CSE 332
Introduction to Visualization

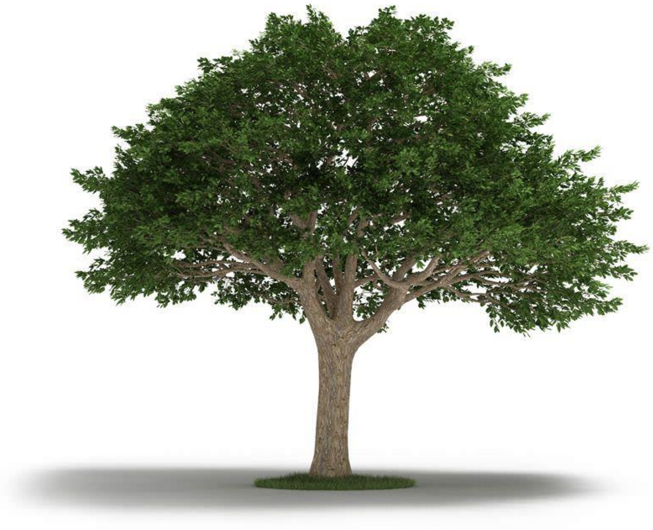
Visualization of Hierarchies

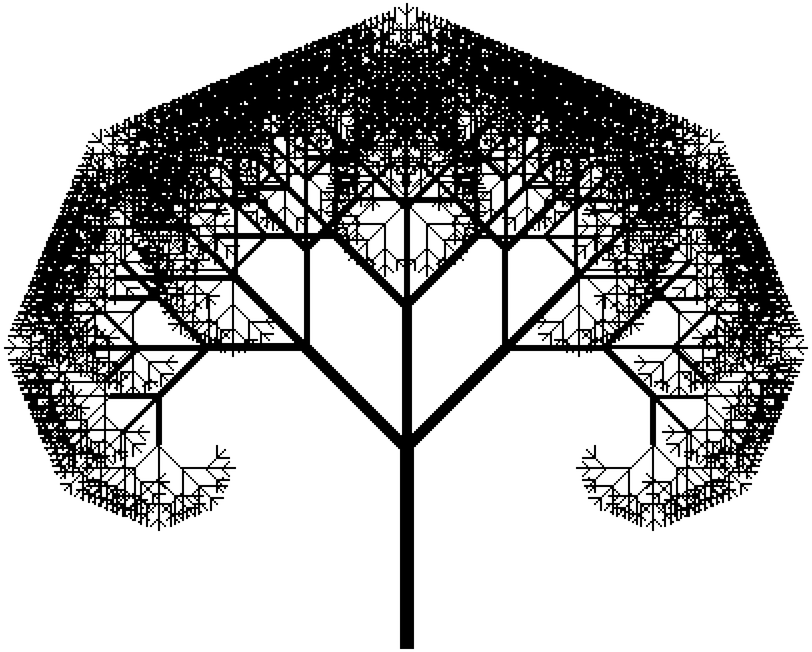
Klaus Mueller

Computer Science Department
Stony Brook University

Lecture	Topic	Projects
1	Intro, schedule, and logistics	
2	Applications of visual analytics, data, and basic tasks	
3	Data preparation and reduction	Project 1 out
4	Data preparation and reduction	
5	Data reduction and similarity metrics	
6	Dimension reduction	
7	Introduction to D3	Project 2 out
8	Bias in visualization	
9	Perception and cognition	
10	Visual design and aesthetics	
11	Cluster and pattern analysis	
12	High-Dimensional data visualization: linear methods	
13	High-D data vis.: non-linear methods, categorical data	Project 3 out
14	Principles of interaction	
15	Visual analytics and the visual sense making process	
16	VA design and evaluation	
17	Visualization of graphs and hierarchies	
18	Visualization of time-varying and time-series data	Project 4 out
19	Midterm	
20	Maps and geo-vis	
21	Computer graphics and volume rendering	
22	Techniques to visualize spatial (3D) data	Project 4 halfway report due
23	Scientific and medical visualization	
24	Scientific and medical visualization	
25	Non-photorealistic rendering	
26	Memorable visualizations, visual embellishments	Project 5 out
27	Infographics design	
28	Projects Hall of Fame demos	

Hierarchies = Trees

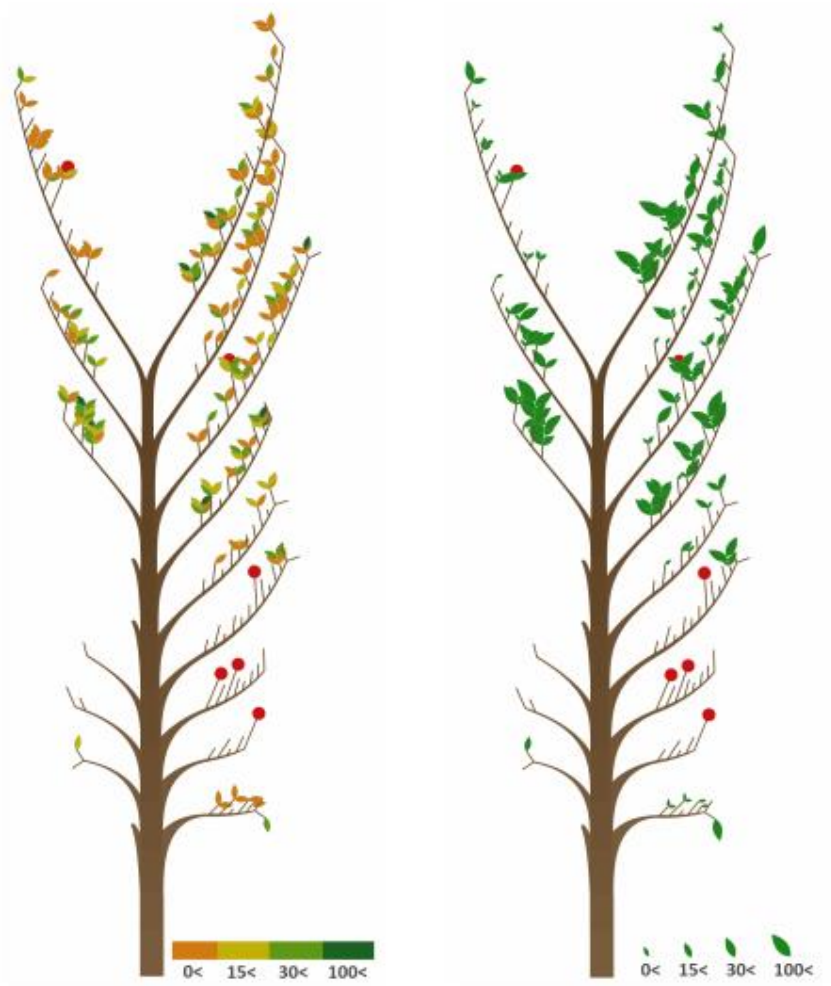




Tree – A Natural Metaphor

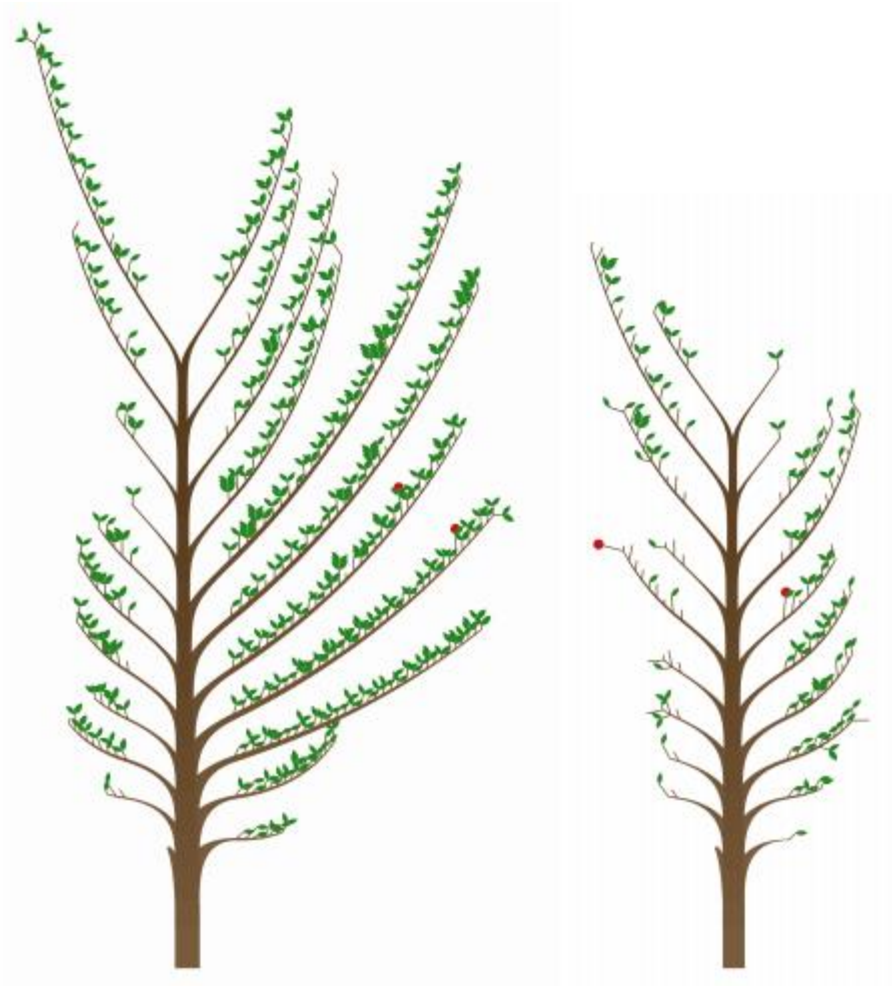
Mapping publications to a tree

- 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

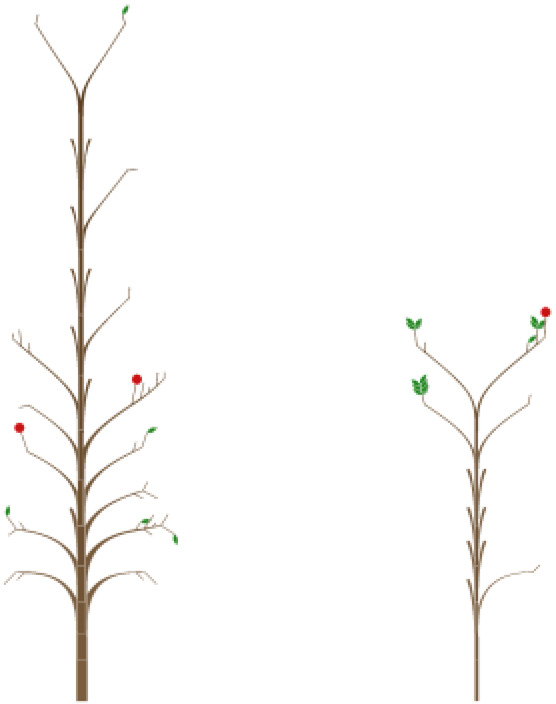


Fung and Ma, Personal VIS 2015

Productive vs. Unproductive Researchers



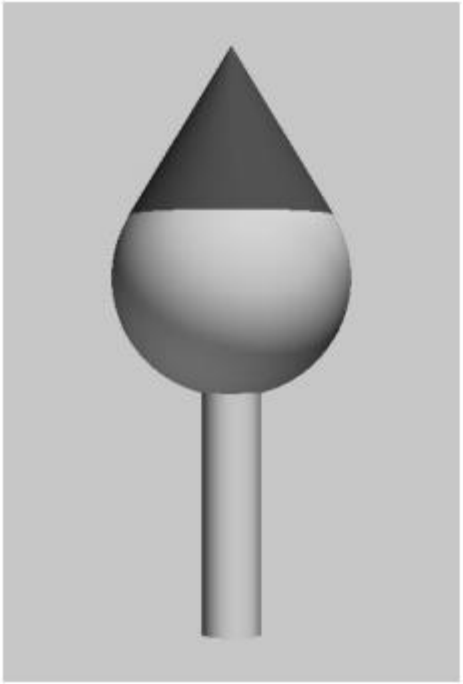
Productive



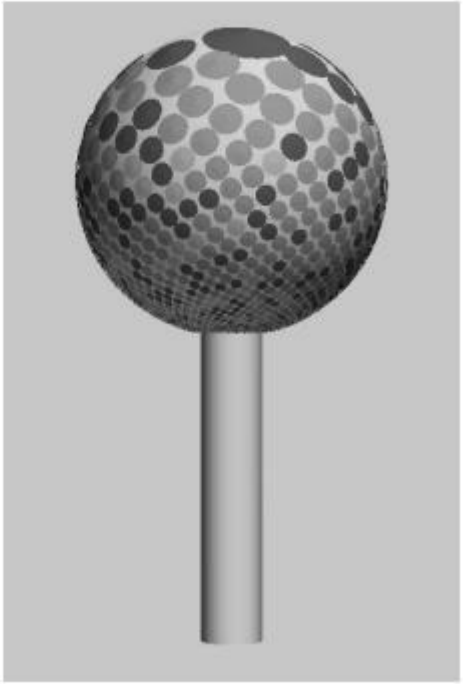
Unproductive

Botanical-Inspired Visualizations

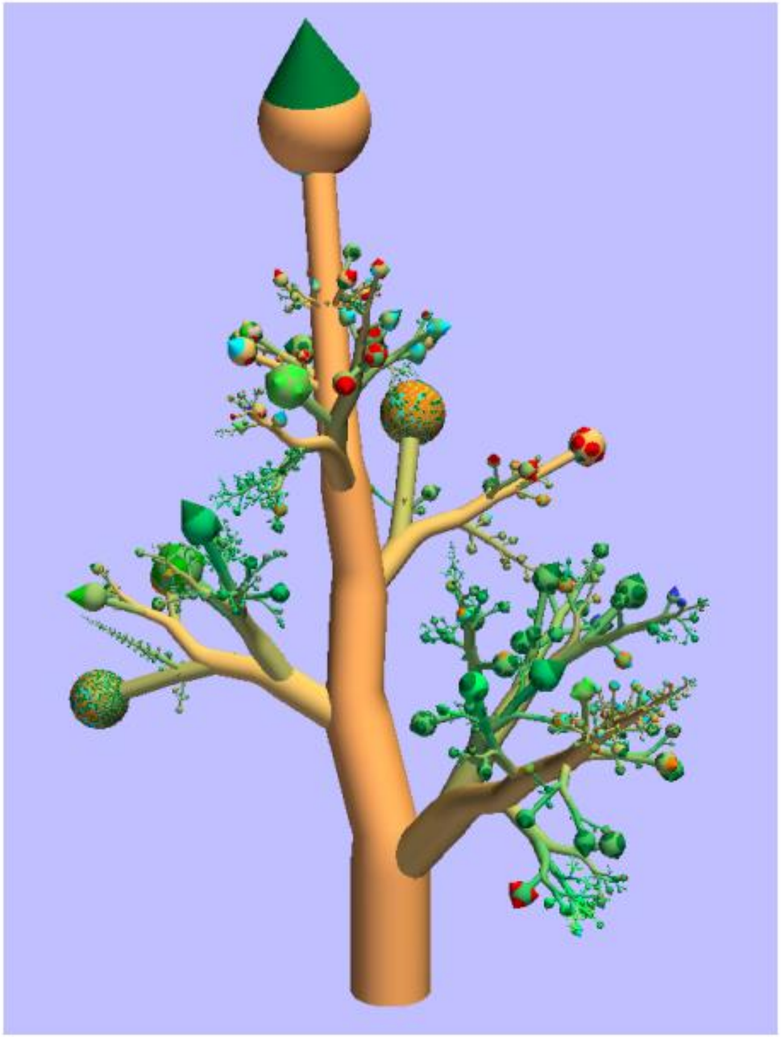
Visualizing hard drives with tree cartoons



one file



many files

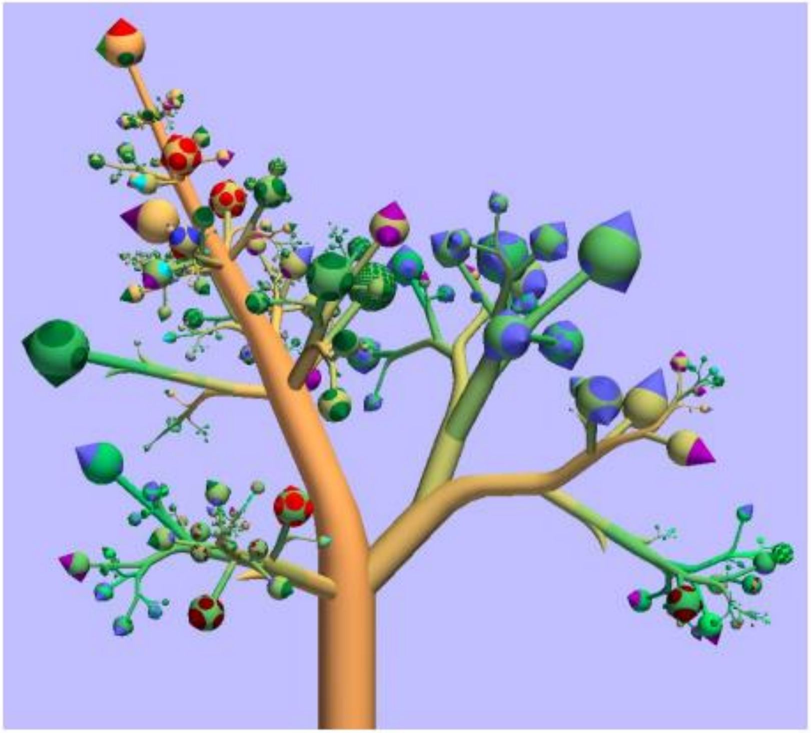


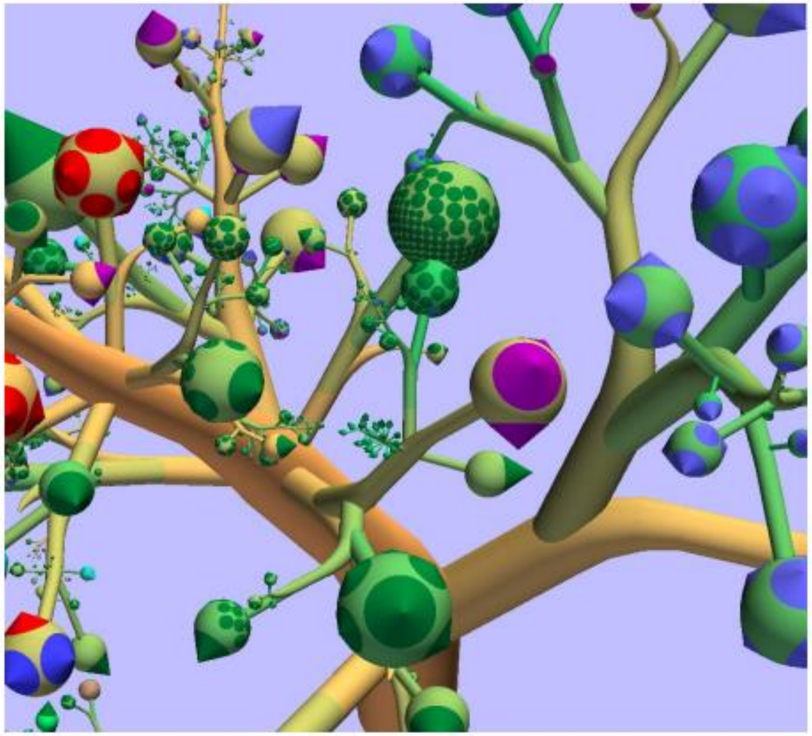
Kleiberg et al., IEEE InfoVis 2001

Botanical-Inspired Visualizations

Color maps to file type

- blue are pdf files, red are image files

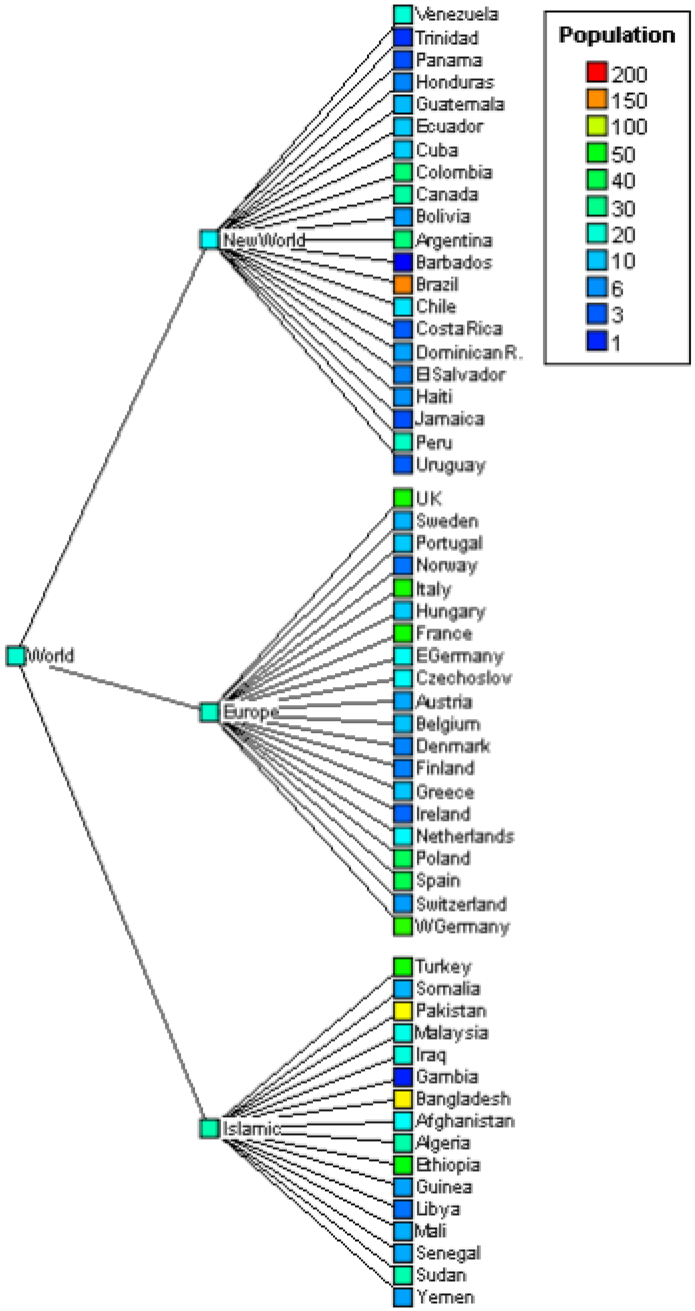




Conventional

Standard Node-Edge layout for a hierarchical network

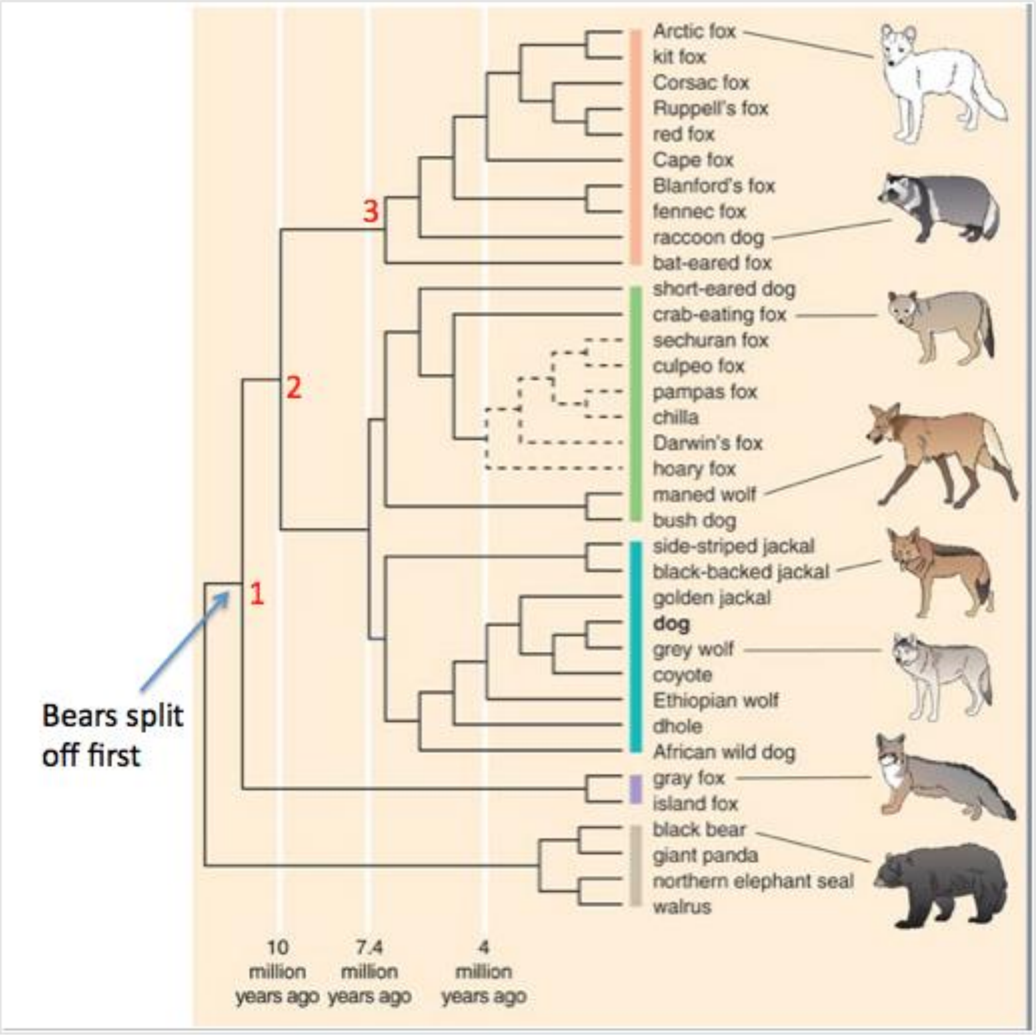
- 3 levels
- color maps to quantitative information (here population)



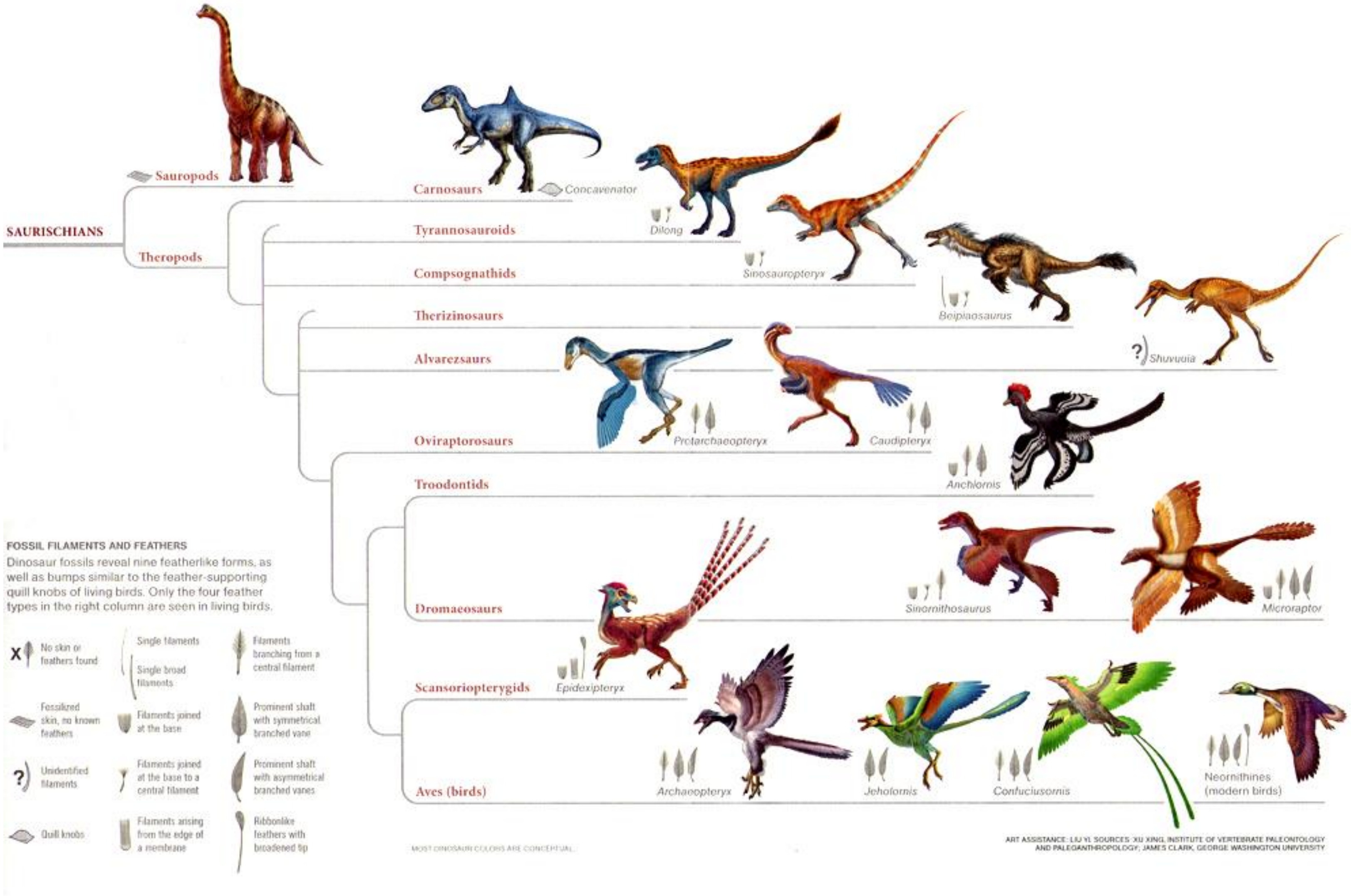
DENDROGRAM

Typically used to depict classification hierarchies

- split-off points visualize proximity


Bears split off first

Birds and Dinosaurs



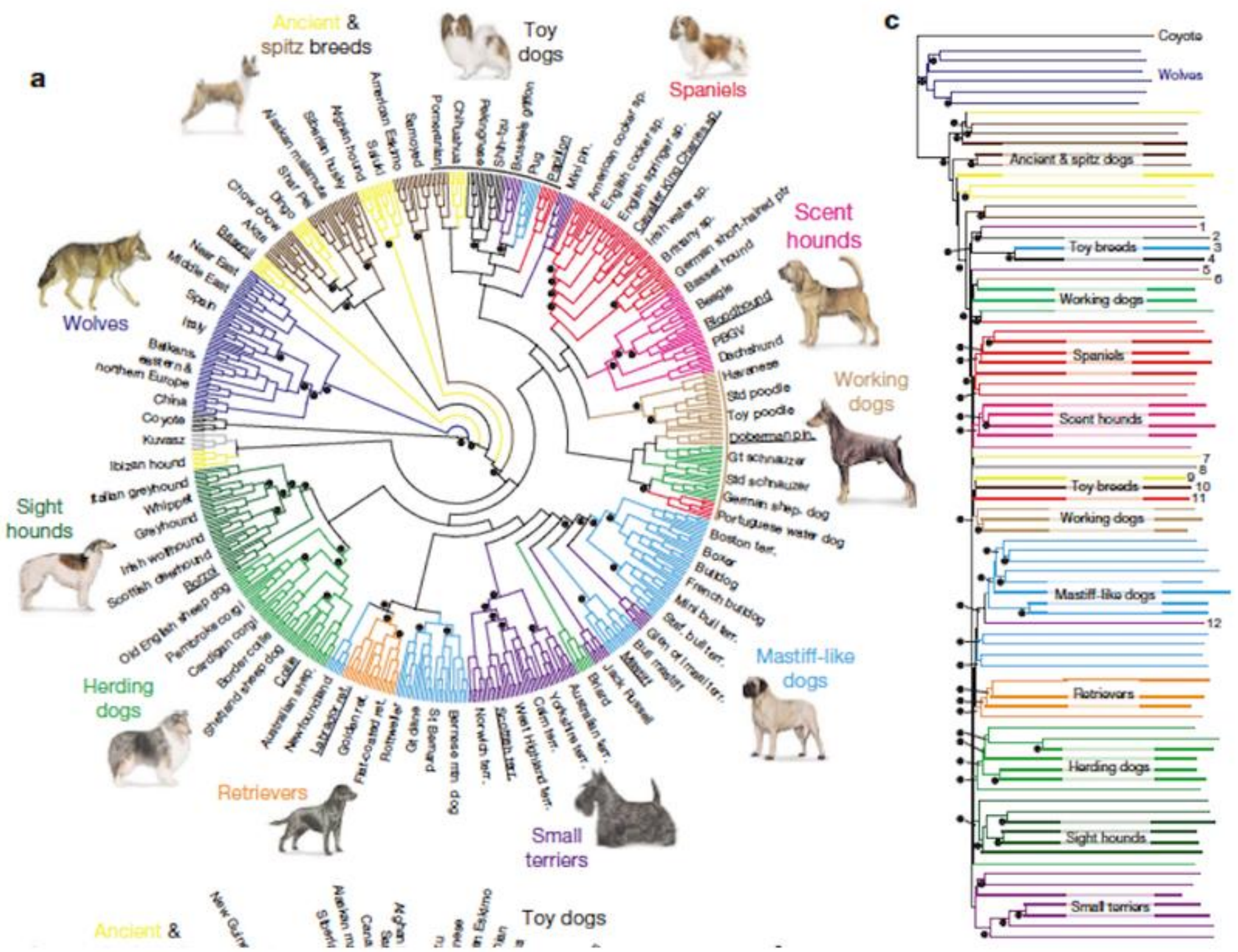
FOSSIL FILAMENTS AND FEATHERS

Dinosaur fossils reveal nine featherlike forms, as well as bumps similar to the feather-supporting quill knobs of living birds. Only the four feather types in the right column are seen in living birds.

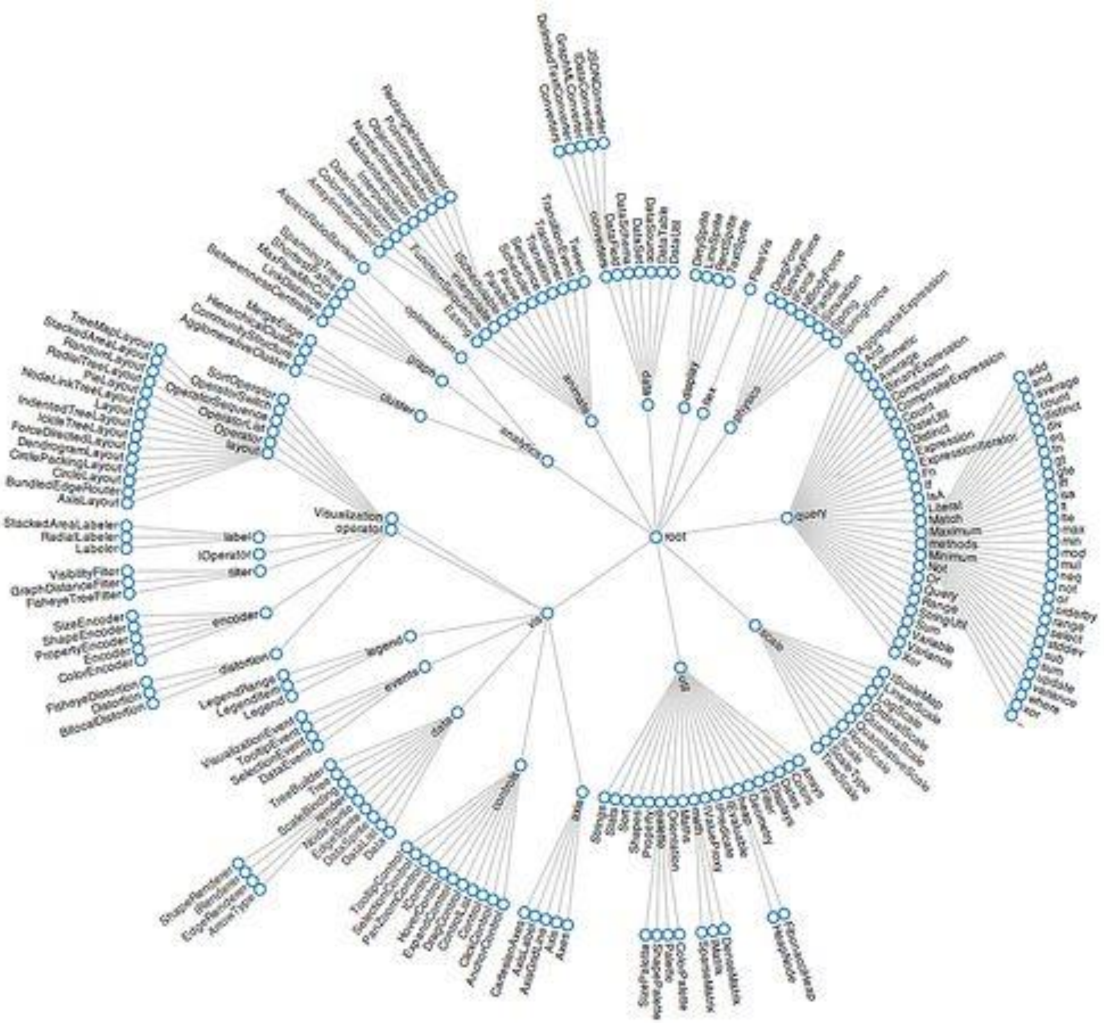
MOST DINOSAUR COLORS ARE CONCEPTUAL.

ART ASSISTANCE: LIU YI. SOURCES: XU XING, INSTITUTE OF VERTEBRATE PALEONTOLOGY AND PALEOANTHROPOLOGY; JAMES CLARK, GEORGE WASHINGTON UNIVERSITY

Circles Are More Space-Efficient



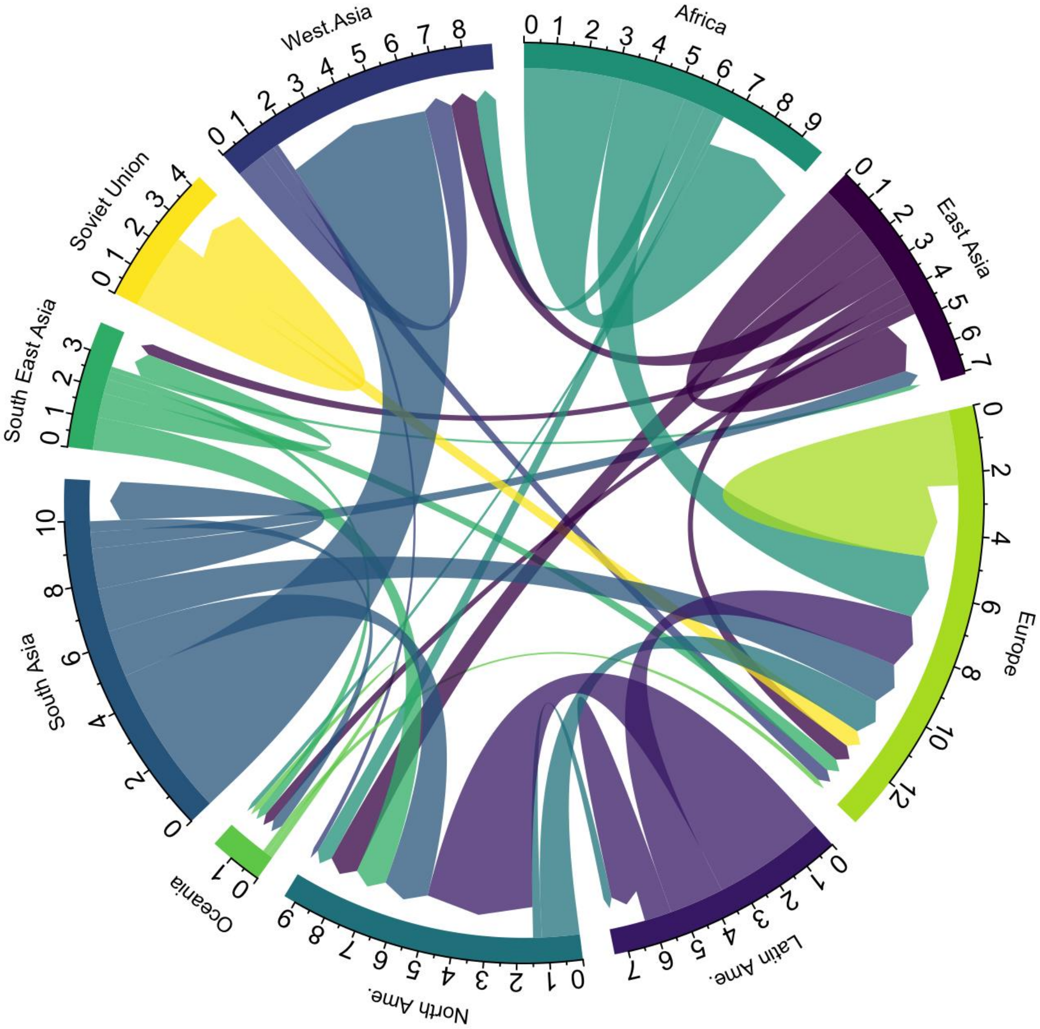
Radial Hierarchical Visualization



Chord Diagrams

Represents flows or connections between several entities

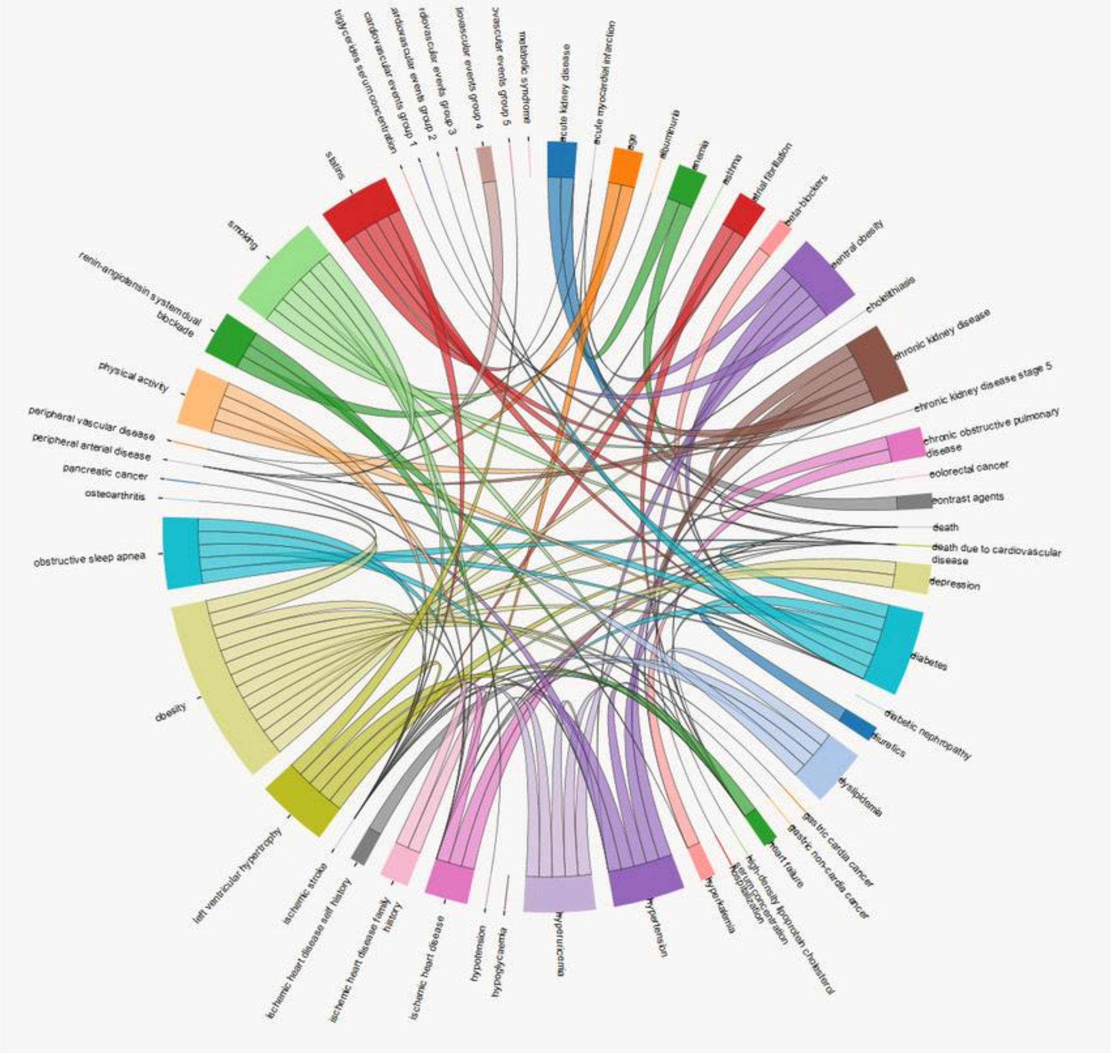
- for example the number of people migrating from one country to another



More Complex Chord Diagram

Can we make it easier to read?

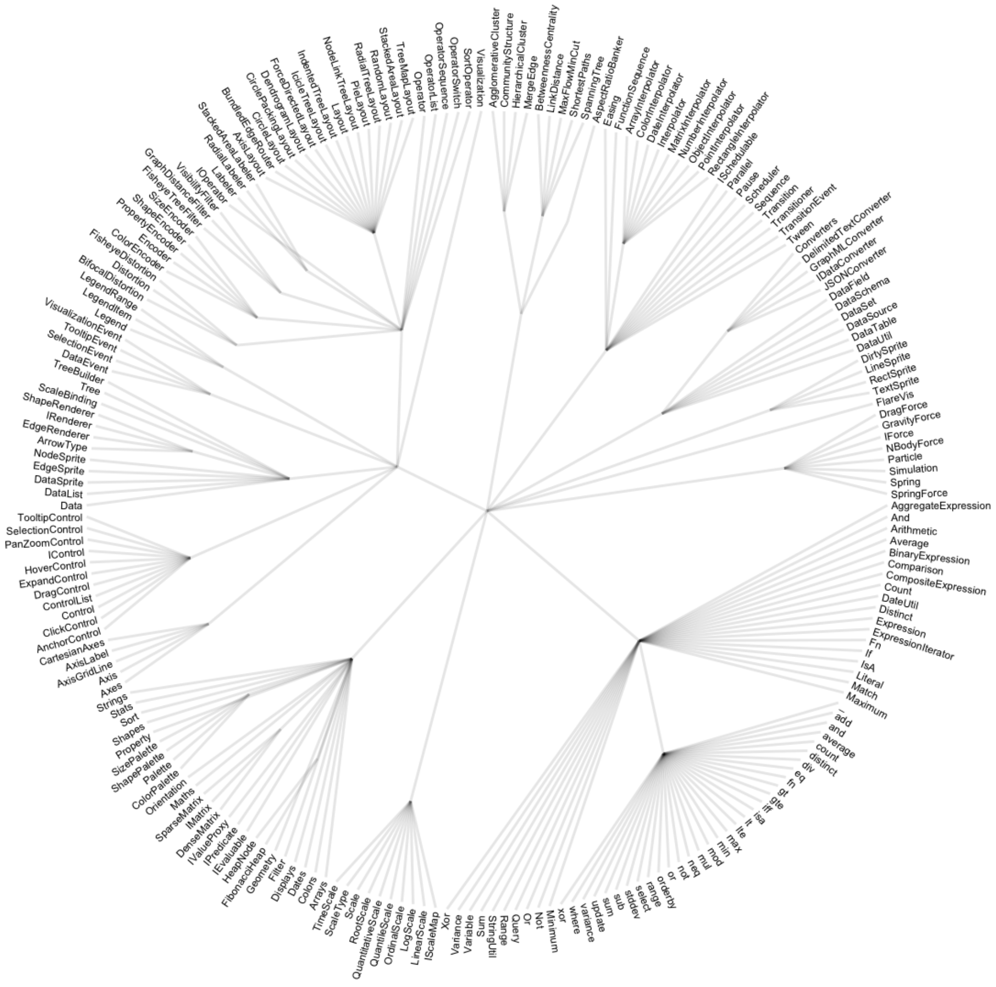
- yes
- via edge bundling



Hierarchical Chord Diagram

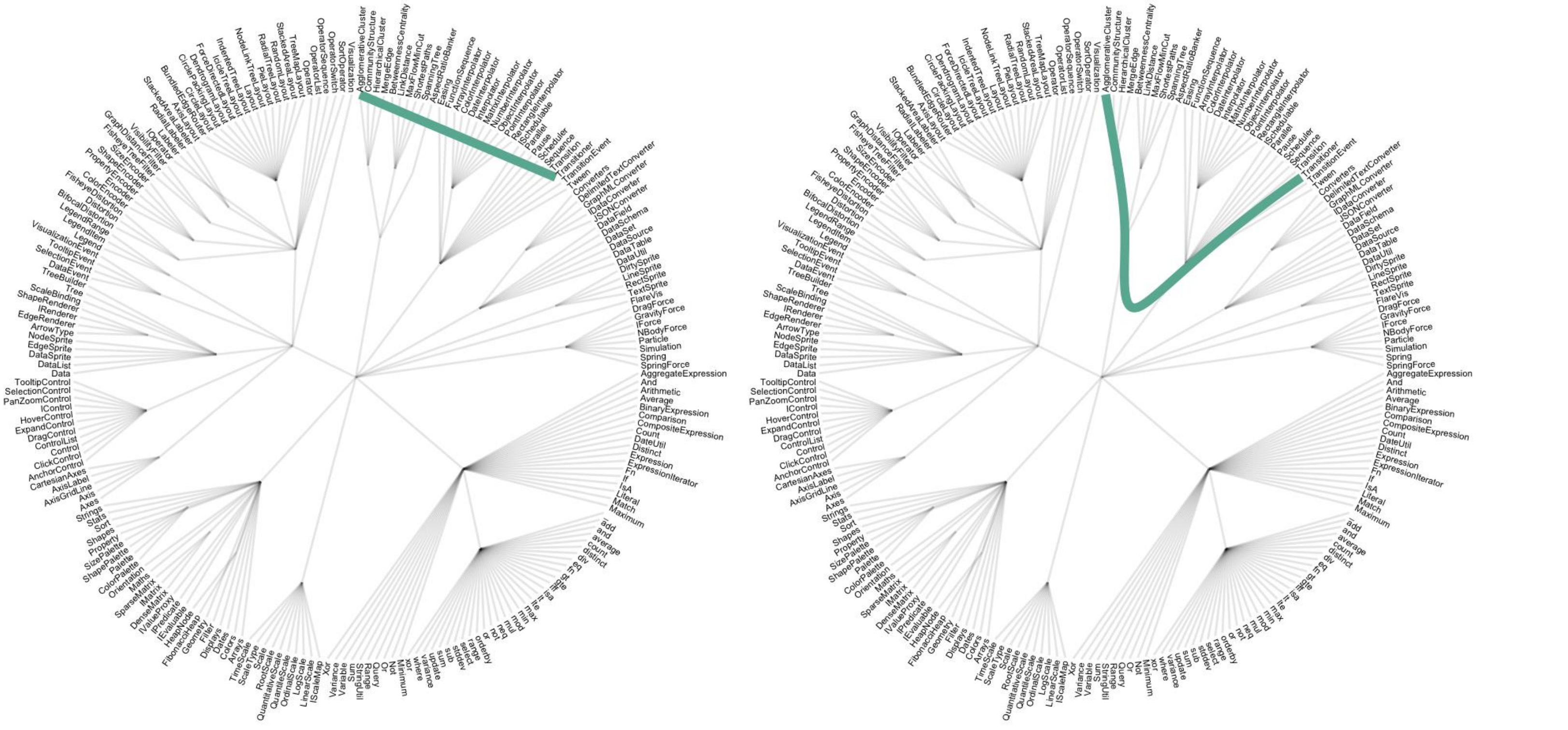
Hierarchy of the Flare ActionScript visualization library

- 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



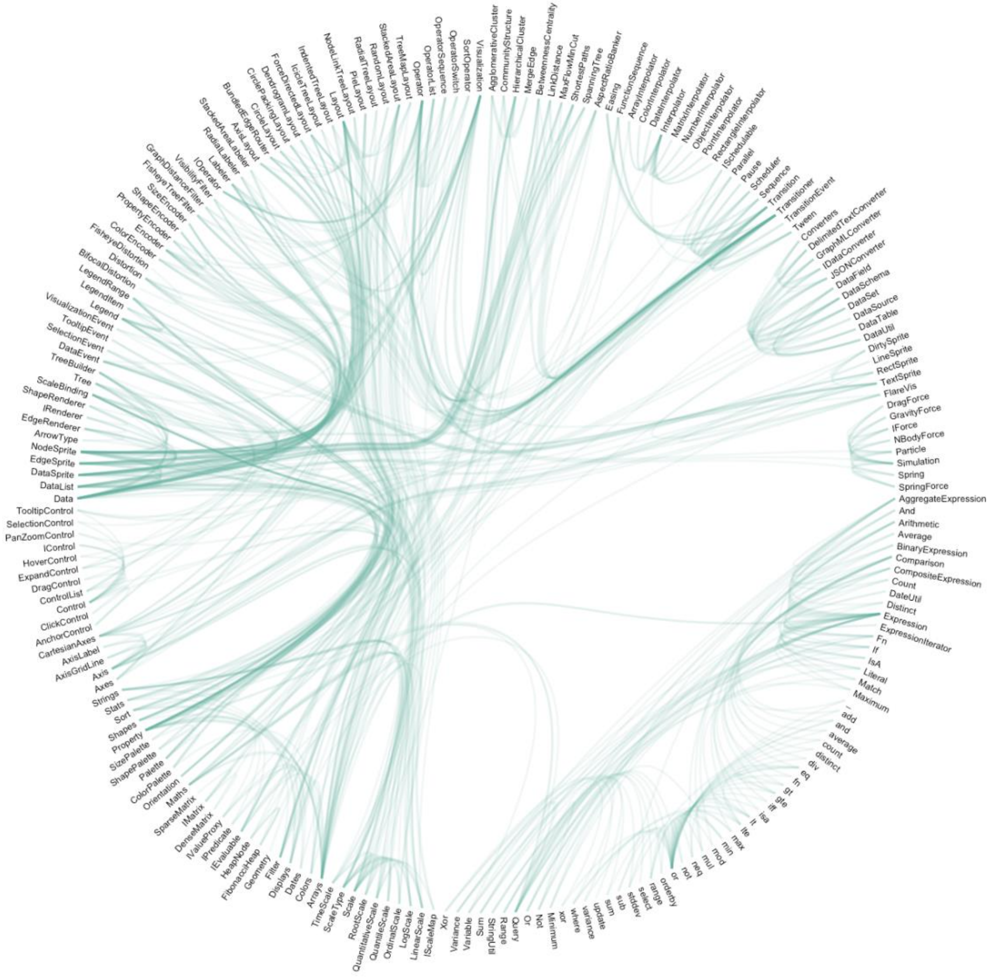
bad: straight line

better: follow a hierarchical edge bundling line

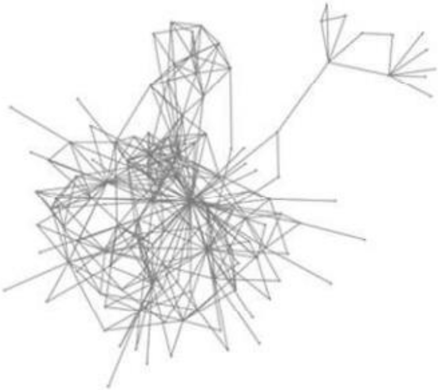
Edge Bundling

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

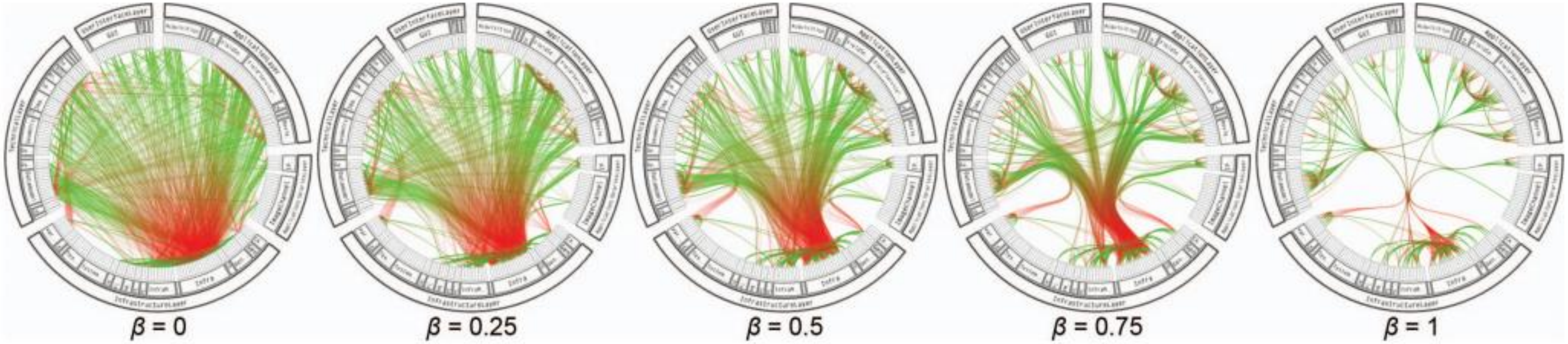


Original Graph

Jia et al., Computer Graphics Forum 2011

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

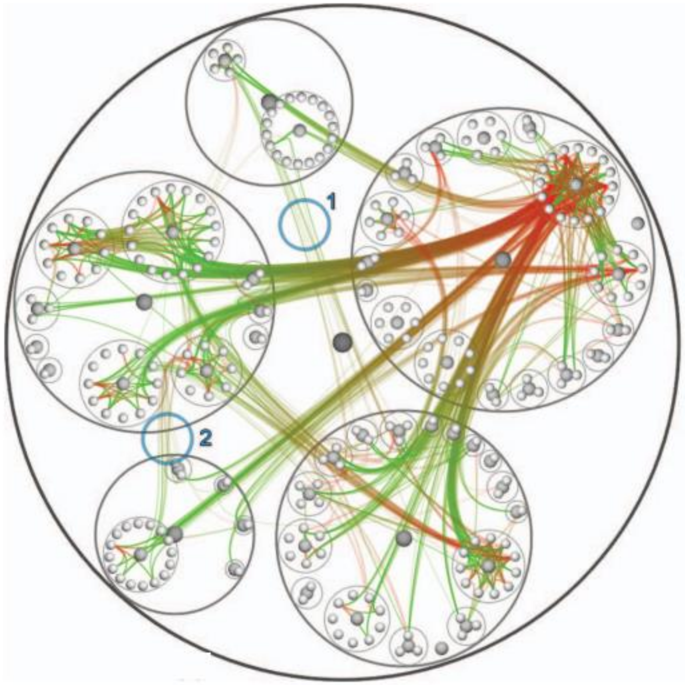
- low values mainly provide low-level, node-to-node connectivity information
- high values provide high-level information

Holten, IEEE TVCG 2006

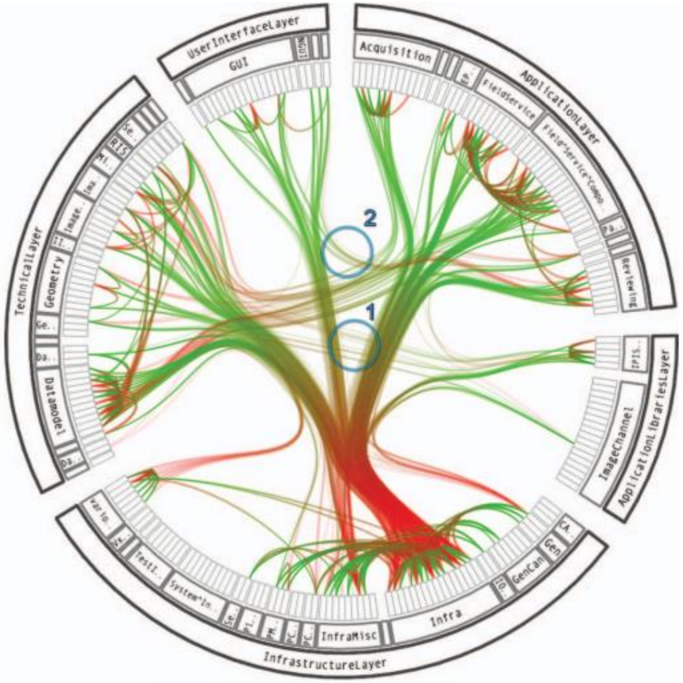
Edge Bundling Example

Software system call graph

- green is caller, red is callee

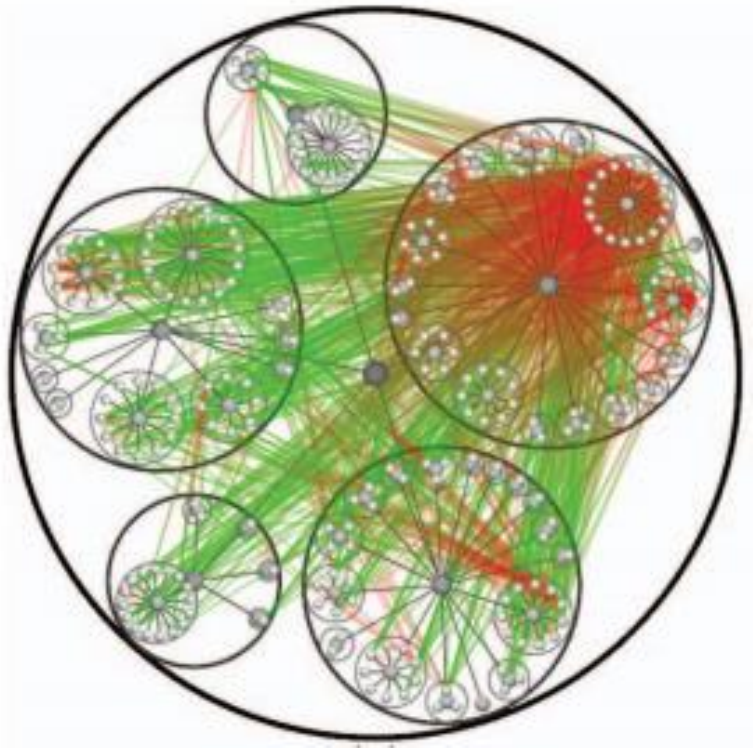


balloon layout (isolated processes)

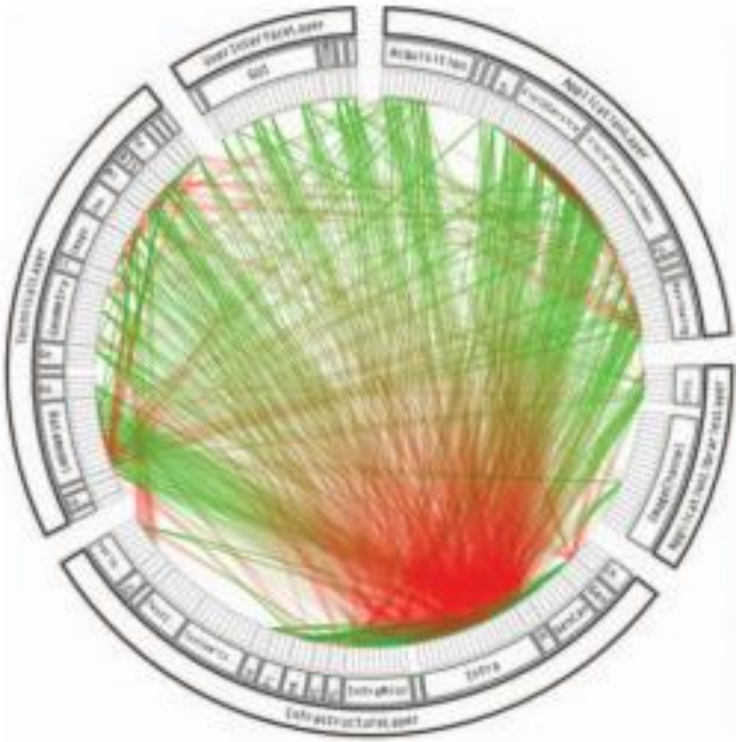


radial layout (more integrated)

Without Edge Bundling



balloon layout

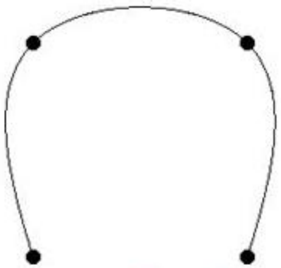


radial layout

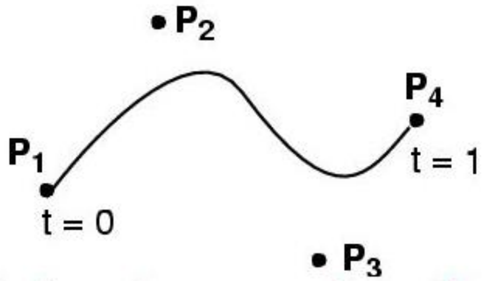
Curved Edges Modeled as Splines

Curved edges are represented as *splines*

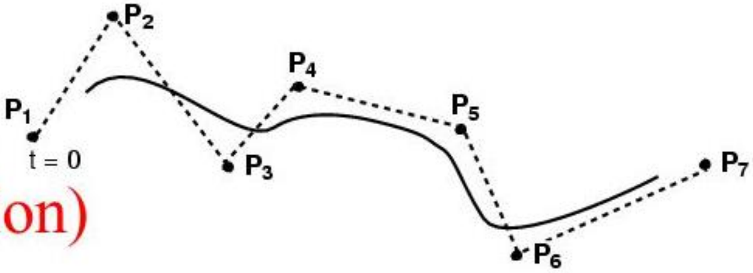
- a spline is a smooth curve defined by some control points
- moving the control points changes the curve



Interpolation



Bézier (approximation)



BSpline (approximation)

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

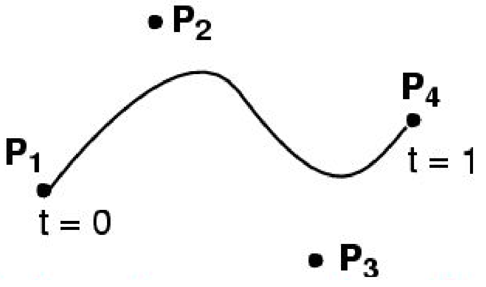
$$x(t) = \frac{1}{6}\begin{bmatrix} P_0 & P_1 & P_2 & P_3 \end{bmatrix} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 0 & 4 \\ -3 & 3 & 3 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} t^3 \\ t^2 \\ t \\ 1 \end{bmatrix}$$

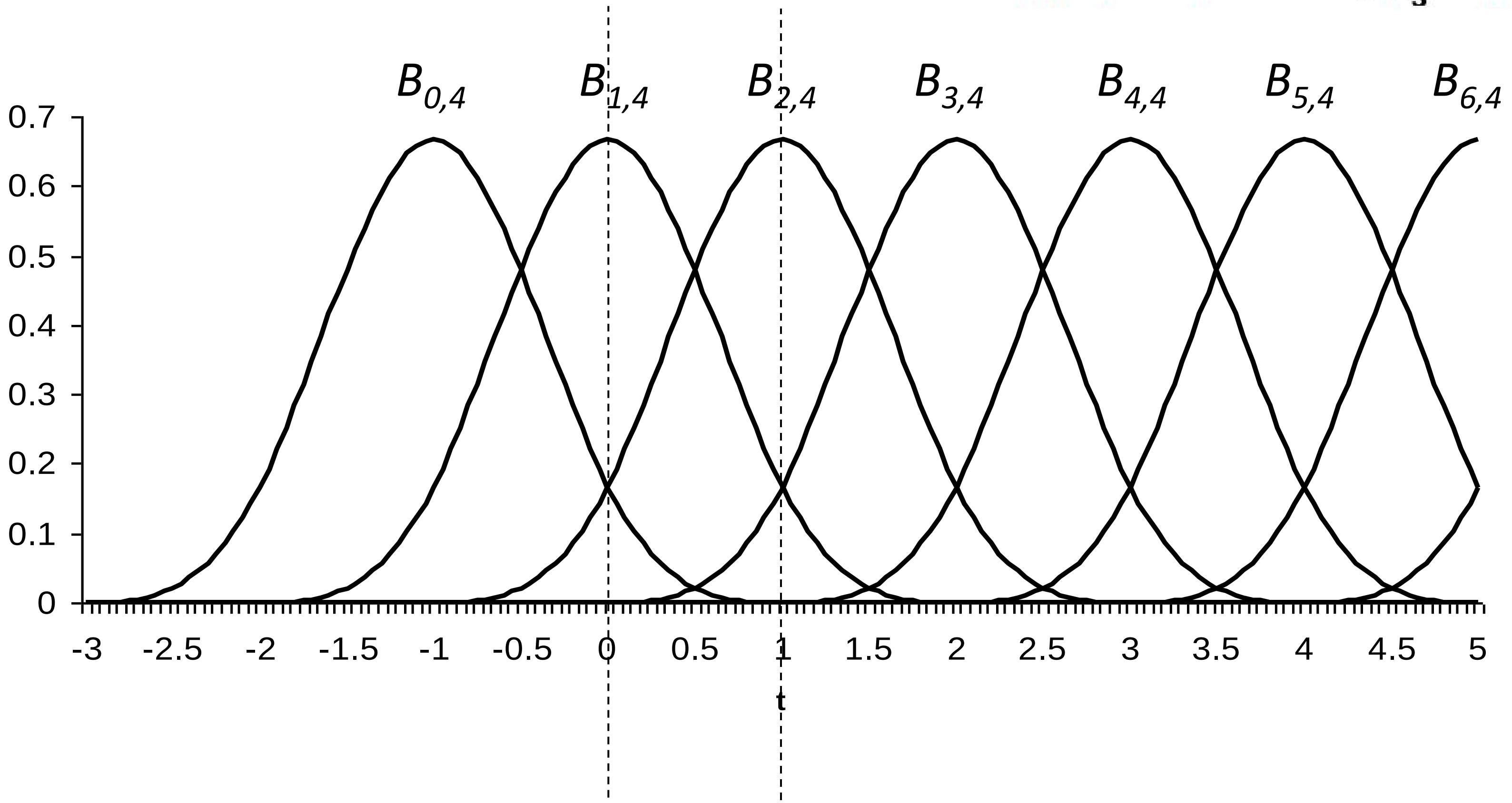
Or in matrix form:

- 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



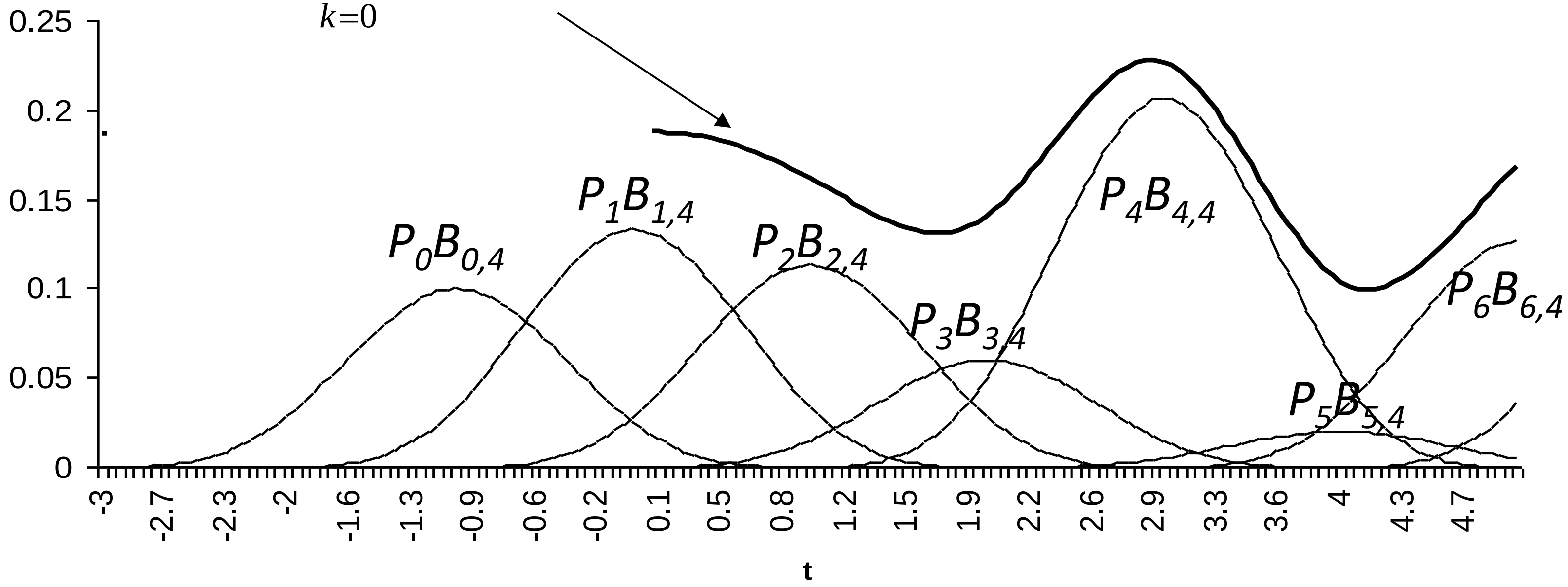


Primer: Uniform Cubic B-Spline

The locations of the control points scale the basis functions

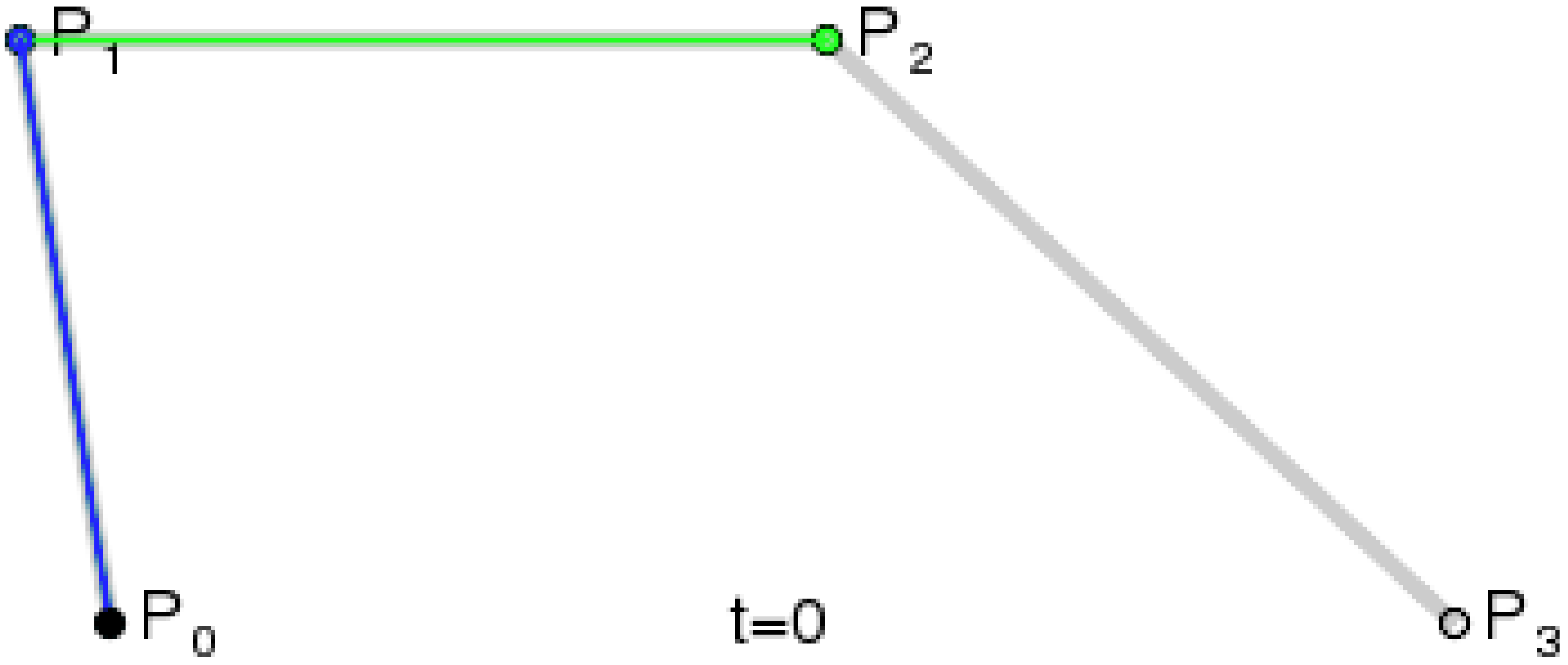
- in this simple example we see a continuous 1D function generated from 6 control points and basis functions

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



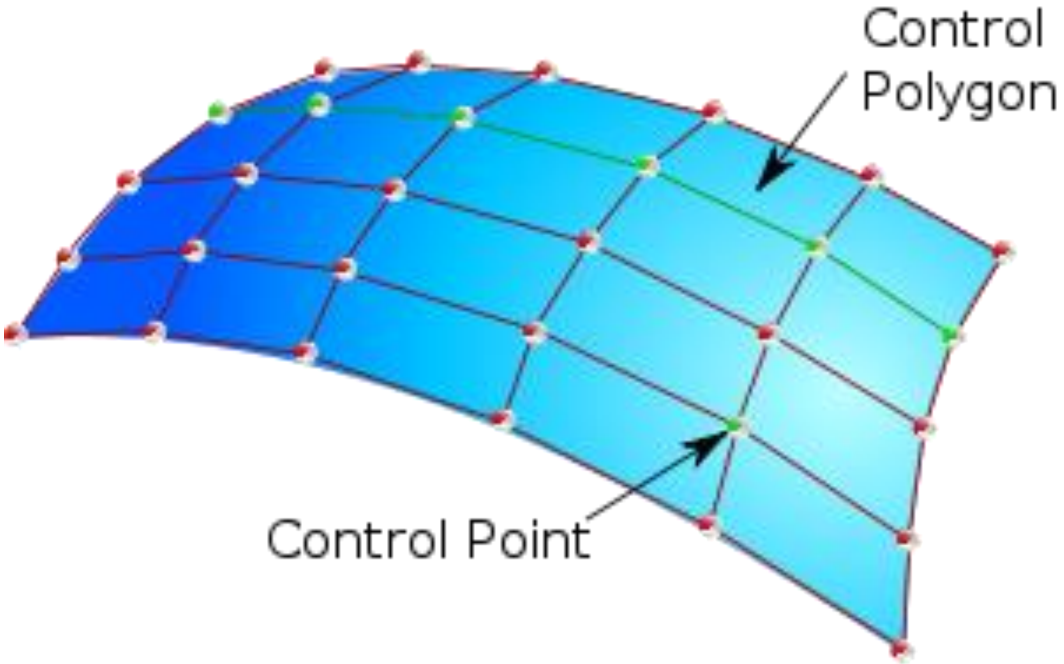
The curve can't start until there are 4 basis functions active

Cubic B-Spline Animated



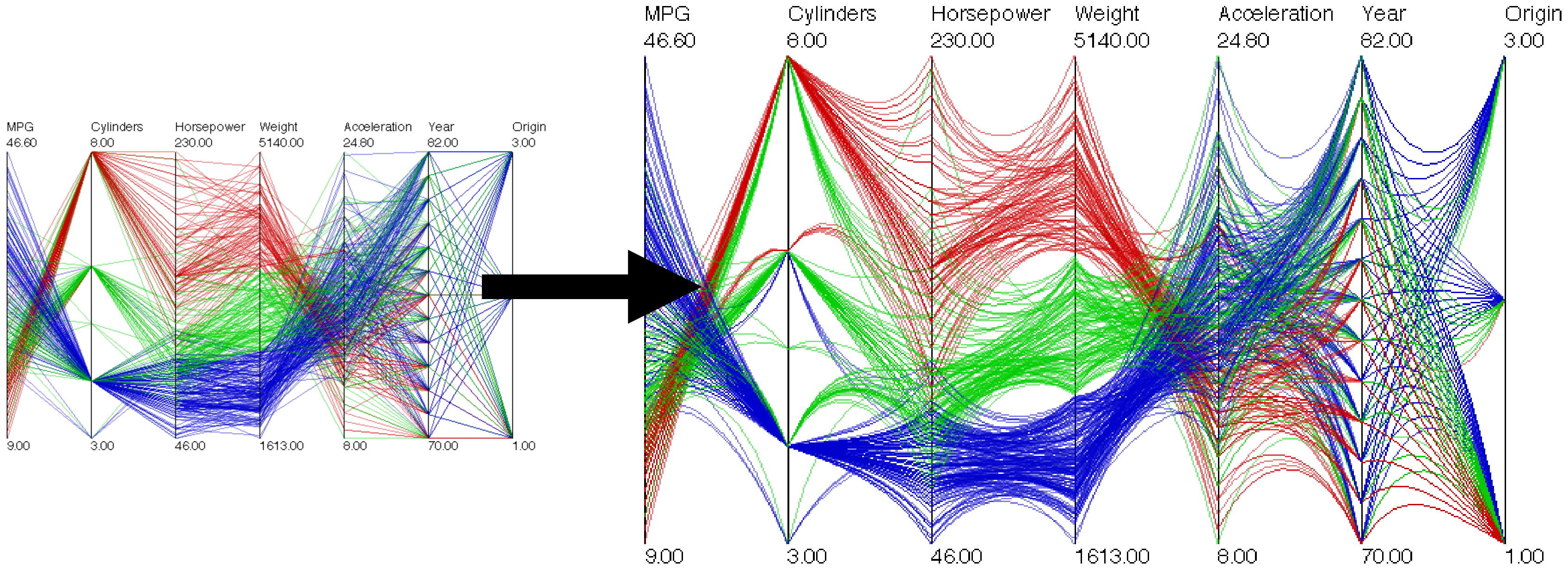
Extension to Surfaces

B-spline surface



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.)

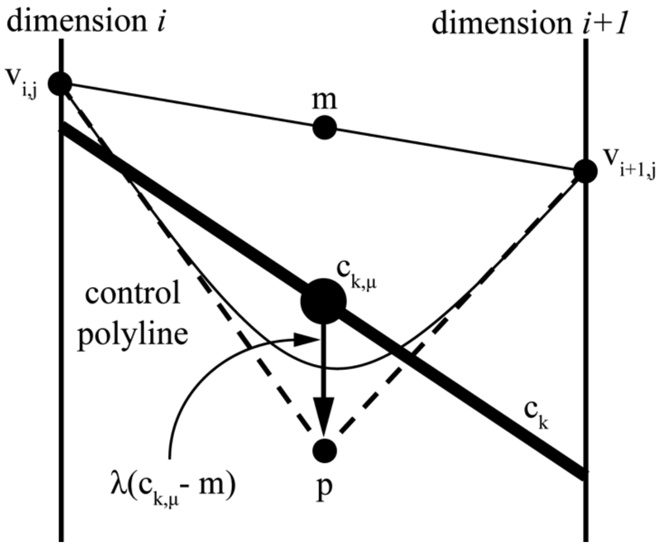
Let m be the mid-point in viewport coordinates of $v_{i,j}$ and $v_{i+1,j}$, end-points of a line segment

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:

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

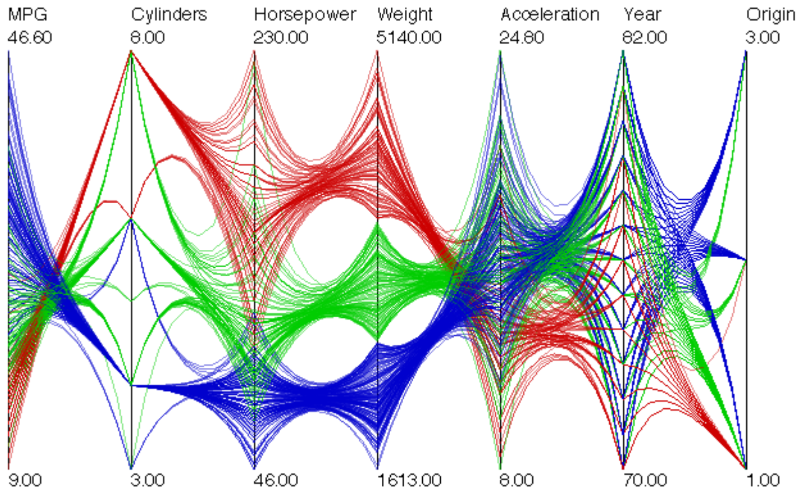


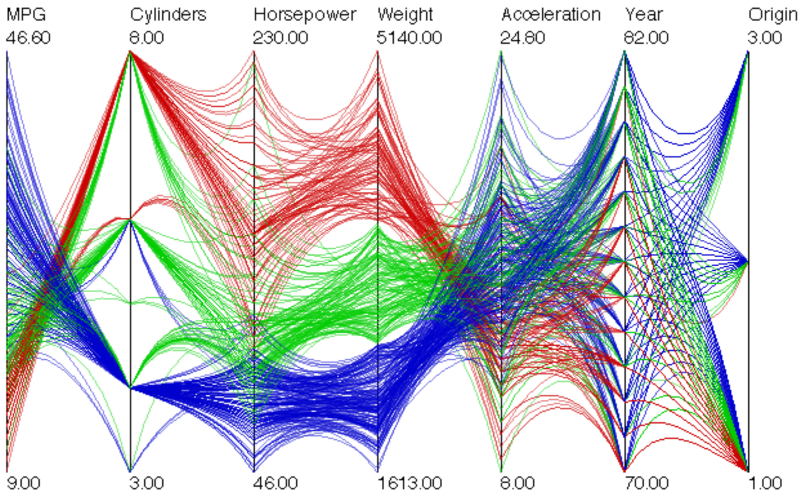
Edge Bundling (cont.)

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:





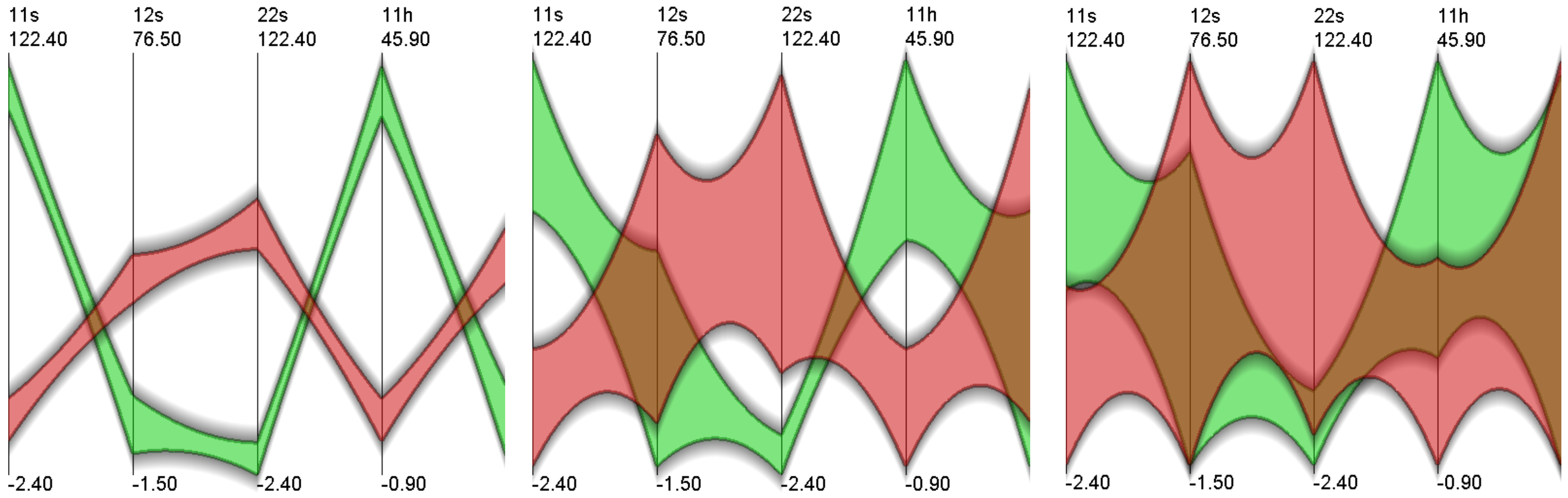
Cluster Rendering

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

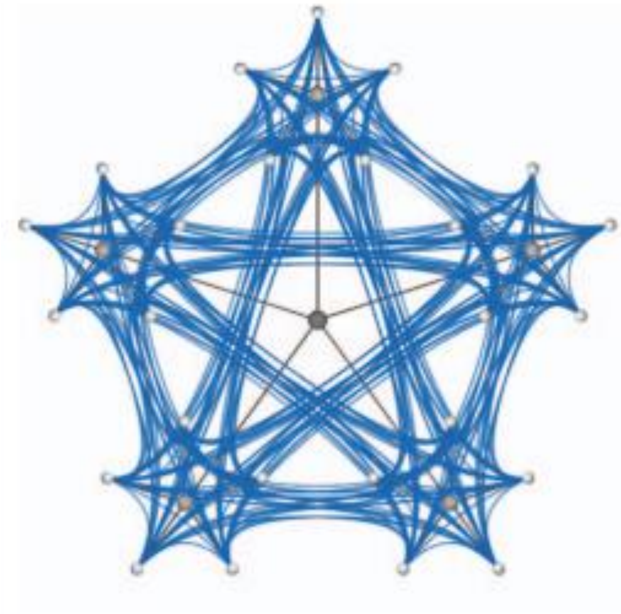
The vertical “spread” can be controlled by the user



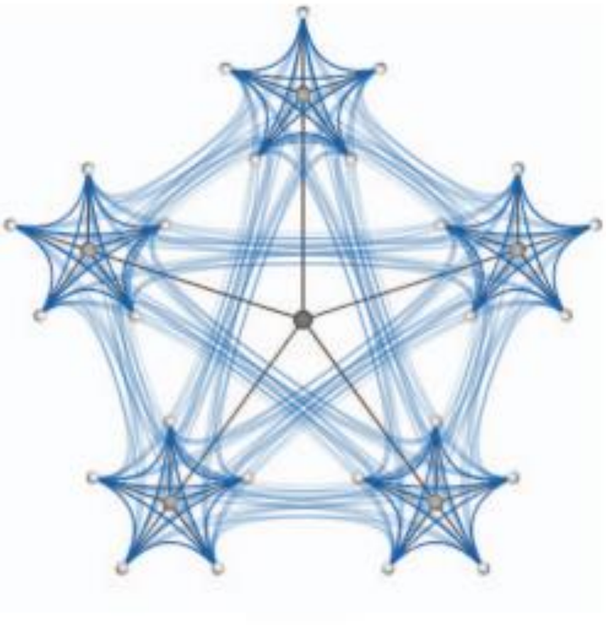
Alpha (Opacity) Blending

Draw curves at different opacities

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



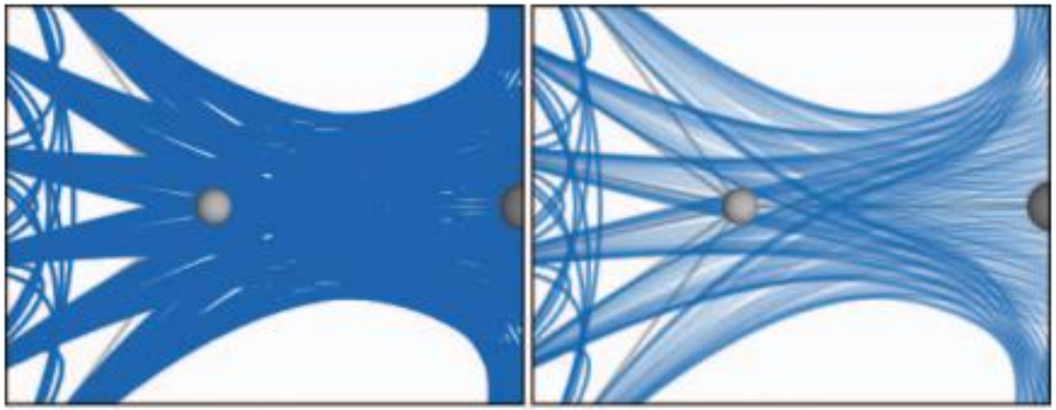
alpha blending disabled



alpha blending enabled

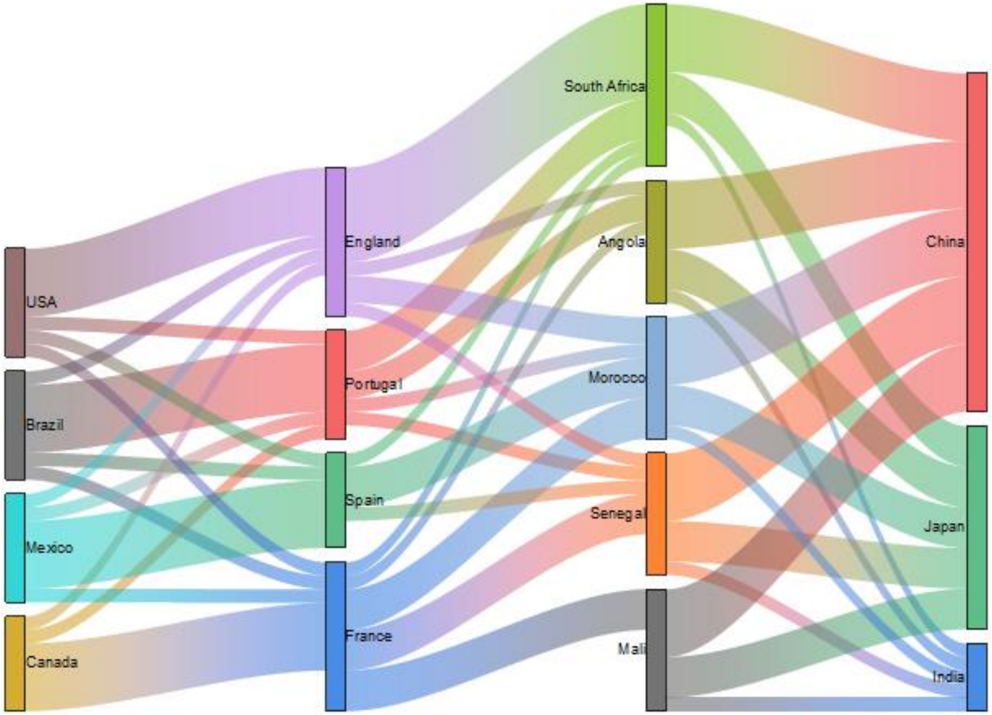
Alpha (Opacity) Blending

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



alpha blending disabled　　alpha blending enabled

Sankey Diagram



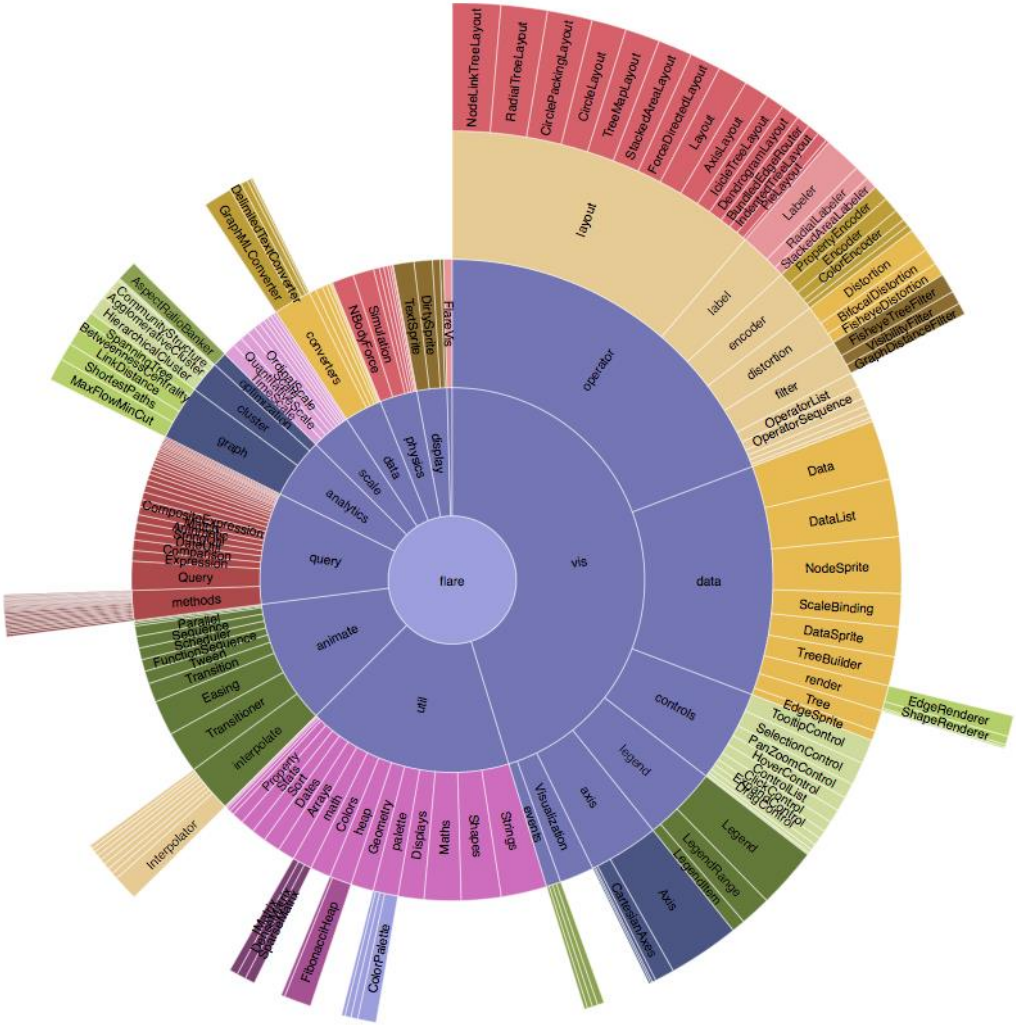
Another bundling technique

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

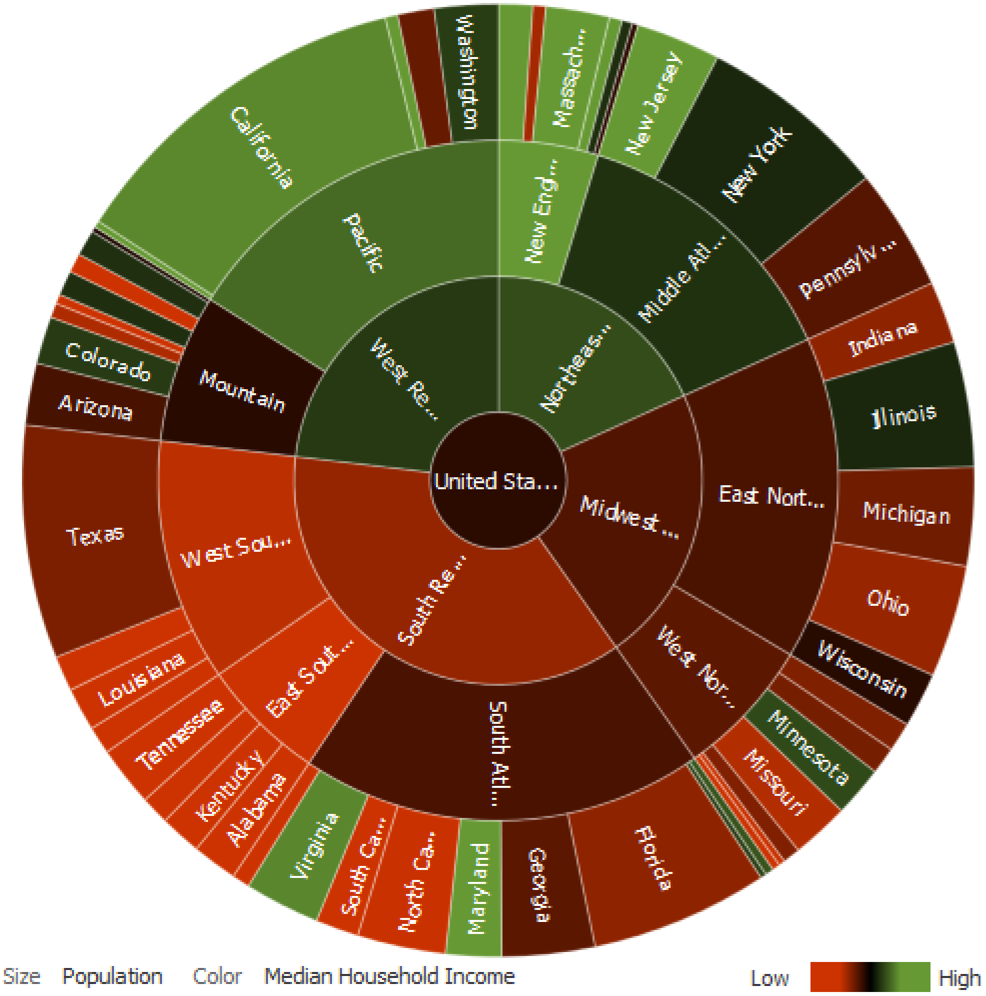
Use cases:

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

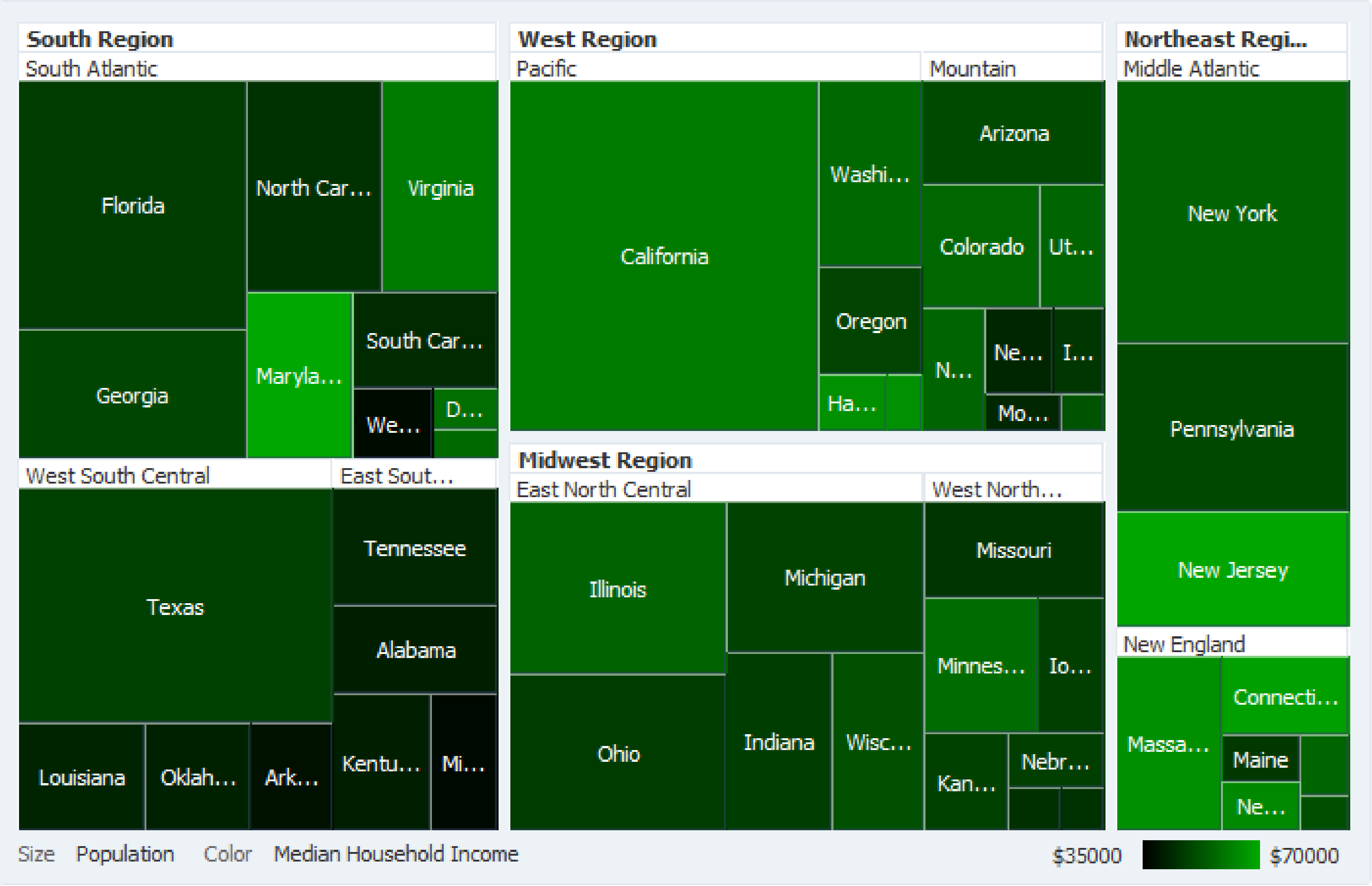
Hierarchies with Sun Burst Displays



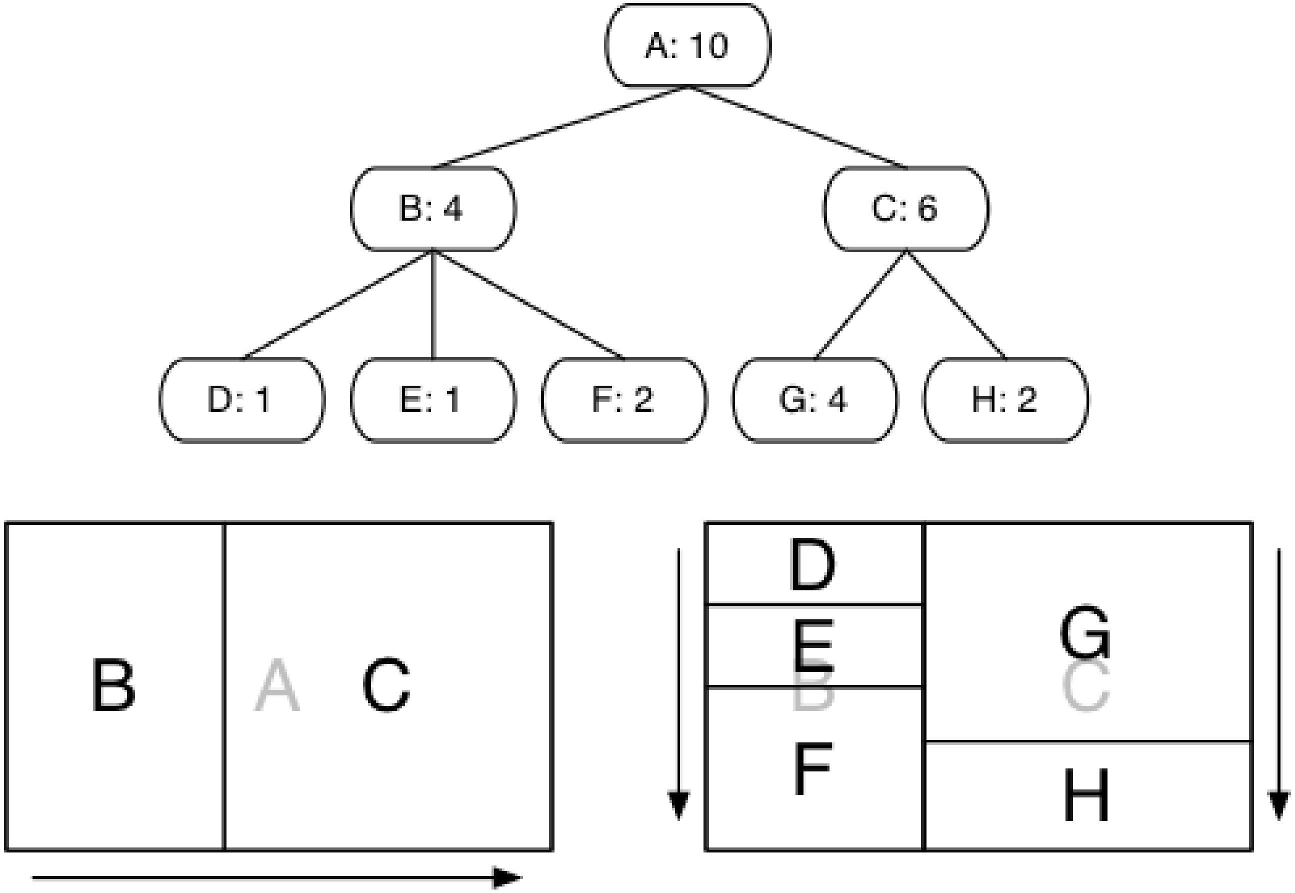
Sunburst with Partition of Unity



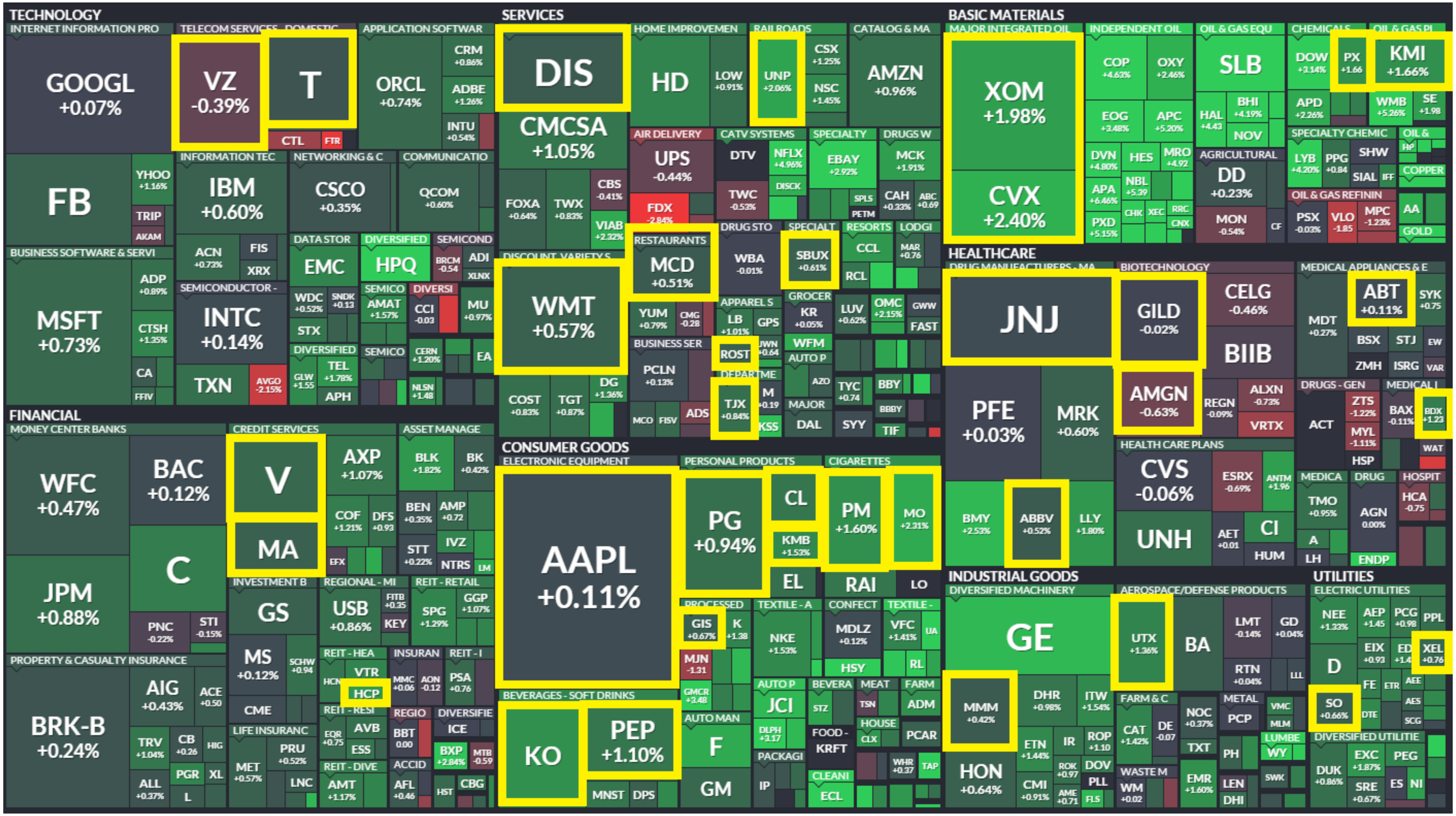
Same Data with Treemap



Treemap Construction

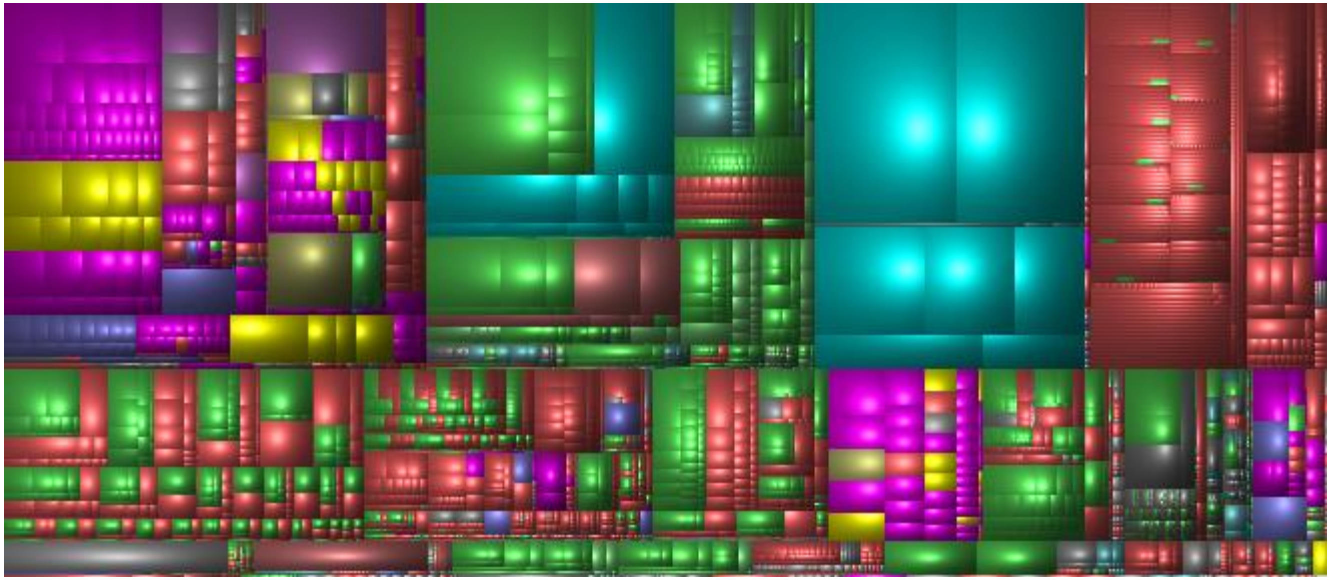


Treemap for Stock Portfolio



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

Cushion Treemap



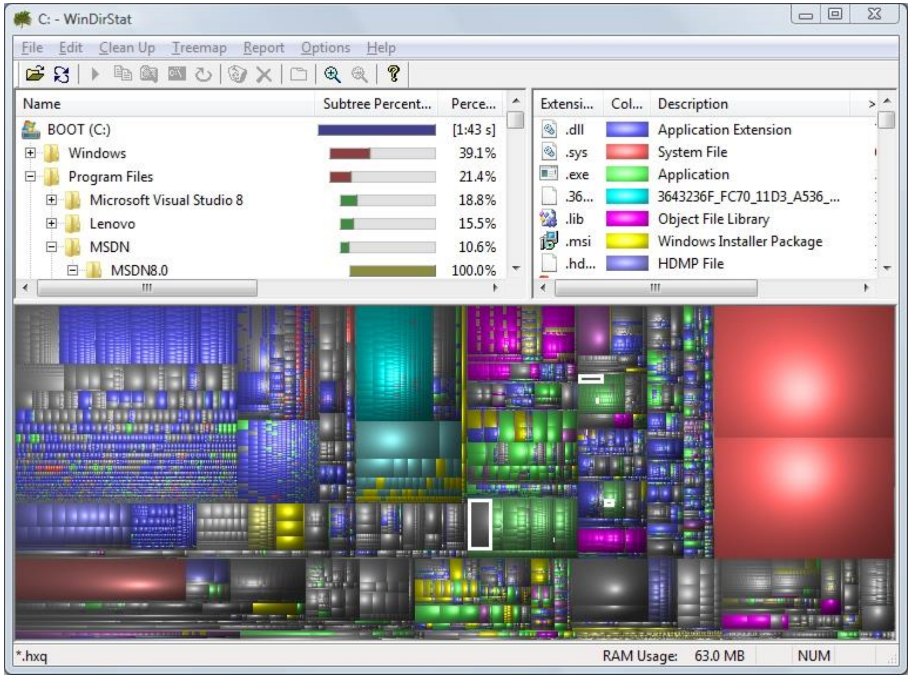
Advantages

- 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

Treemap for Disk Drives

Used in programs like

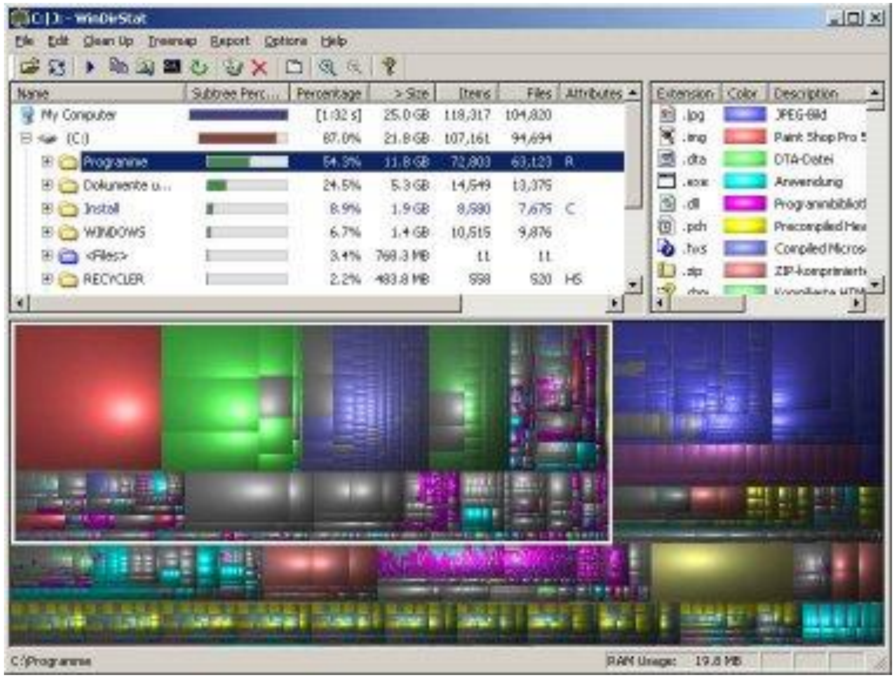
- WinDirStat (Windows)
- KDirStat (Linux)
- DiskInventory (Mac)



Name	Subtree Percent...	Perce...
BOOT (C:)		[1:43 s]
Windows		39.1%
Program Files		21.4%
Microsoft Visual Studio 8		18.8%
Lenovo		15.5%
MSDN		10.6%
MSDN8.0		100.0%

Extensi...	Col...	Description
.dll	🟦	Application Extension
.sys	🟥	System File
.exe	🟩	Application
.36...	🟦	3643236F_FC70_11D3_A536_...
.lib	🟪	Object File Library
.msi	🟨	Windows Installer Package
.hd...	🟦	HDMP File

*.hxq RAM Usage: 63.0 MB NUM

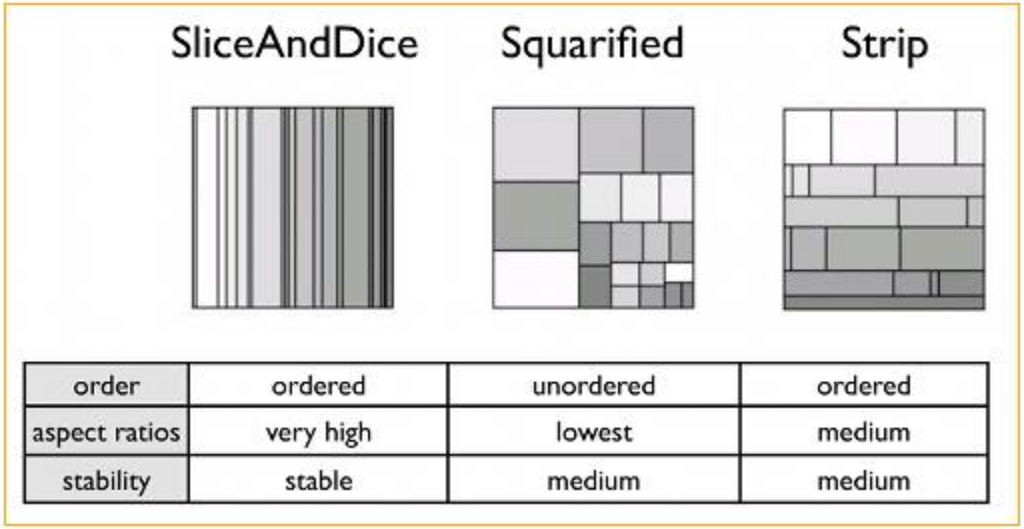


Name	Subtree Perc...	Percentage	> Size	Items	Files	Attributes
My Computer		[1:32 s]	25.0 GB	118,317	104,820	
(C:)		87.0%	21.8 GB	107,161	94,694	
Programme		54.3%	11.8 GB	72,803	63,123	R
Dokumente u...		24.5%	5.3 GB	14,549	13,376	
Install		8.9%	1.9 GB	8,580	7,675	C
WINDOWS		6.7%	1.4 GB	10,515	9,876	
<Files>		3.4%	768.3 MB	11	11	
RECYCLER		2.2%	483.8 MB	598	520	HS

Extension	Color	Description
.jpg	🟦	JPEG-Bild
.img	🟥	Paint Shop Pro 5
.dta	🟩	DTA-Datei
.exe	🟦	Anwendung
.dll	🟪	Programmbibliot
.pch	🟨	Precompiled Hea
.hxs	🟦	Compiled Micros
.zip	🟥	ZIP-komprimierte
.chm	🟩	Kompilierte HTM

C:\Programme RAM Usage: 19.8 MB

Treemap Variations



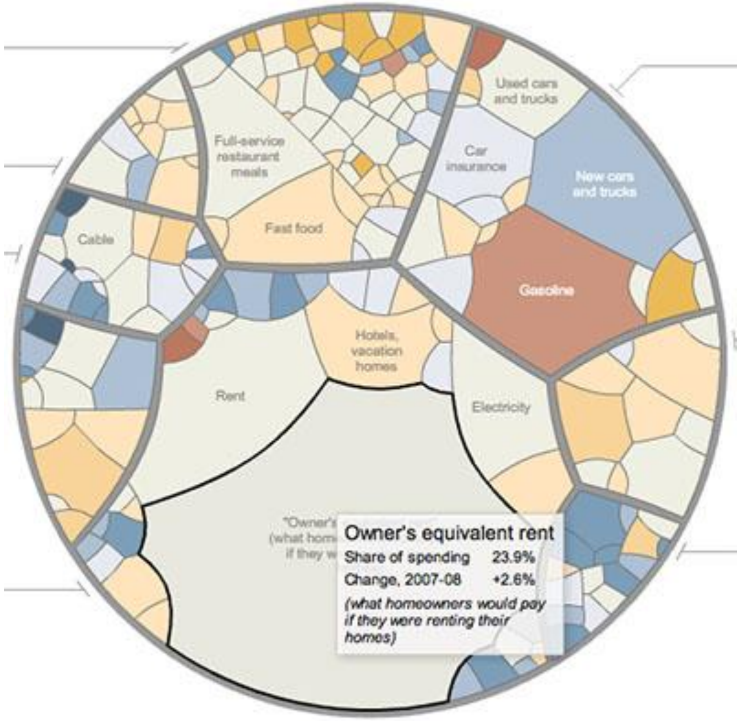
order	ordered	unordered	ordered
aspect ratios	very high	lowest	medium
stability	stable	medium	medium

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

Voronoi treemap

- based on Voronoi tesselation



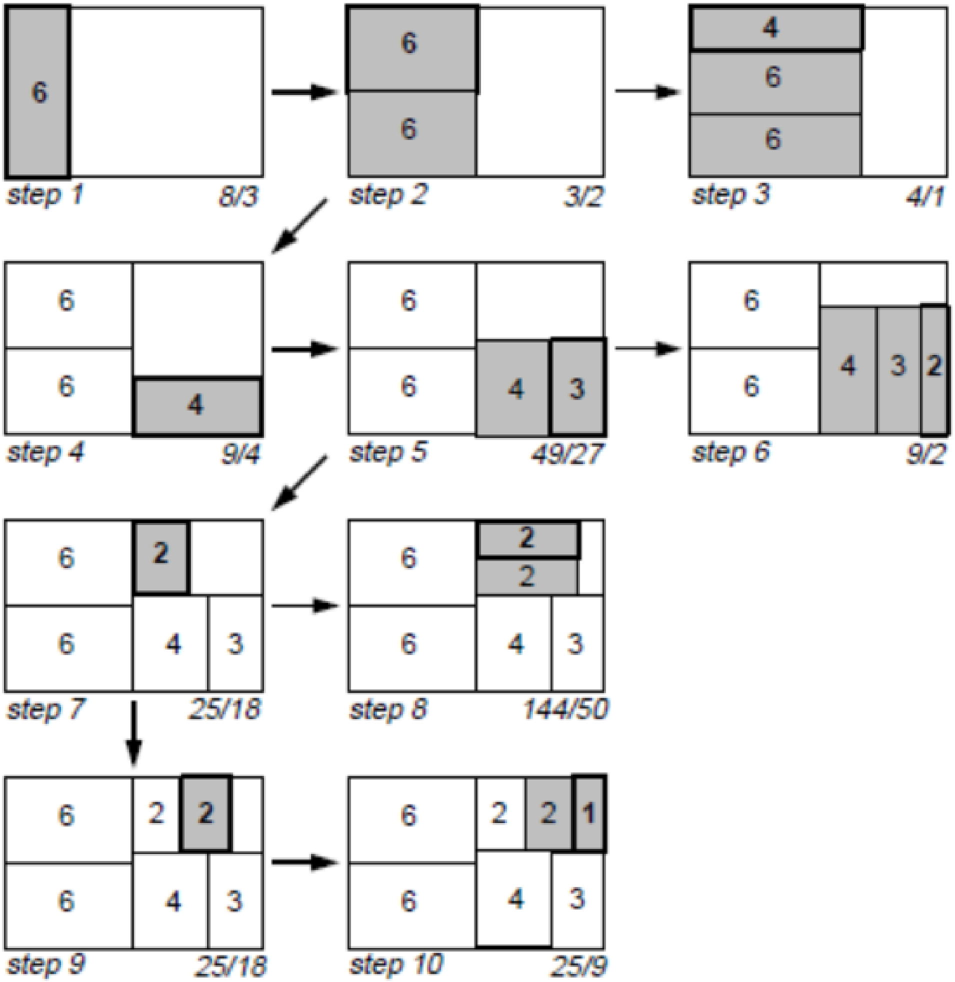
Constructing a Squarified Tree map

Optimization criterion

- keep aspect ratio of boxes close to 1

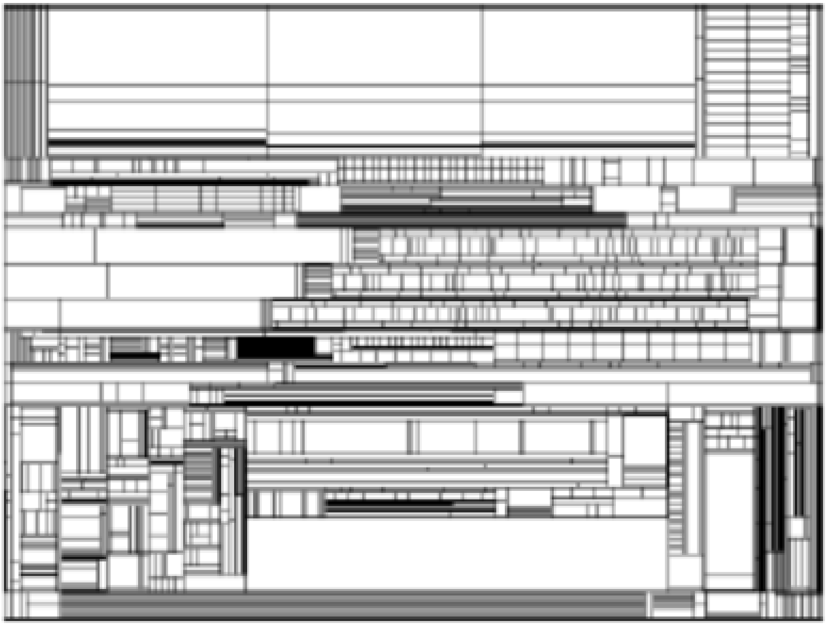
Sequence:

- steps 1, 2, 4, 5, 7, 9, 10
- steps 3, 6, 8 would increase the aspect ratios of the boxes
- start a new row

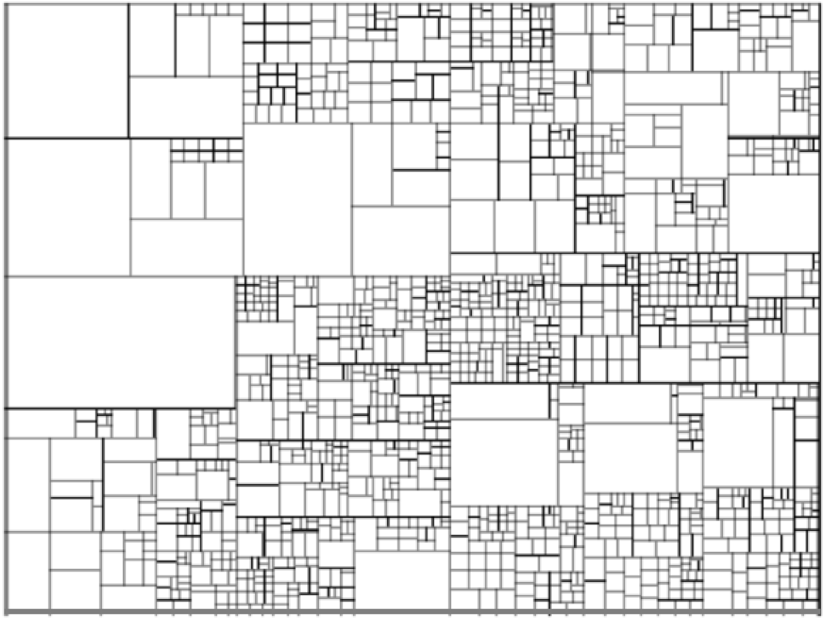


Bruls, Huizing, Van Wijk, (2000). Squarified Treemaps. *Data Visualization.* Springer

Squarified Treemap Example



standard layout



squarified

Can greatly improve

- ability to compare the magnitude of different leaf nodes
- at the same time maintain some level of the original hierarchy