Graphs

- Introduction to graphs
- NetworkX
- Basic operations and problems
  - Webcrawler using a graph and not a matrix
- Drawing large graphs

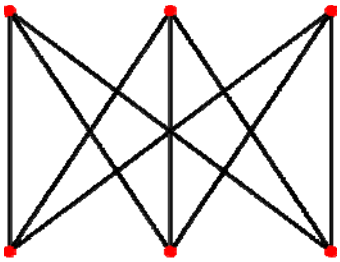
References

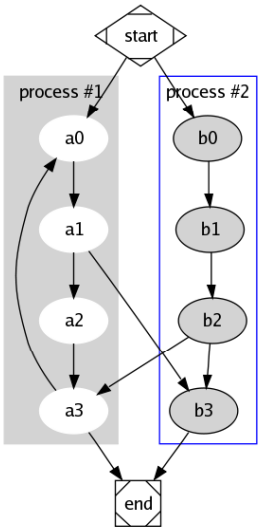
- http://networkx.lanl.gov/tutorial/index.html - tutorial
- http://people.hofstra.edu/geotrans/eng/ch2en/meth2en/ch2m1en.html - gives formal definitions
- http://en.wikipedia.org/wiki/Graph_theory - history and links

Graphs

- Mathematical model for studying pair-wise connectivity among entities
- A graph is a discrete structure consisting of
  - nodes (also called vertices, sites) and
  - edges between nodes (also called links, connectors, arcs)
- Graphs are used to model problems in almost every discipline

Two graphs





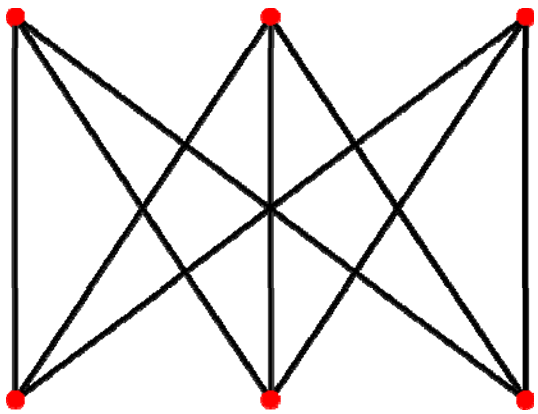
3 Monday, March 30, 2009

	NODES	EDGES
Communication network	computers	fiber optic cable
Financial stocks	currency	transactions
Transportation	street intersections	highways
Protein interaction networks	proteins	protein-protein interactions
Internet	web pages	hyperlinks
Social networks	people	friendships
Software systems	functions	function calls
Games board	positions	legal moves
Chemical compounds	molecules	chemical bonds

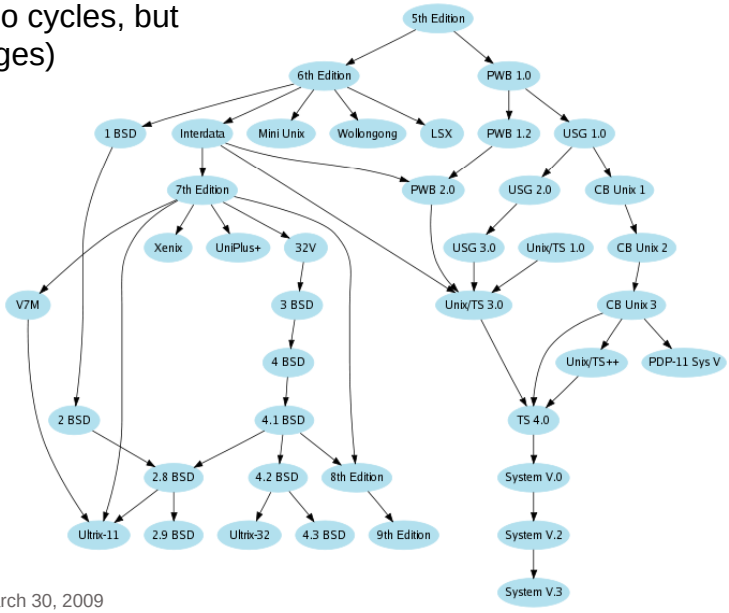
4 Monday, March 30, 2009

	NODES	EDGES
Communication network	computers	fiber optic cable
Financial stocks	currency	transactions
Transportation	street intersections	highways
Protein interaction networks	**proteins**	**protein-protein interactions**
Internet	web pages	hyperlinks
Social networks	people	friendships
Software systems	functions	function calls
Games board	positions	legal moves
Chemical compounds	molecules	chemical bonds

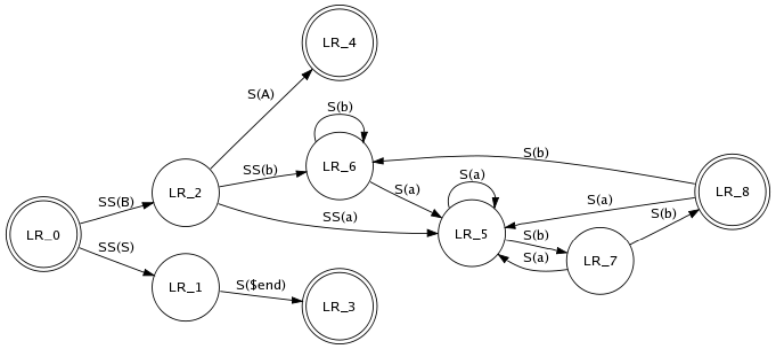
Graph consisting of 6 nodes and 9 edges
Edges have no directions,



Edges have directions
(there are no cycles, but
shortcut edges)



Directed graph with cycles and self-loops
(edges have labels, nodes have labels)



Clicker Question:
Which of the following is not naturally modeled by a graph?

A. **Facebook users and their friends**
B. **Spread of a disease**
C. **Weather prediction**
D. **Course requirements and prerequisites**
E. **Electric power distribution system**

Some terminology

- Graphs have a finite number of nodes, but they can grow and shrink
- Undirected and directed graphs
- Weights on edges: weighted and unweighted graphs
- Nodes generally have names
- At most one edge between a pair of nodes: simple graph (otherwise a multi-graph)
- self-loops are often excluded as possible edges

More terminology

- A tree is a special type of graph
- Undirected graph
  - a node is "adjacent" to other nodes
  - these nodes are also called its neighbors
  - A tree has n-1 edges and connects all nodes
- Directed graph
  - a node has edges going out and edges coming in (in-degree and out-degree of a node)
  - adjacency is often the number of edges leaving the node
- Path between two nodes – sequence of edges connecting them