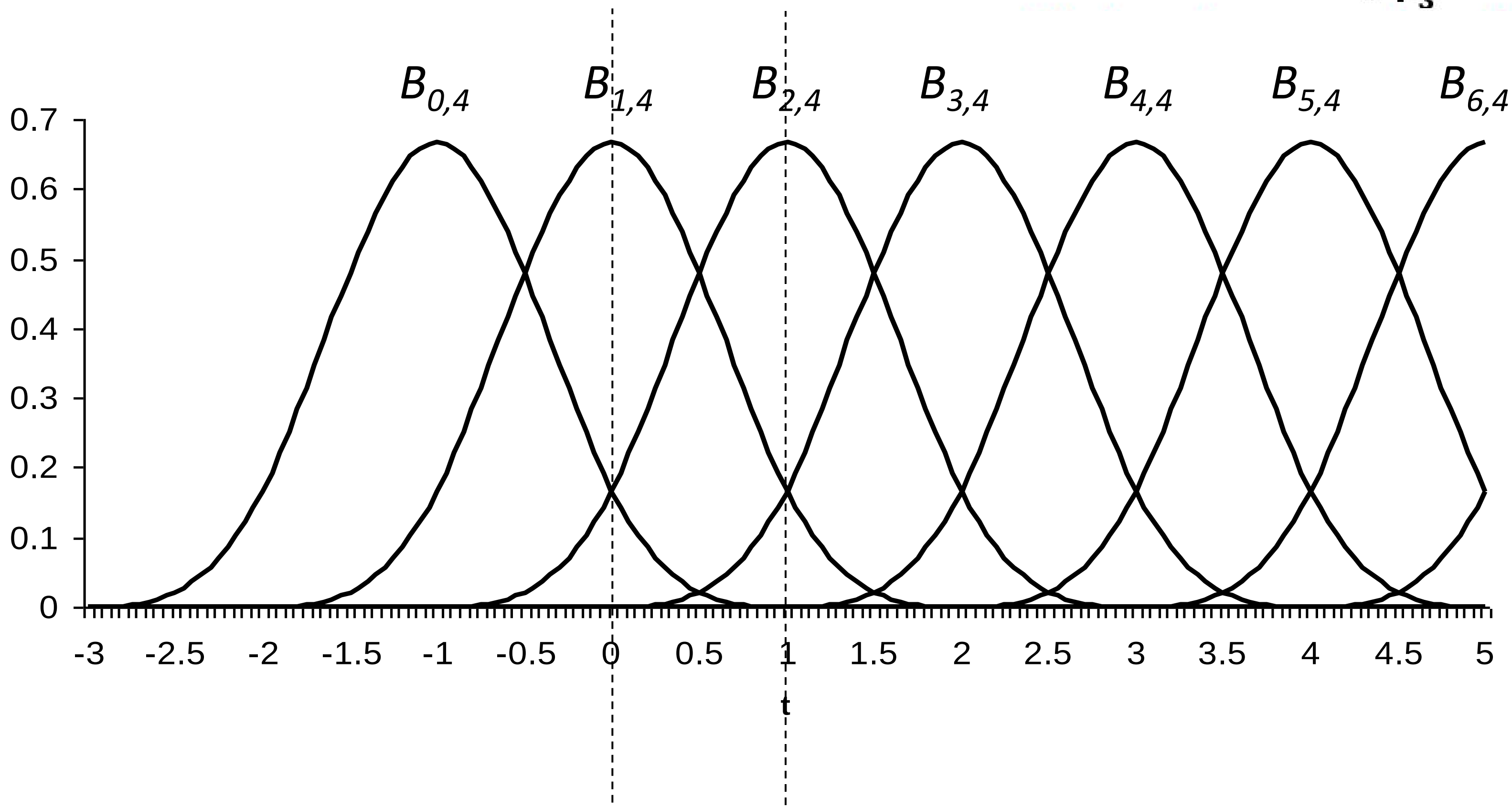
Or in matrix form:

- t is the *parametric variable*
- defined on [0,1]

Primer: Uniform Cubic B-Spline

Four basis functions B must be active to define the B-Spline curve



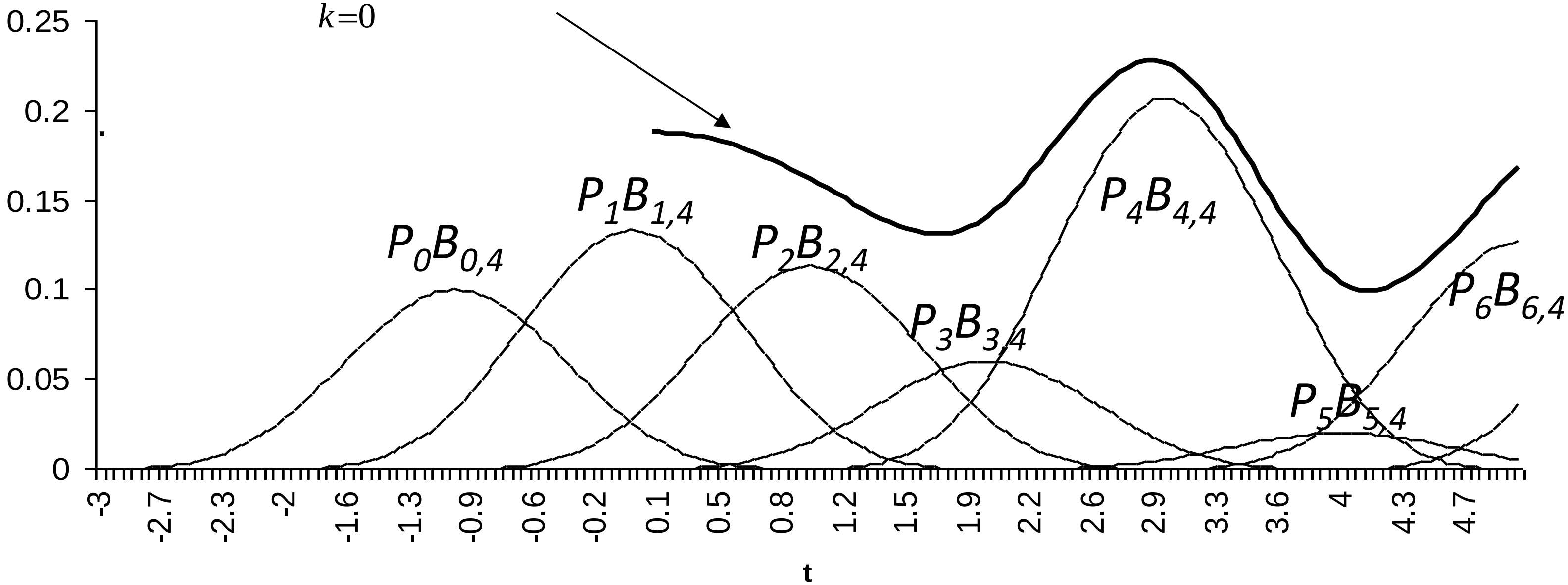


Primer: Uniform Cubic B-Spline

The locations of the control points scale the basis functions

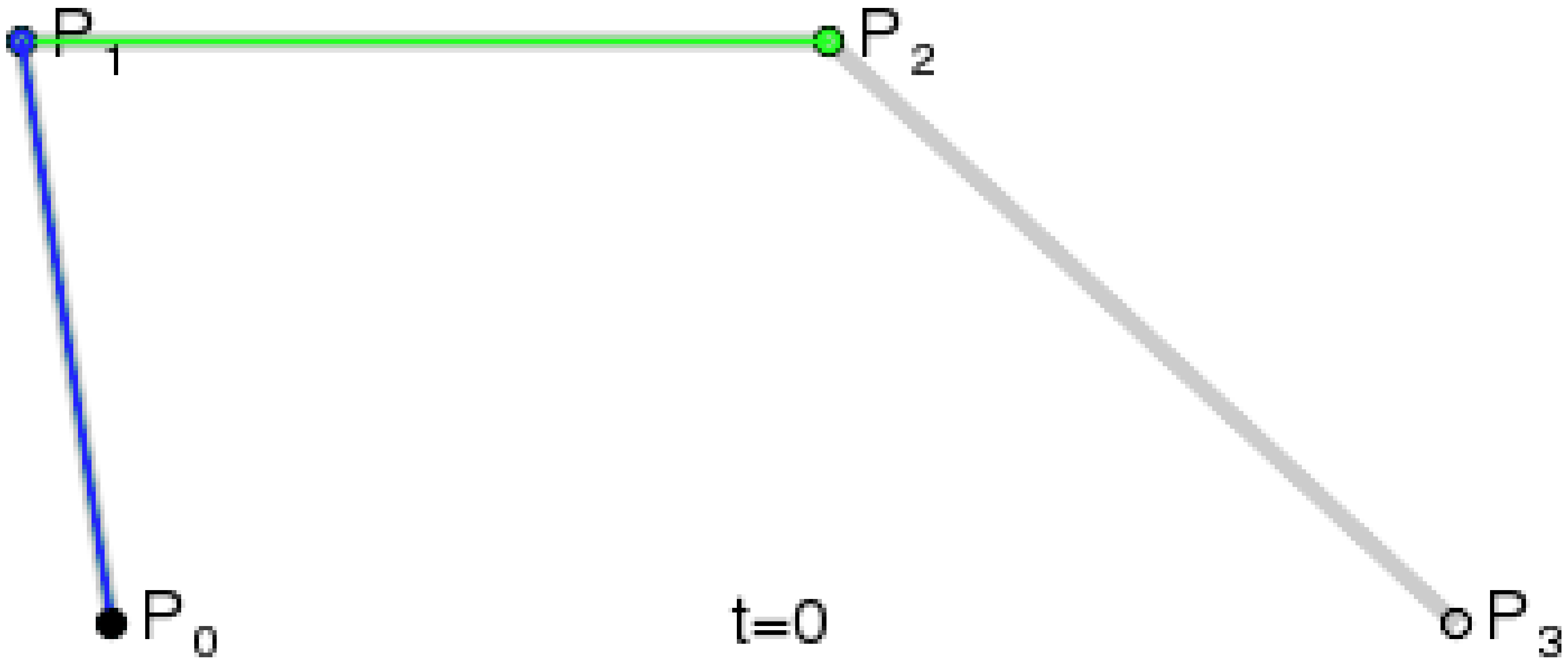
- in this simple example we see a continuous 1D function generated from 6 control points and basis functions

$$X(t) = \sum_{k=0}^{n} P_k B_{k,4}(t)$$



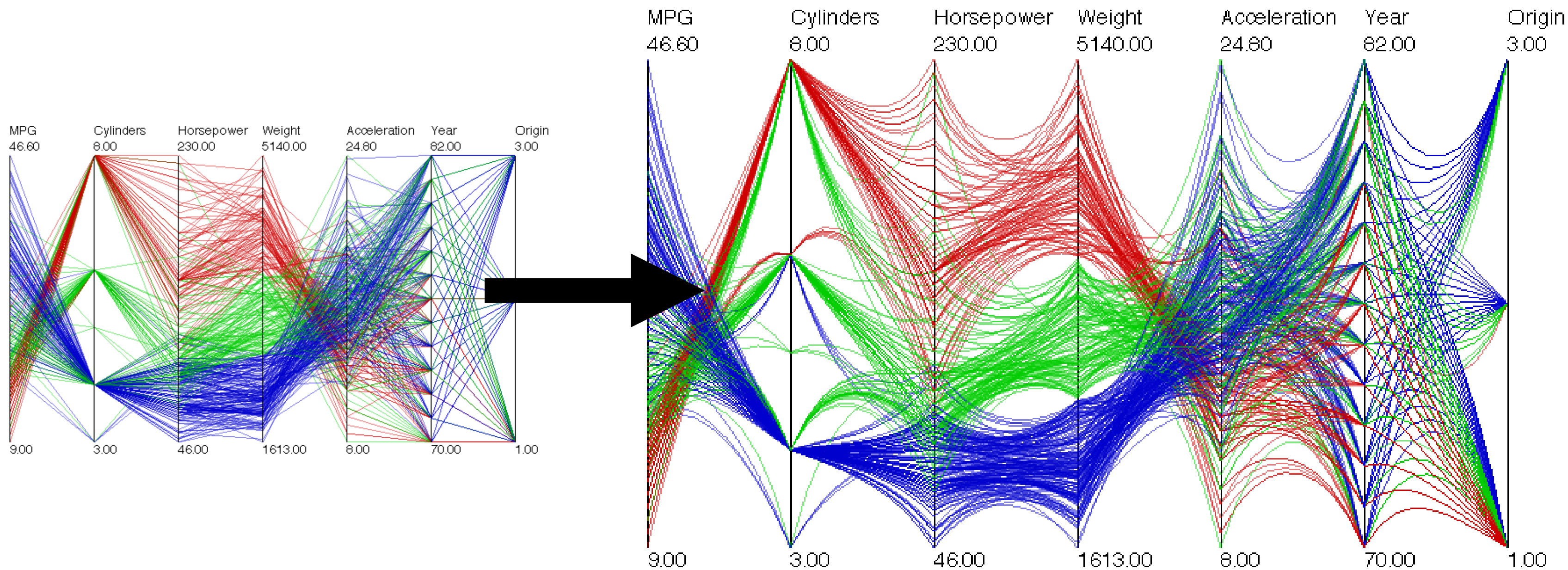
The curve can’t start until there are 4 basis functions active

Cubic B-Spline Animated



Application to Parallel Coordinates

One straightforward way of reducing clutter is to replace polylines with polycurves:



Each line segment is replaced with an end-point interpolating, quadratic B-spline. A tension parameter can be controlled by the user.

McDonnell and Mueller, Computer Graphics Forum, 2008

Edge Bundling (cont.)

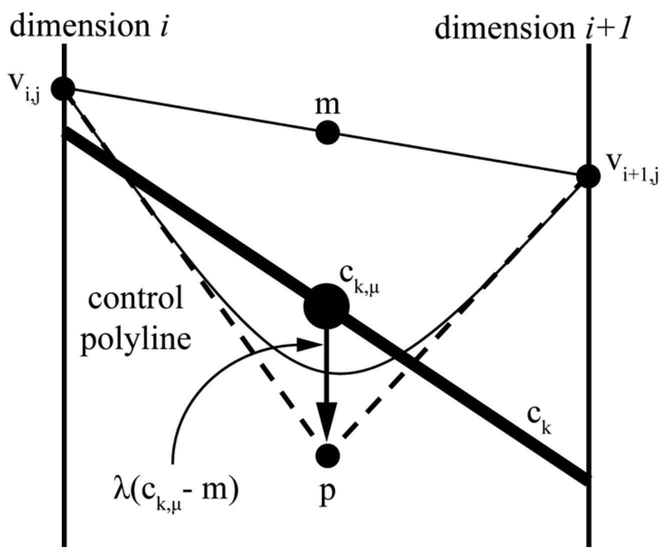
Let m be the mid-point in viewport coordinates of $v_{i,j}$ and $v_{i+1,j}$, end-points of a line segment

Let c_k be the cluster to which this segment belongs and $c_{k,\mu}$ be its mid-point in viewport coordinates

Let λ and β be tension parameters (usually λ = 0.75) and 0 ≤ β ≤ 1 is set by the user

The control points of the spline are given by:

- $(-1, v_{i,j})$
- $(0, \beta m + (1 - \beta)p)$
- $(1, v_{i+1,j})$

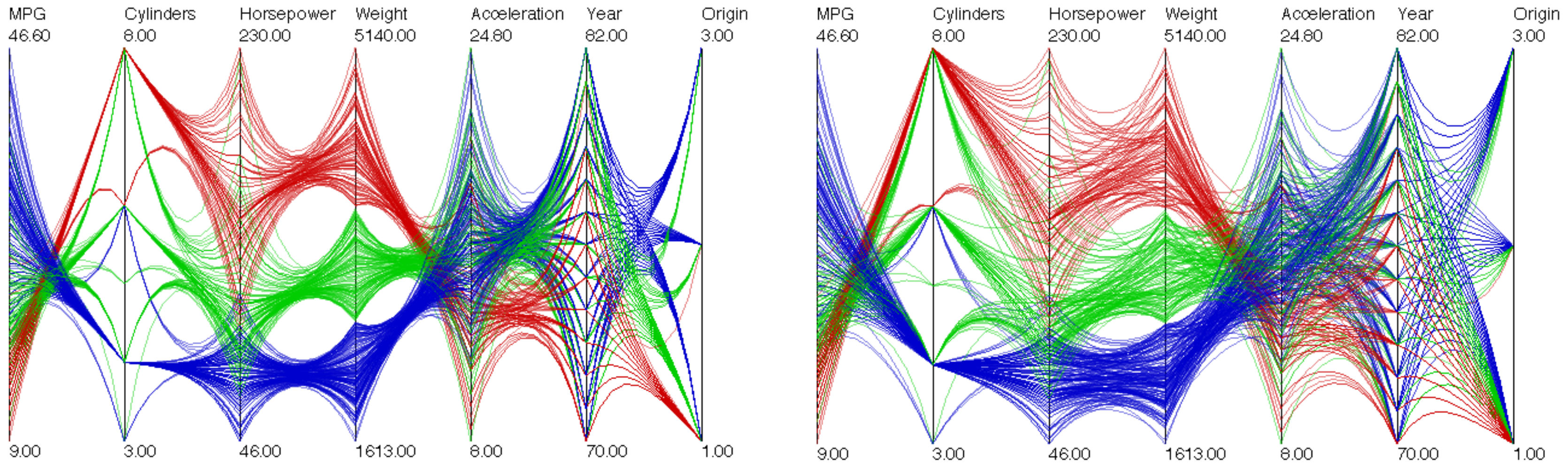


Edge Bundling (cont.)

The tension can be changed to control the amount of clutter reduction

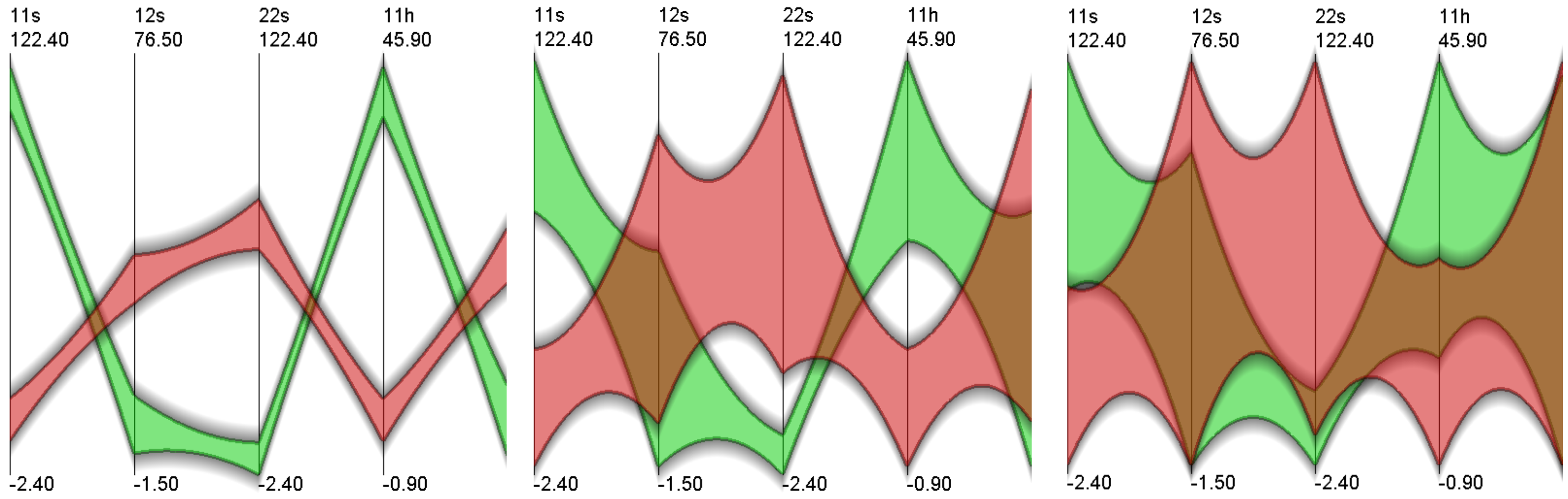
In our implementation, the λ parameter is fixed, but the β parameter can be changed in the GUI

Examples of medium and low tension, respectively:



Cluster Rendering

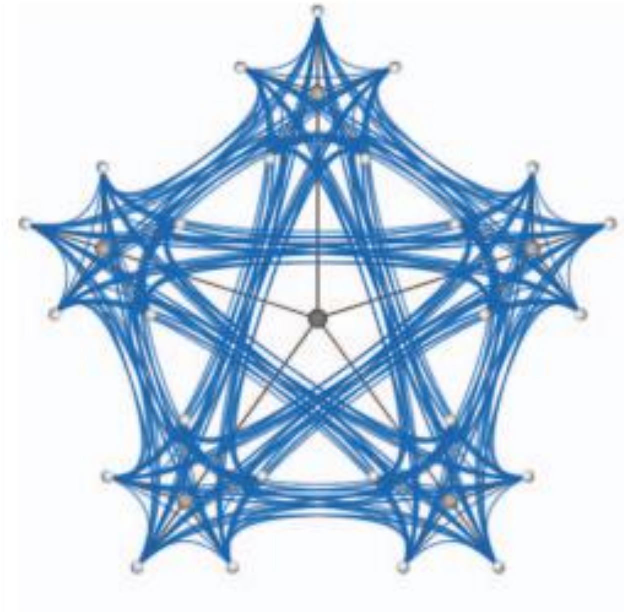
- Recall that clusters are often rendered as heavy line segments on top of the dataset
- In IPC we render the clusters as polygonal meshes
- They help to show the ranges of each cluster along axes
- The vertical "spread" can be controlled by the user



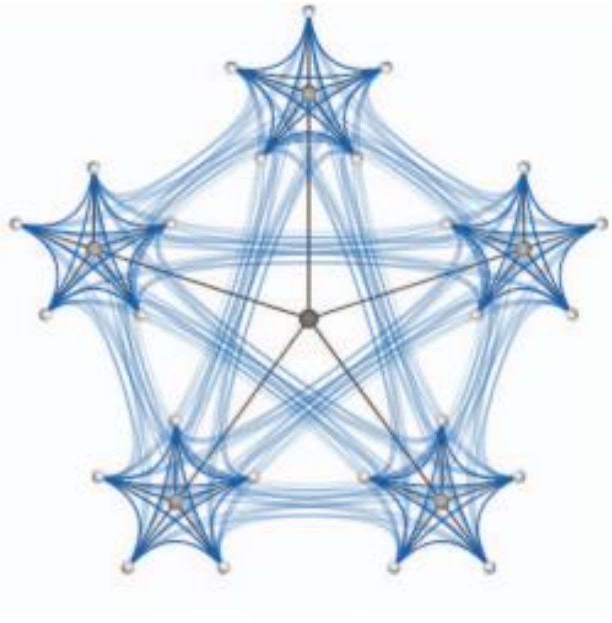
Alpha (Opacity) Blending

Draw curves at different opacities

- long curves: low opacities (high transparencies)
- short curves: high opacity (makes short curves visible)



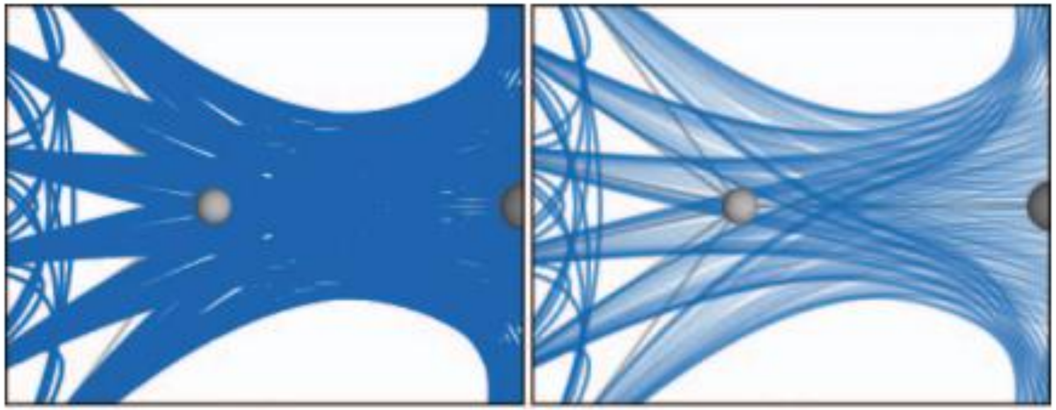
alpha blending disabled



alpha blending enabled

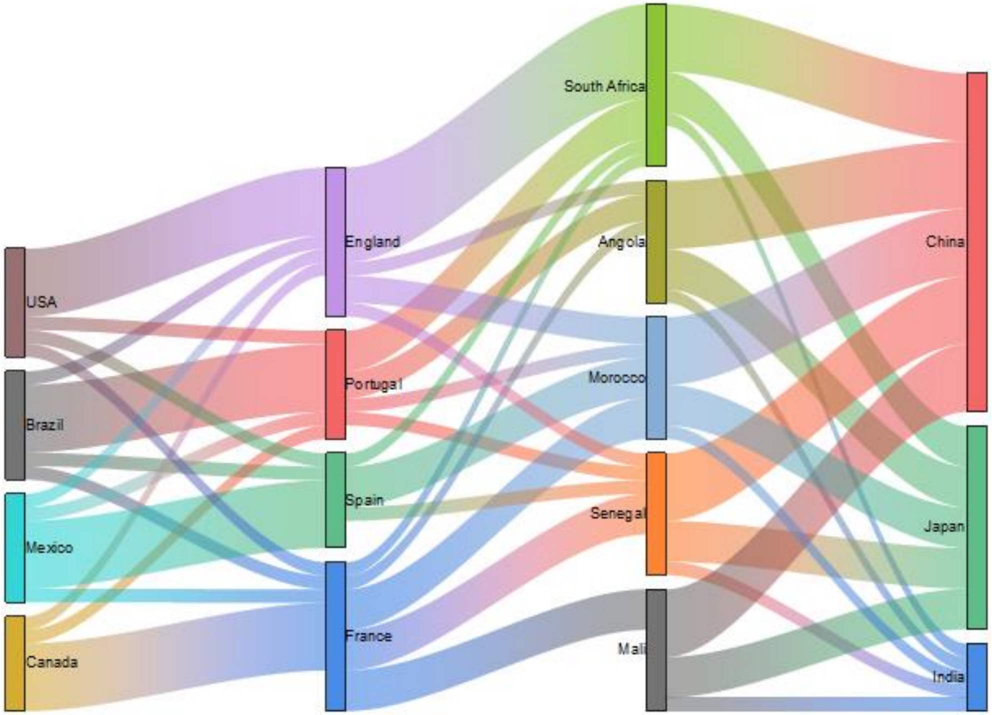
Alpha (Opacity) Blending

Alpha blending also enables visualization of sub-bundles and differentiation of lines



alpha blending disabled   alpha blending enabled

Sankey Diagram



Another bundling technique

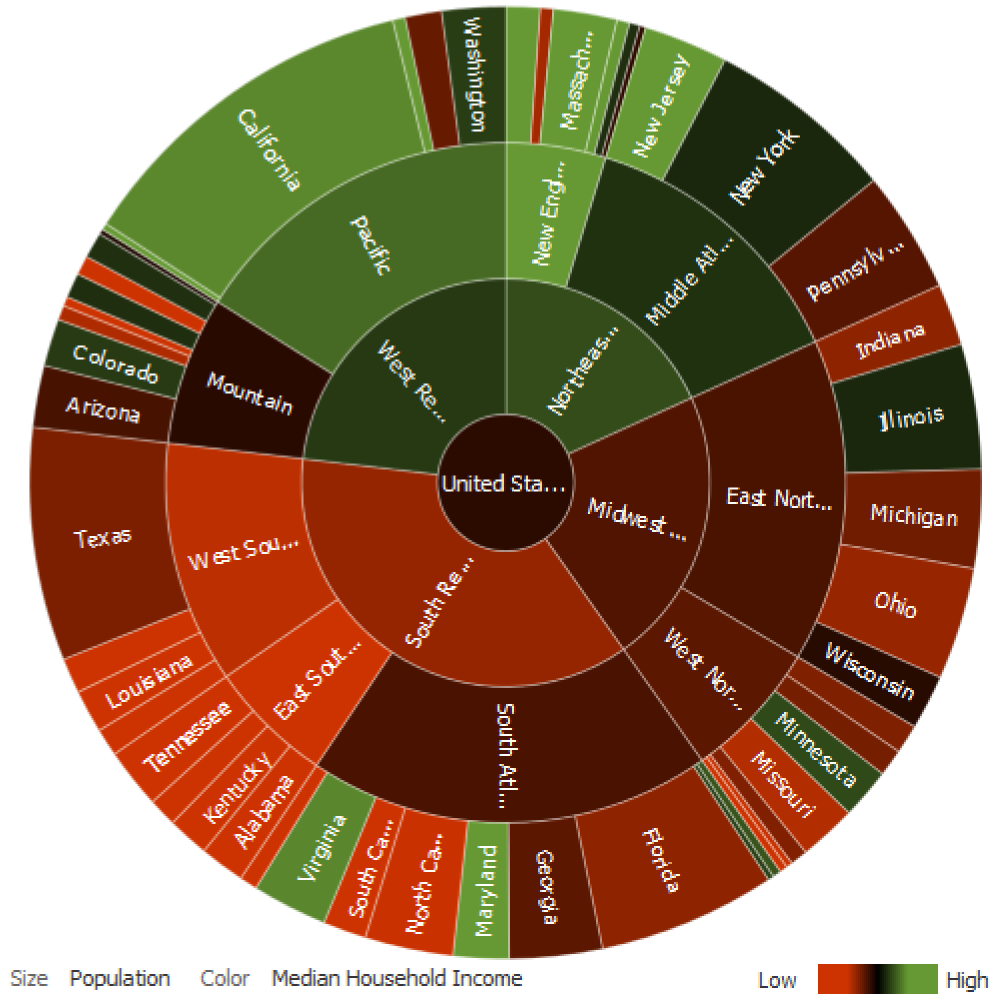
- flow diagram
- the width of the arrows is proportional to the flow rate

Use cases:

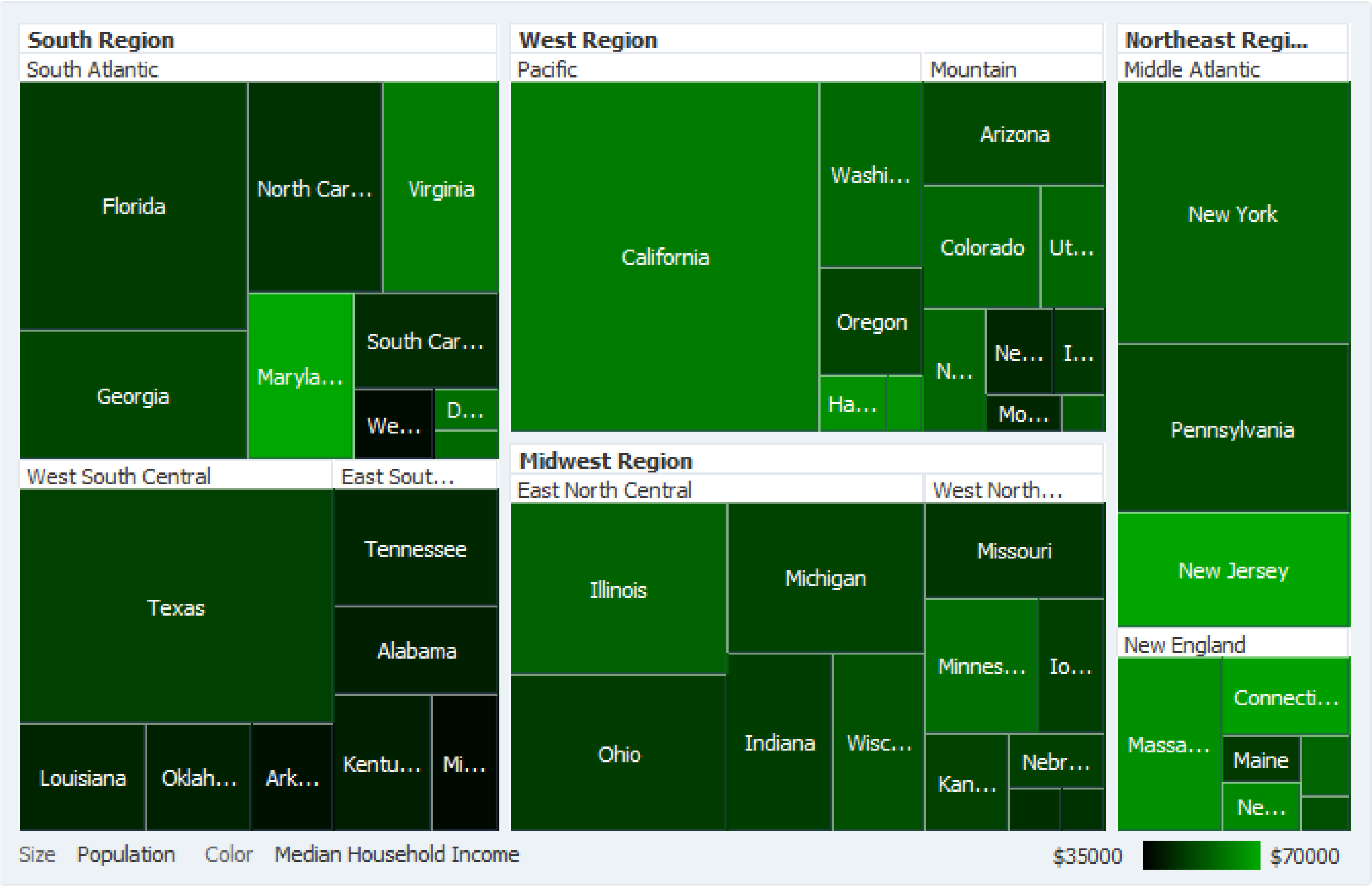
- where money came from and went to (budgets, contributions)
- flows of energy from source to destination
- flows of goods from place to place

Hierarchies with Sun Burst Displays

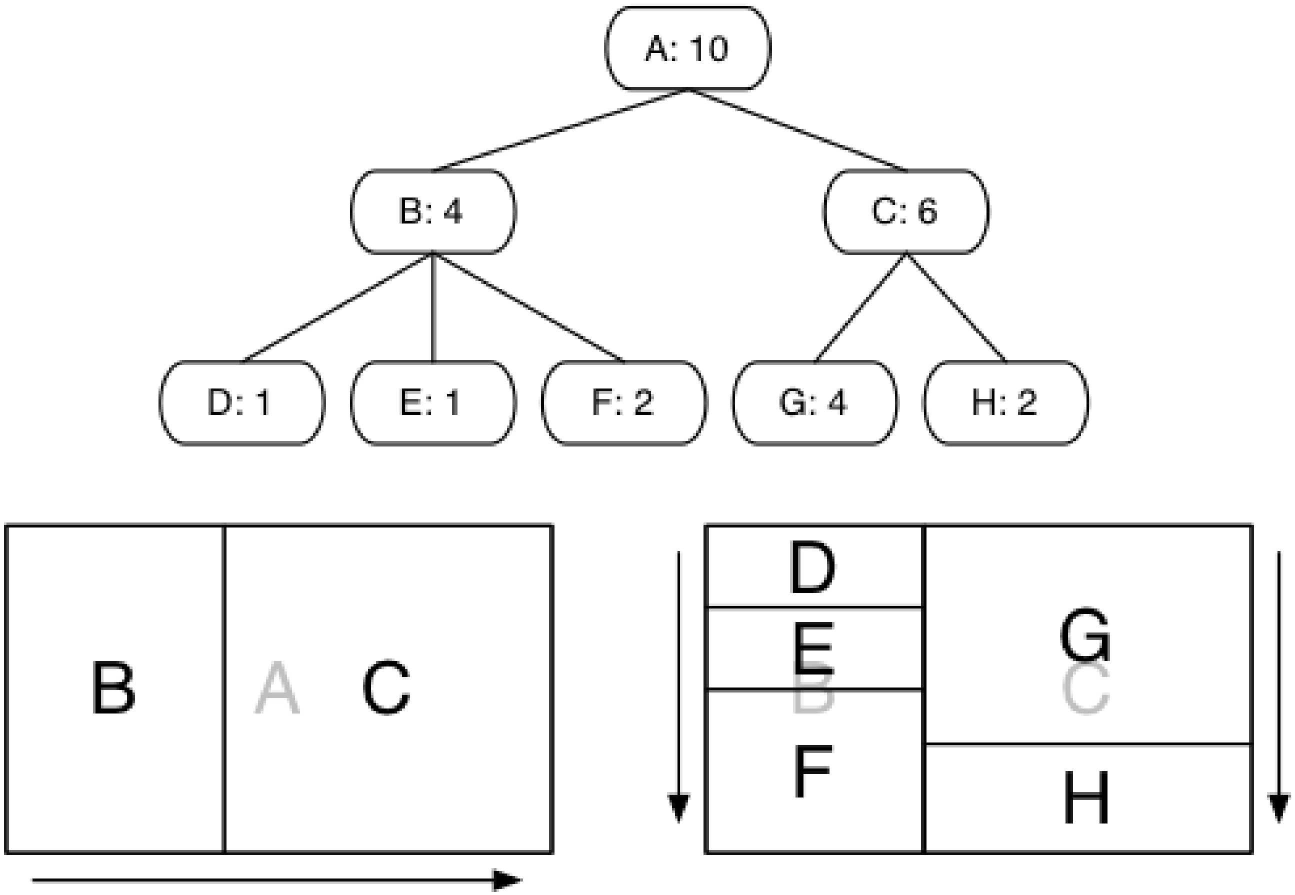
Sunburst with Partition of Unity



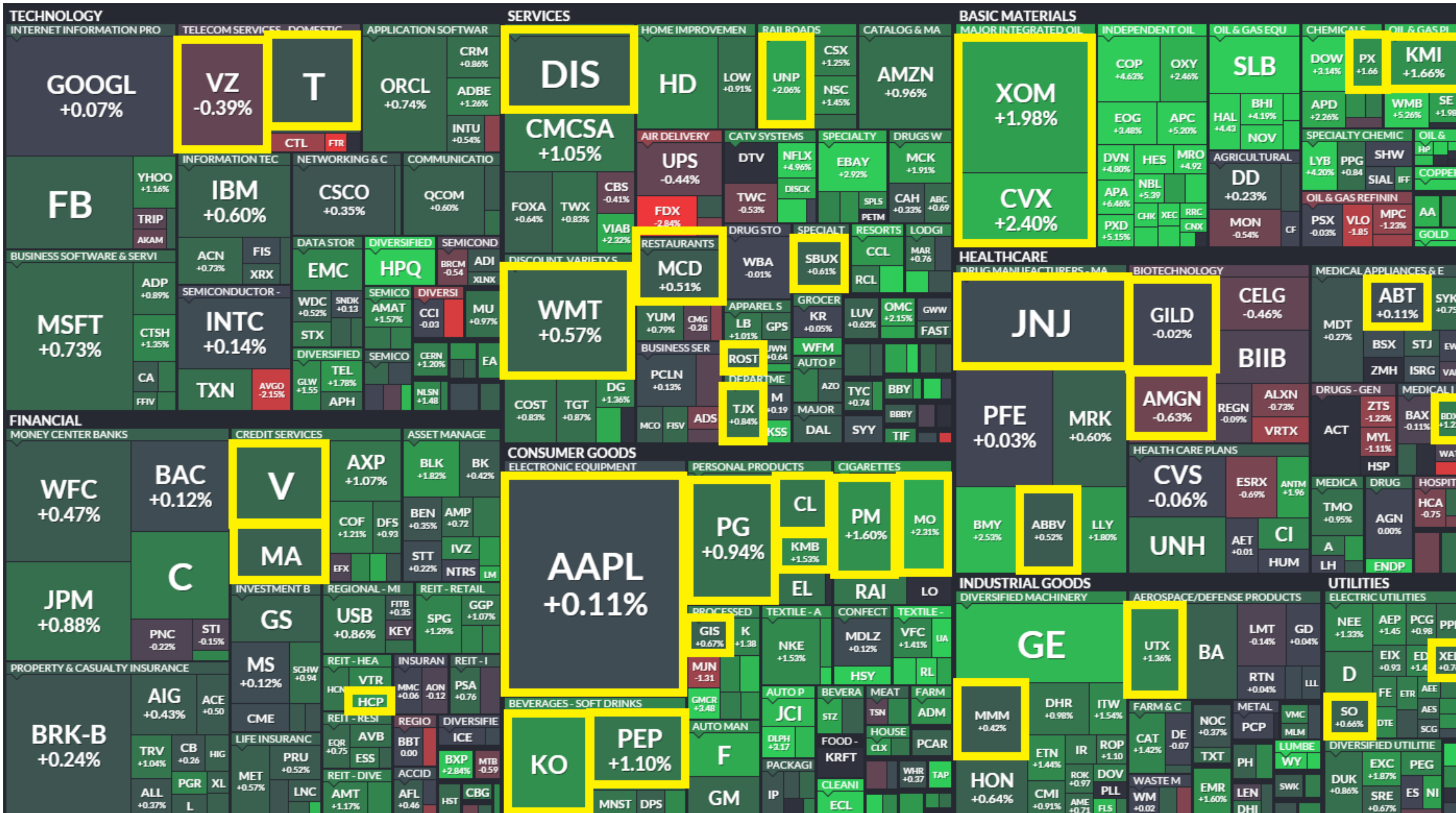
Same Data with Treemap



Treemap Constructon

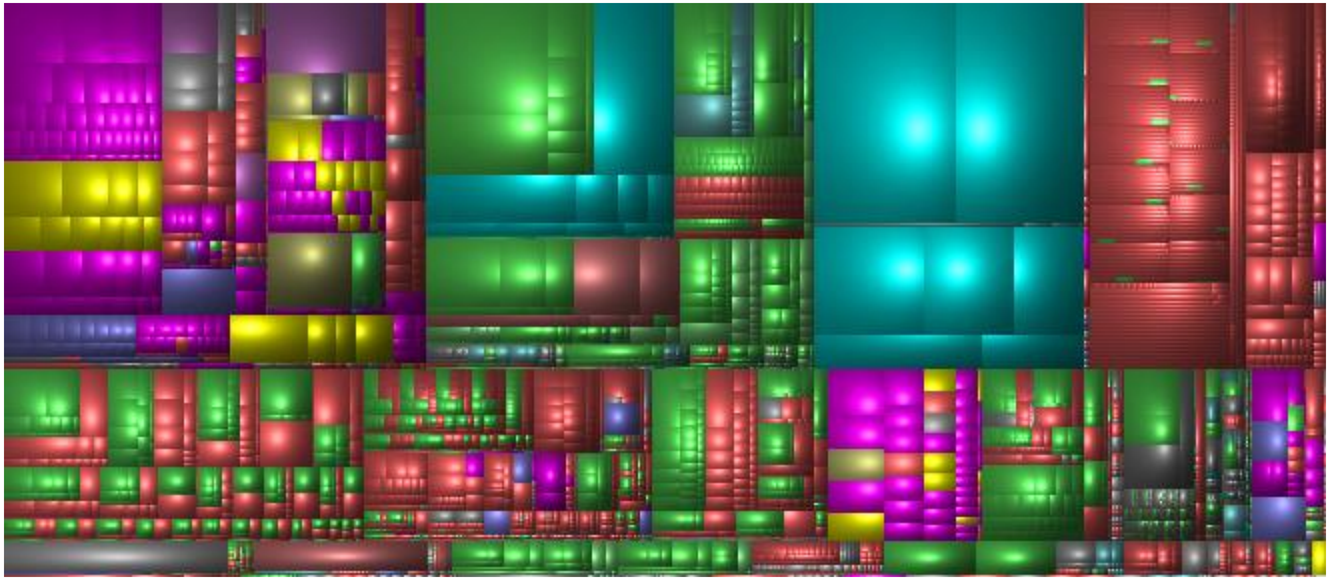


Treemap for Stock Portfolio



Size is mapped to market cap, yellow boxes are investor's holdings

Cushion Treemap



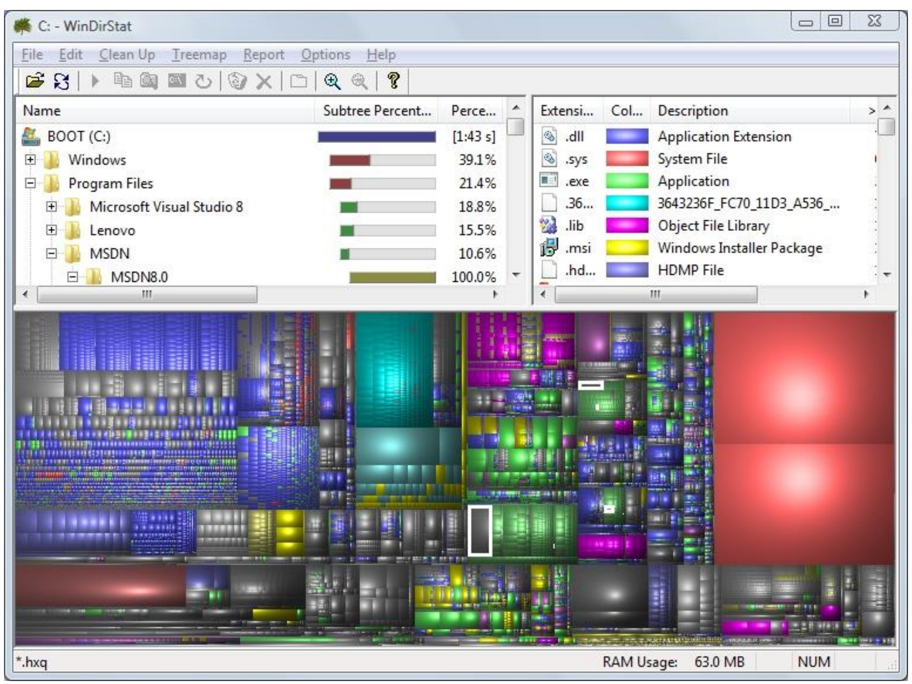
Advantages

- due to perceived discontinuity in texture between nodes, lines are no longer necessary to separate nodes
- more of the space can be used for the actual node display
- much smaller nodes can be shown than in a flat treemap

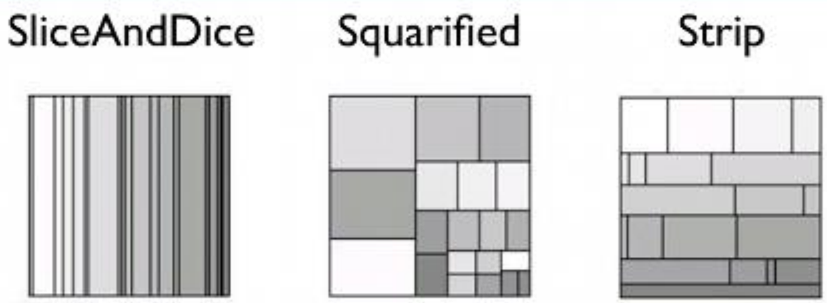
Tree map for Disk Drives

Used in programs like

- WinDirStat (Windows)
- KDirStat (Linux)
- DiskInventory (Mac)



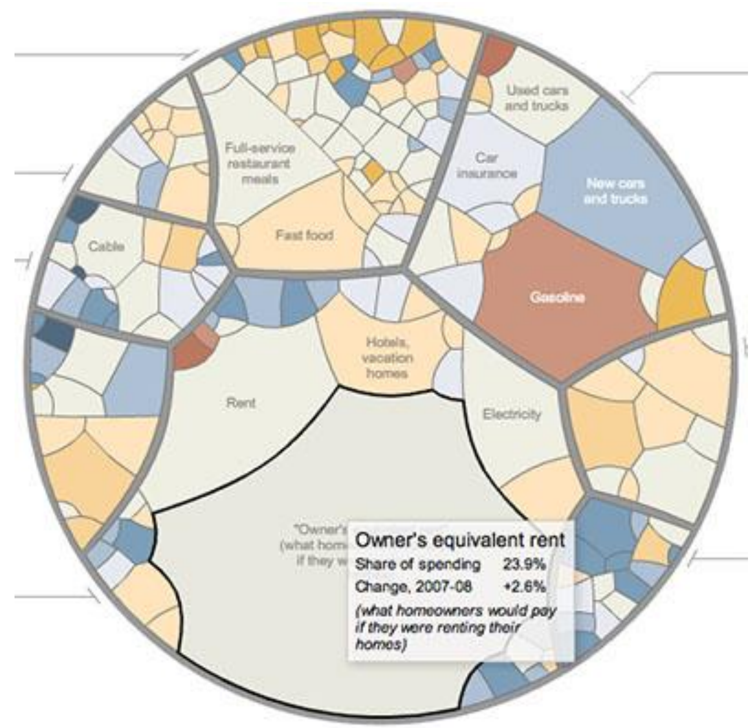
Treemap variations



order	ordered	unordered	ordered
aspect ratios	very high	lowest	medium
stability	stable	medium	medium

Voronoi treemap

- based on Voronoi tesselation



Squarified treemap is preferred

- it’s difficult to visually compare long slivery tiles with tiles that have a more even aspect ratio
- a squarified treemap makes the map more globally comparable