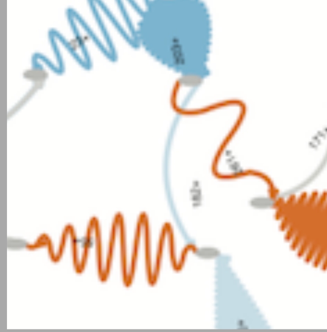
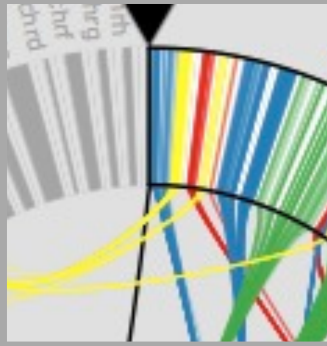
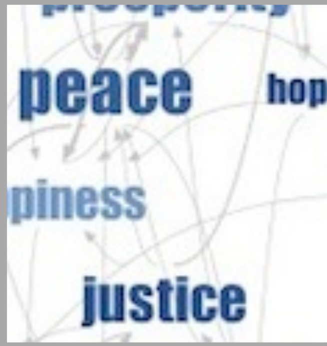
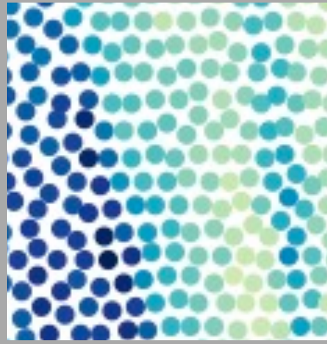
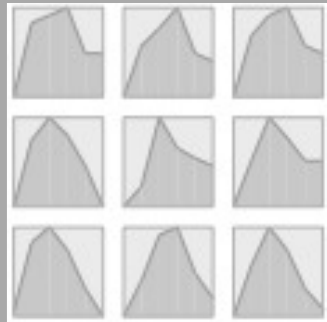
cs6964 | March 1 2012

TREES & GRAPHS

Miriah Meyer

University of Utah

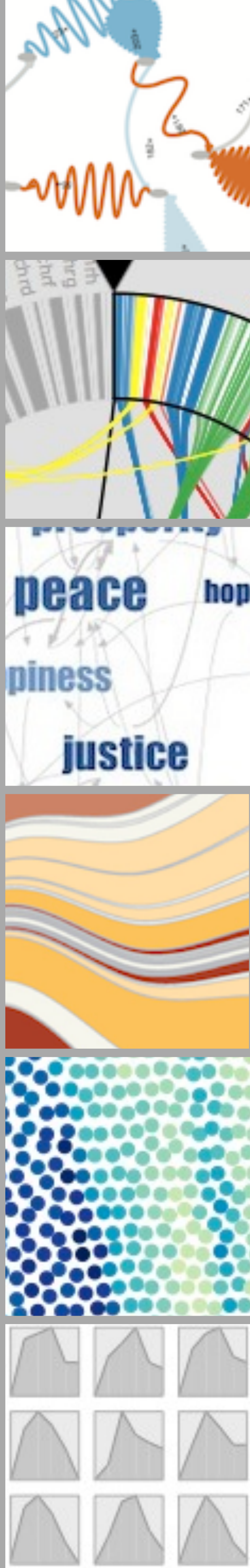




cs6964 | March 1 2012

TREES & GRAPHS

Miriah Meyer
University of Utah

slide acknowledgements:
Hanspeter Pfister, Harvard University
Jeff Heer, Stanford University
Tamara Munzner, UBC



administrivia

feb 14-23 : proposal meetings

march 7 : presentation topics due

march 9 : proposals due

march 27-april 3 : project updates

april 5-24 : paper presentations

may 1 : final project presentations

may 3 : process books due

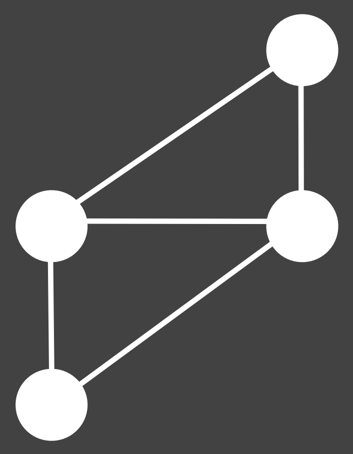
- **definitions**

- **visualizing trees**
  - indented
  - node link
  - enclosure
  - layered

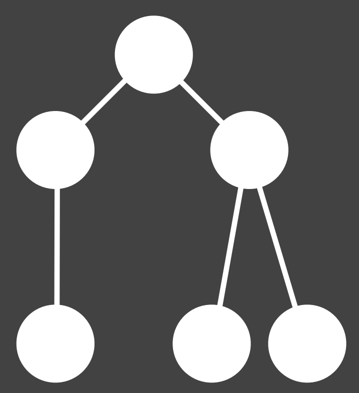
- **visualizing graphs**
  - node link
  - matrix
  - network summarizations

GRAPHS & TREES

- **graphs**
  - model relations amount data
  - *nodes* and *edges*



- **trees**
  - graphs with hierarchical structure
    - ***connected graph with N-1 edges***
  - nodes as *parents* and *children*



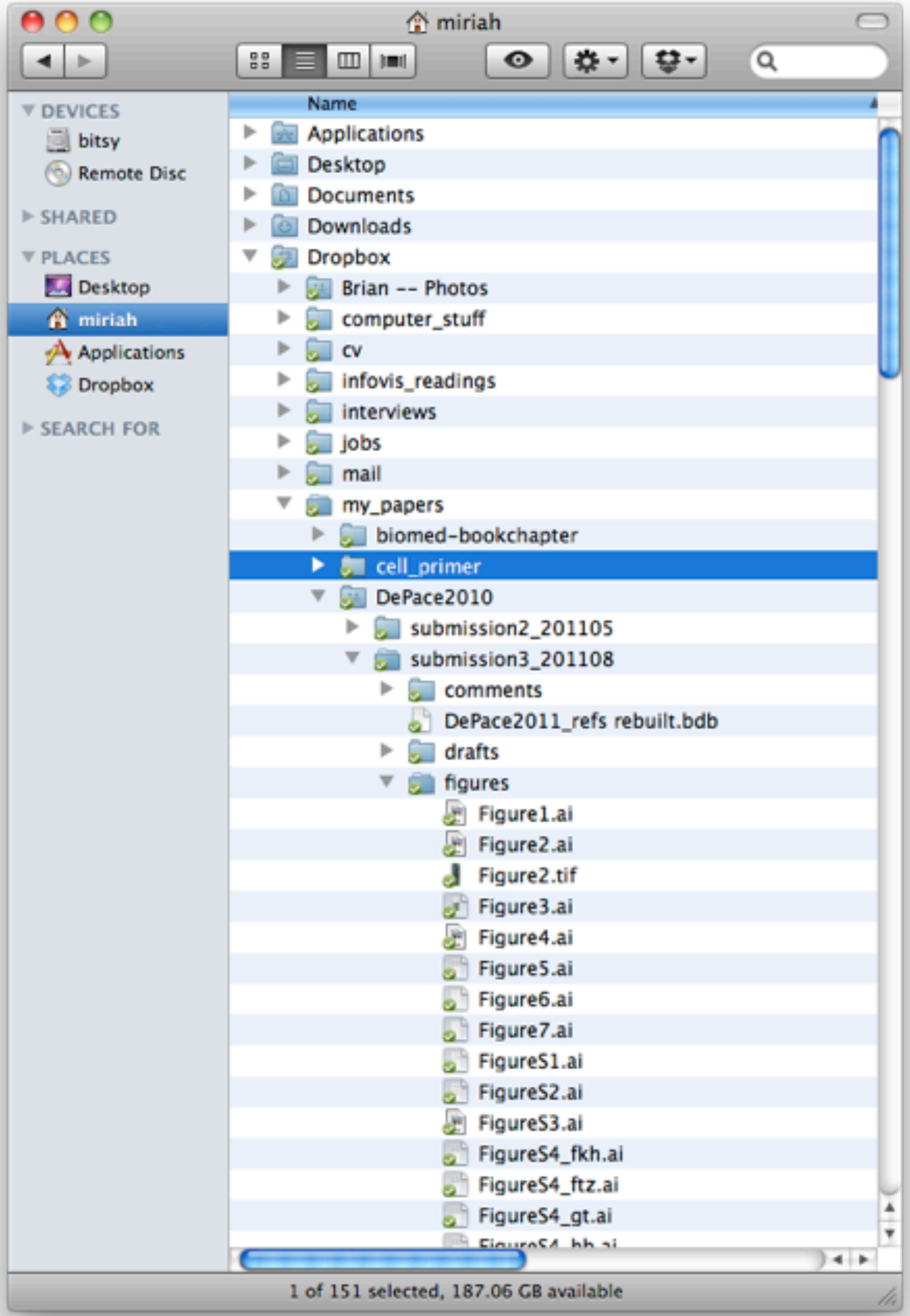
SPATIAL LAYOUT

- **primary concern of graph drawing is the spatial layout of nodes and edges**
- **often (but not always) the goal is to effectively depict the graph structure**
  - connectivity, path-following
  - network distance
  - clustering
  - ordering (e.g., hierarchy level)

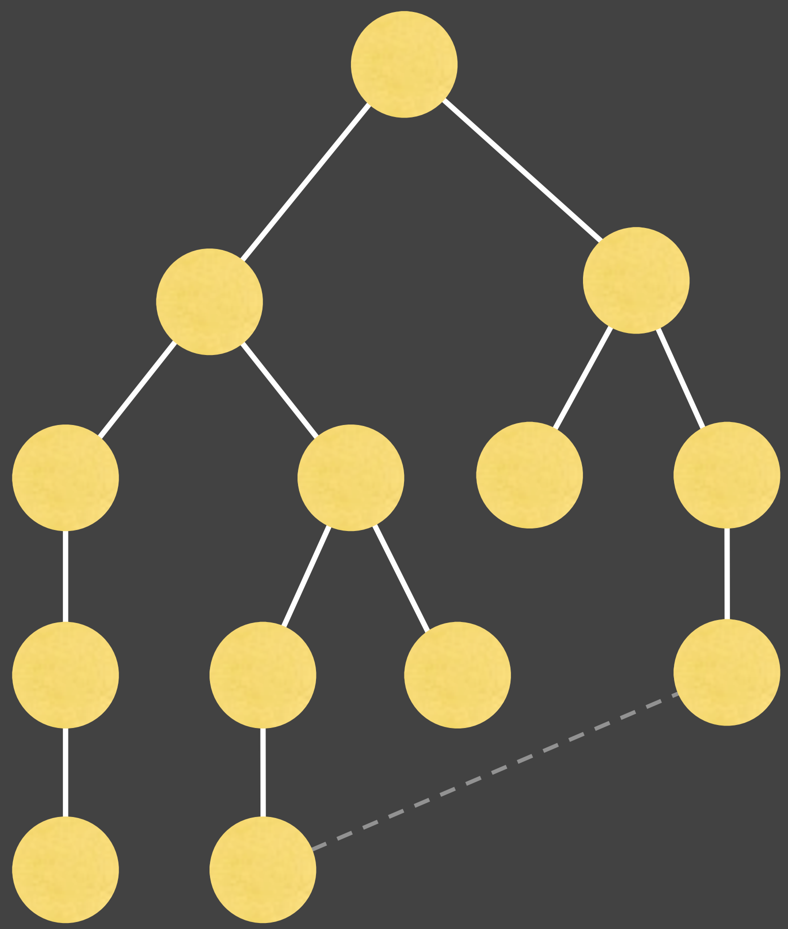
VISUALIZING TREES

- **recursion makes it elegant and fast to draw trees**
- **approaches:**
  - indentation
  - node link
  - enclosure
  - layering

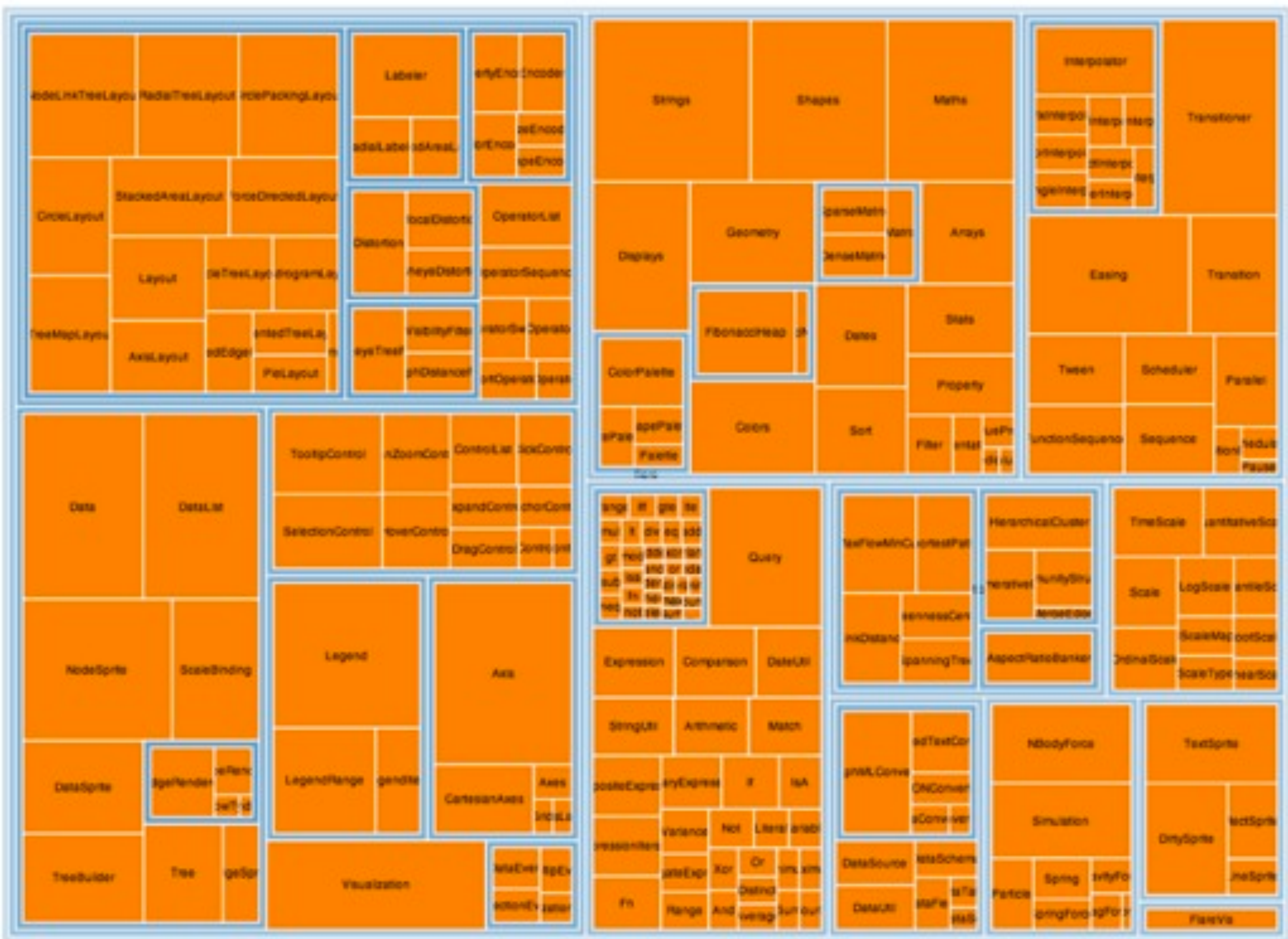
INDENTATION

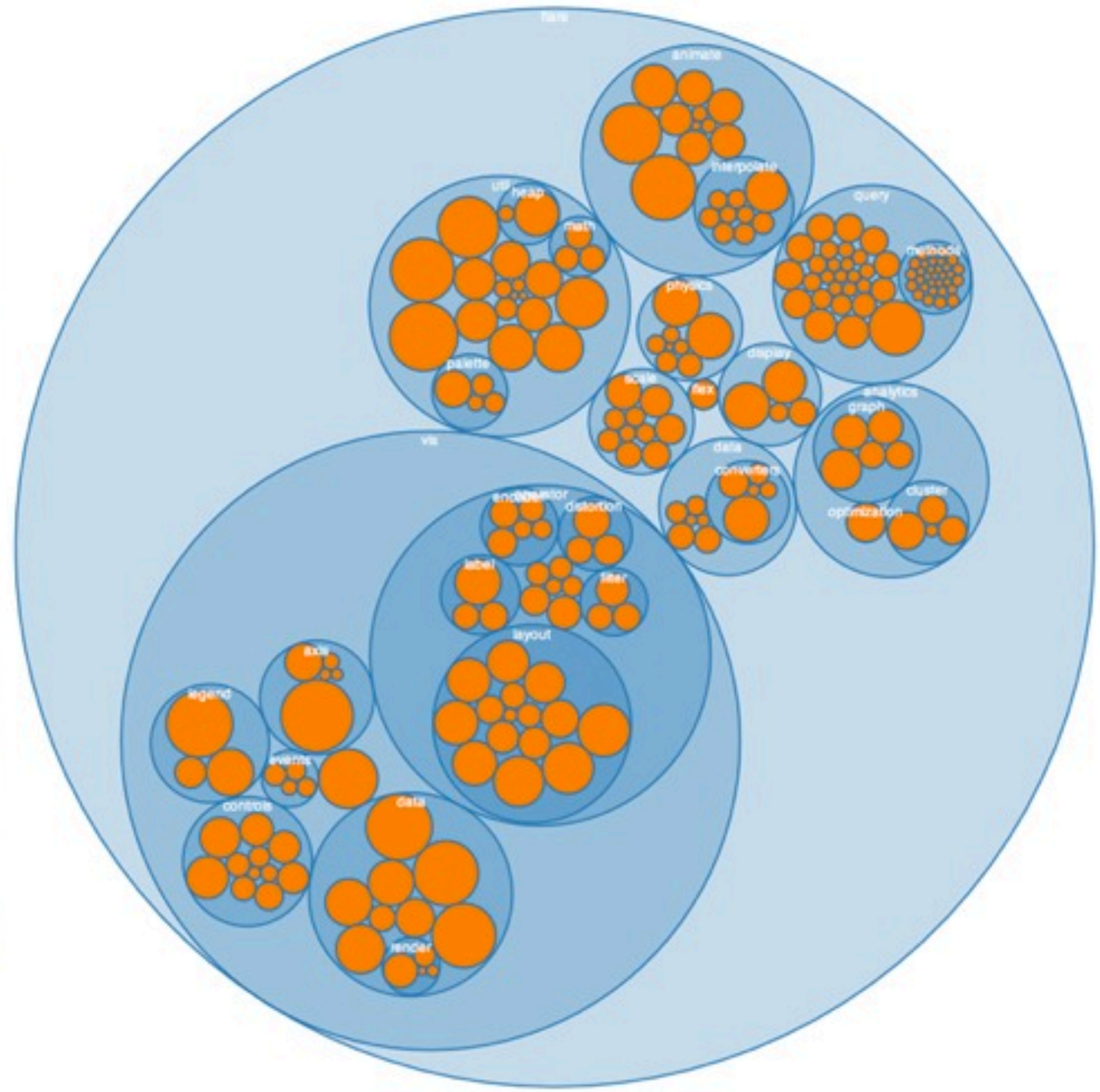


NODE-LINK: Reingold-Tilford



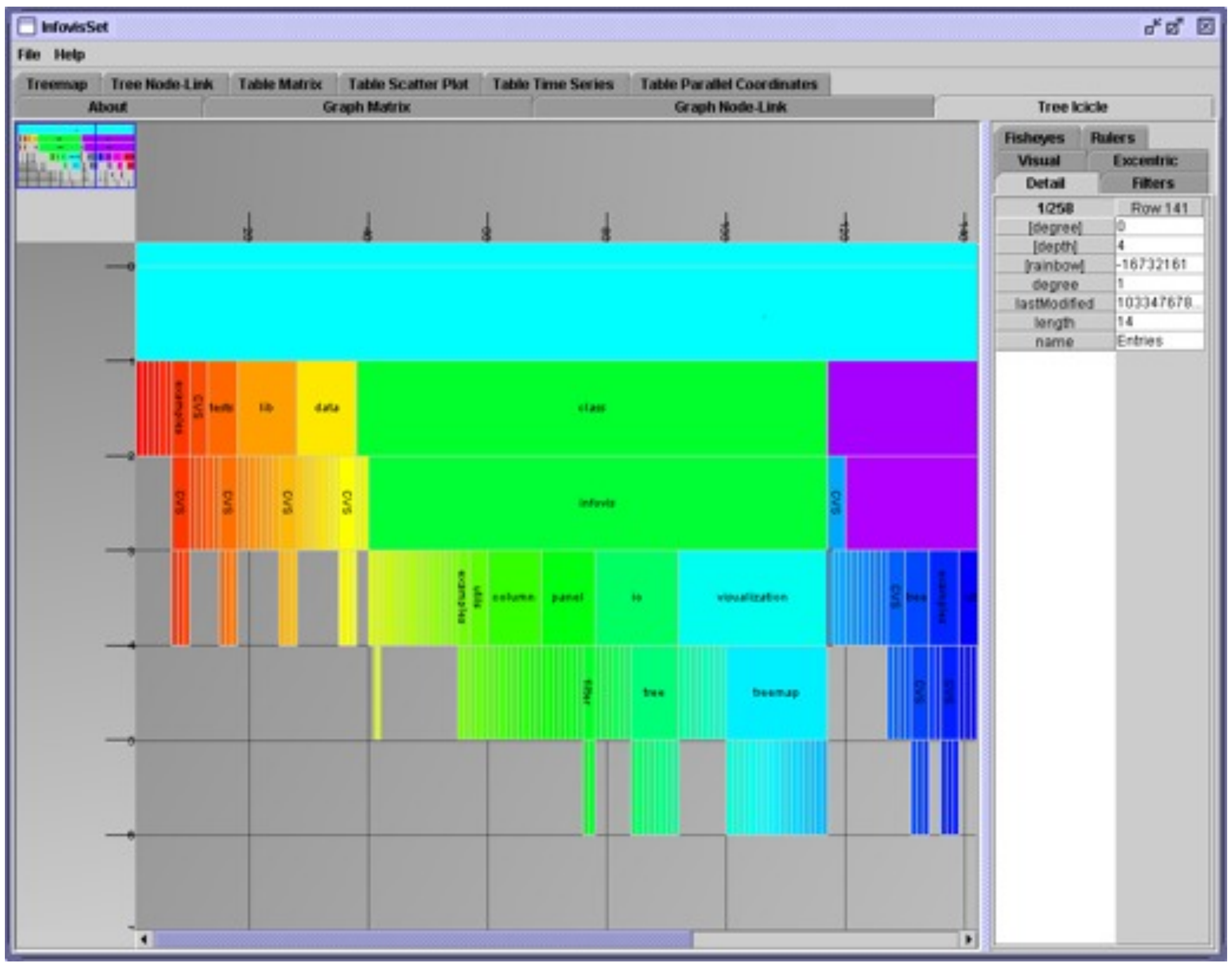
ENCLOSURE

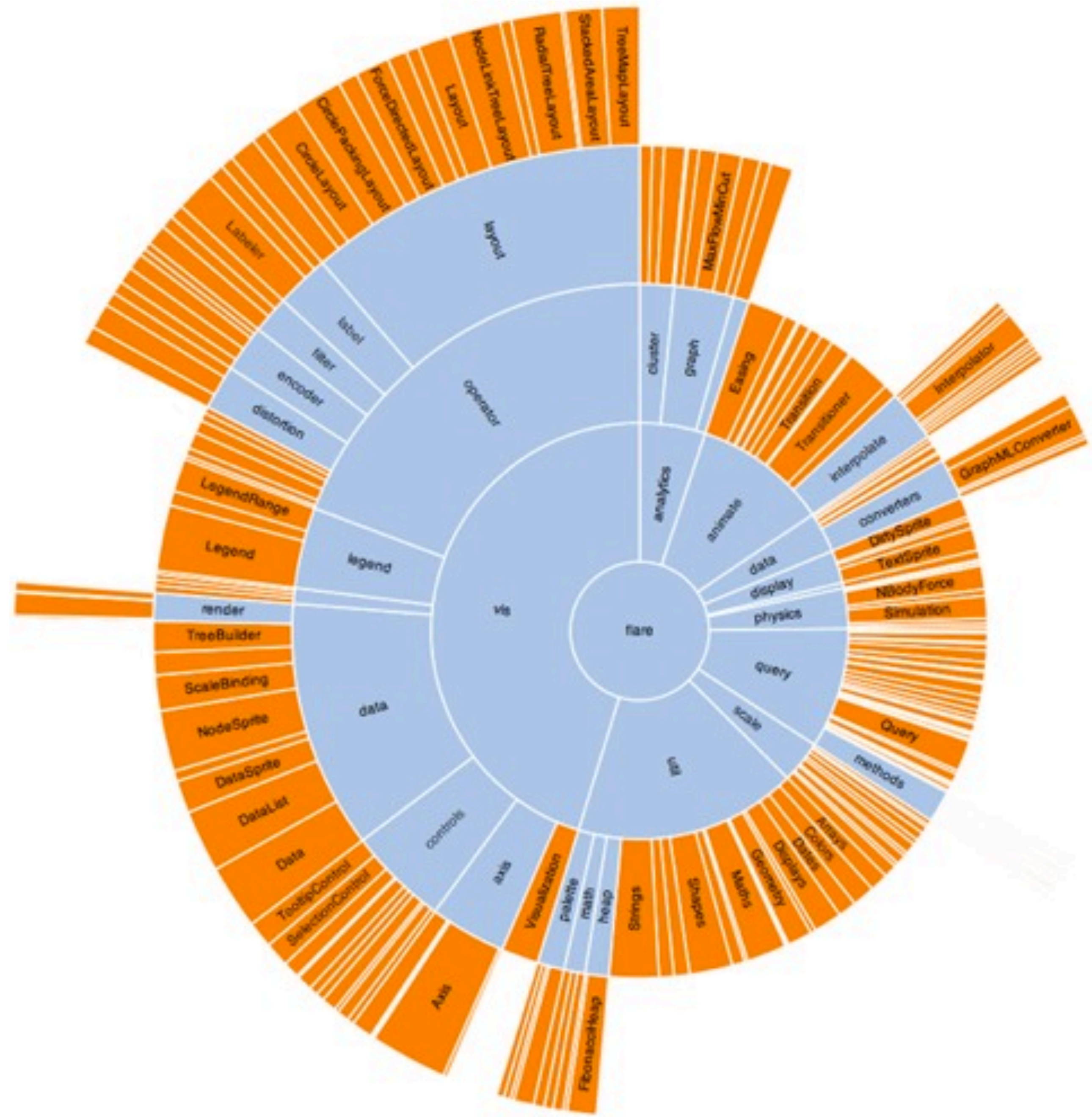




http://hci.stanford.edu/jheer/files/zoo/

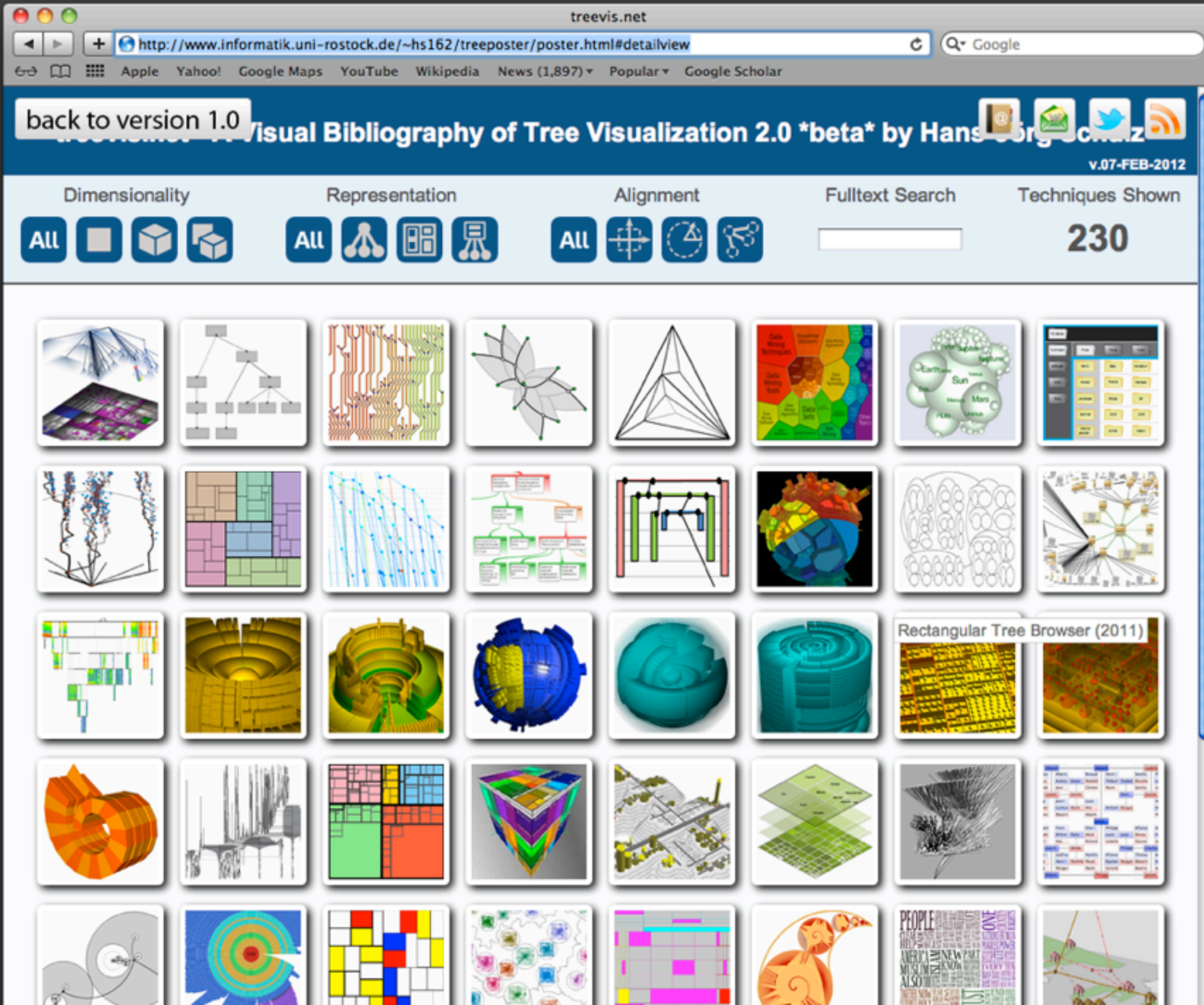
LAYERED





http://ivtk.sourceforge.net/

http://hci.stanford.edu/jheer/files/zoo/


treevis.net
http://www.informatik.uni-rostock.de/~hs162/treeposter/poster.html#detailview
Google
Apple Yahoo! Google Maps YouTube Wikipedia News (1,897) Popular Google Scholar
back to version 1.0
isual Bibliography of Tree Visualization 2.0 *beta* by Hans
v.07-FEB-2012
Dimensionality
Representation
Alignment
Fulltext Search
Techniques Shown
All
All
All
230
Rectangular Tree Browser (2011)

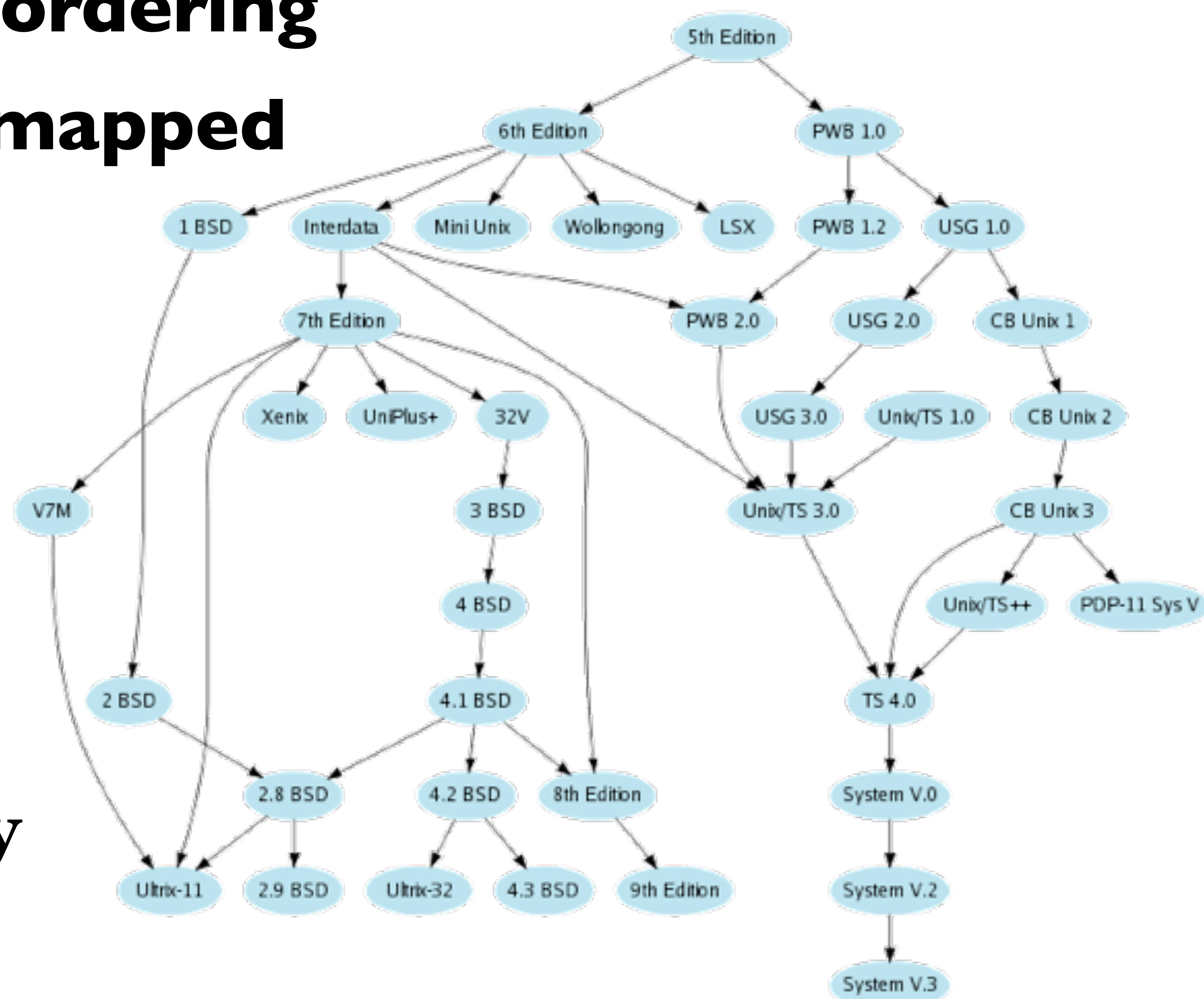
VISUALIZING GRAPHS

GRAPH LAYOUTS

- **node link layouts**
  - layered / Sugiyama
  - force directed
  - other
- **matrix layouts**
- **attribute based layouts**

SUGIYAMA-TYPE LAYOUT

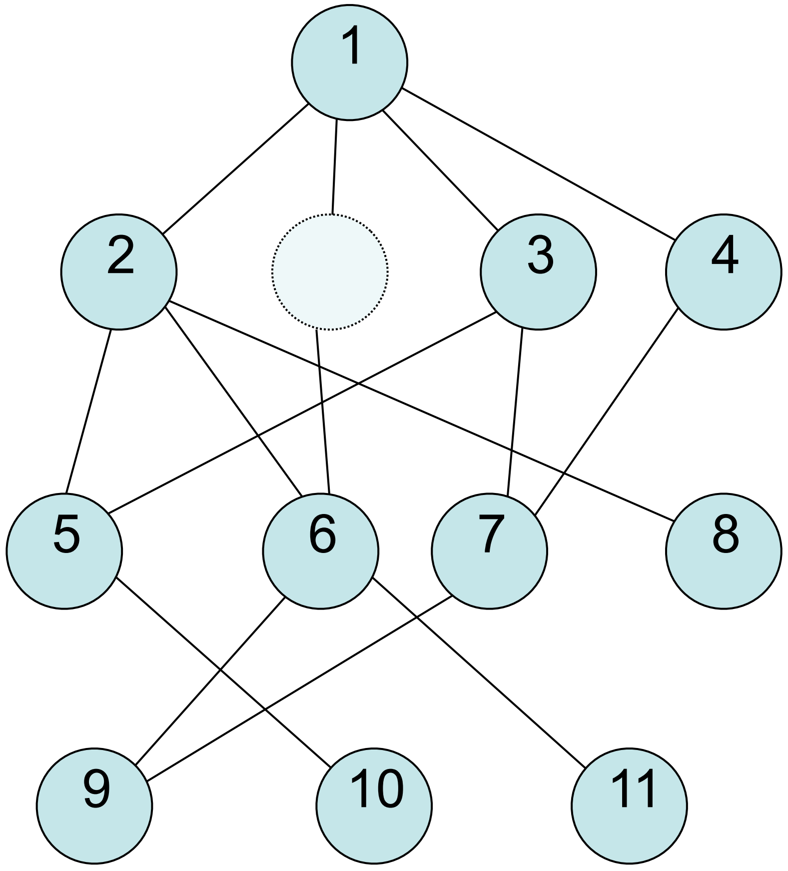
- **great for graphs that have an intrinsic ordering**
- **'depth' in graph mapped to one axis**



UNIX ancestry

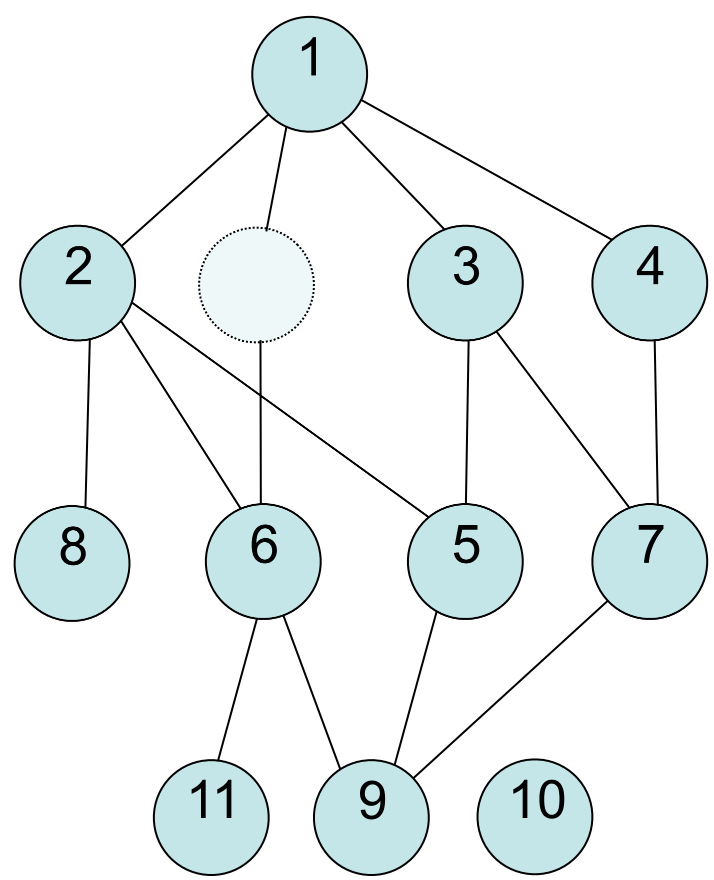
SUGIYAMA STEP 1

- **create layering of graph**
  - from domain specific knowledge
  - longest path from root
  - algorithmically determine best layering (NP-Hard)
- **dummy nodes for long edges**



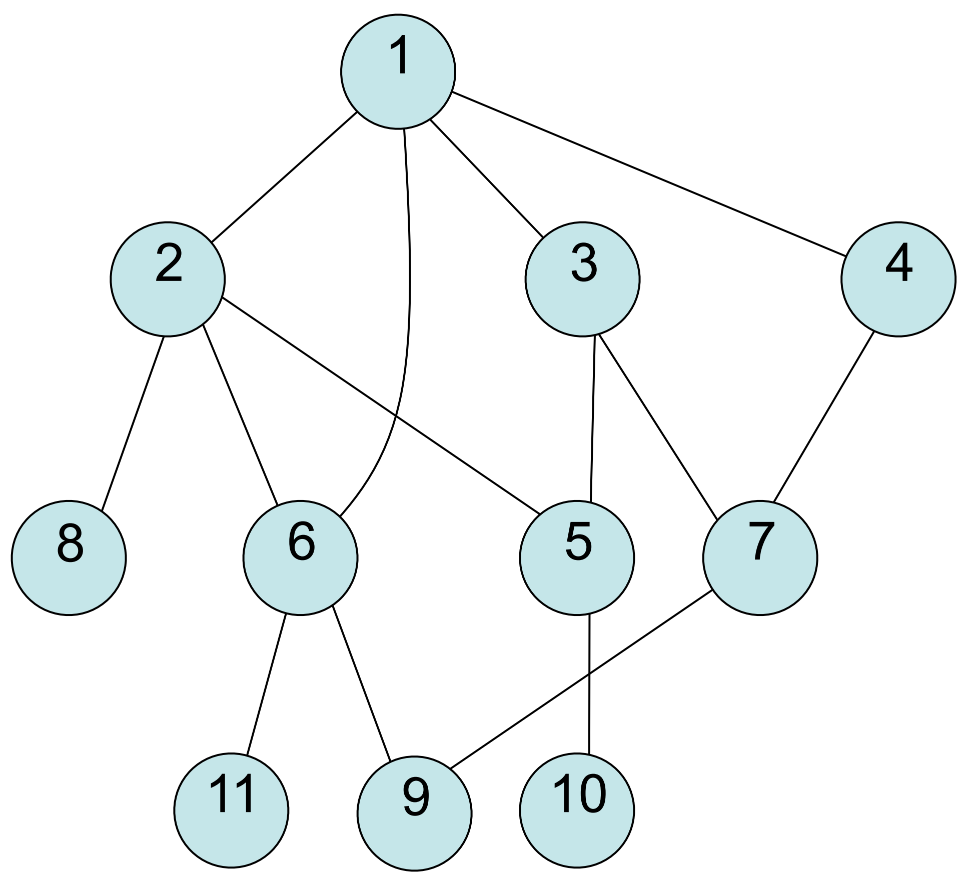
SUGIYAMA STEP 2

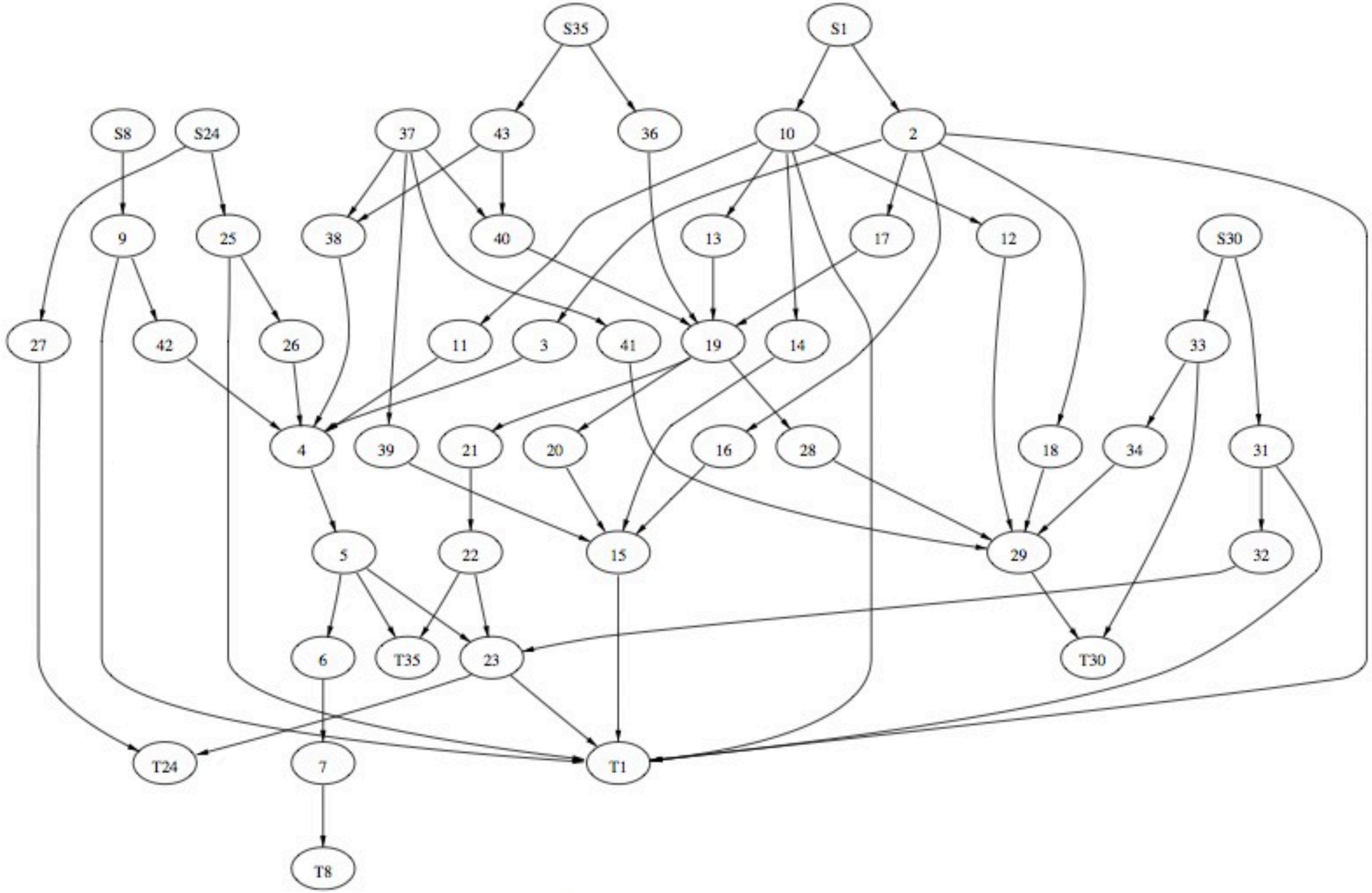
- **minimize crossings layer by layer (NP-hard)**
- **numerous heuristics available**



SUGIYAMA STEP 3

- **final assignment of x-coordinates**
- **routing of edges**





Gansner 1993

SUGIYAMA

+ nice, readable top down flow

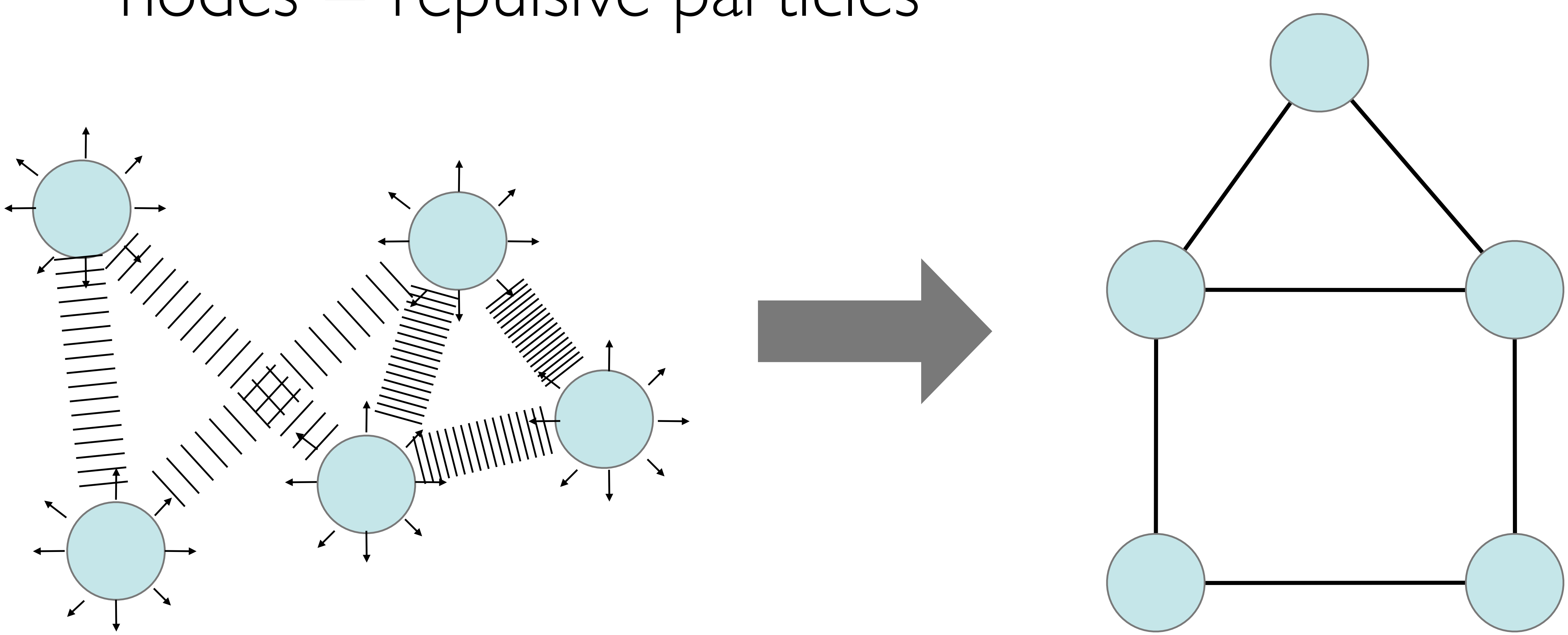
+ relatively fast (depending on heuristic used for crossing minimization)

- **not really suitable for graphs that don't have an intrinsic top down structure**

- **hard to implement**
  - use free graphviz lib instead: http://www.graphviz.org

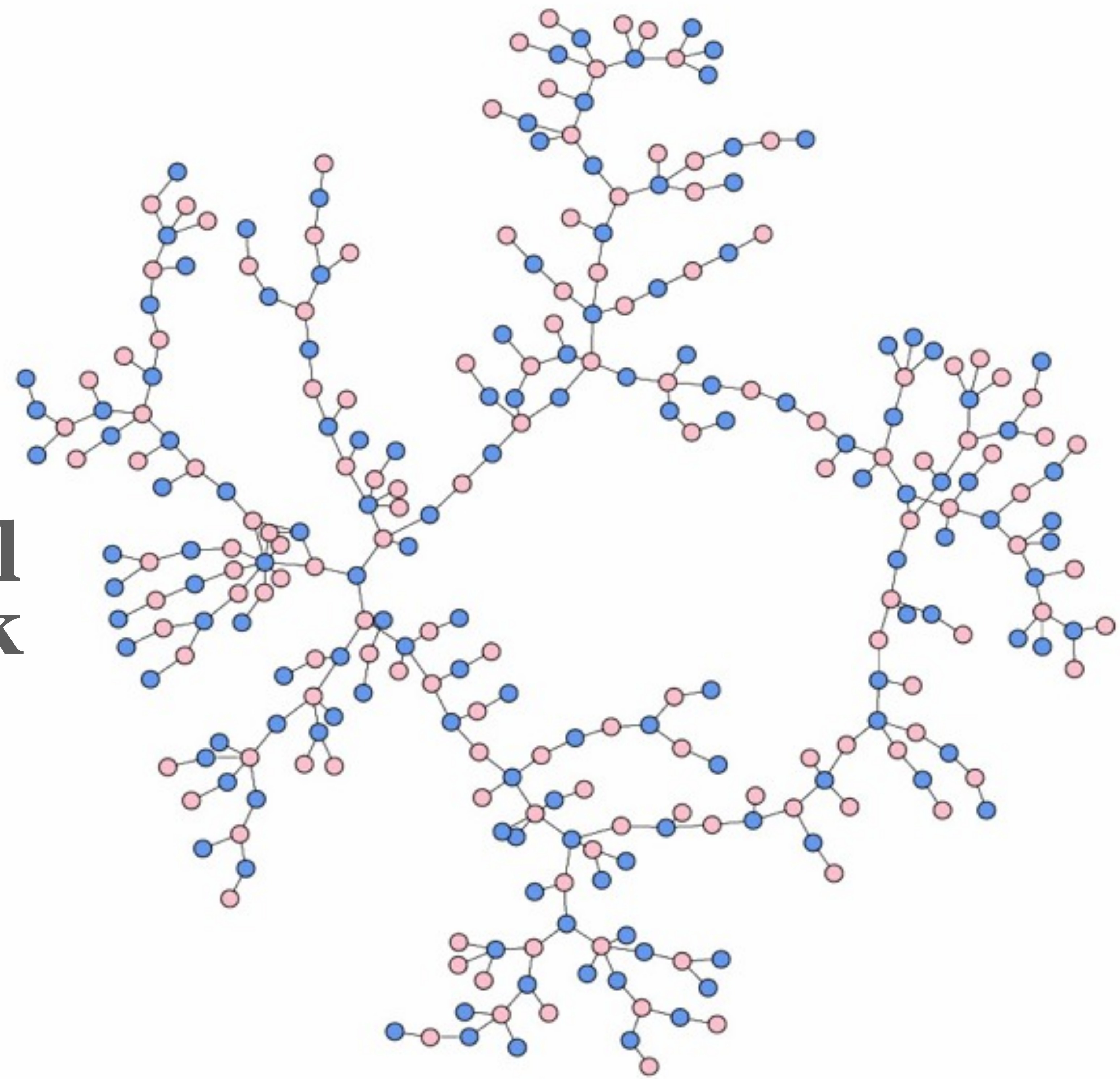
FORCE DIRECTED LAYOUT

- **no intrinsic layering, now what?**
- **physics model**
  - edges = springs
  - nodes = repulsive particles



AESTHETIC RESULTS

highschool dating network



FORCE MODEL

- **many variations, but usually physical analogy:**
  - **repulsion** : $f_R(d) = C_R * m_1 * m_2 / d^2$
    - *m_1, m_2 are node masses*
    - *d is distance between nodes*
  - **attraction** : $f_A(d) = C_A * (d - L)$
    - *L is the rest length of the spring*
    - *i.e. Hooke's Law*
- **total force on a node x with position x'**
  - $\sum^{\text{neighbors}(x)} : f_A(||x'-y'||) * (x'-y') + -f_R(||x'-y'||) * (x'-y')$

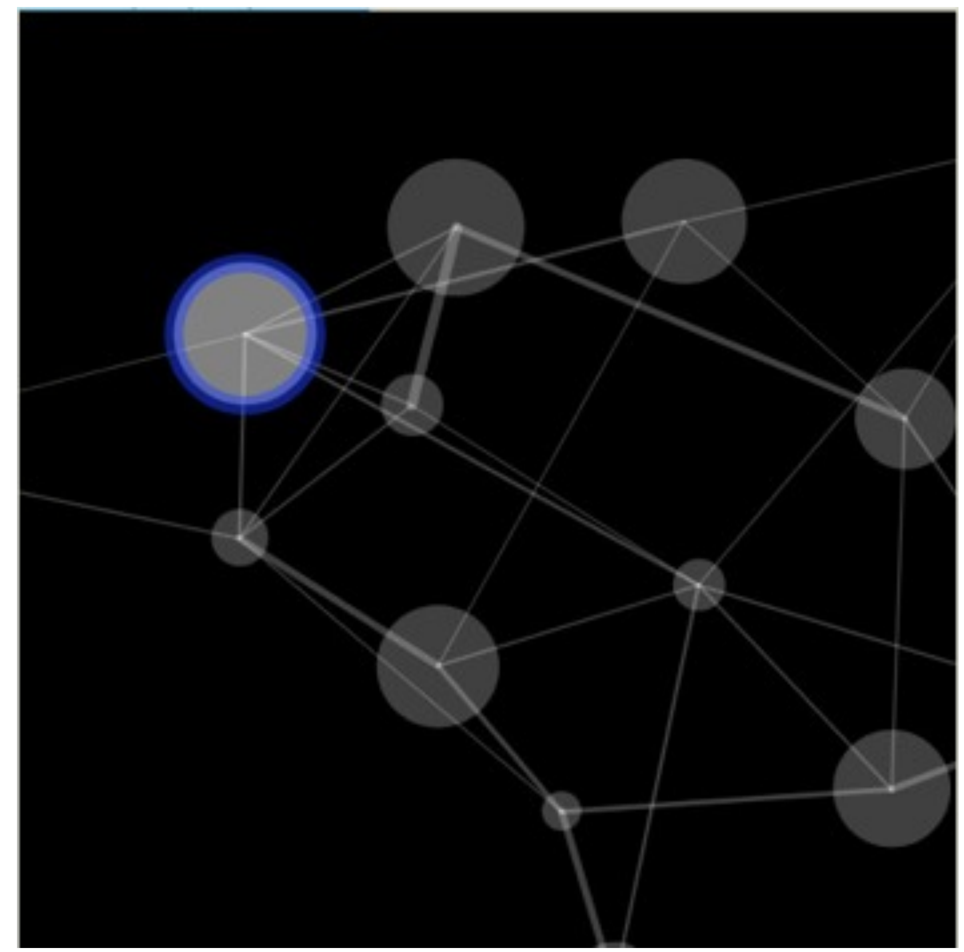
ALGORITHM

- **start from random layout**
- **(global) loop:**
  - for every node pair compute repulsive force
  - for every edge compute attractive force
  - accumulate forces per node
  - update each node position in direction of accumulated force
- **stop when layout is ‘good enough’**

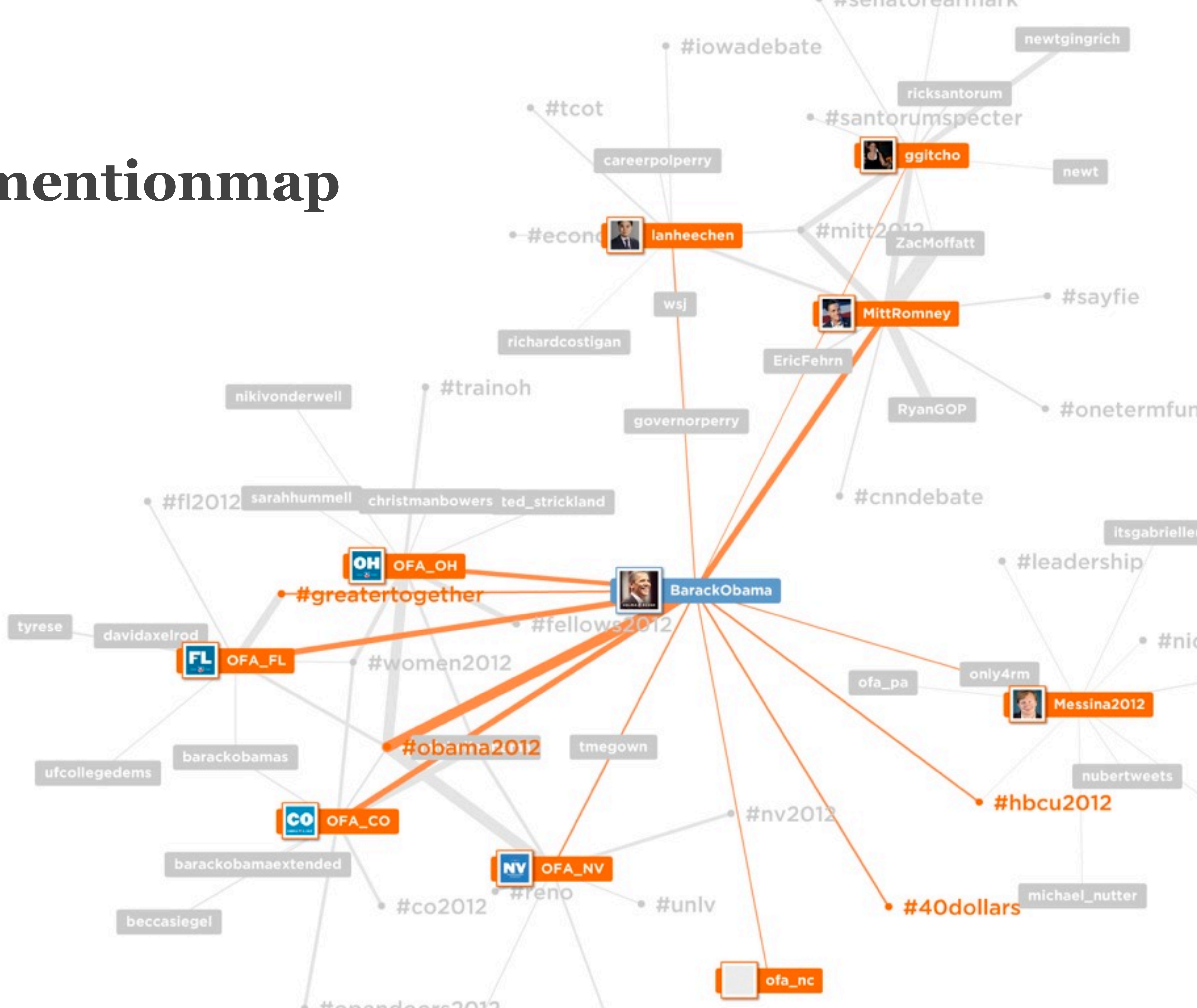
FORCE DIRECTED LAYOUTS

+ **very flexible, aesthetic layouts on many types of graphs**
+ **can add custom forces**
+ **relatively easy to implement**

- **repulsion loop is $O(n^2)$ per iteration**
  - can speed up to O(N log N) using quadtree or k-d tree
- **prone to local minima**
  - can use simulated annealing



mentionmap


#iowadebate
newtgingrich
ricksantorum
#tcot
#santorumspecter
ggitcho
newt
careerpolperry
lanheechen
#mitt2012
ZacMoffatt
wsj
MittRomney
#sayfie
richardcostigan
EricFehrn
nikivonderwell
#trainoh
governorperry
RyanGOP
#cnndebate
#fl2012
sarahhummell
christmanbowers
ted_strickland
#leadership
OFA_OH
#greatertogether
BarackObama
tyrese
davidaxelrod
#fellows2012
OFA_FL
#women2012
ofa_pa
only4rm
Messina2012
#obama2012
tmegown
barackobamas
ufcollegedems
nubertweets
#hbcu2012
OFA_CO
#nv2012
barackobamaextended
OFA_NV
#reno
#unlv
michael_nutter
#40dollars
beccasiegel
#co2012
ofa_nc

vizster

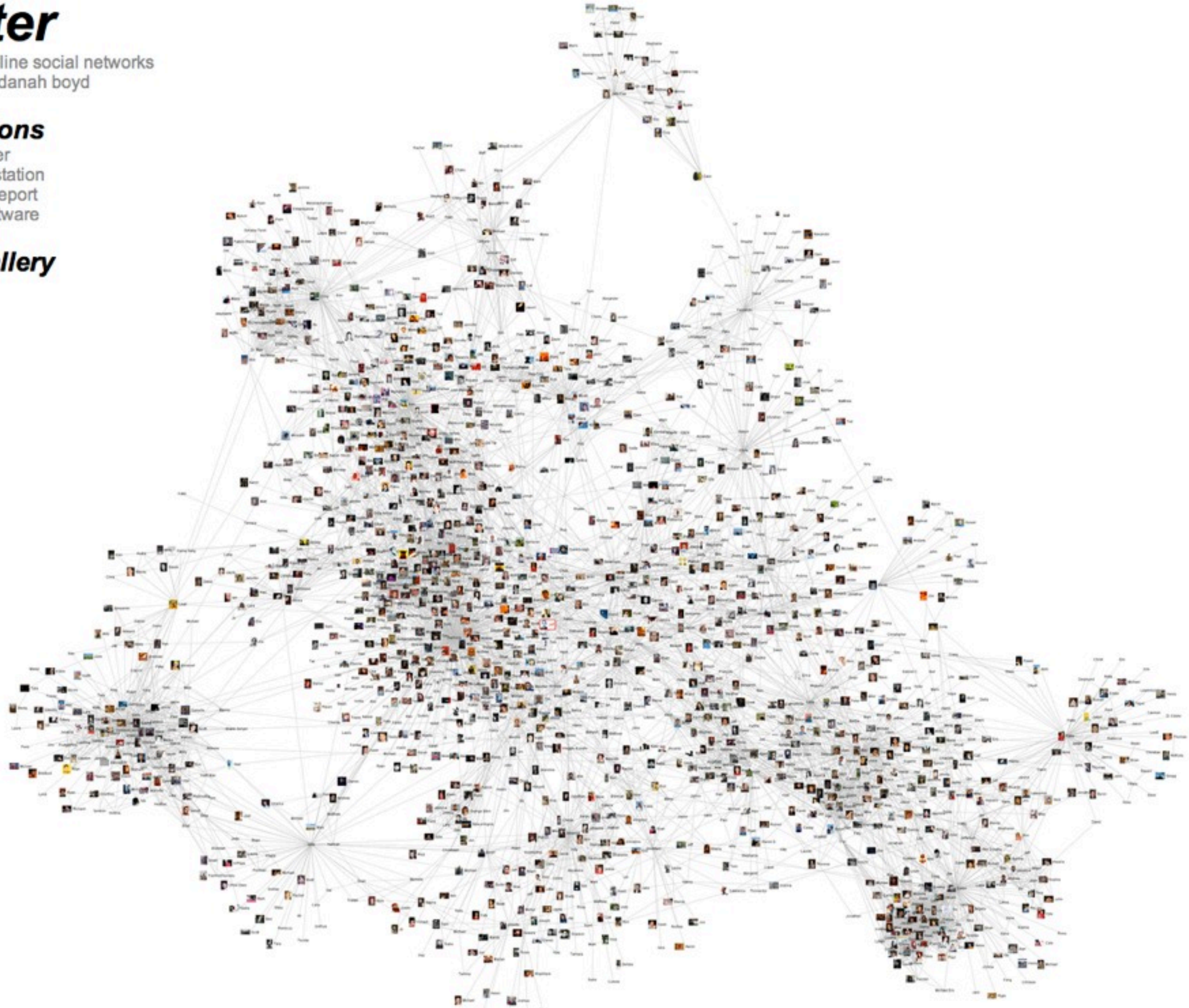
visualizing online social networks
jeffrey heer + danah boyd

publications

research paper
video demonstation
early design report
download software

photo gallery

egocentric
community
linkage
search
x-ray 1
x-ray 2



Gephi, an open source graph visualization and manipulation software

http://gephi.org/

hookes law

Apple Yahoo! Google Maps YouTube Wikipedia News (1,949) Popular Google Scholar

Download Blog Store Wiki Forum Support Bug tracker

Gephi
makes graphs handy

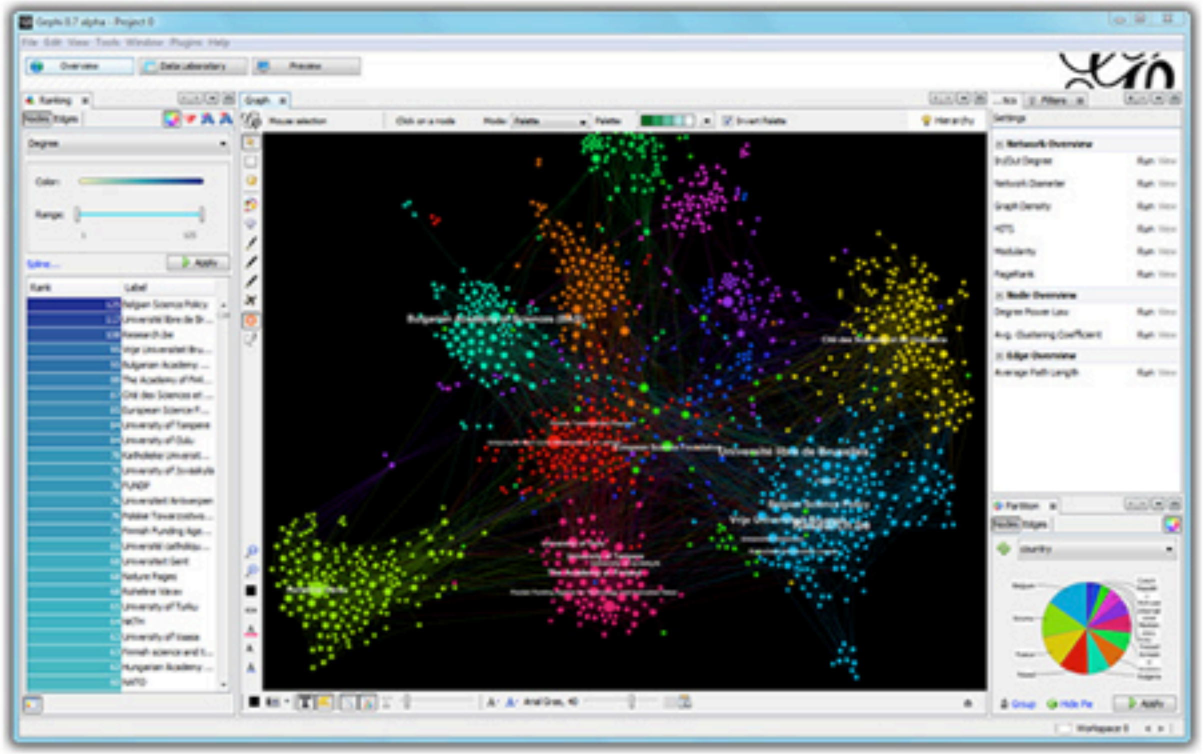
Home Features Learn Develop Plugins Consortium

The Open Graph Viz Platform

Gephi is an interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs.

Runs on Windows, Linux and Mac OS X. Gephi is open-source and free.

Learn More on Gephi Platform »



Download FREE
Gephi 0.8 beta

Release Notes | System Requirements

- Features
- Quick start
- Screenshots
- Videos

Gephi 0.8 beta has been released! Discover a new Preview and dynamic features, start building commercial applications with the new open source license.

Learn More »

APPLICATIONS

- **Exploratory Data Analysis**: intuition-oriented analysis by networks manipulations in real time.
- **Link Analysis**: revealing the underlying structures of associations between objects, in particular in scale-free networks.
- **Social Network Analysis**: easy creation of social data connectors to map community organizations and small-world networks.
- **Biological Network analysis**: representing patterns of biological data.
- **Poster creation**: scientific work promotion with hi-quality printable maps.

Learn More »

"Like Photoshop™ for graphs.

— the Community

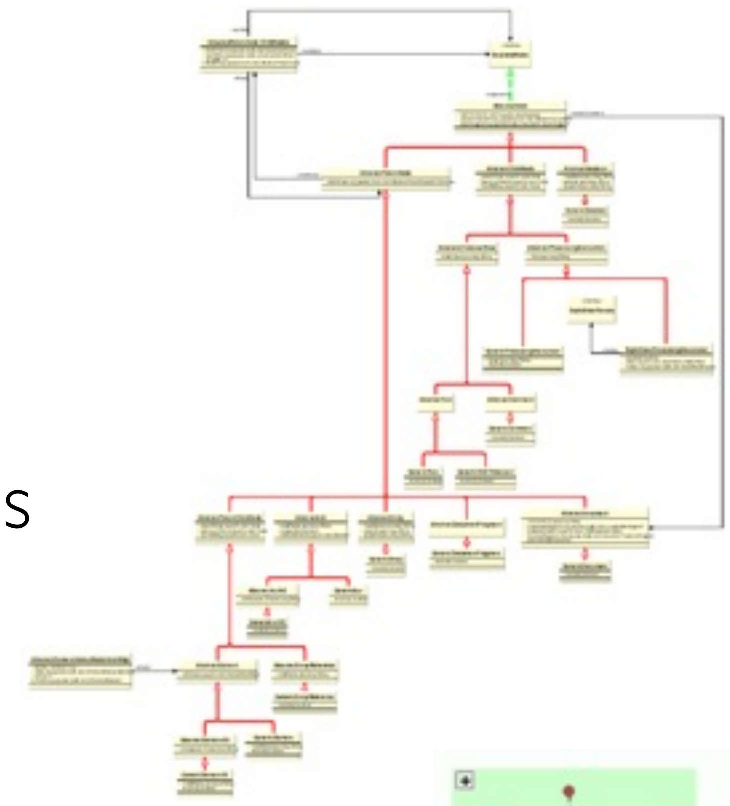
LATEST NEWS

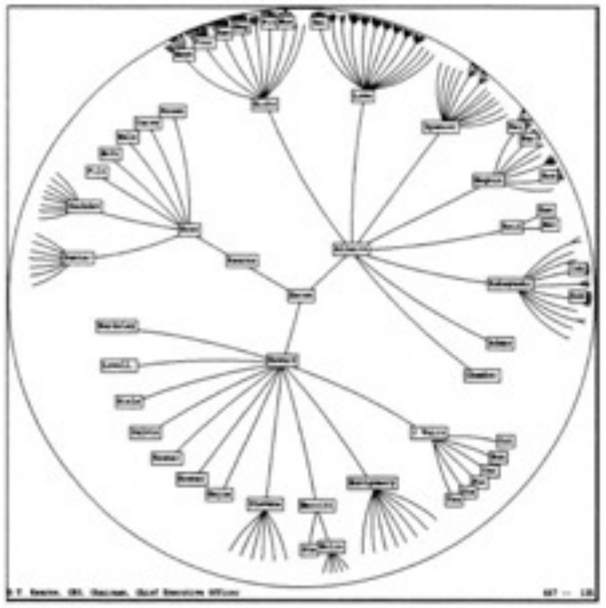
- Weekly news
  February 27, 2012
- Annual report 2011
  February 25, 2012
- Gephi-Neo4j presentation at FOSDEM
  February 20, 2012
- Gephi meet-up #4 in Berlin
  February 2, 2012
- Introducing the Gephi Plugins Bootcamp
  January 12, 2012

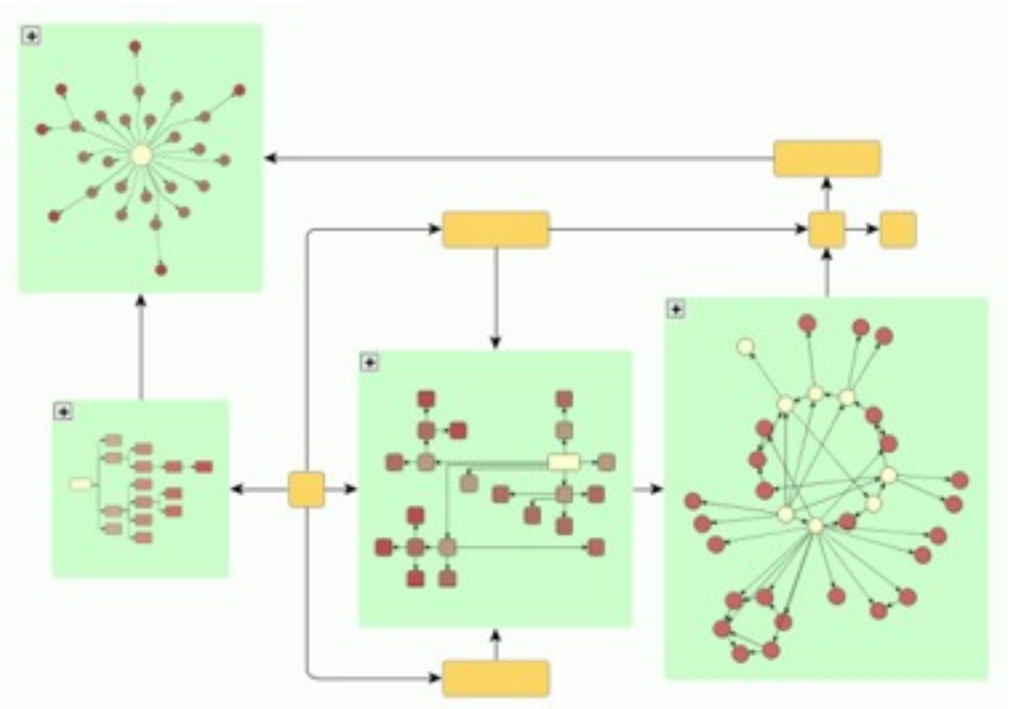
PAPERS

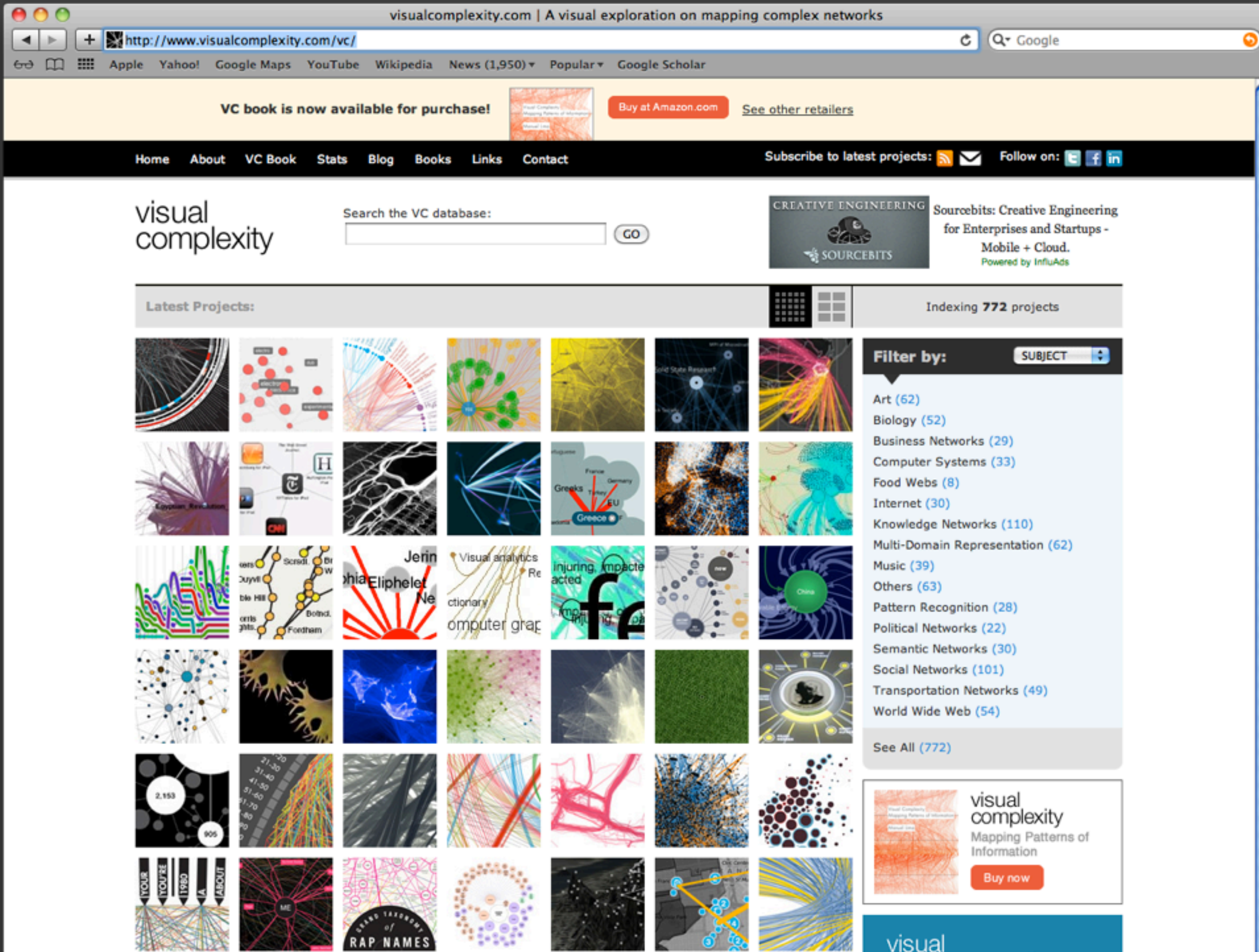
OTHER NODE LINK LAYOUTS

- **orthogonal**
  - great for UML diagrams
  - algorithmically complex
- **circular layouts**
  - emphasizes ring topologies
  - used in social network diagrams
- **nested layouts**
  - recursively apply layout algorithms
  - great for graphs with hierarchical structure

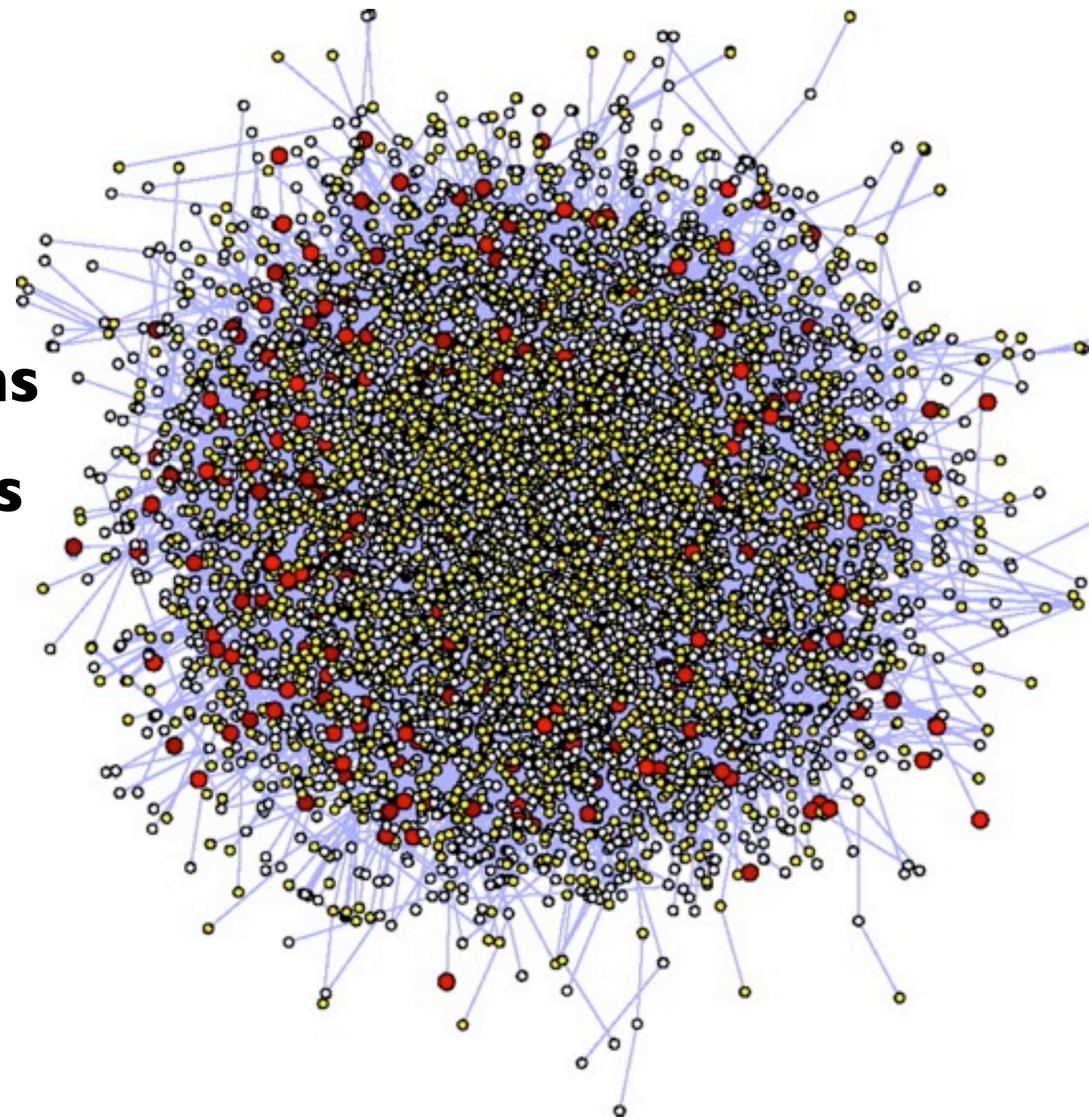







visualcomplexity.com | A visual exploration on mapping complex networks
http://www.visualcomplexity.com/vc/
Google
Apple Yahoo! Google Maps YouTube Wikipedia News (1,950) Popular Google Scholar
VC book is now available for purchase!
Buy at Amazon.com
See other retailers
Home About VC Book Stats Blog Books Links Contact
Subscribe to latest projects:
Follow on:
visual complexity
Search the VC database:
GO
CREATIVE ENGINEERING
SOURCEBITS
Sourcebits: Creative Engineering for Enterprises and Startups - Mobile + Cloud.
Powered by InfluAds
Latest Projects:
Indexing 772 projects
Filter by:
SUBJECT
Art (62)
Biology (52)
Business Networks (29)
Computer Systems (33)
Food Webs (8)
Internet (30)
Knowledge Networks (110)
Multi-Domain Representation (62)
Music (39)
Others (63)
Pattern Recognition (28)
Political Networks (22)
Semantic Networks (30)
Social Networks (101)
Transportation Networks (49)
World Wide Web (54)
See All (772)
visual complexity
Mapping Patterns of Information
Buy now
visual
RAP NAMES

NODE LINK LAYOUTS



+ understandable visual mapping

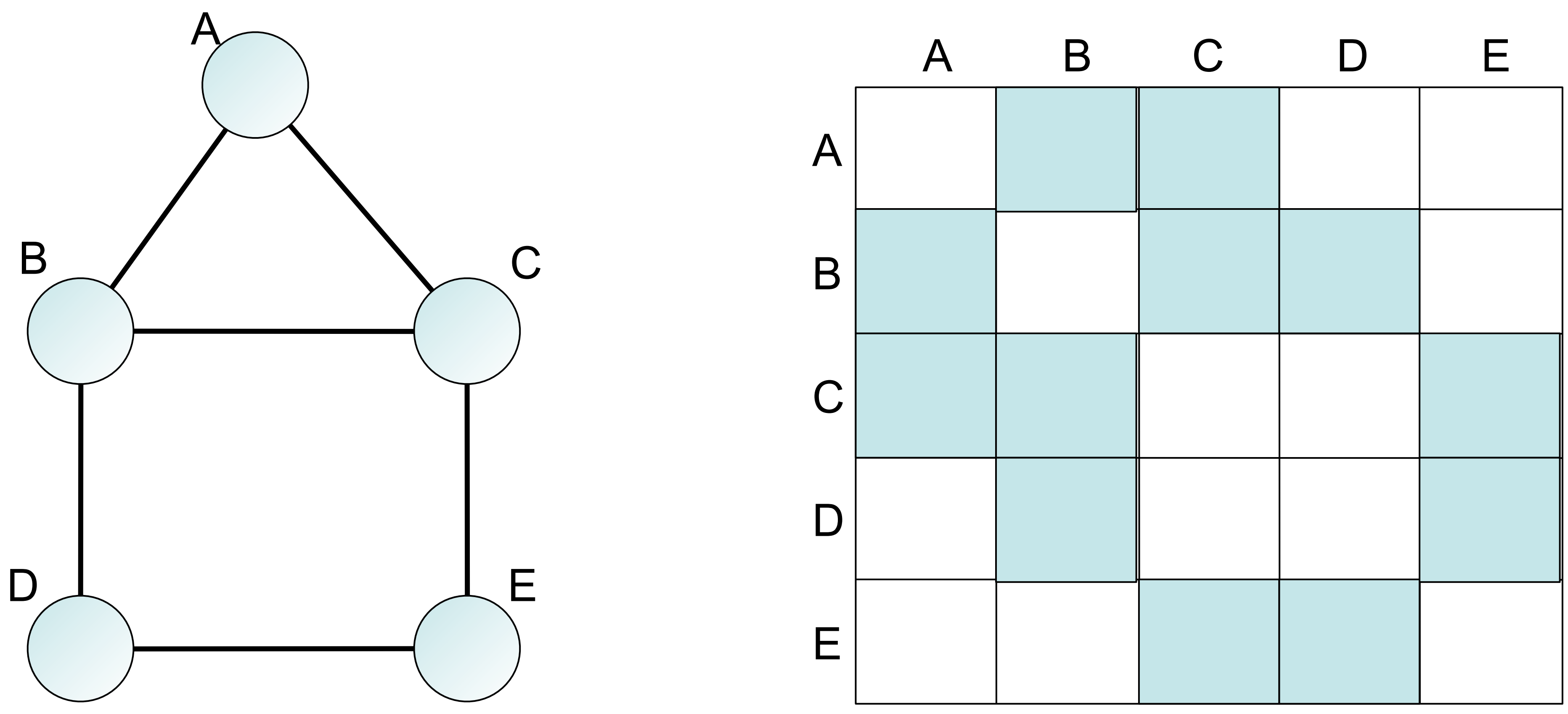
+ can show overall structure, clusters, paths

+ flexible, many variations

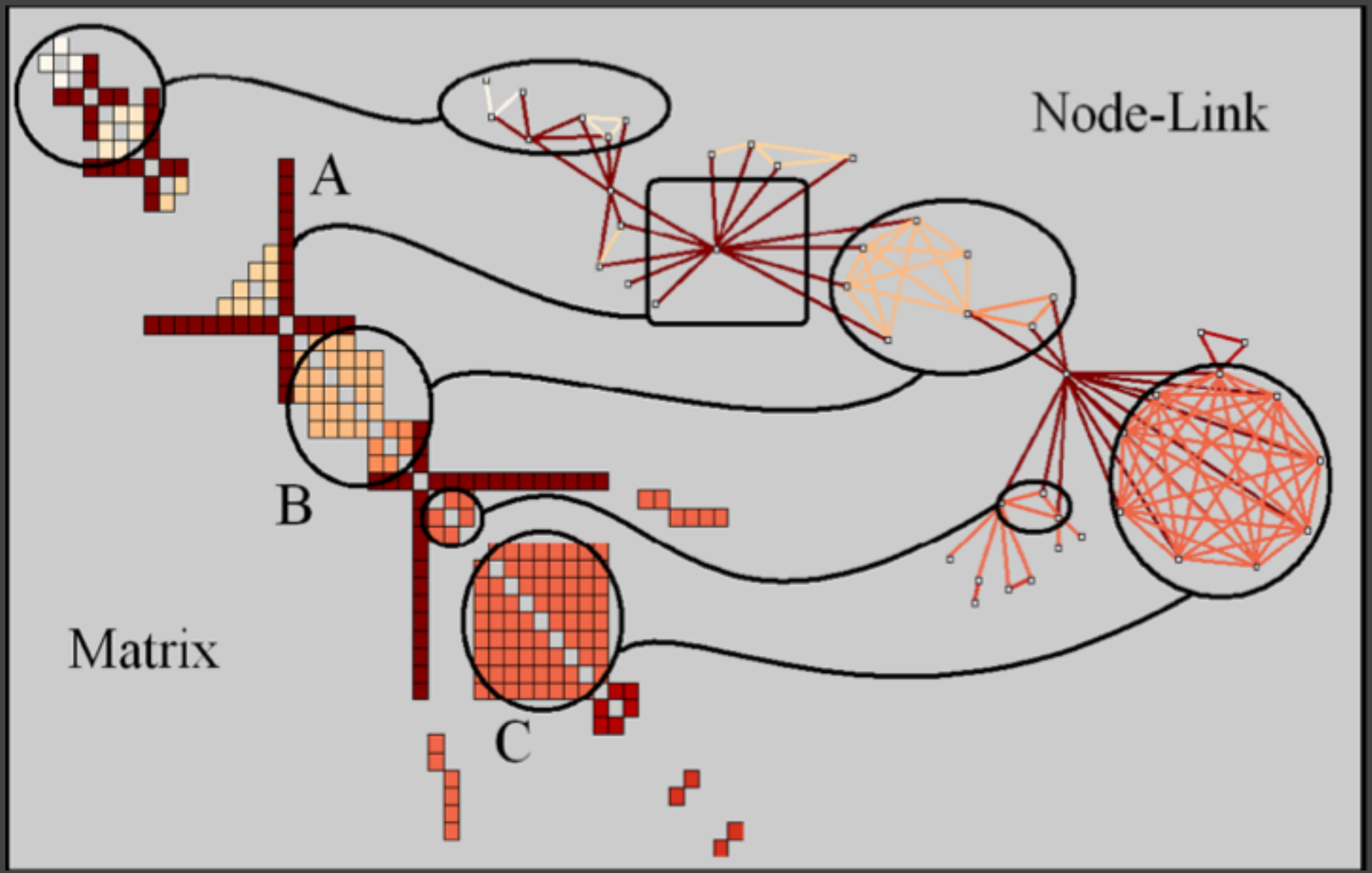
- **all but the most trivial algorithms are > $O(N^2)$**
- **not good for dense graphs**
  - hairball problem!

MATRIX LAYOUTS

- **instead of node link diagram, use adjacency matrix representation**



SPOTTING PATTERNS IN MATRICES



Henry 2006

MATRIX REPRESENTATIONS

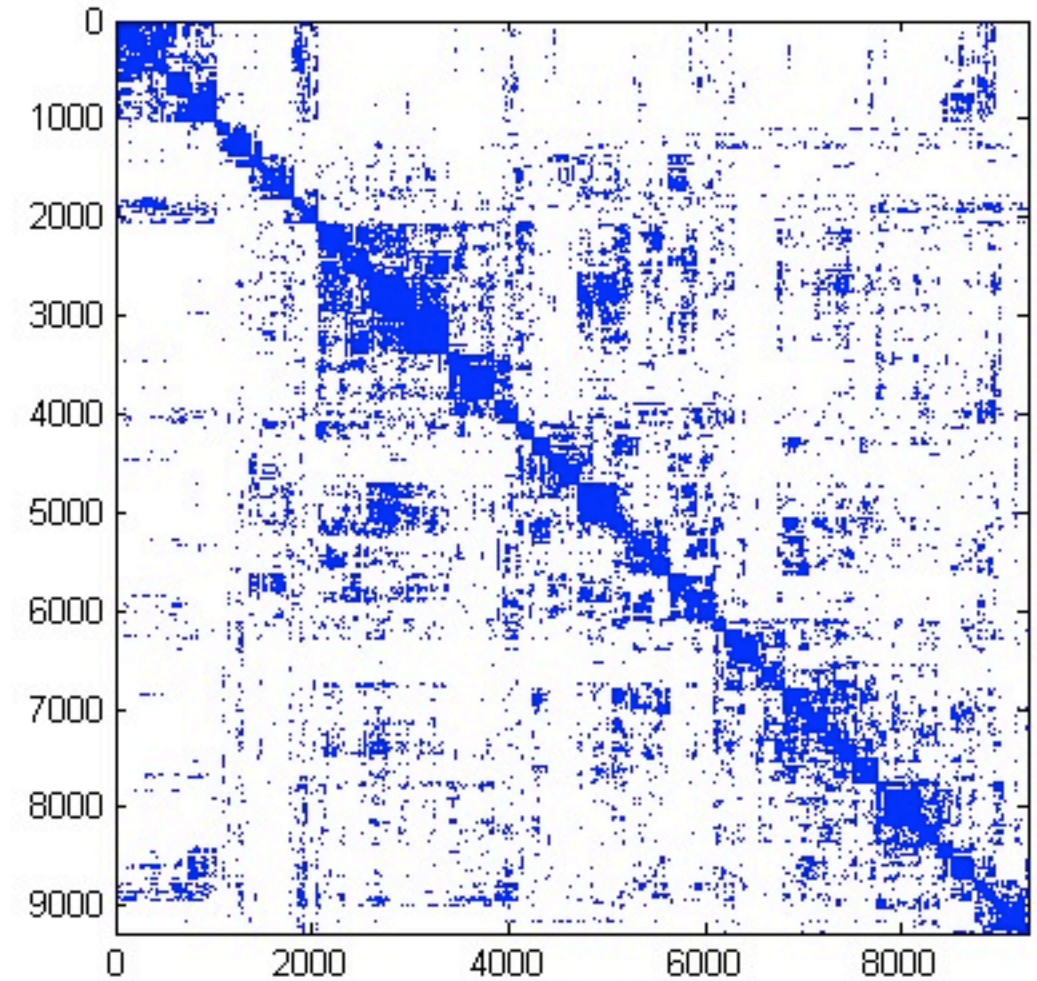
+ great for dense graphs

+ visually scalable

+ can spot clusters

- abstract visualization

- hard to follow paths

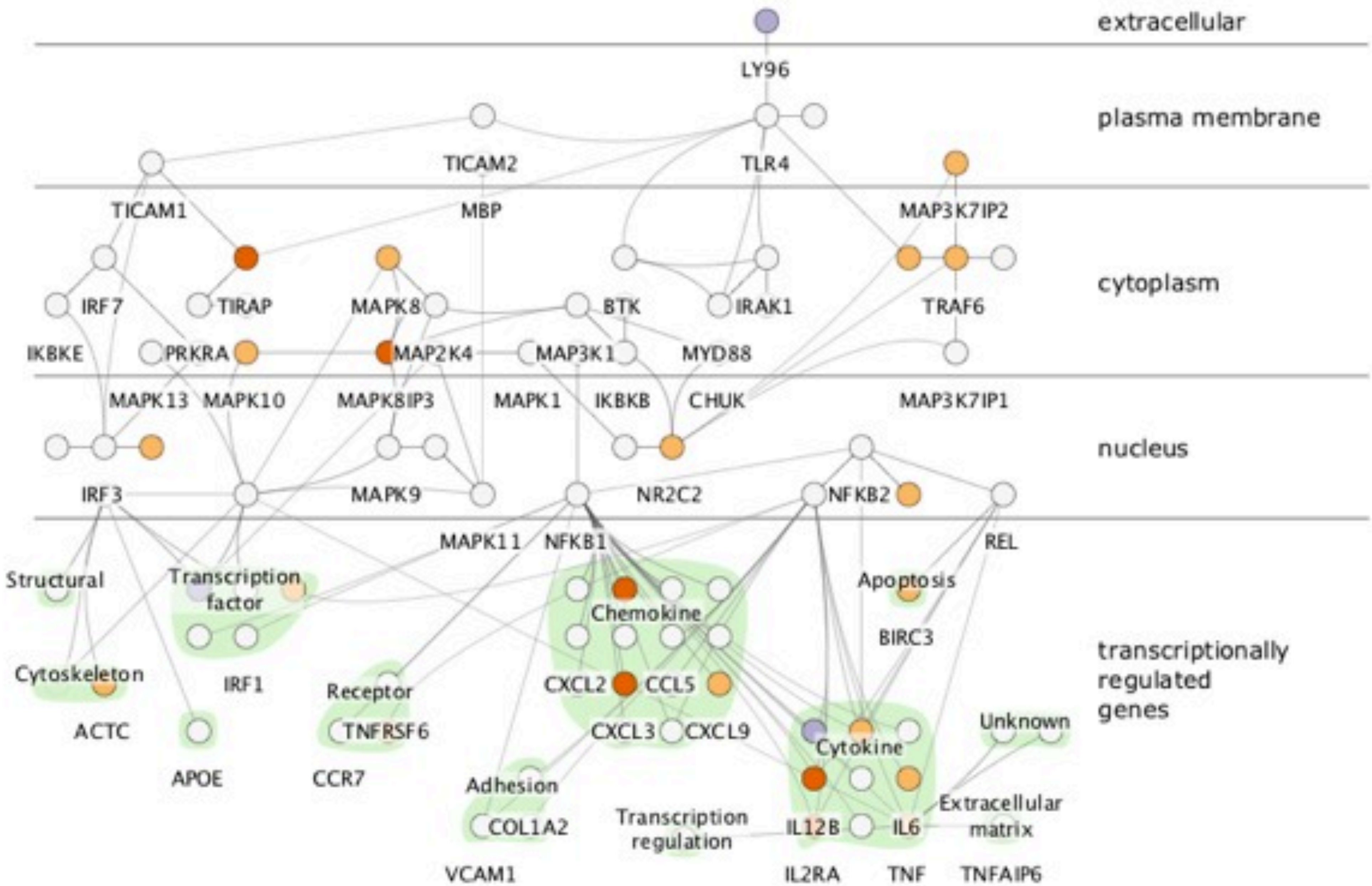


ALTERNATIVE LAYOUT

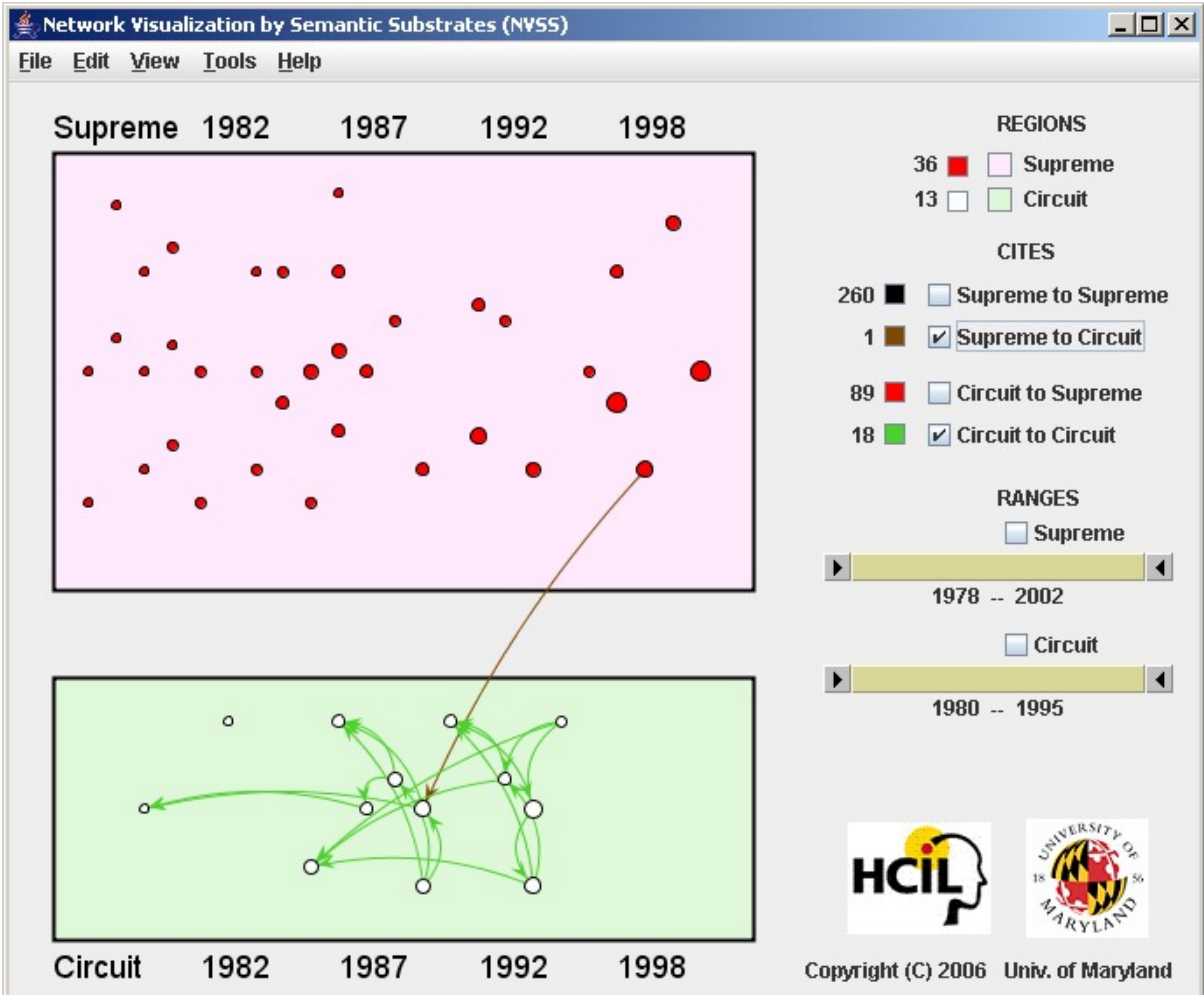
ATTRIBUTE-DRIVEN LAYOUT

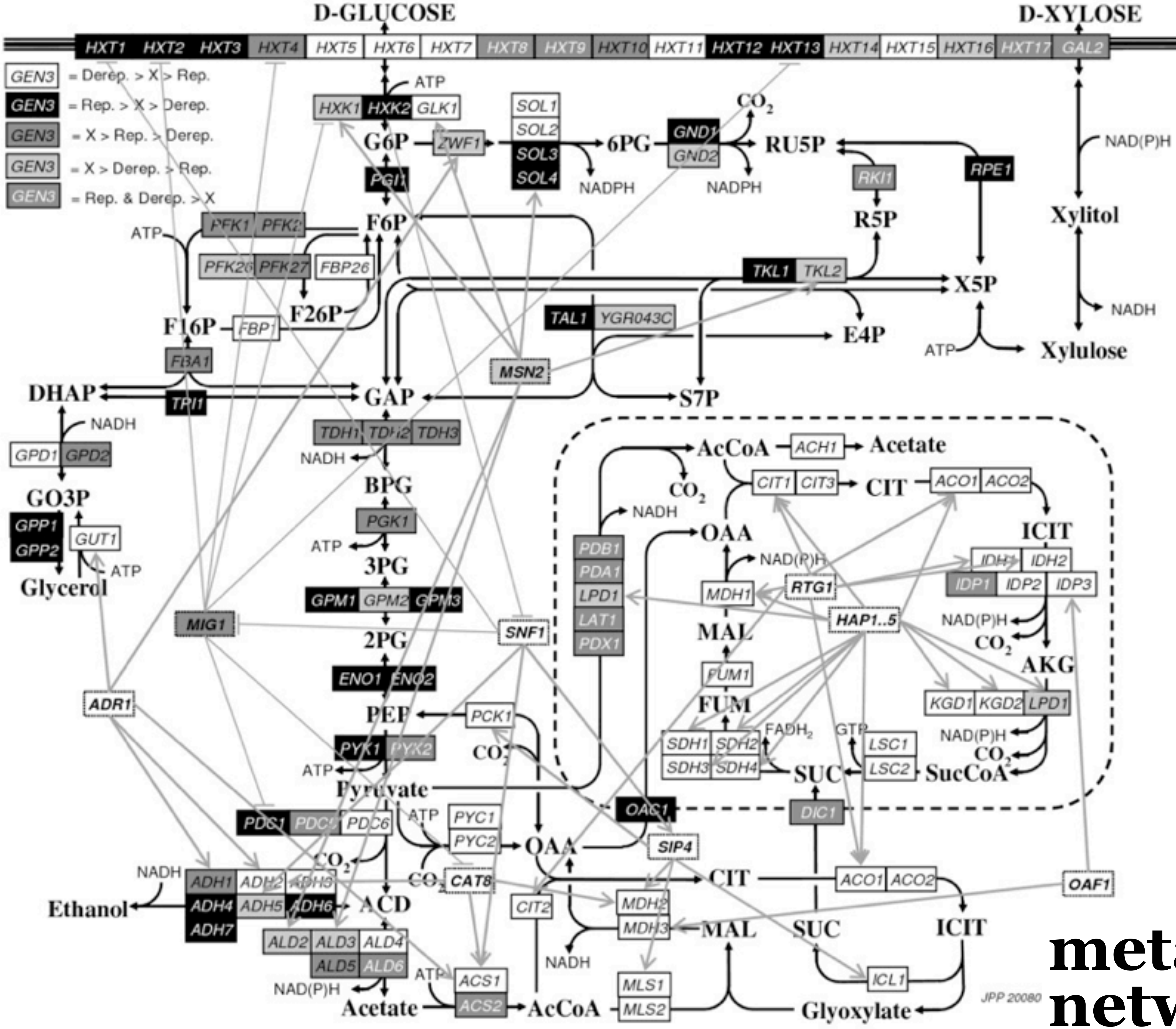
- **large node-link diagrams get messy!**
- **are there additional structures we can exploit?**
- **idea: use data attributes to perform layout**
  - e.g., scatterplot based on node values
- **dynamic queries and/or brushing can be used to enhance perception of connectivity**

cerebral

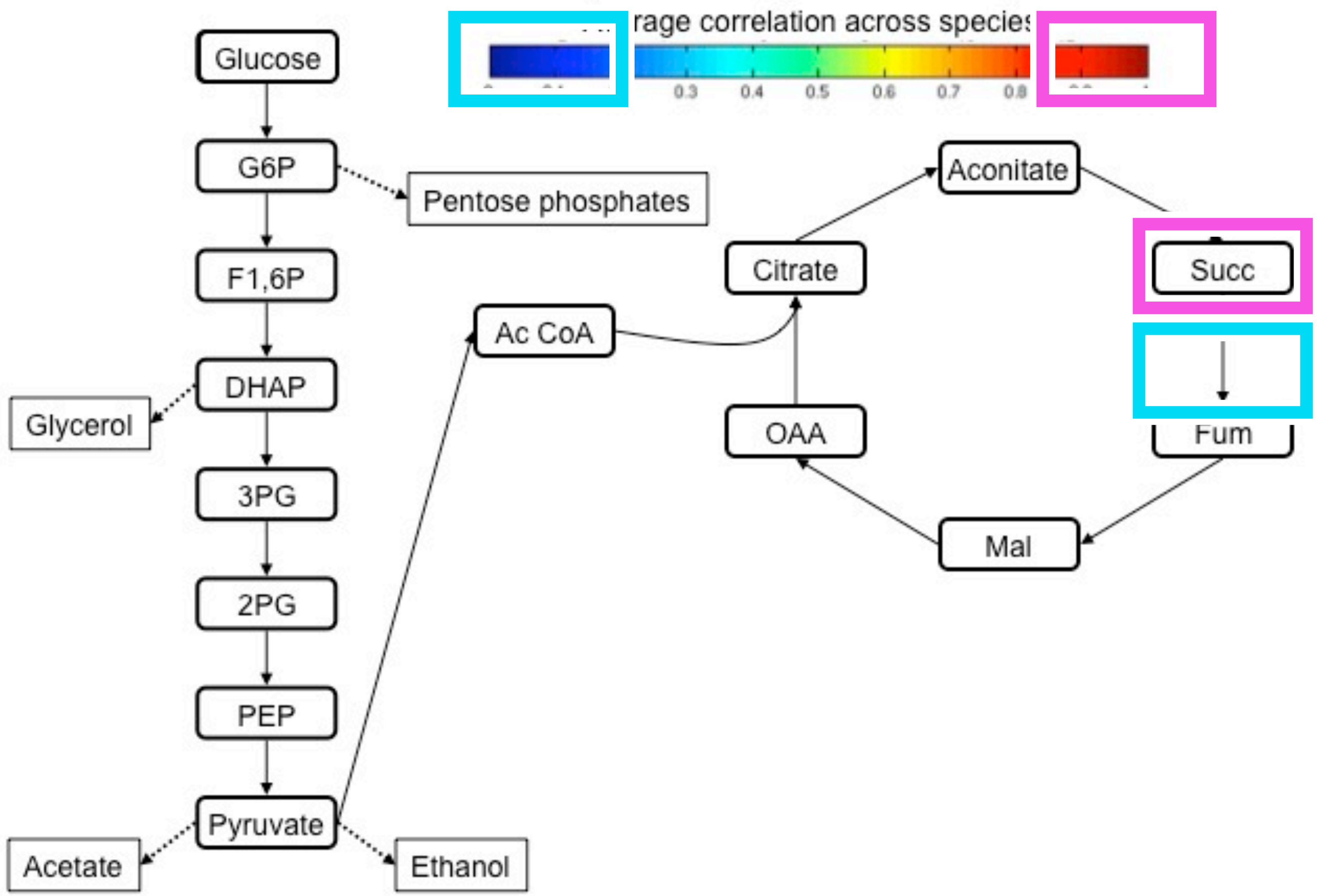

extracellular
LY96
plasma membrane
TICAM2
TLR4
TICAM1
MBP
MAP3K7IP2
cytoplasm
IRF7
TIRAP
MAPK8
BTK
IRAK1
TRAF6
IKBKE
PRKRA
MAP2K4
MAP3K1
MYD88
MAPK13
MAPK10
MAPK8IP3
MAPK1
IKBKB
CHUK
MAP3K7IP1
nucleus
IRF3
MAPK9
NR2C2
NFKB2
MAPK11
NFKB1
REL
Structural
Transcription factor
Chemokine
Apoptosis
BIRC3
transcriptionally regulated genes
Cytoskeleton
IRF1
Receptor
CXCL2
CCL5
ACTC
TNFRSF6
CXCL3
CXCL9
Unknown
Cytokine
APOE
CCR7
Adhesion
Extracellular matrix
COL1A2
Transcription regulation
IL12B
IL6
VCAM1
IL2RA
TNF
TNFAIP6

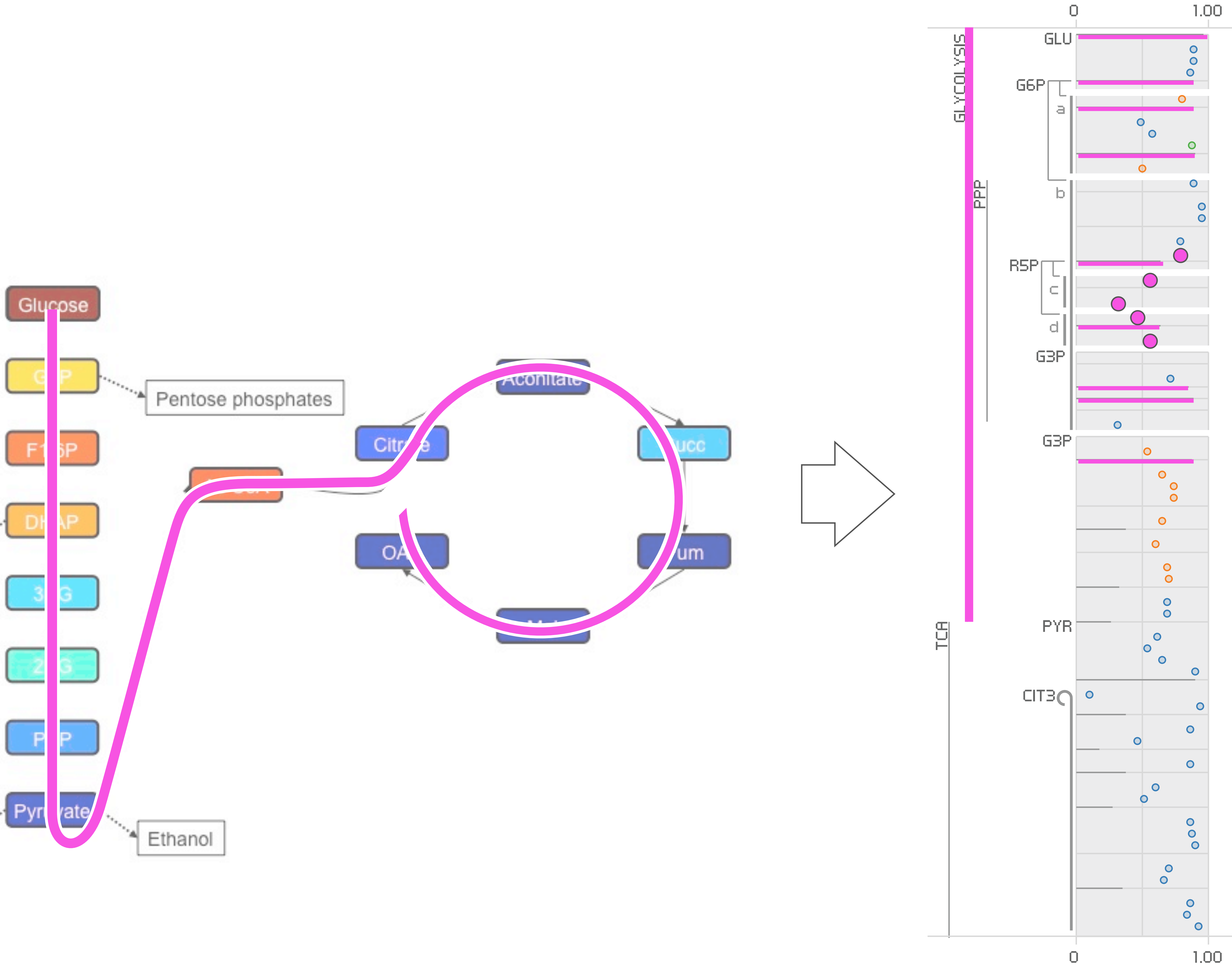
semantic substrates





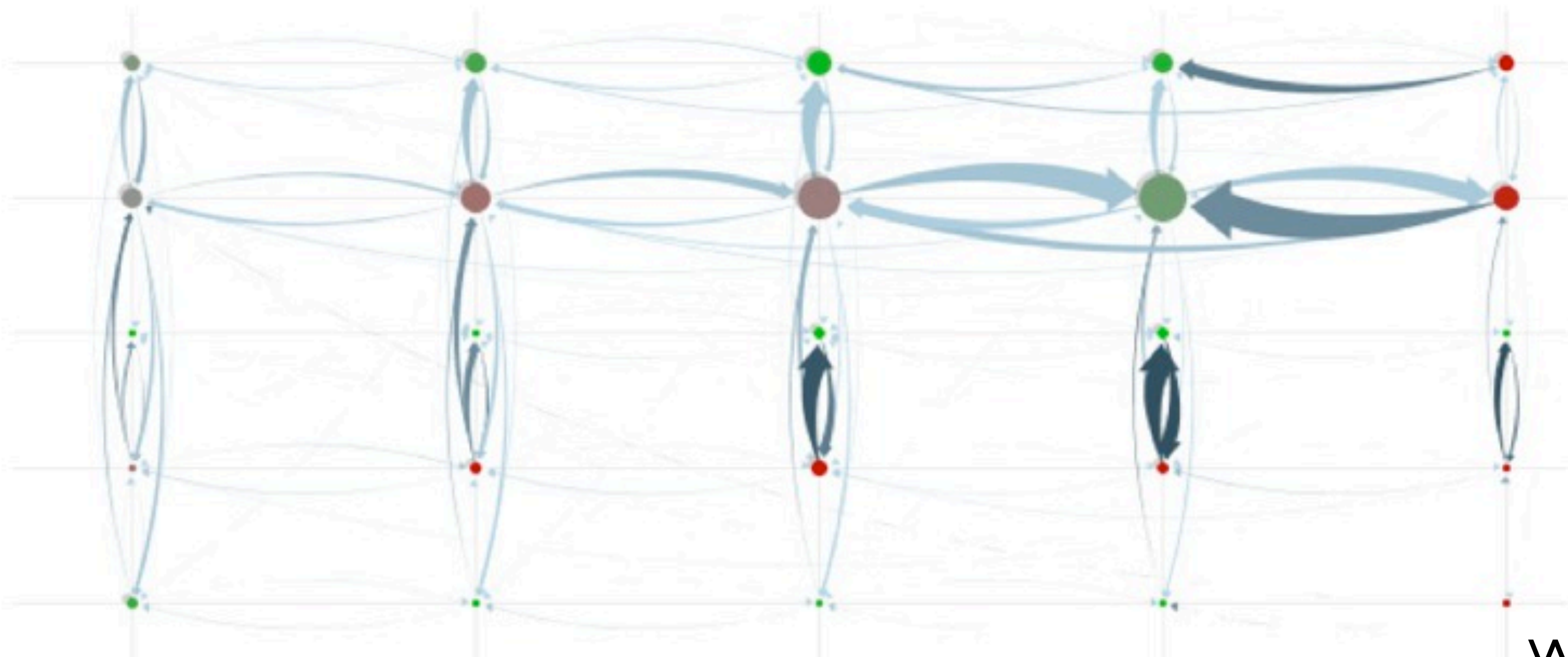
metabolic network


rage correlation across specie:
0.3
0.4
0.5
0.6
0.7
0.8
Glucose
G6P
Pentose phosphates
F1,6P
DHAP
Glycerol
3PG
2PG
PEP
Pyruvate
Acetate
Ethanol
Ac CoA
Citrate
Aconitate
Succ
Fum
Mal
OAA


Glucose
Pentose phosphates
Aconitate
Ethanol
0
1.00
GLU
G6P
a
b
R5P
c
d
G3P
G3P
PYR
CIT3
GLYCOLYSIS
PPP
TCA
0
1.00

PIVOT GRAPH

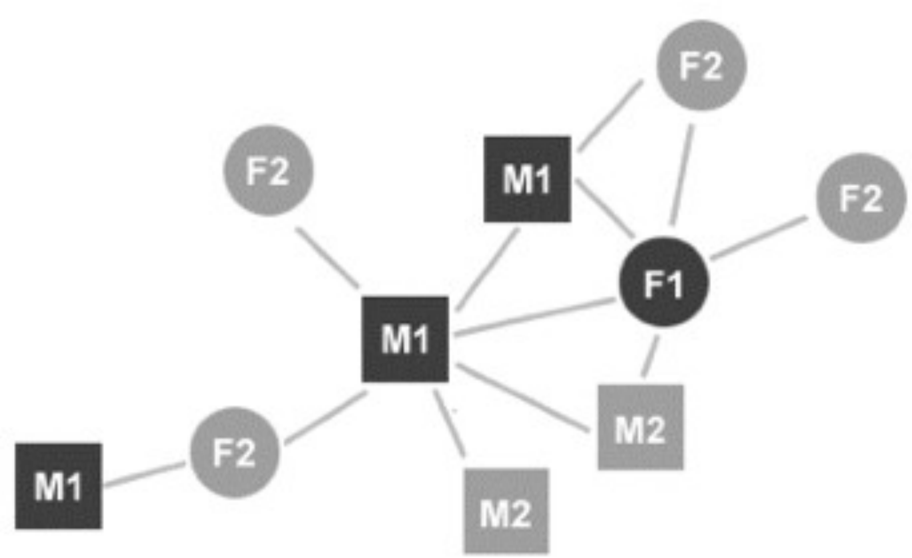
- **task abstraction**
  - show relationship between node attributes and connections in a multi-attribute graph
- **data abstraction**
  - relational dataset
    - *nodes (and edges) have multiple discrete attributes*
  - rollup and selection transformations



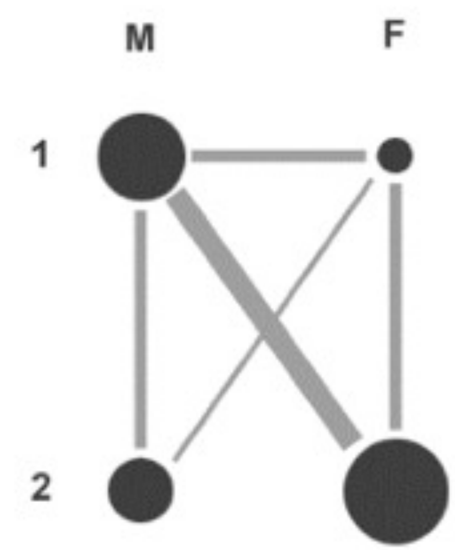
Wattenberg 2006

VISUAL ENCODING

- **line (1D) or grid (2D) layout**
  - area subdivided by number of values for an attribute
- **number of nodes based on attribute count, not original graph node count**
- **size of nodes and edges related to number of aggregated original nodes and edges**
- **scalability through abstraction, not layout algorithm**



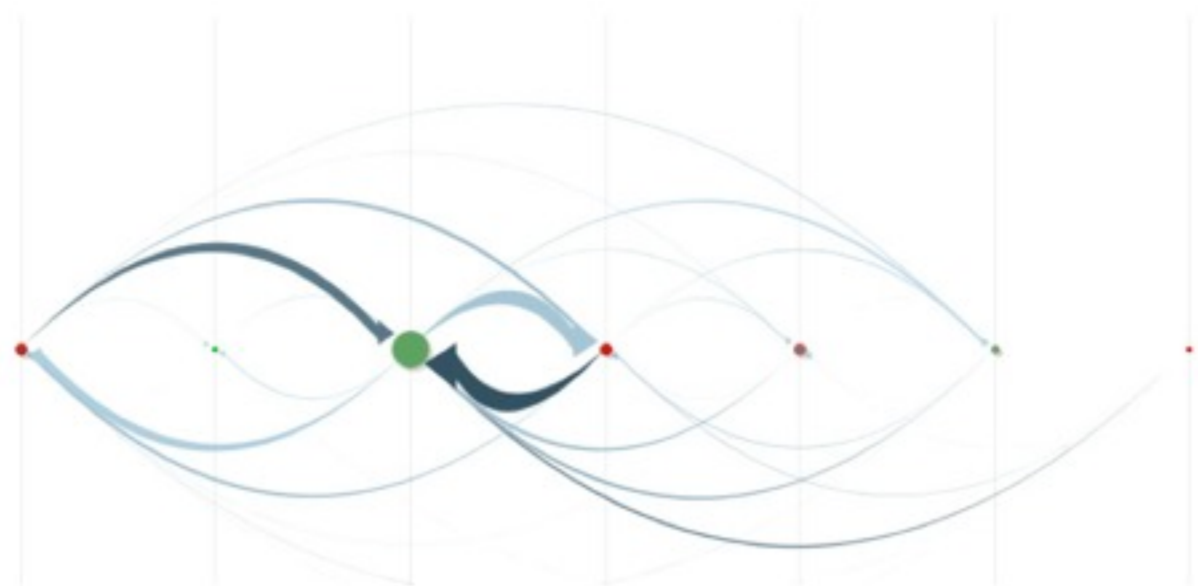
Node and Link Diagram

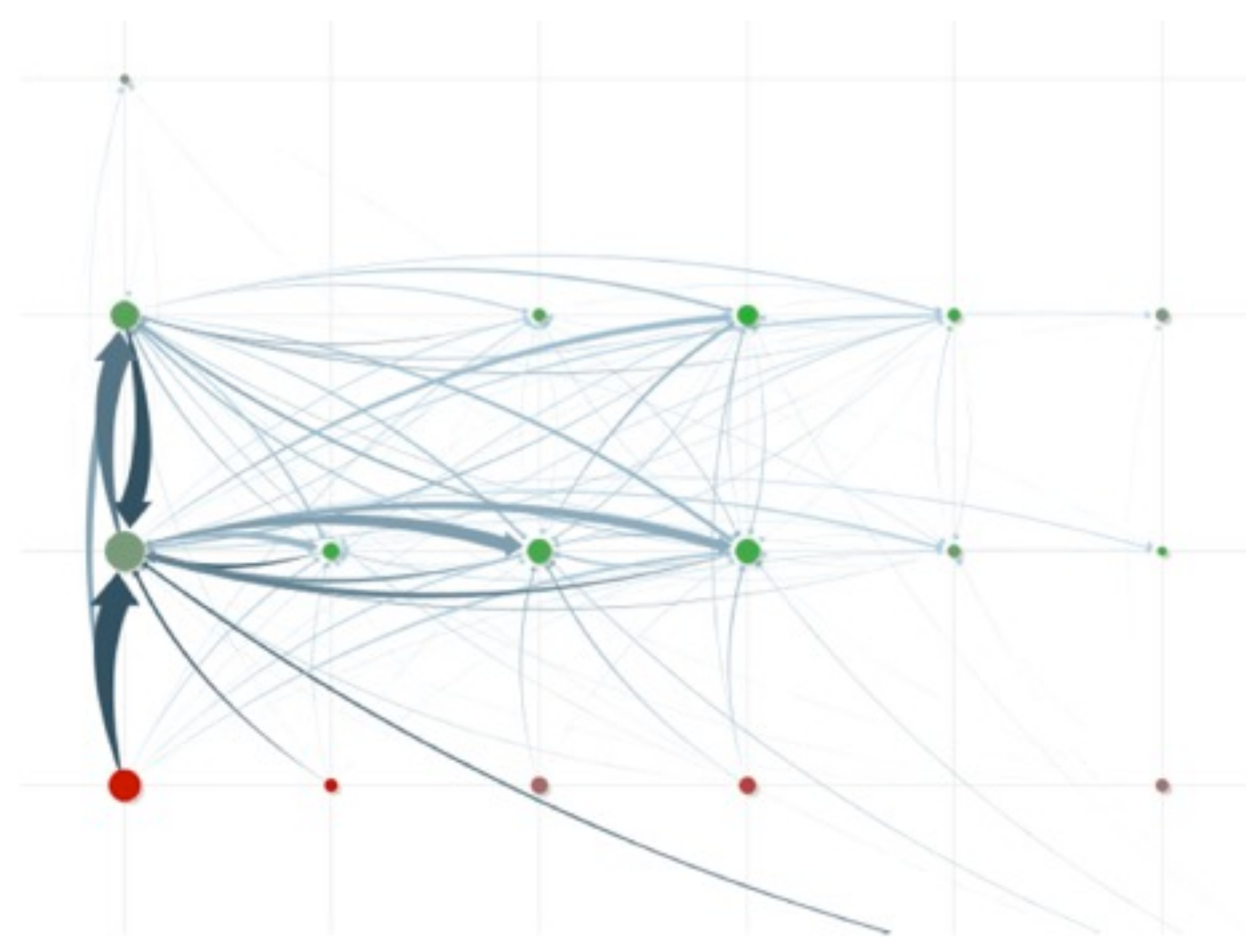


PivotGraph Roll-up

VISUAL ENCODING

- **line for 1D rollup, or grid for 2D case**

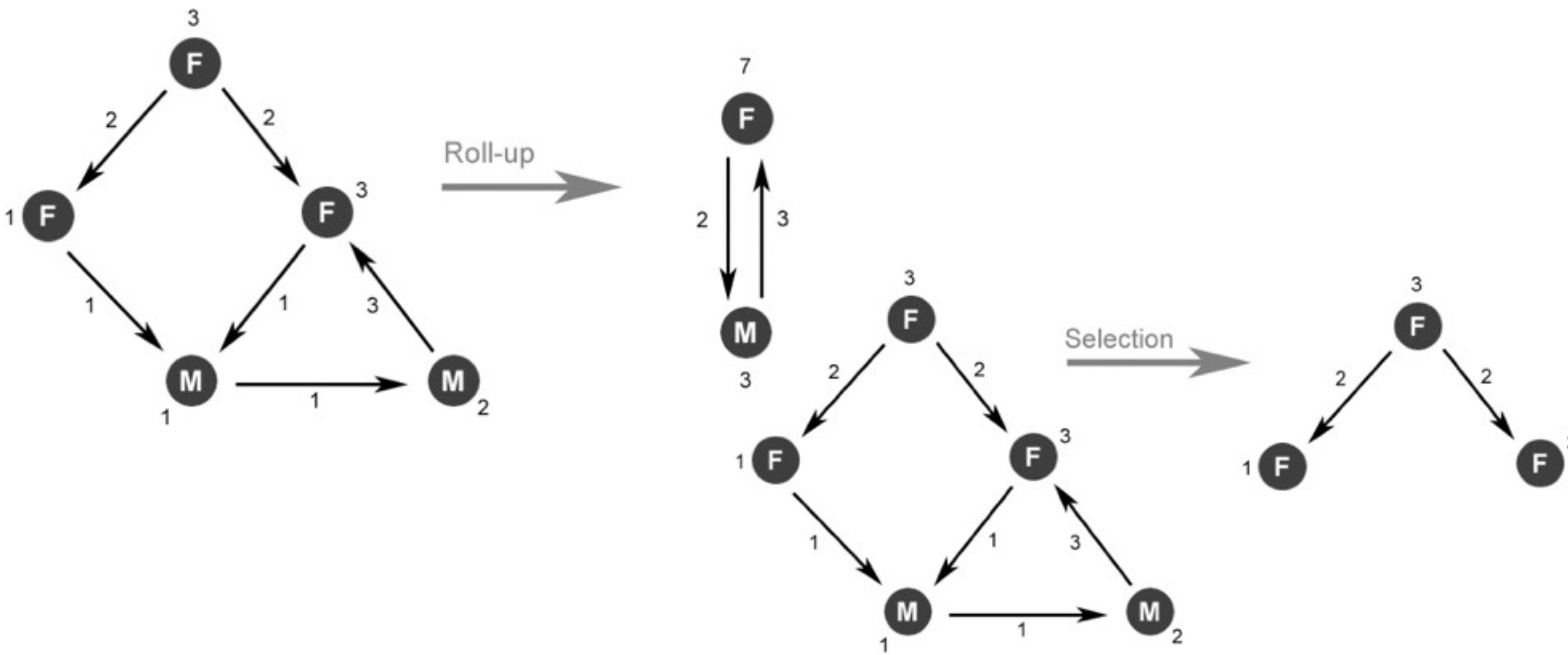




Wattenberg 2006

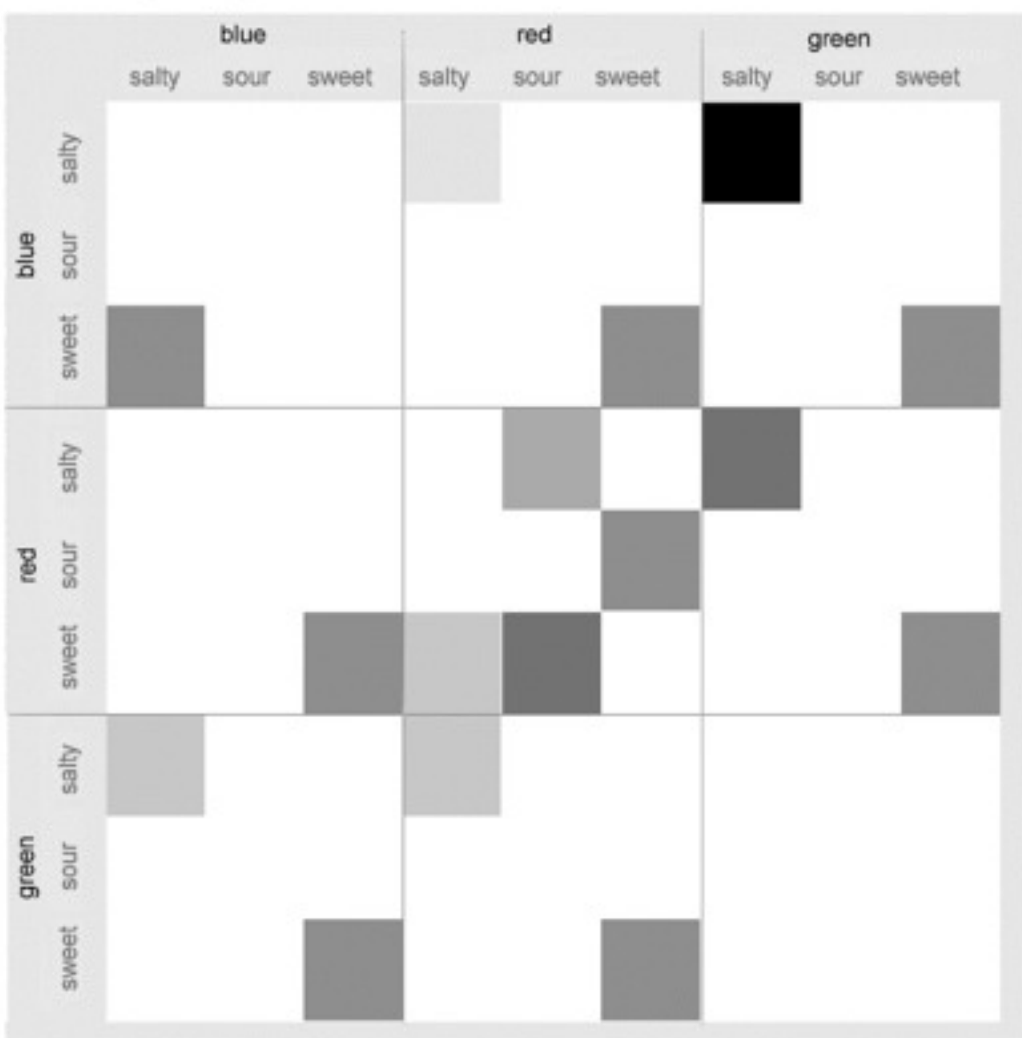
INTERACTION

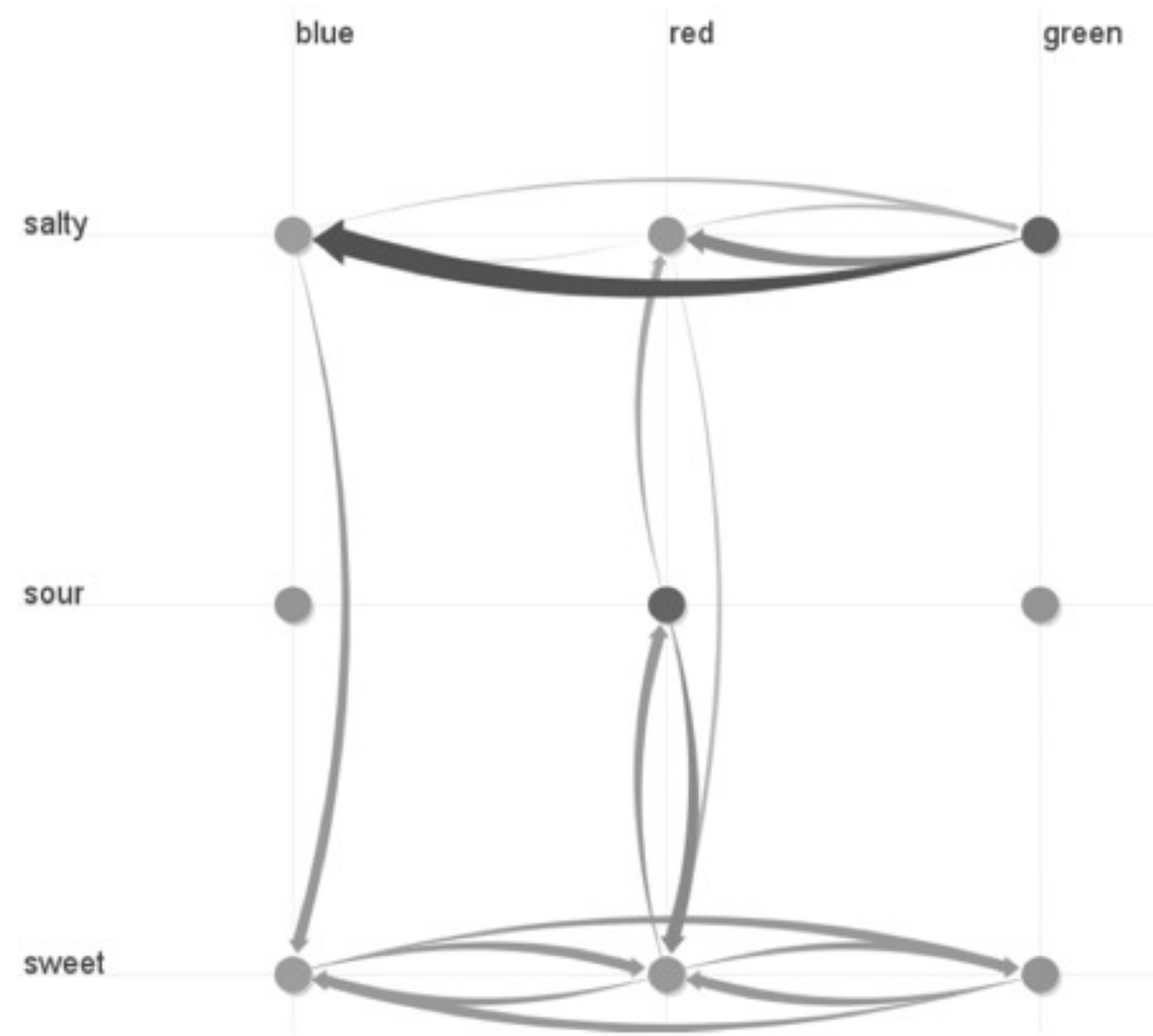
- **changing rollup/selection choices**
- **animated transitions between states**



PIVOT GRAPH

- **in general, more compact than matrix representation**





CRITIQUE: what do you think?

CRITIQUE

- **strengths**
  - focus on multi-attribute graphs
  - abstraction matches focus
  - scales to large graphs
  - nice component of multiview system
- **weaknesses**
  - graph topology not always preserved through rollup and selection
  - only two attributes
  - doesn’t support continuous variables
  - multivariate edges?

RECAP

- **TREES**
  - **indentation**
    - *simple, effective for small trees*
  - **node link** and **layered**
    - *looks good but needs exponential space*
  - **enclosure** (treemaps)
    - *great for size related tasks but suffer in structure related tasks*
- **GRAPHS**
  - **node link**
    - *familiar, but problematic for dense graphs*
  - **adjacency matrices**
    - *abstract, hard to follow paths*
  - **attribute-driven**
    - *not always possible*
- **TAKE HOME MESSAGE: no ultimate solution**

exercise

GRAPH DRAWING EXERCISE

adjacency matrix

	1	2	3	4	5	6	7	8	9	10
1	0	0	1	0	0	1	1	0	0	0
2	0	0	1	0	0	1	0	1	1	0
3	1	1	0	0	0	0	0	0	0	1
4	0	0	0	0	1	0	1	0	1	0
5	0	0	0	1	0	0	0	1	0	0
6	1	1	0	0	0	0	0	0	1	1
7	1	0	0	1	0	0	0	1	0	0
8	0	1	0	0	1	0	1	0	0	0
9	0	1	0	1	0	1	0	0	0	0
10	0	0	1	0	0	1	0	0	0	0

create an aesthetically pleasing
node-link diagram
representation

L16: Maps

REQUIRED READING

The Cartographic Journal *Vol. 45 No. 1 pp. 32–42 February 2008*

REFEREED PAPER

Unfolding the Earth: Myriahedral Projections

Jarke J. van Wijk

Dept. of Mathematics and Computer Science, Technische Universiteit Eindhoven, Eindhoven, The Netherlands
Email: vanwijk@win.tue.nl

Myriahedral projections are a new class of methods for mapping the earth. The globe is projected on a myriahedron, a polyhedron with a very large number of faces. Next, this polyhedron is cut open and unfolded. The resulting maps have a large number of interrupts, but are (almost) conformal and conserve areas. A general approach is presented to decide where to cut the globe, followed by three different types of solution. These follow from the use of meshes based on the standard graticule, the use of recursively subdivided polyhedra and meshes derived from the geography of the earth. A number of examples are presented, including maps for tutorial purposes, optimal foldouts of Platonic solids, and a map of the coastline of the earth.

INTRODUCTION

Mapping the earth is an old and intensively studied problem. For about two thousand years, the challenge to show the round earth on a flat surface has attracted many cartographers, mathematicians, and inventors, and hundreds of solutions have been developed. There are several reasons for this high interest. First of all, the geography of the earth itself is interesting for all its inhabitants. Secondly, there are no perfect solutions possible such that the surface

are obtained by using different myriahedra and choices for the edges to be cut, which are described in three separate sections. The use of graticule-based meshes, recursively subdivided polyhedra, and geographically aligned meshes lead to different maps, each with their own strengths. Finally, the results are discussed.

BACKGROUND

WORLDMAPPER: the world as you've never seen it before

Danny Dorling, Anna Barford and Mark Newman

Abstract—This paper describes the Worldmapper Project, which makes use of novel visualization techniques to represent a broad variety of social and economic data about the countries of the world. The goal of the project is to use the map projections known as cartograms to depict comparisons and relations between different territories, and its execution raises many interesting design challenges that were not all apparent at the outset. We discuss the approaches taken towards these challenges, some of which may have considerably broad application. We conclude by commenting on the positive initial response to the worldmapper images published on the web, which we believe is due, at least in part, to the particular effectiveness of the cartogram as a tool for communicating quantitative geographic data.

Index Terms—Geographic Visualization, Computer Graphics, Worldmapper, Data Visualization, Social Visualization, Cartogram.

1 Introduction

This paper describes the Worldmapper Project, whose aim is to communicate to the widest possible audience, but particularly to school and university students, how parts of the world relate to each other and the implications of these relationships for society. This is achieved by mapping quantitative geographic data using an unusual cartographic projection, with the resulting maps being made freely available on the Internet.

The Worldmapper Project has been made possible by the recent development of a new algorithm to create area cartograms, maps in which the sizes of territories (or other regions) are scaled in proportion to some variable of interest, such as population, income, or energy consumption. This algorithm's development is complemented by the recent release by several United Nations agencies of international data of a much higher quality than has previously been made available, as well as recent geographical theories that suggest places can better be understood globally in relation to one another than when studied separately.

the proportion of the world total of the quantity of interest that belongs to that territory. Thus Figure 1 above, for example, shows the territories of the world with sizes proportional to net emigration from those territories, i.e., to the numbers of people who have left less the numbers who have entered in the fifty years prior to the year 2000. Territories with positive net immigration, such as the United States of America, receive a negative score by this measure, but since it is not possible to give a territory negative area on a map, such territories are given area zero in this case (see 2.1 below for an alternative approach to this issue). The map of net emigration shows from where more people have left than entered in recent years. Simultaneously it also shows proportionally how many more people are involved in the subject being mapped in each place. For example, the large blue territory dominating the top left hand corner of Figure 1 is Mexico, which has the largest net emigration of any territory represented in the map.

1.1 Rationale

World population statistics can be difficult to understand and are usually presented in the form of statistical tables (when they are

Flow Map Layout via Spiral Trees

Kevin Verbeek, Kevin Buchin, and Bettina Speckmann

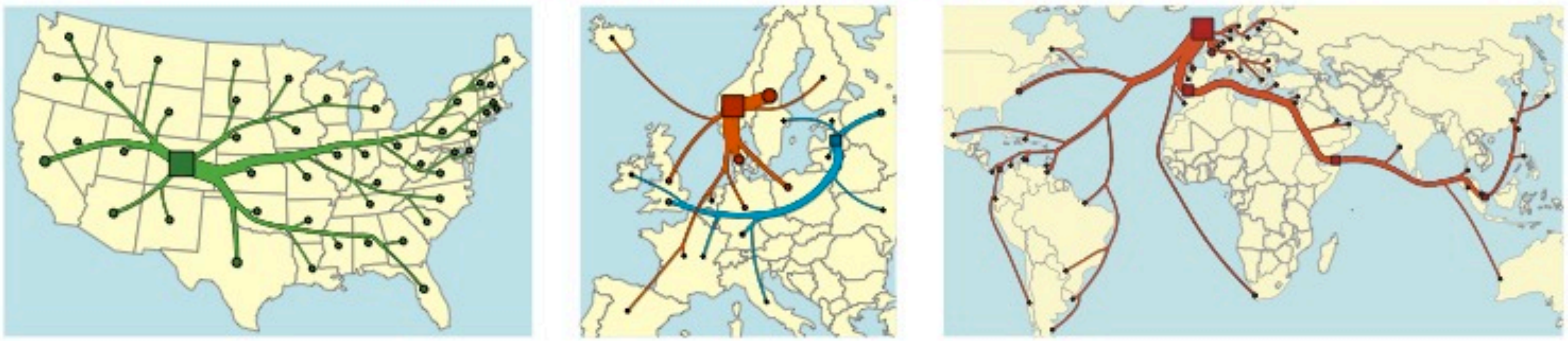


Fig. 1. Flow maps: Migration from Colorado, migration from Norway and Latvia, whisky exports from Scotland.

Abstract—Flow maps are thematic maps that visualize the movement of objects, such as people or goods, between geographic regions. One or more sources are connected to several targets by lines whose thickness corresponds to the amount of flow between a source and a target. Good flow maps reduce visual clutter by merging (bundling) lines smoothly and by avoiding self-intersections. Most flow maps are still drawn by hand and only few automated methods exist. Some of the known algorithms do not support edge-bundling and those that do, cannot guarantee crossing-free flows. We present a new algorithmic method that uses edge-bundling and computes crossing-free flows of high visual quality. Our method is based on so-called *spiral trees*, a novel type of Steiner tree which uses logarithmic spirals. Spiral trees naturally induce a clustering on the targets and smoothly bundle lines. Our flows can also avoid obstacles, such as map features, region outlines, or even the targets. We demonstrate our approach with extensive experiments.

Index Terms—Flow maps, Automated Cartography, Spiral Trees.

1 Introduction

Flow maps are thematic maps that visualize the movement of objects, such as people or goods, between geographic regions [6, 20]. So called *distributive flow maps* are used to depict quantitative data, for example, the amount of wine exported by France or the magnitude of migration between countries. One or more sources are connected to several targets by lines whose thickness corresponds to the amount of flow between a source and a target.

A (distributive) flow map typically consists of one or more *flow trees* which are drawn atop a base map. A flow tree is a single-source flow, that is, it connects a single source (the root) to several targets (the leaves). The widths of the flow lines of a flow tree are scaled proportionally (linearly) to the values they represent. When a flow line (trunk) separates into several smaller lines (branches) the width

Specifically, every flow tree is plane, that is, crossing free. A good flow tree avoids its own nodes, that is, all branches have a minimum length so that their widths can be interpreted. In addition, the main branches of a flow tree grow as straight and smooth as possible, to make it easy to follow them. When a flow map consists of several flow trees then the crossings between different trees should be minimized [6]. Thin branches should be drawn atop thick ones while avoiding a "weaving" effect. Furthermore, flow maps often try to avoid covering important map features with flow trees to aid recognizability.

Flow maps that depict trade often route edges along actual shipping routes. In that case a moderate distortion of the underlying geography is admissible. In contrast, flow maps that depict abstract data or data which is not linked to specific routes, such as migration or internet