Phys 597A, CMPS 497E
Graphs and Networks in Systems Biology

Lecturer: Réka Albert

ralbert@phys.psu.edu

122 Davey Laboratory

Networks, networks everywhere

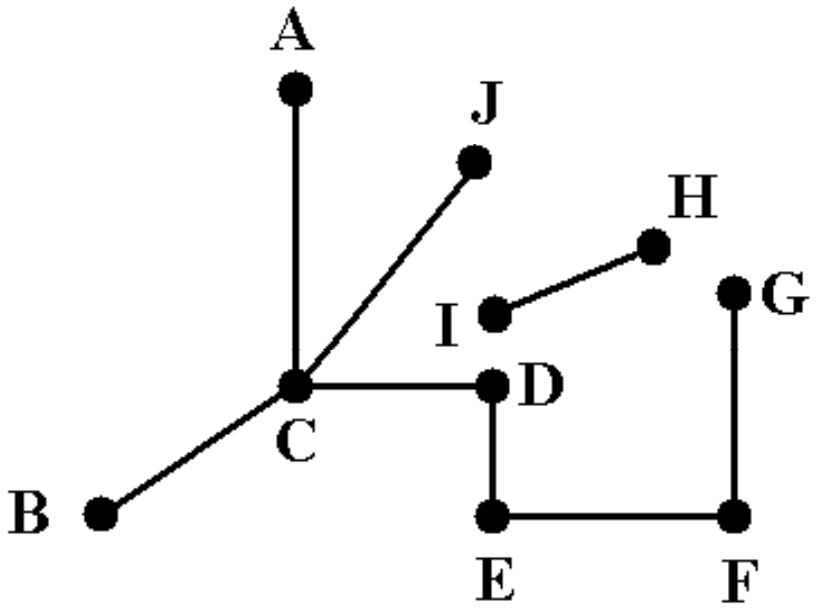
- Network infrastructure, social networking
- Network - a tool for understanding complex systems
- Many non-identical elements connected by diverse interactions
- E.g. interaction networks within cells: protein interactions, chemical reactions, gene regulation
- Graph measures provide information on interaction graphs
- Network models explain and predict properties of graph classes
- Network topology influences network robustness and the dynamics of flows
- E.g. dynamics of molecular interaction networks determines the behavior of cells.
- Understand emergent properties – synchronization, phase transitions, homeostasis

Suggested reading on networks

1. A.-L. Barabási, Linked: The new science of networks.
2. D. J. Watts, Six degrees: The science of a connected age.
3. M. Newman, A.-L. Barabási, D. J. Watts (eds.), The Structure and Dynamics of Networks.
4. G. Caldarelli, Scale-Free Networks: Complex webs in nature and technology.
5. F. Chung, L. Lu, Complex graphs and Networks.
6. R. Pastor-Satorras, A. Vespignani, Evolution and Structure of the Internet: A Statistical Physics Approach.

7. Center for Complex Network Research webpage http://www.barabasilab.com/
8. 24 July 2009 Science special section on Complex Systems and Networks.

Definition of graphs (networks)



Network (graph): a set of nodes connected by edges

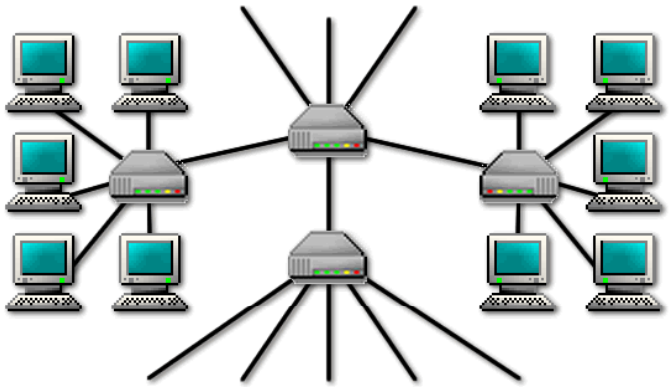
Nodes (vertices): A, B, C…

Edges (links): AC, BC, CD, CJ …

The spatial arrangement of nodes and edges does not matter.

Can be augmented by additional node and edge information.

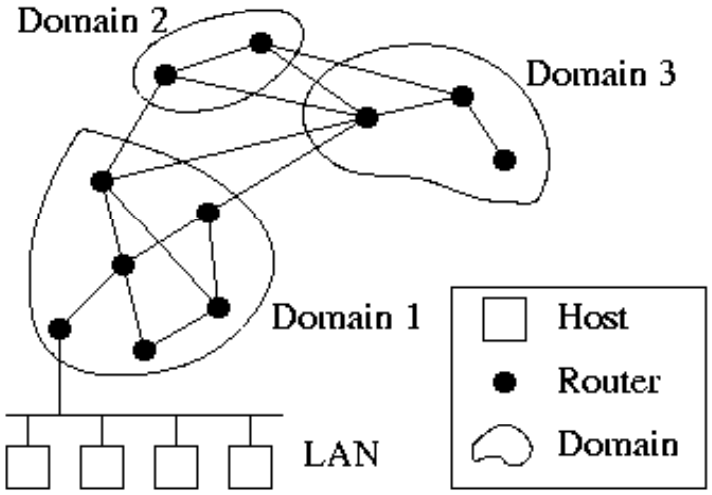
Many complex systems have an underlying network topology



Internet, router level

- nodes: routers, hosts
- edges: wires, cables, wireless

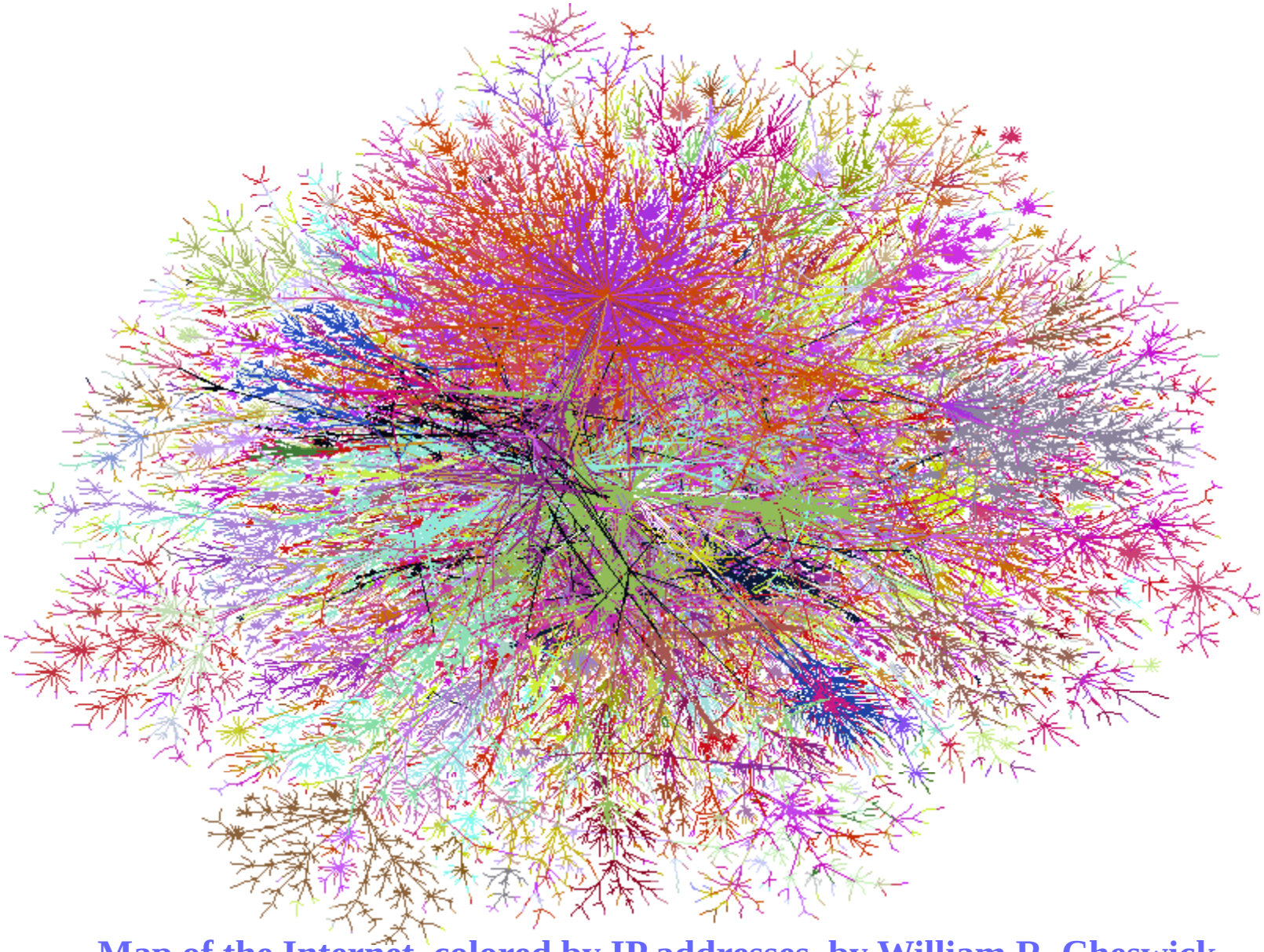
Q: Which edges are static and which change?



Internet, domain level

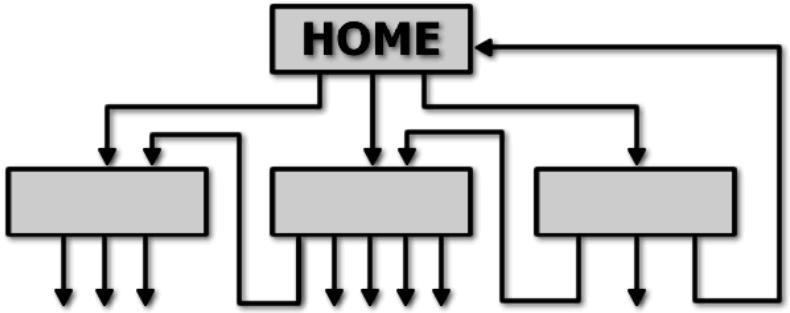
- nodes: domains (ISPs)
- edges: gateway protocols
- Undirected

Q: What is the nature of edges?



Map of the Internet, colored by IP addresses, by William R. Cheswick

The World Wide Web is the higher level of the Internet



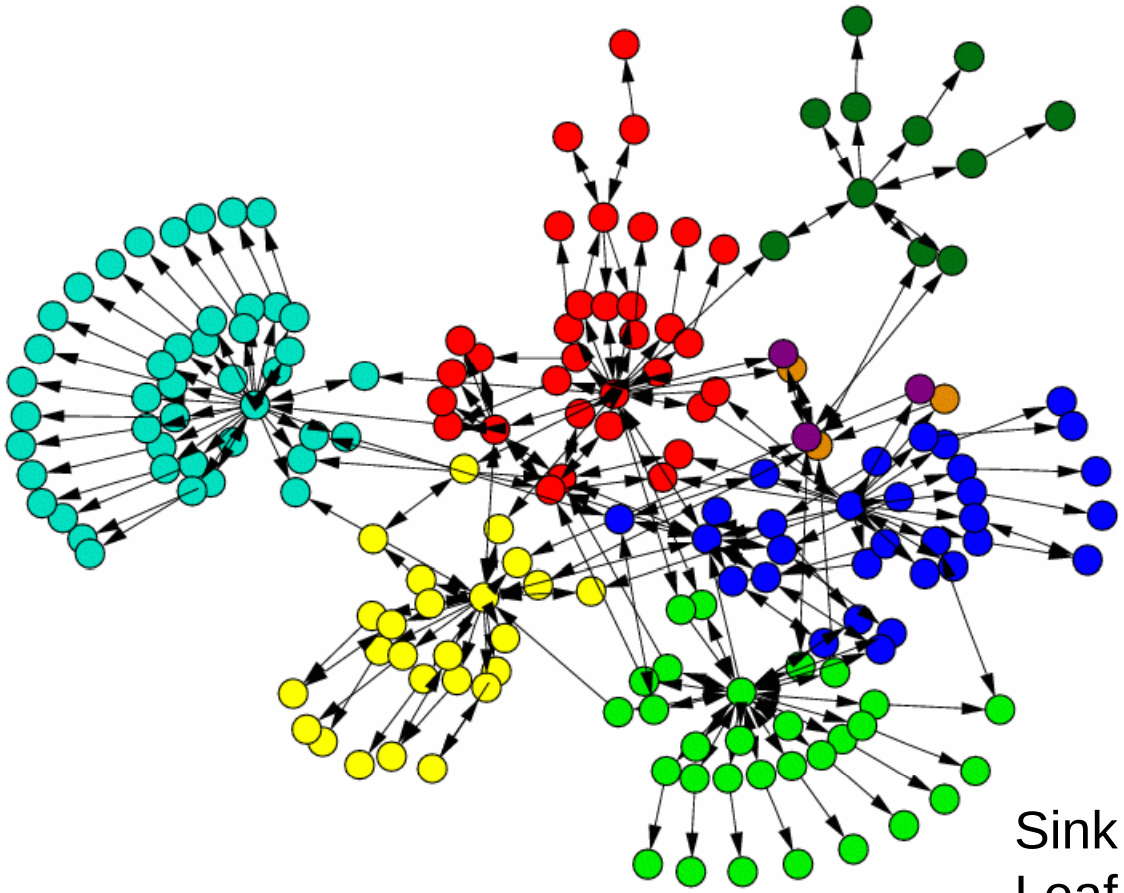
- nodes: webpages
- edges: hyperlinks - directed

The WWW is the largest network with topological information available.

The size of the WWW has surpassed 30 billion nodes, it is increasing.

Search engines can index only a fraction of the Web.

Structure of a website

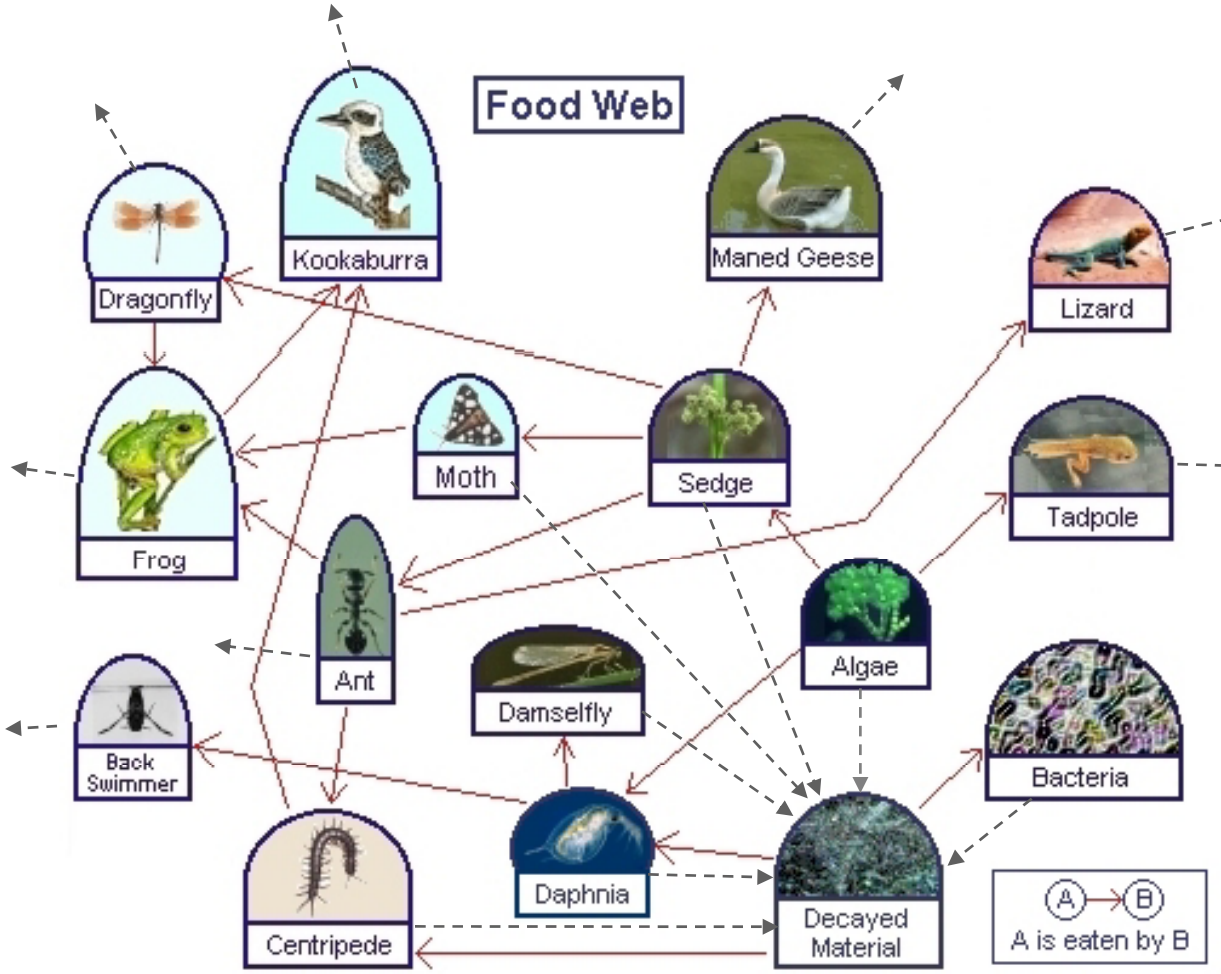


Color: inferred node clustering

Sink nodes, Leaf nodes

M. E. J. Newman and M. Girvan, *Phys. Rev. E* (2004)

Food webs describe the energy flow within species



Nodes: species

Edges: nutrient flow

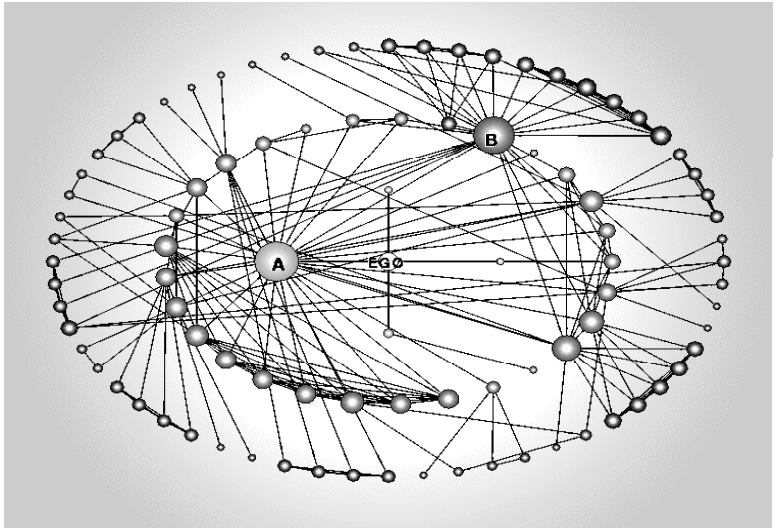
predator-prey relationship →

decay - - - ▸

Directed edges

Q: In what ways is this food web different from a food chain?

Social systems can be regarded as networks



- nodes: individuals
- edges: social interaction
- “six degrees of separation”:

  the social distance between people is small

actor collaboration

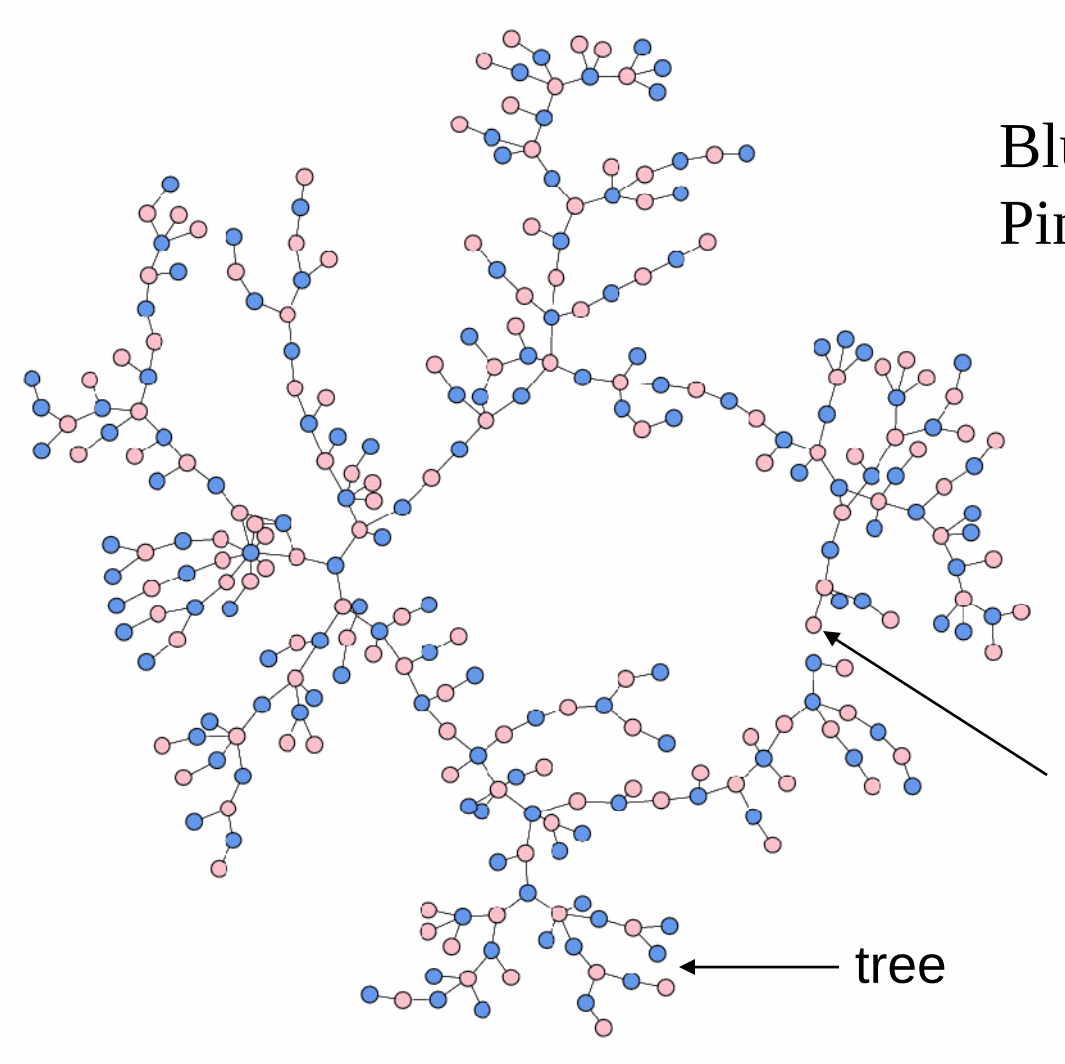
- nodes: actors
- edges: cast jointly

scientific coauthorship

- nodes: scientists
- edges: wrote a paper

Q: Can you propose an alternative network based on actor/movie or author/paper information?

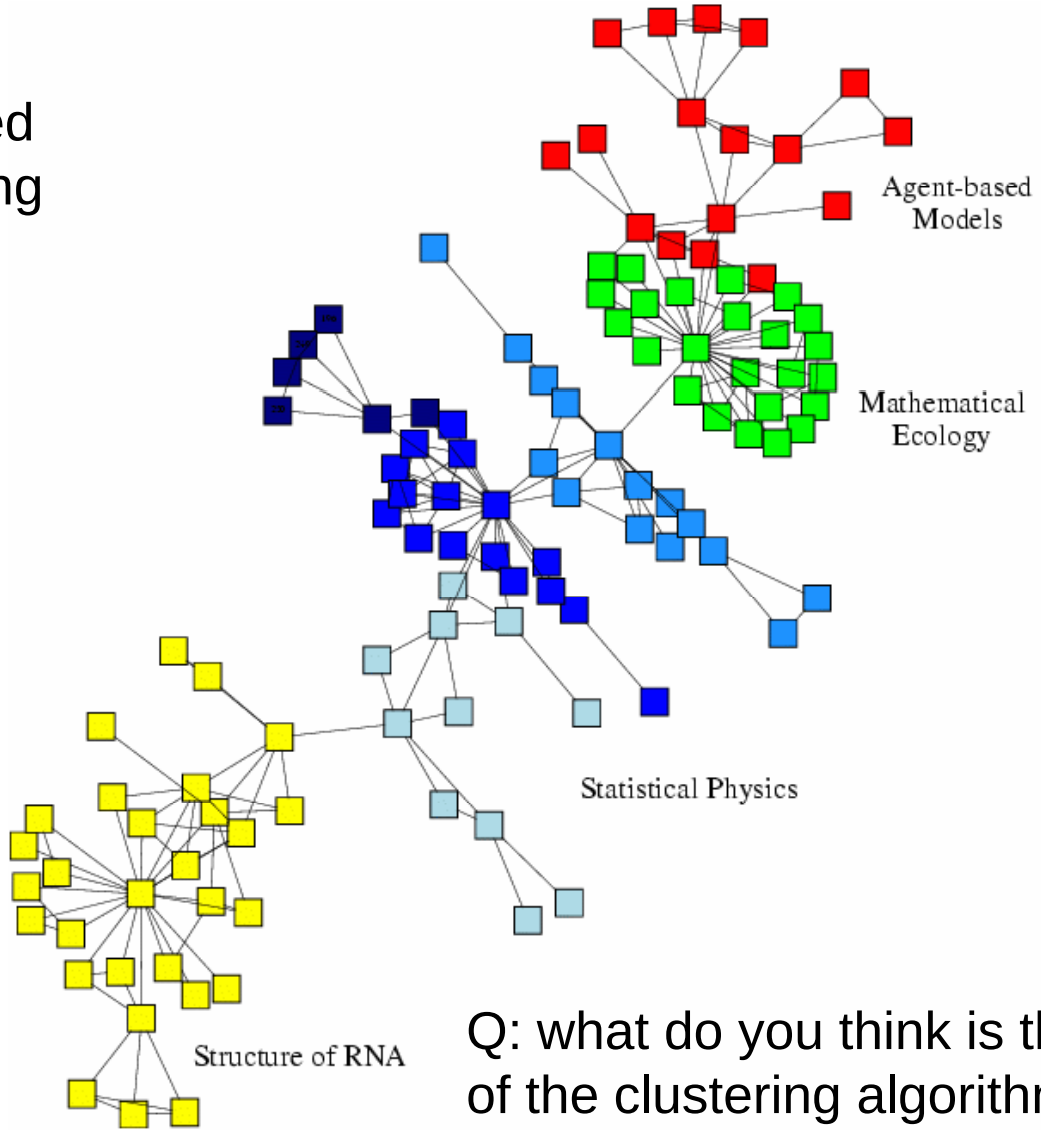
Dating network in a high-school



Q: does it surprise you that the network is connected?

Collaborations at the Santa Fe Institute

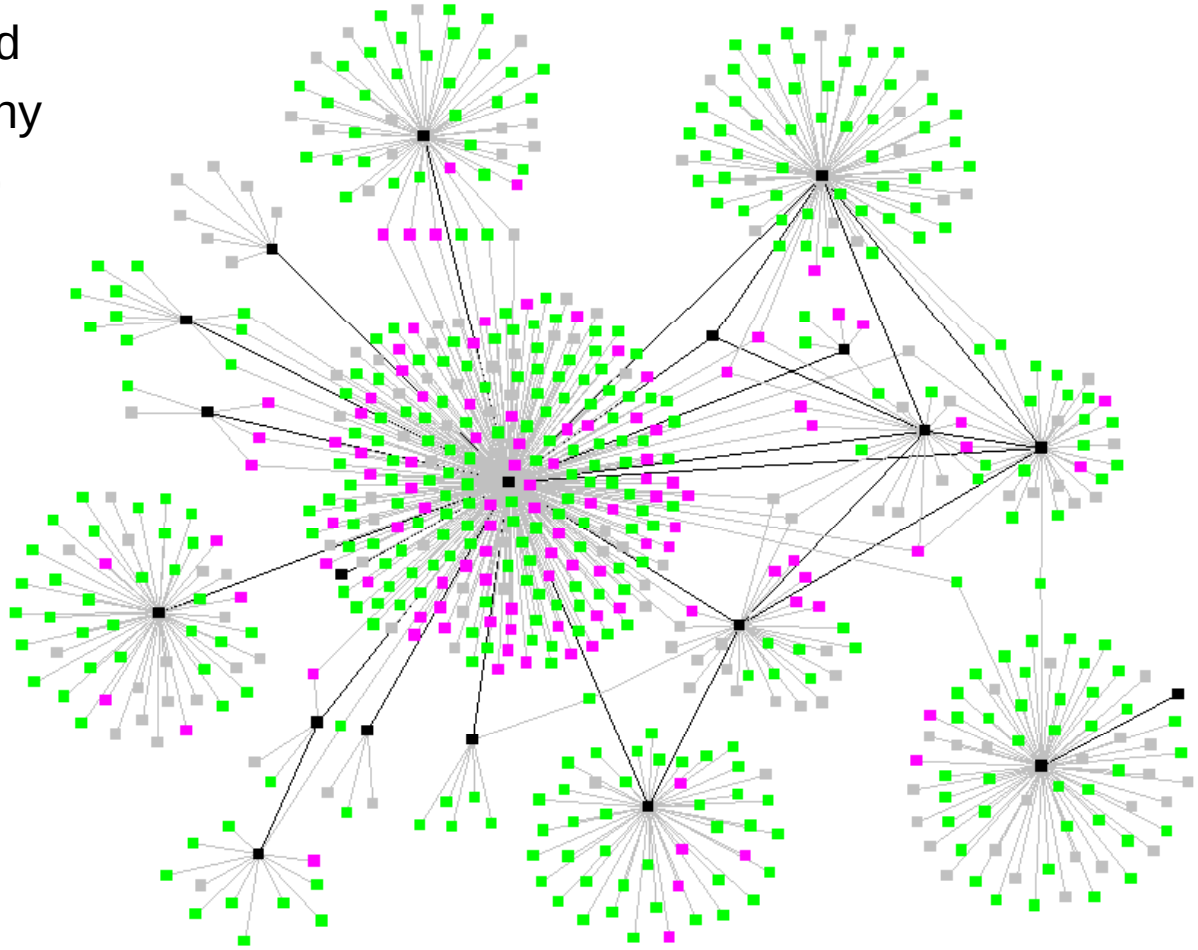
Color: inferred node clustering



Q: what do you think is the basic idea of the clustering algorithm?

Spread of disease in a social network

- black: diseased
- pink: infected
- green: healthy
- grey: no info

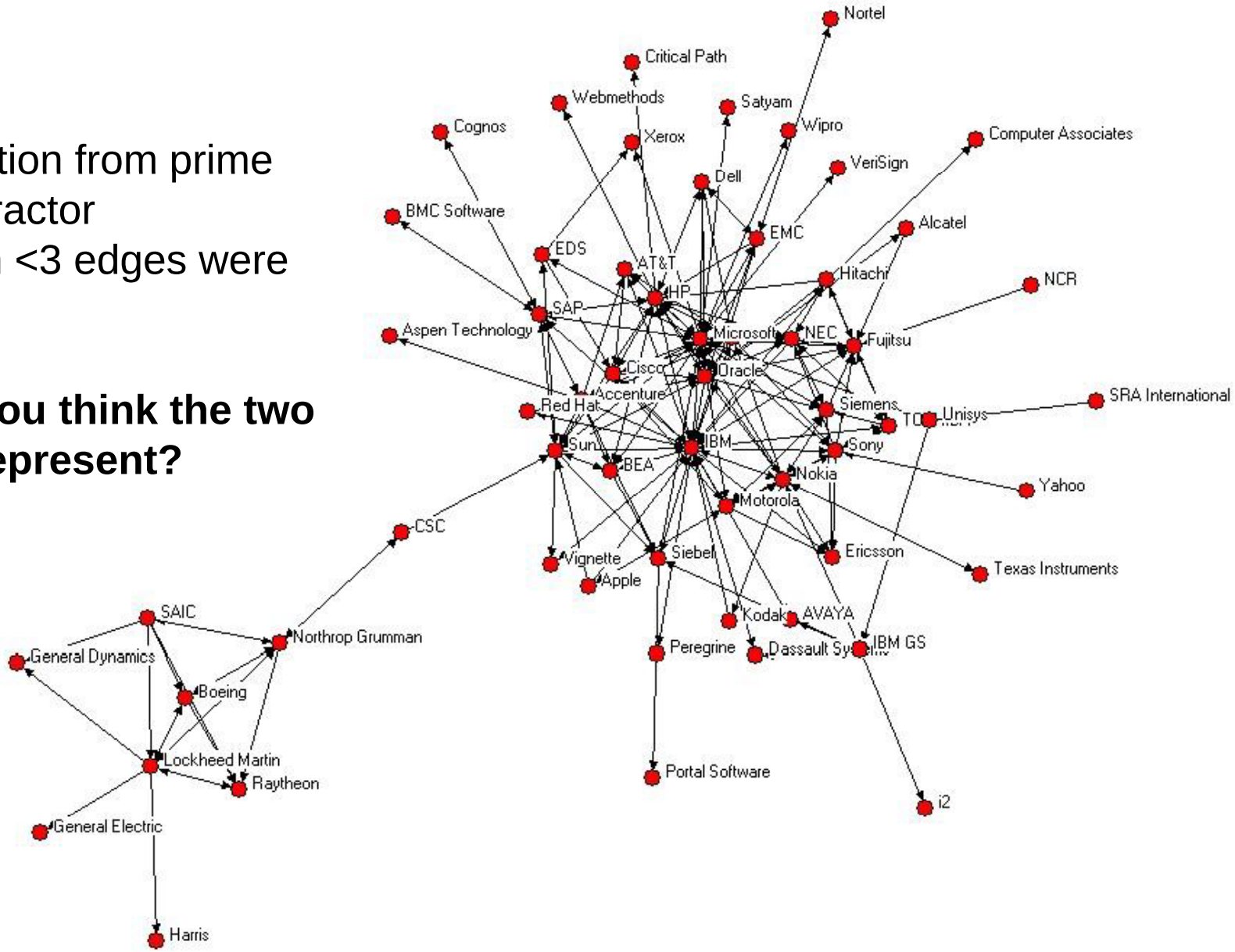


Q: Where do you think the network mapping started?

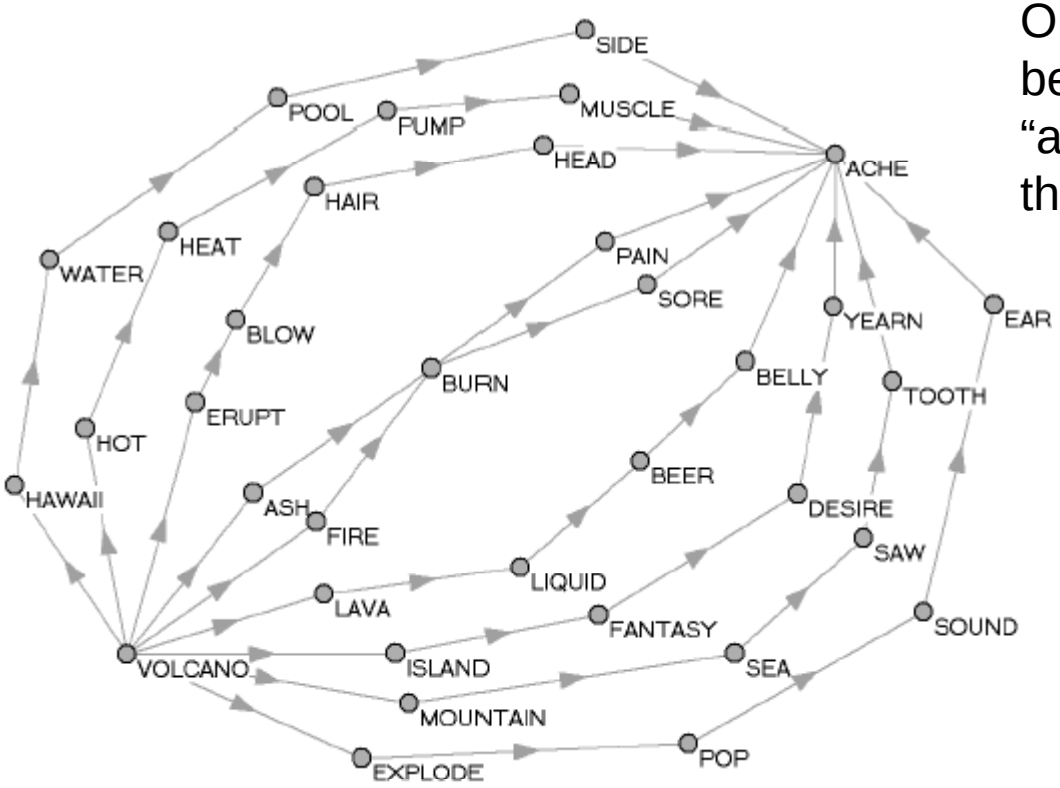
Business alliances among IT firms

Edge direction from prime to subcontractor
Nodes with <3 edges were filtered out

What do you think the two clusters represent?



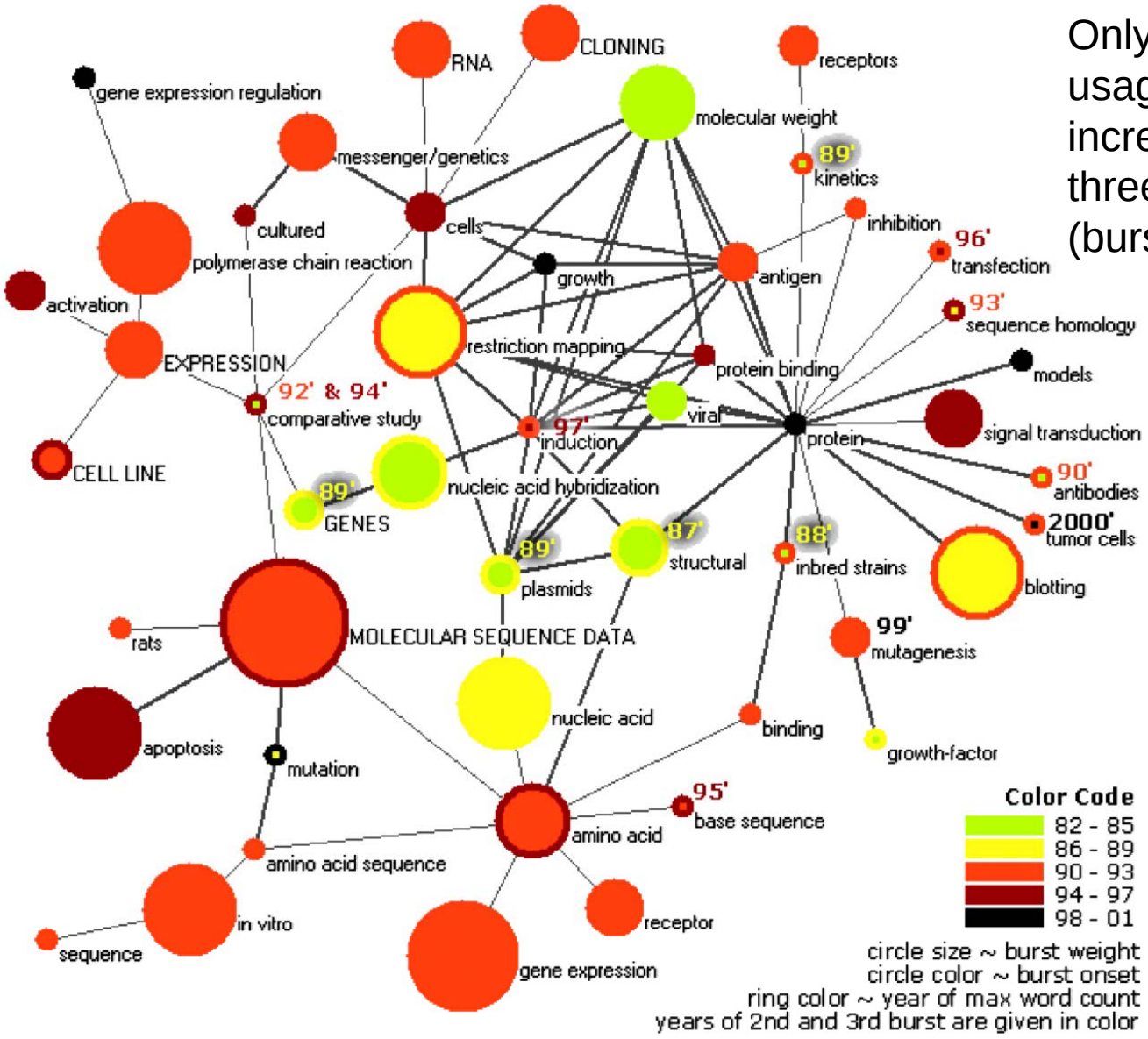
Network of free semantic associations



Only words in paths between “volcano” and “ache” are included in the illustration

Based on the University of South Florida Word Association, Rhyme and Word Fragment Norms

Word co-usage network in PNAS publications



Only words whose usage significantly increased during a three-year period (burst) are included

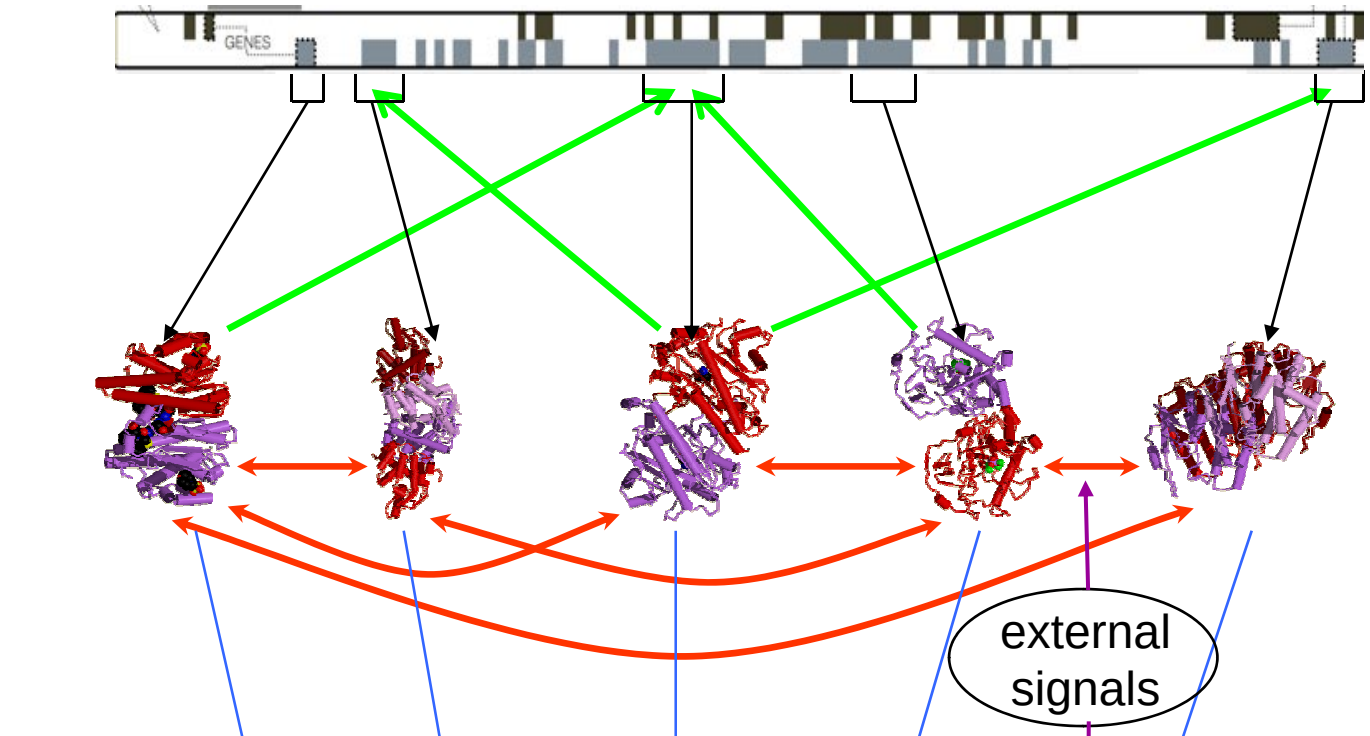
Examples we have seen so far:

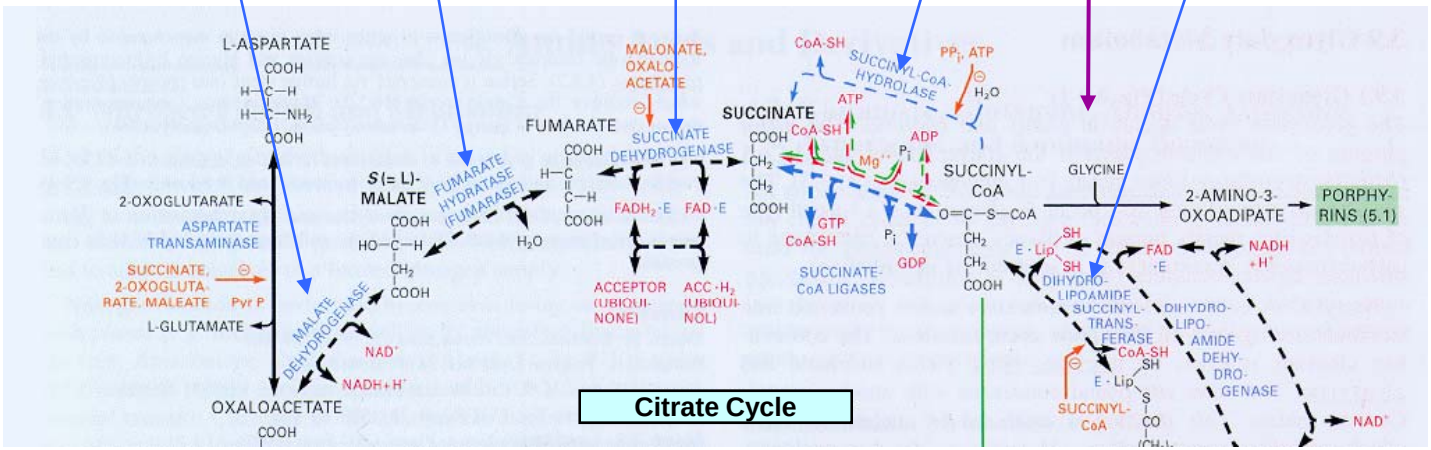
- Internet
- World Wide Web
- Food web
- Social network
- Business network
- Semantic network

Focus on the nodes of these networks. In which examples did the nodes represent single entities, and in which did they represent groups?

What additional information do you think is necessary in the latter case?

Many non-identical elements, diverse interactions





GENOME

gene regulation

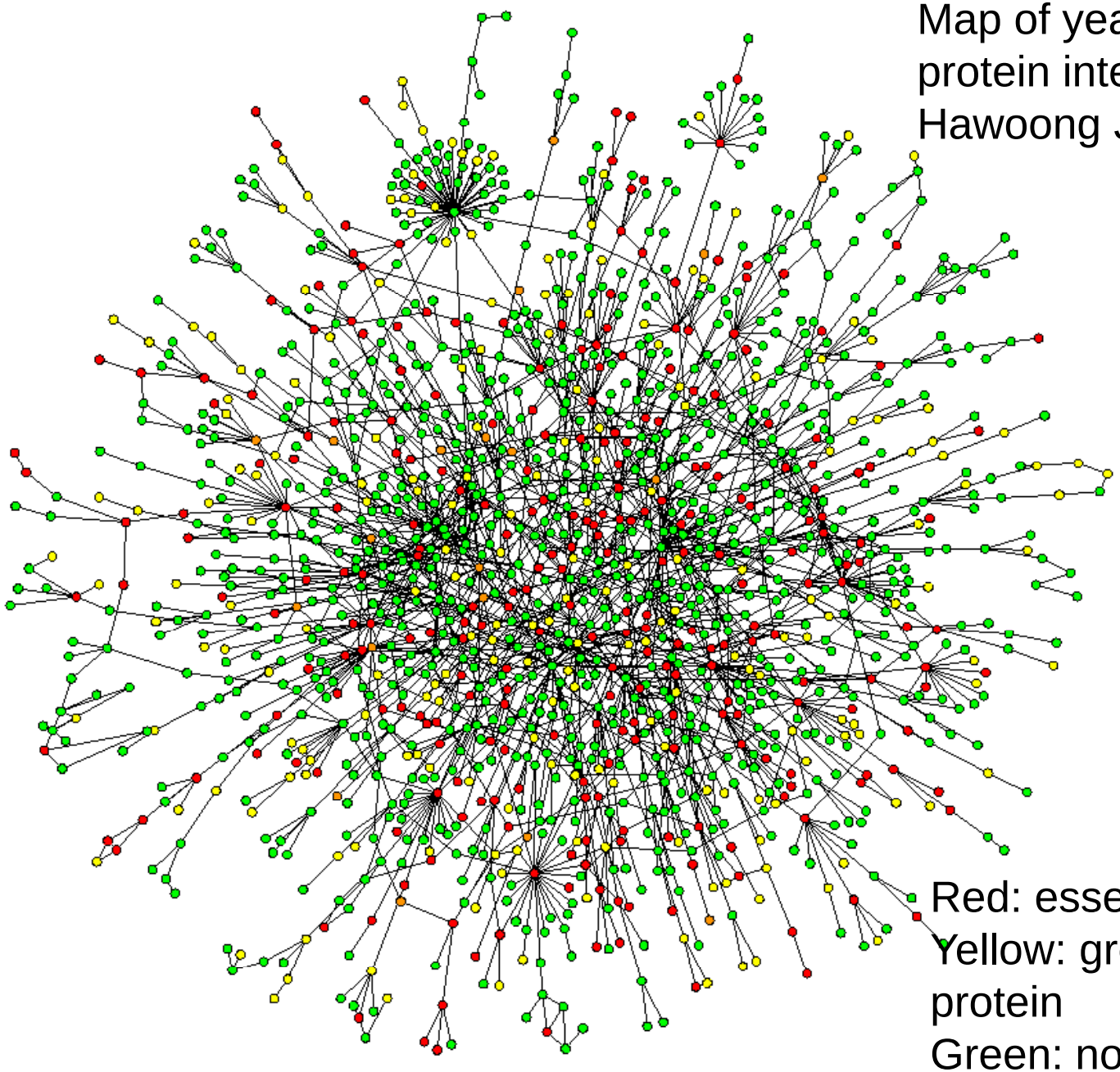
PROTEOME

protein-protein interactions

signal transduction

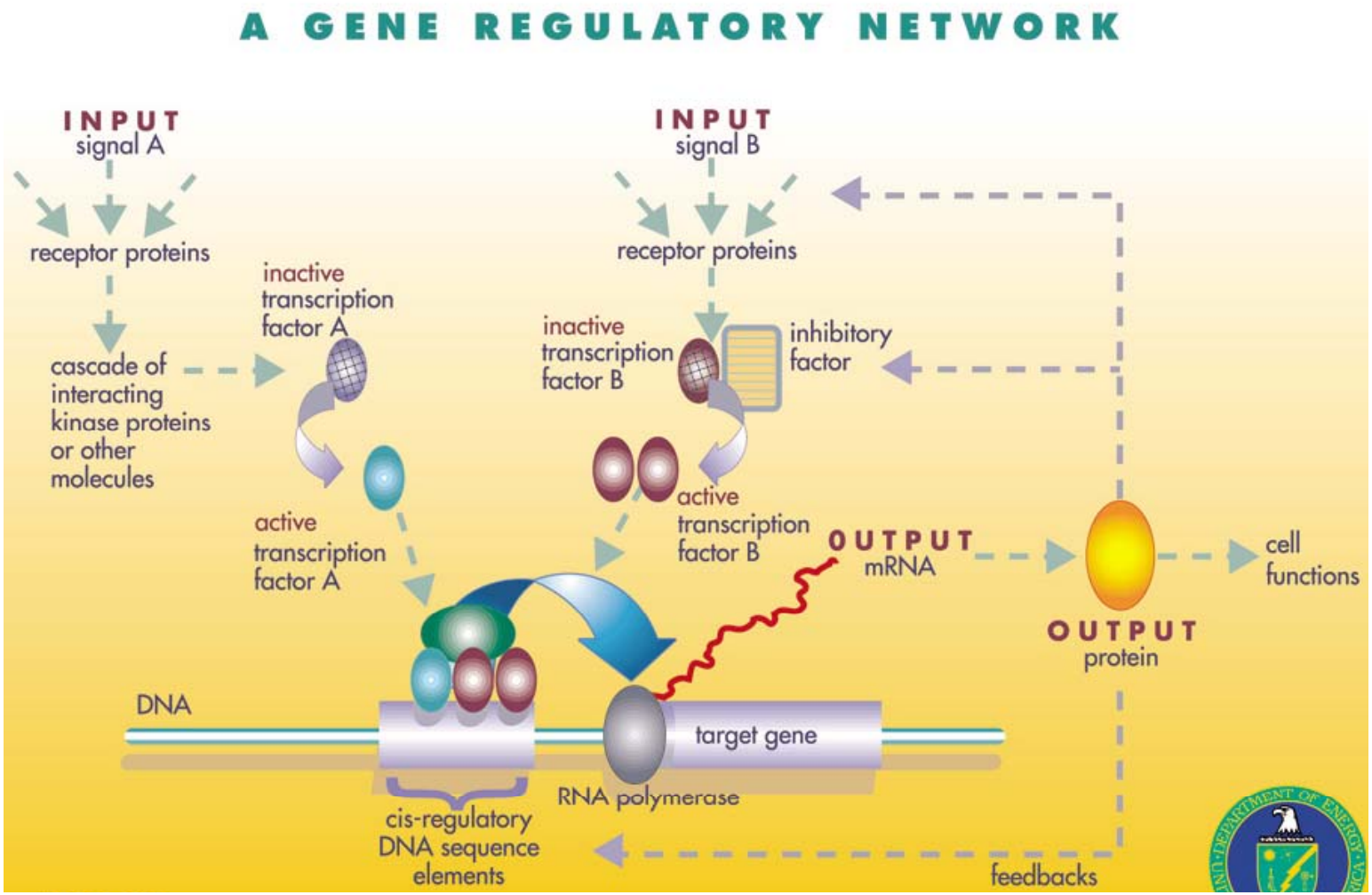
METABOLISM

Bio-chemical reactions



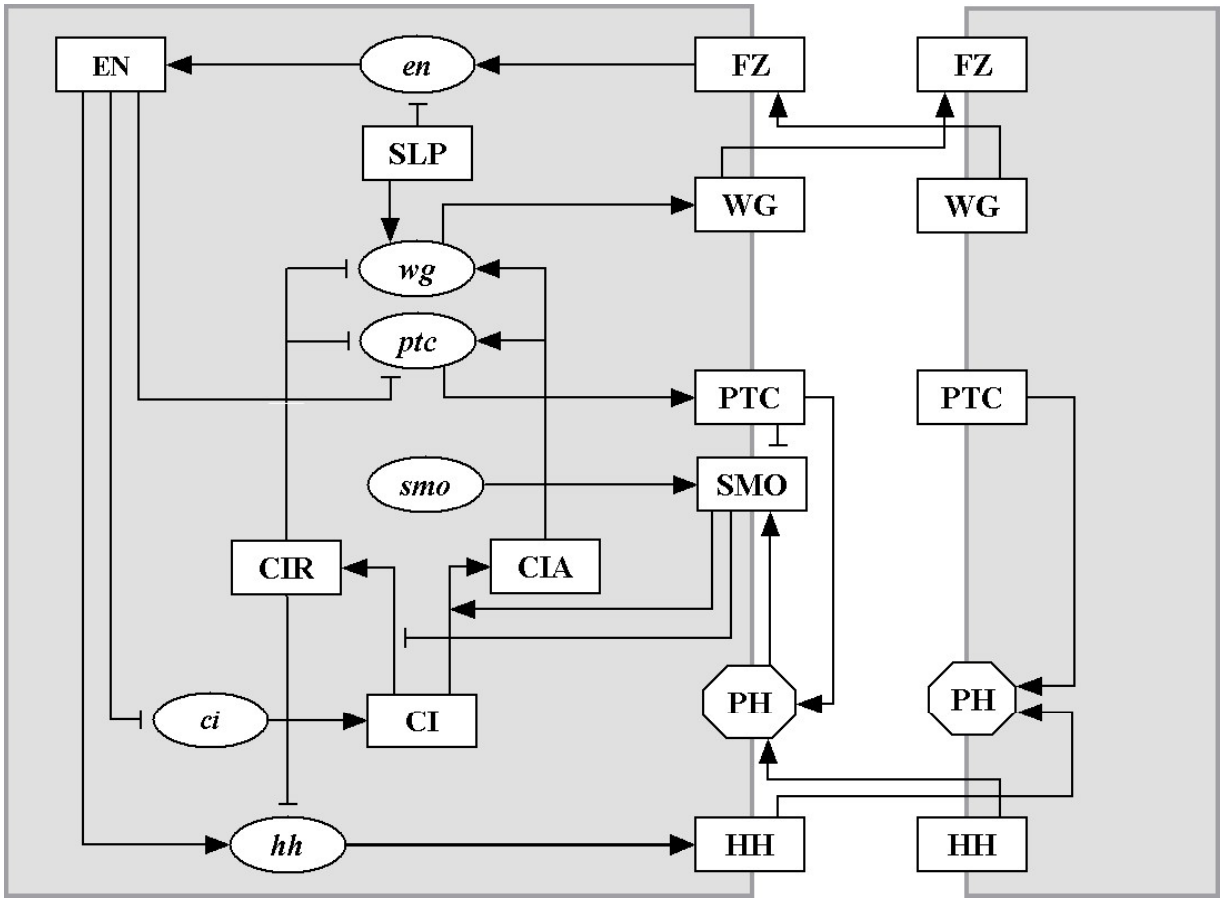
Map of yeast protein-protein interactions, by Hawoong Jeong

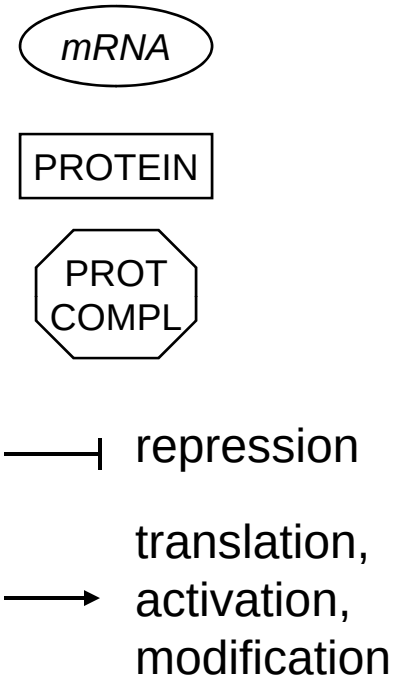
Red: essential protein
Yellow: growth- affecting protein
Green: non-essential protein



Ex: Draw an alternative, less pictorial network representation

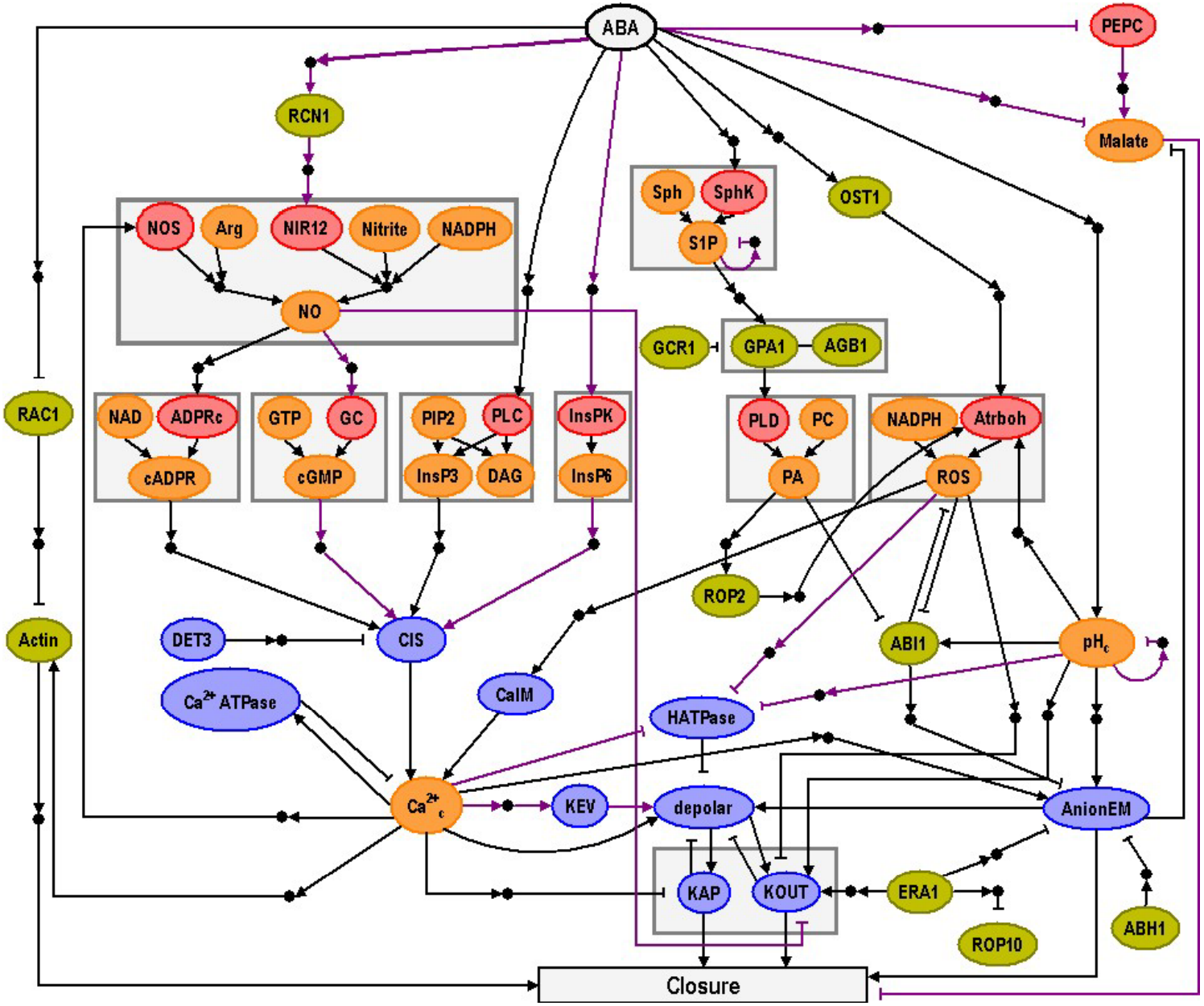
Interaction network of the Drosophila segment polarity genes





Albert, Othmer
J. Theor. Bio.
2003

ABA signal transduction network



Red: enzymes
Blue: transport
Orange: small molecules
Green: sign. transd. proteins
Black points: unknown intermediary nodes
Purple edges: inferences

Li, Assmann, Albert, PLoS Biology 2006

Why study networks?

• It is increasingly recognized that complex systems cannot be described in a reductionist view.

• Understanding the behavior of such systems starts with understanding the topology of the corresponding network.

• Topological information is fundamental in constructing realistic models for the function of the network.

• Network - related questions:

How do we determine or infer network topology ?

How can we quantitatively describe large networks?

How did networks get to be the way they are?

What are the consequences of a specific network organization?