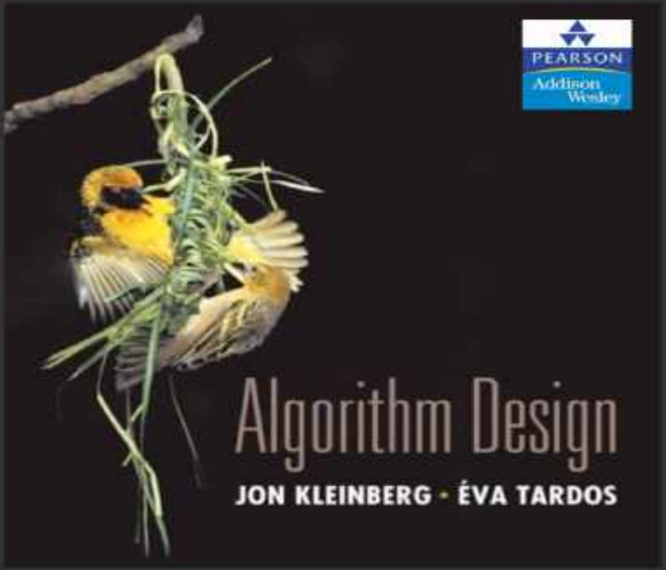
Graphs: Concepts, Representation and Examples

Jussara M Almeida



Lecture slides by Kevin Wayne

http://www.cs.princeton.edu/~wayne/kleinberg-tardos

3. Graphs

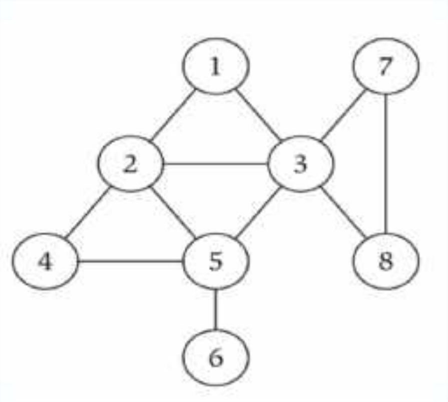
- *basic definitions and applications*
- *graph connectivity and graph traversal*
- *testing bipartiteness*
- *connectivity in directed graphs*
- *DAGs and topological ordering*

Last updated on 7/29/17 6:09 PM

Undirected graphs

Notation. $G = (V, E)$

- V = nodes (or vertices).
- E = edges (or arcs) between pairs of nodes.
- Captures pairwise relationship between objects.
- Graph size parameters: $n = |V|, m = |E|$.

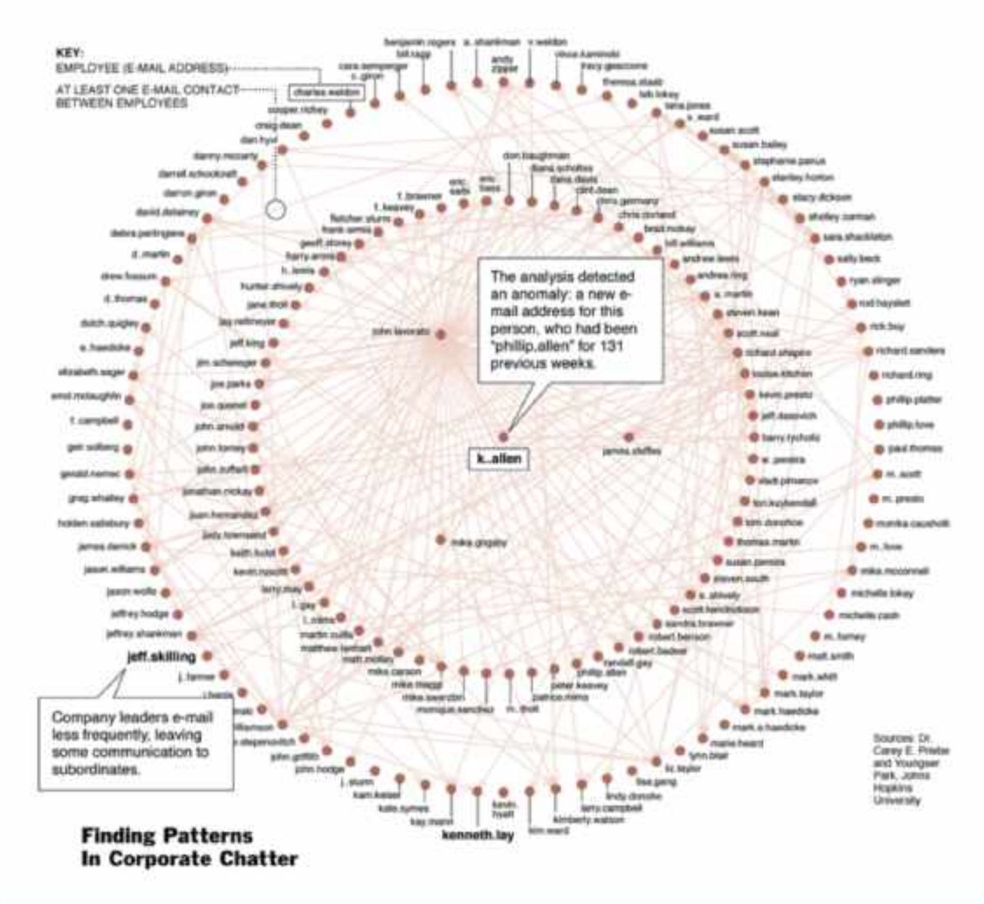


$V = \{ 1, 2, 3, 4, 5, 6, 7, 8 \}$

$E = \{ 1\text{–}2, 1\text{–}3, 2\text{–}3, 2\text{–}4, 2\text{–}5, 3\text{–}5, 3\text{–}7, 3\text{–}8, 4\text{–}5, 5\text{–}6, 7\text{–}8 \}$

$m = 11, n = 8$

One week of Enron emails



Framingham heart study

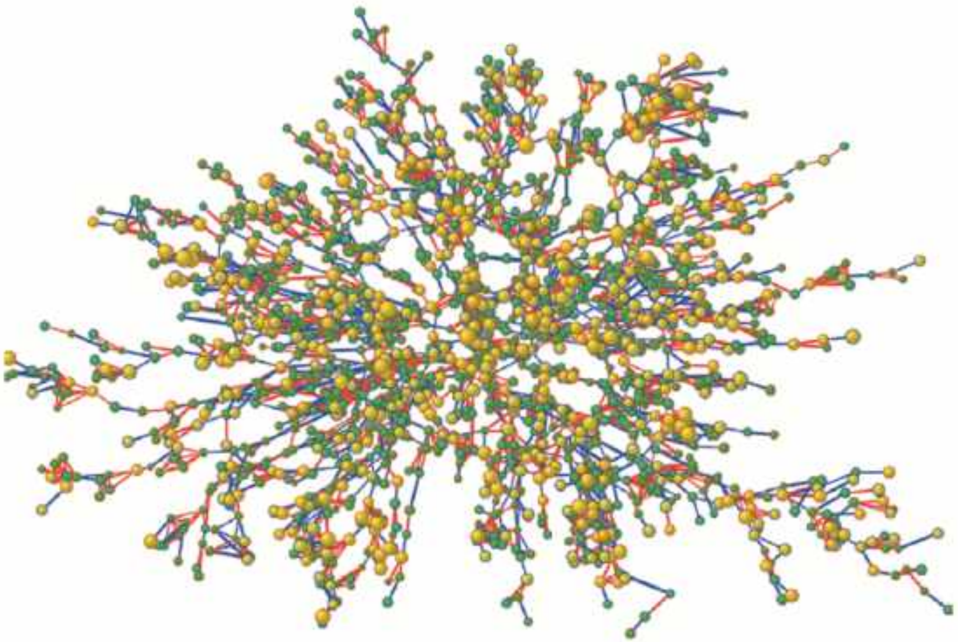


Figure 1. **Largest Connected Subcomponent of the Social Network in the Framingham Heart Study in the Year 2000.**

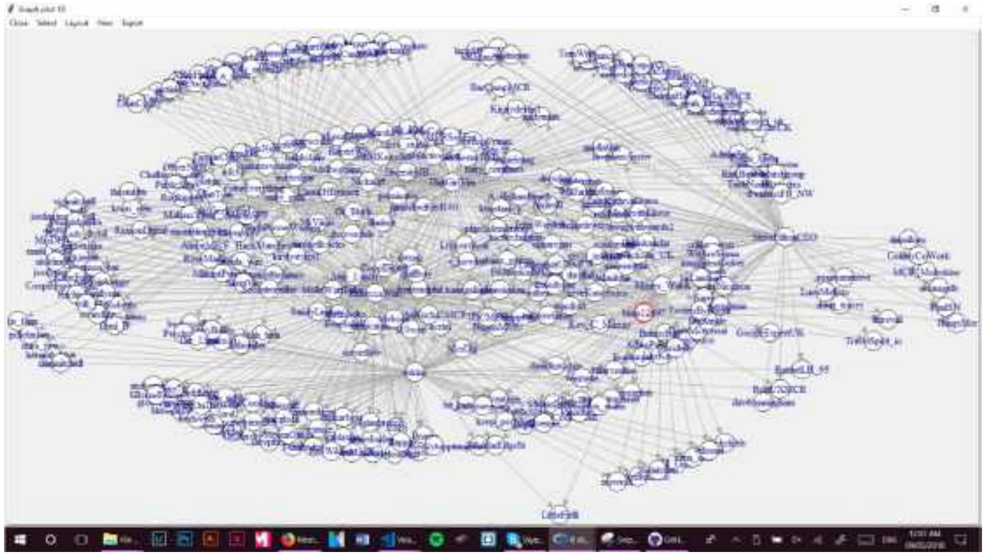
Each circle (node) represents one person in the data set. There are 2200 persons in this subcomponent of the social network. Circles with red borders denote women, and circles with blue borders denote men. The size of each circle is proportional to the person's body-mass index. The interior color of the circles indicates the person's obesity status: yellow denotes an obese person (body-mass index, ≥30) and green denotes a nonobese person. The colors of the ties between the nodes indicate the relationship between them: purple denotes a friendship or marital tie and orange denotes a familial tie.

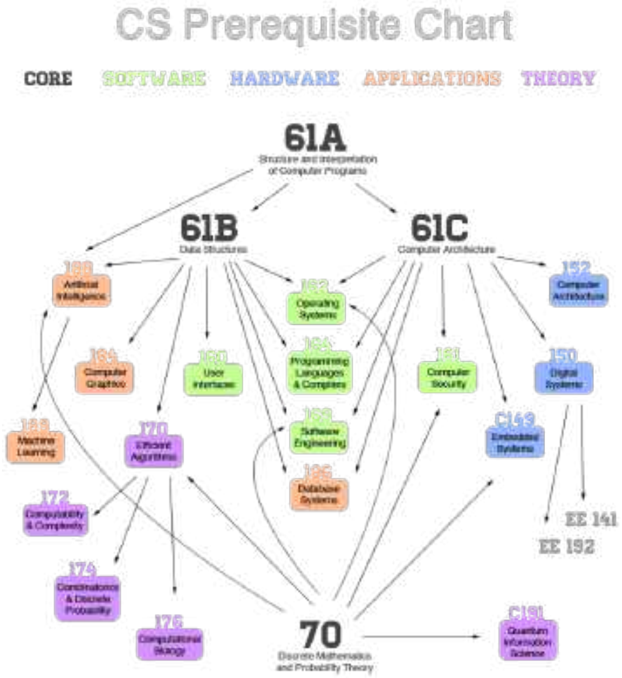
"The Spread of Obesity in a Large Social Network over 32 Years" by Christakis and Fowler in New England Journal of Medicine, 2007

Directed graph

Directed graph: Asymetric relationship between two nodes; edges have direction

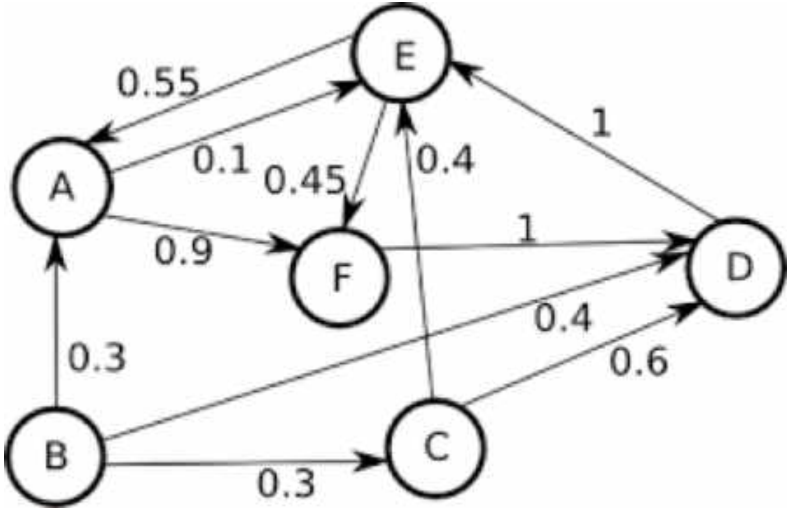
Direction may capture dependencies





Weighted graph

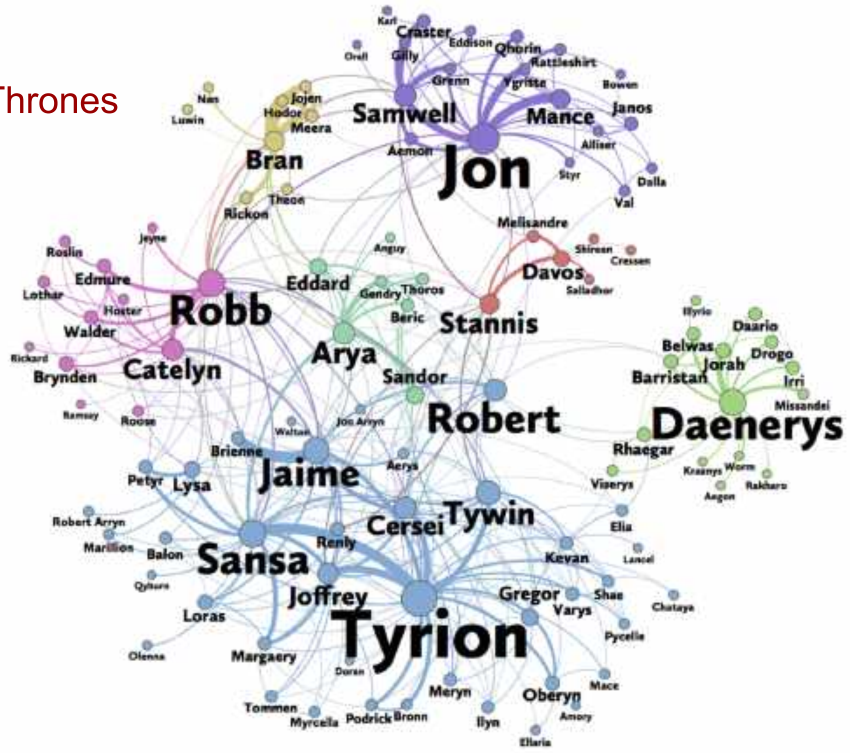
Weights assigned to edges capture strength of relationship.



Weighted graph

Weights may be assigned to nodes

Network of Thrones



Node: character

Edge: two characters mentioned within 15 words in book

Size of node: importance (PageRank)

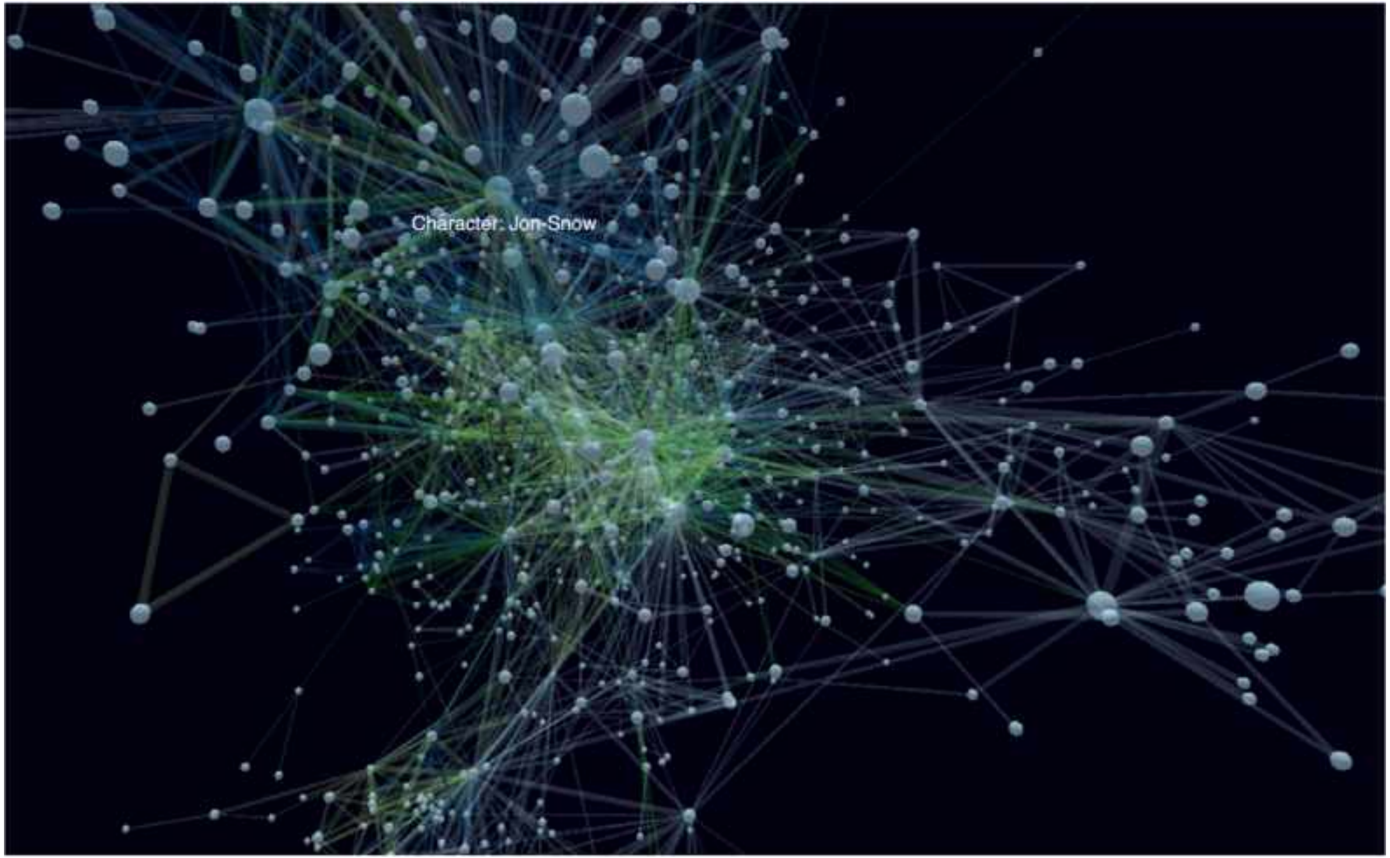
Size of label: importance (betweenness)

Node color: community

Source: A. Beveridge and J. Shan, "Network of Thrones," *Math Horizons Magazine* , Vol. 23, No. 4 (2016), pp. 18-22

Weighted graph

Network of Thrones



Source: A. Beveridge and J. Shan, "Network of Thrones," *Math Horizons Magazine* , Vol. 23, No. 4 (2016), pp. 18-22

Some graph applications

graph	node	edge
communication	telephone, computer	fiber optic cable
circuit	gate, register, processor	wire
mechanical	joint	rod, beam, spring
financial	stock, currency	transactions
transportation	street intersection, airport	highway, airway route
internet	class C network	connection
game	board position	legal move
social relationship	person, actor	friendship, movie cast
neural network	neuron	synapse
protein network	protein	protein-protein interaction
molecule	atom	bond

Graph traversal

- One of the most fundamental graph operations is that of traversing a sequence of nodes connected by edges
  - Ex: user browsing Web pages
  - Rumor passing by word-of-mouth
  - Airline passenger with multiple connections

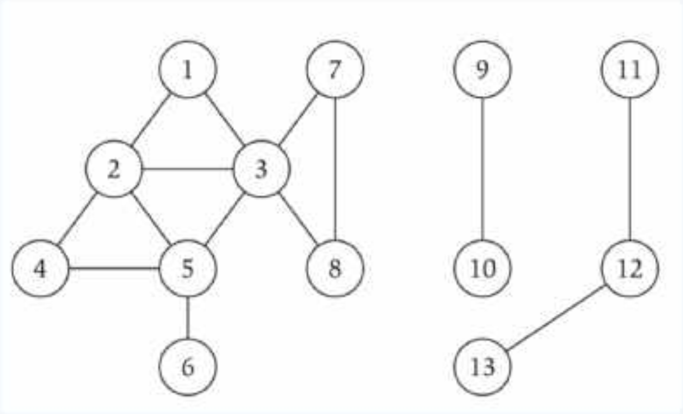
Transportation networks: Small number of hubs ; short paths between cities

Paths and connectivity

Def. A path in an undirected graph $G = (V, E)$ is a sequence of nodes $v_1, v_2, \ldots, v_k$ with the property that each consecutive pair v_{i-1}, v_i is joined by an edge in E.

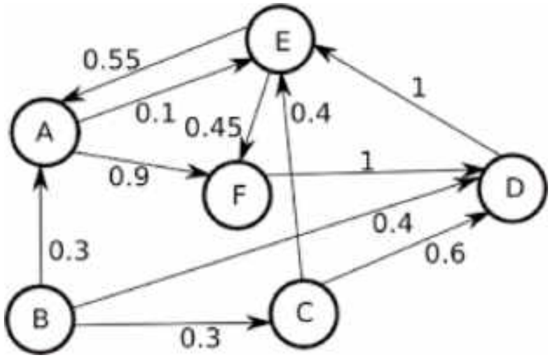
Def. A path is simple if all nodes are distinct.

Def. An undirected graph is connected if for every pair of nodes u and v, there is a path between u and v.



Connectivity in directed graphs

In a directed path, it is possible for u to have a path to v, while v has no path to u

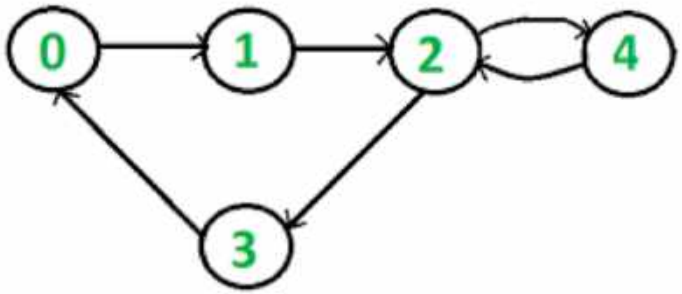


There is a path from B to every other node;
There is no path from any node back to B

Connectivity in directed graphs

A directed graph is strongly connected if, for every two nodes u and v,

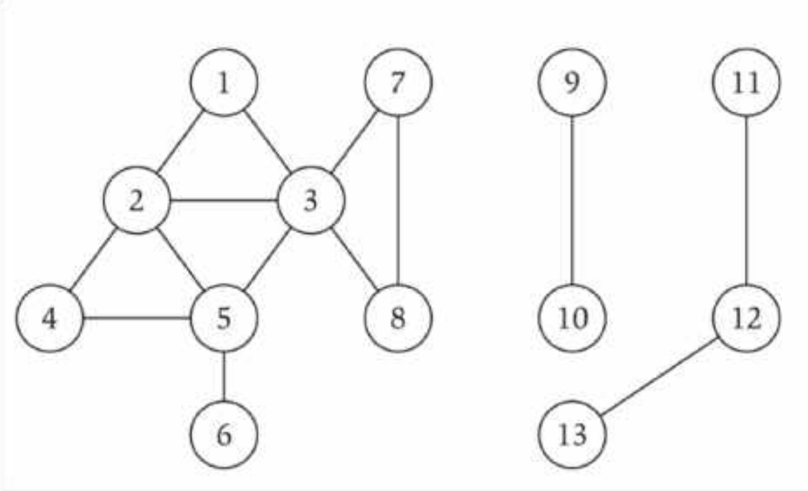
- there is a path from u to v and
- there is a path from v to u



Strongly Connected

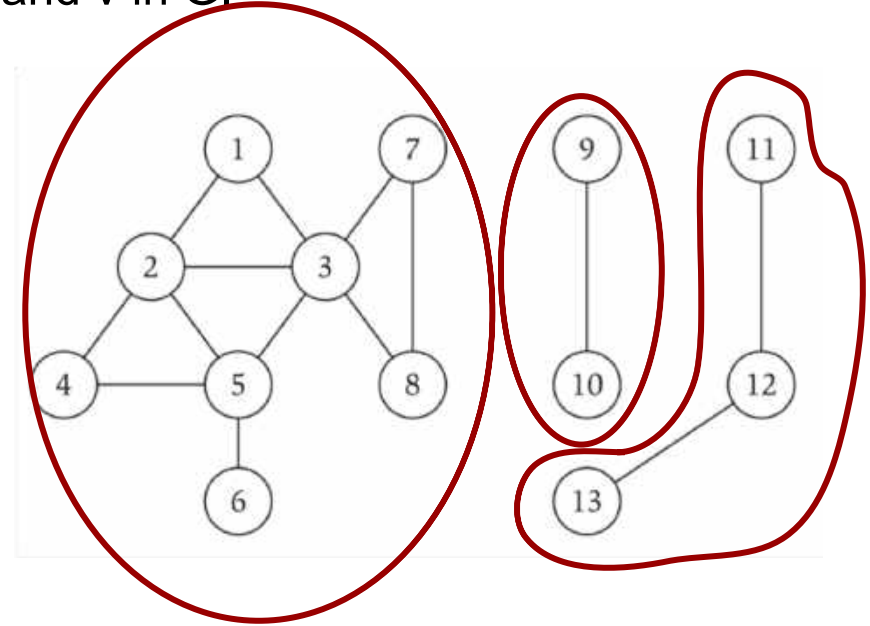
Connected components

- A connected component of an undirected graph G = (V, E) is a subgraph G' = (V', E'), where $V' \subset V$ and $E' \subset E$ such that, for every $u, v \in V'$, there is a path between u and v in G.

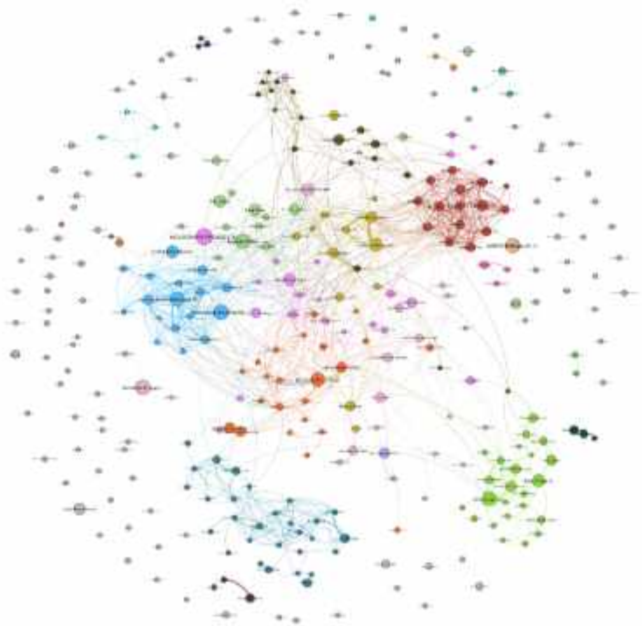


Connected components

- A connected component of an undirected graph G = (V, E) is a subgraph G' = (V', E'), where V' ⊂ V and E' ⊂ E such that, for every u, v ∈ V', there is a path between u and v in G.



Connected components



(b) Election Campaign Network

Figure 4: Group networks (nodes are groups and edges connect groups with users in common).

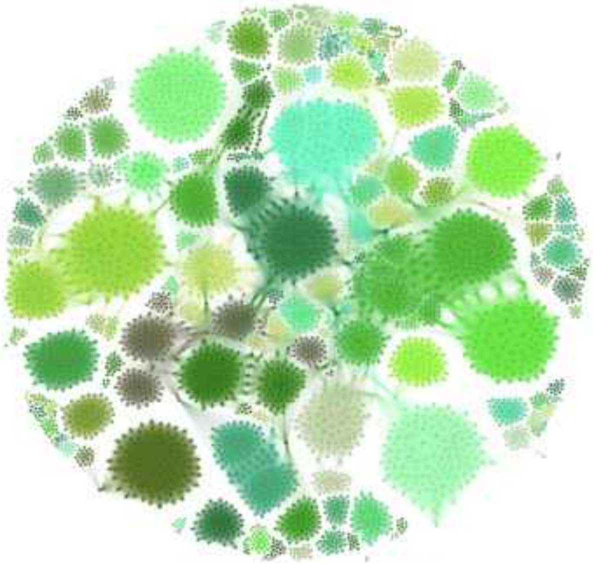
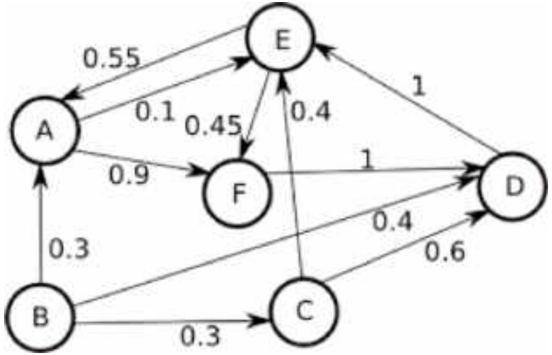


Figure 5: User network (nodes are users and edges connect users with group in common. Subgraph of election period).

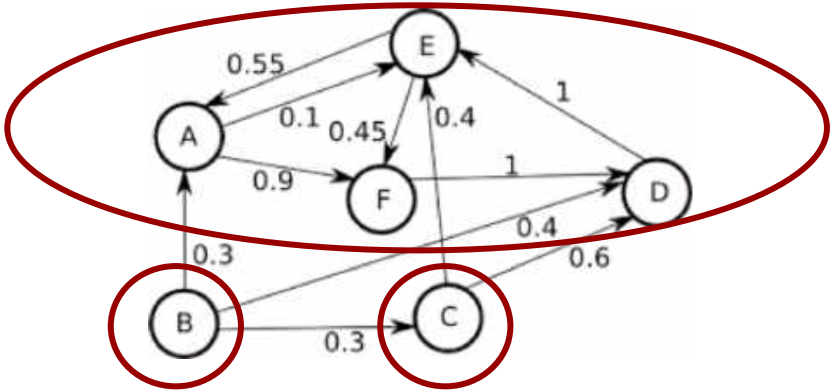
Strongly connected components

- A strongly connected component of a directed graph G = (V, E) is a subgraph G' = (V', E'), where V' ⊂ V and E' ⊂ E such that, for every u, v ∈ V', there are paths between u and v and between v and u in G.



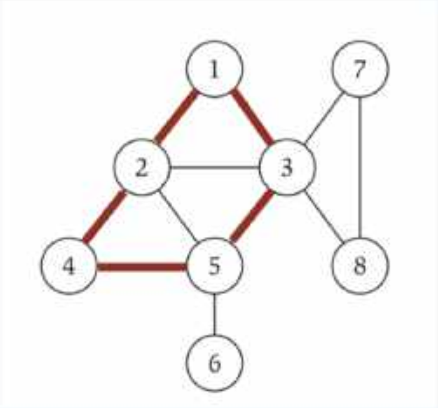
Strongly connected components

- A strongly connected component of a directed graph G = (V, E) is a subgraph G' = (V', E'), where V' ⊂ V and E' ⊂ E such that, for every u, v ∈ V', there are paths between u and v and between v and u in G.



Cycles

Def. A cycle is a path v_1, v_2, …, v_k in which $v_1 = v_k$, $k > 2$, and the first $k - 1$ nodes are all distinct.



cycle C = 1-2-4-5-3-1

Trees

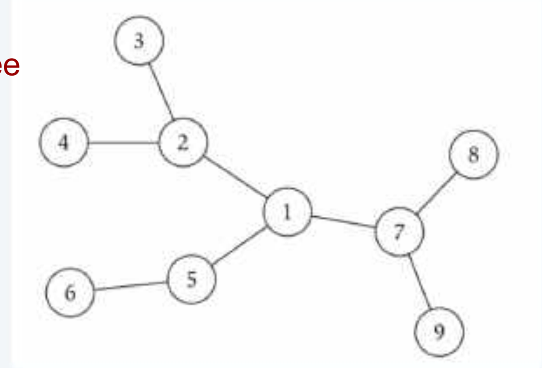
Def. An undirected graph is a tree if it is connected and does not contain a cycle.

Theorem. Let G be an undirected graph on n nodes. Any two of the following statements imply the third:

- G is connected.
- G does not contain a cycle.
- G has $n-1$ edges.

Deleting any edge from a tree will disconnect it!

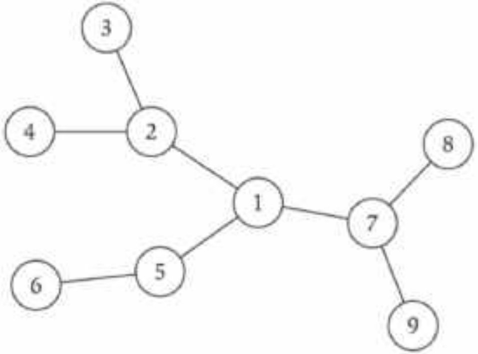
Any tree with n nodes has exactly n-1 edges



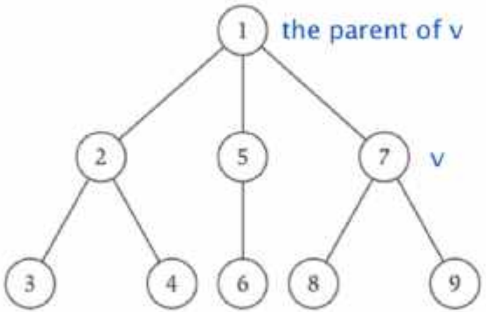
Rooted trees

Given a tree T, choose a root node r and orient each edge away from r.

Importance. Models hierarchical structure.



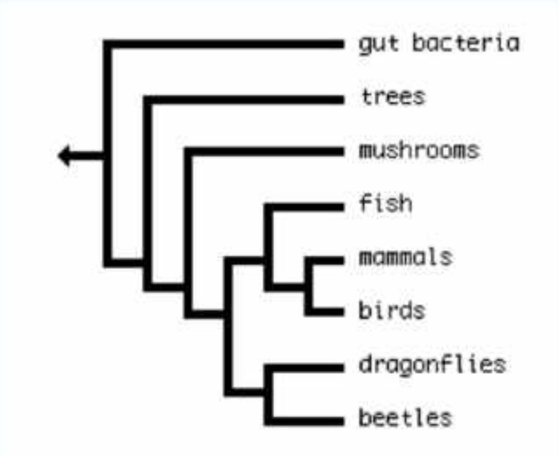
a tree



the same tree, rooted at 1

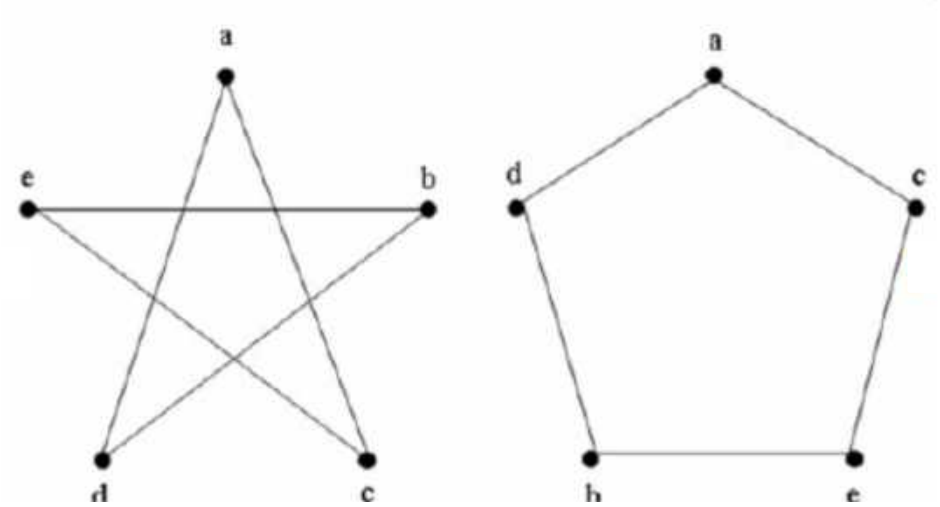
Phylogeny trees

Describe evolutionary history of species.



Isomorphic graphs

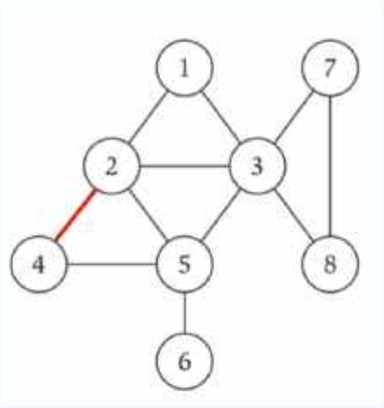
- Two graphs that contain the same number of vertices connected in the same way are said to be isomorphic
- Formally: two graphs G and H, both with vertices V = {1, 2, …, n}, are isomorphic if there exists a permutation p of V such that
  (u,v) is an edge in G if and only if (p(u), p(v)) is an edge in H



Graph representation: adjacency matrix

Adjacency matrix. n-by-n matrix with $A_{uv} = 1$ if (u, v) is an edge.

- Two representations of each edge.
- Space proportional to n^2.
- Checking if (u, v) is an edge takes $\Theta(1)$ time.
- Identifying all edges takes $\Theta(n^2)$ time.



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

Graph representation: adjacency lists

Adjacency lists. Node-indexed array of lists.

- Two representations of each edge.
- Space is $\Theta(m + n)$.
- Checking if (u, v) is an edge takes $O(degree(u))$ time.
- Identifying all edges takes $\Theta(m + n)$ time.

degree = number of neighbors of u