

Network Evolution

Statistics of Networks

Comparing Networks

Networks in Cellular Biology

A. Metabolic Pathways

B. Regulatory Networks

C. Signaling Pathways

D. Protein Interaction Networks - PIN

Empirical Facts

Dynamics on Networks (models)

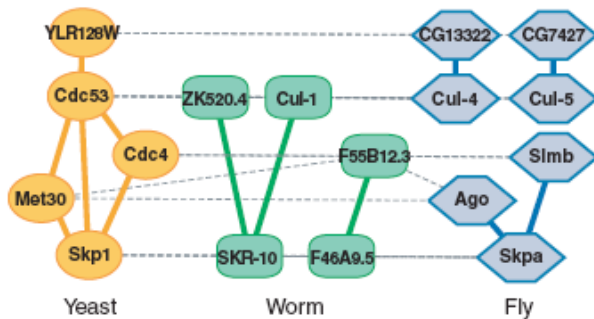
Models of Network Evolution

Network Alignment & Motifs

Barabasi & Oltvai, 2004, Sharan & Ideker, 2006

1. Are nodes/edges labelled?
2. Which operations are allowed?
3. Pair/Multiple?

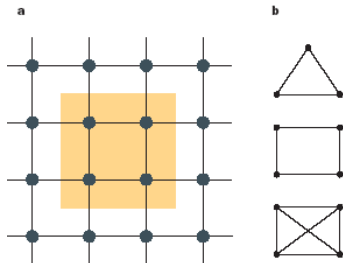
•Network Alignment



•Network Search

Find (approximately) a network within a set of others.

•Motifs



•Network integration

Combine a set of networks to one large network.

Network Description and Statistics I

Barabasi & Oltvai, 2004

- Degree/Indegree/Outdegree
- Shortest Path $Dist(i, j)$
- Mean Path Length
- Diameter: $Max_{i,j} \{Dist(i, j)\}$
- Clustering Coefficient - $C_I = 2T_I / n_I(n_I - 1)$

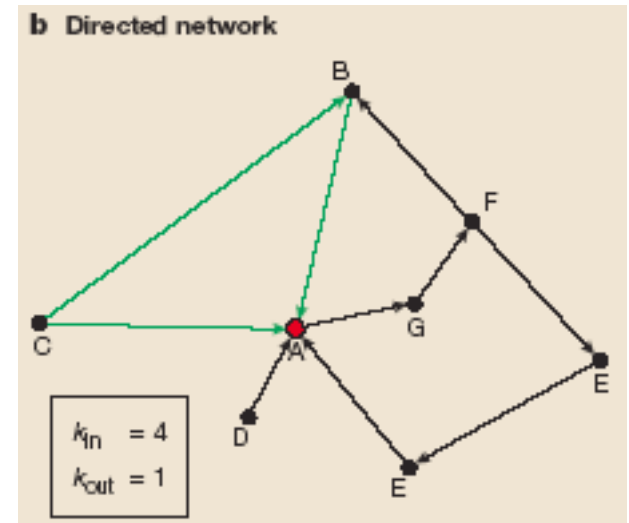
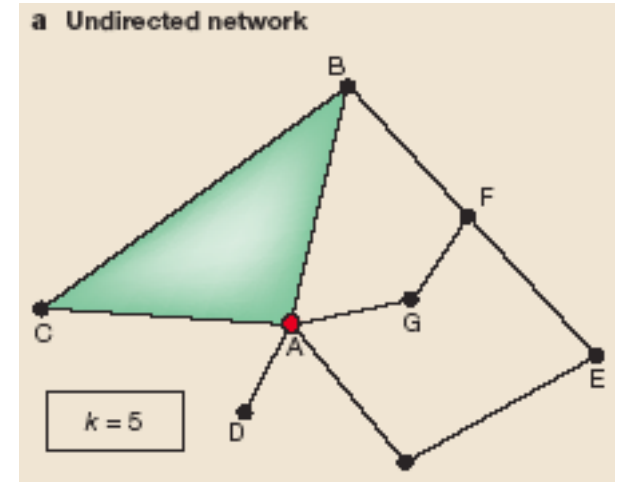
$$C_A = 1/10$$

- Degree Distribution - $P(k)$
- Scale Free Networks $P(k) \sim k^{-\gamma}$ $\gamma > 2$
- Hubs: multiply connected nodes

The lower γ , the more hubs.

Small World Property:

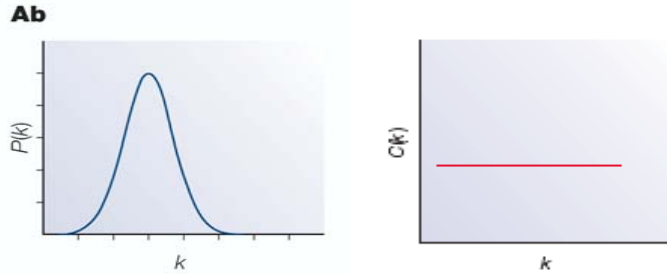
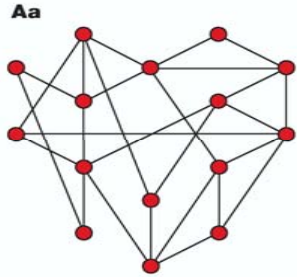
Graph connected and path lengths small



Network Description and Statistics II

Barabasi & Oltvai, 2004

A. Random Networks [Erdos and Rényi (1959, 1960)]



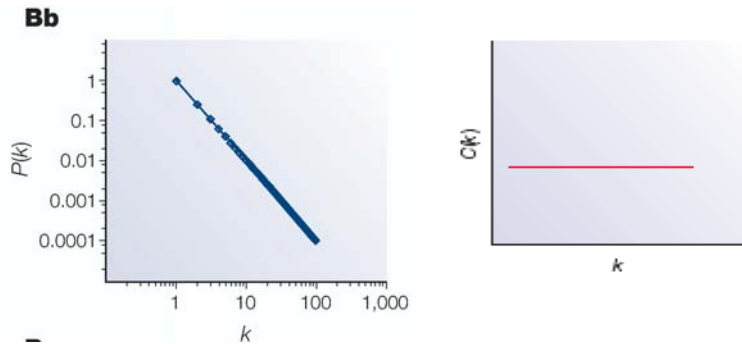
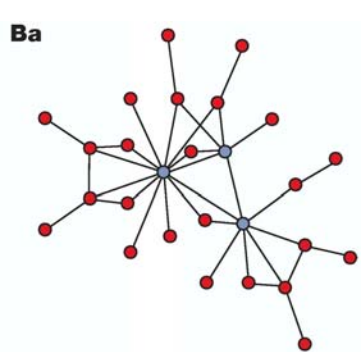
$$P(k) = \frac{e^{-\bar{k}} \bar{k}^k}{k!}$$

Mean path length $\sim \ln(k)$

Phase transition:

Connected if: $p \geq \ln(k)/k$

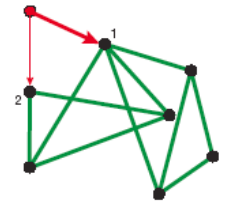
B. Scale Free [Price, 1965 & Barabasi, 1999]



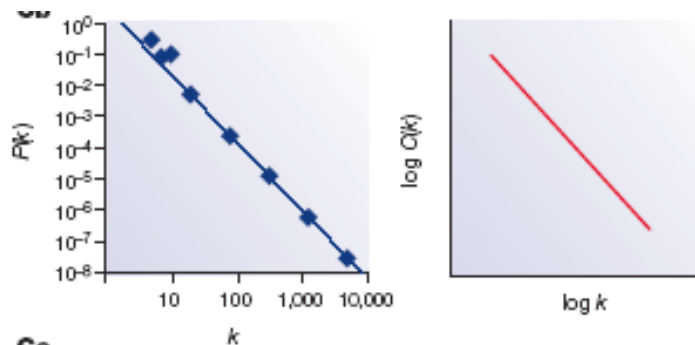
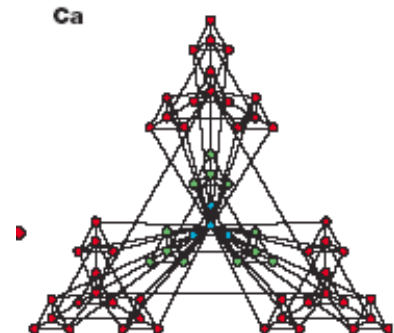
$$P(k) \sim k^{-\gamma}, k \gg 1, 2 < \gamma$$

Mean path length $\sim \ln \ln(k)$

Preferential attachment. Add proportionally to connectedness



C. Hierarchical

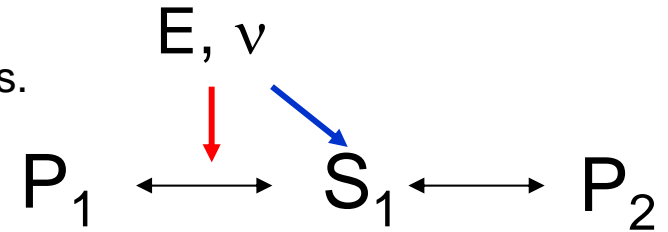


Copy smaller graphs and let them keep their connections.

Control Coefficients

(Heinrich & Schuster: Regulation of Cellular Systems. 1996)

Flux J_j (edges) – Enzyme conc., E_k (edges), S – internal nodes.



Flux Control Coefficient – FCC:

$$C_{E_k}^{J_j} = \left(\frac{E_k}{J_j} \frac{\Delta J_j}{\Delta E_k} \right)_{\Delta E_k \rightarrow 0} = \frac{E_k}{J_j} \frac{\partial J_j}{\partial E_k} = \frac{\partial \ln(J_j)}{\partial \ln(E_k)}$$

Kacser & Burns, 73



$$C_{v_k}^{J_j} = \left(\frac{v_k}{J_j} \frac{\Delta J_j}{\Delta v_k} \right)_{\Delta v_k \rightarrow 0} = \frac{v_k}{J_j} \frac{\partial J_j}{\partial v_k} = \frac{\partial \ln(J_j)}{\partial \ln(v_k)}$$

Heinrich & Rapoport, 73-74

FCC: gluconeogenesis from lactate

| | |
|------------------------|-----|
| Pyruvate transport | .01 |
| Pyruvate carboxylase | .83 |
| Oxaloacetate transport | .04 |
| PEOCK | .08 |



A Model for Network Inference

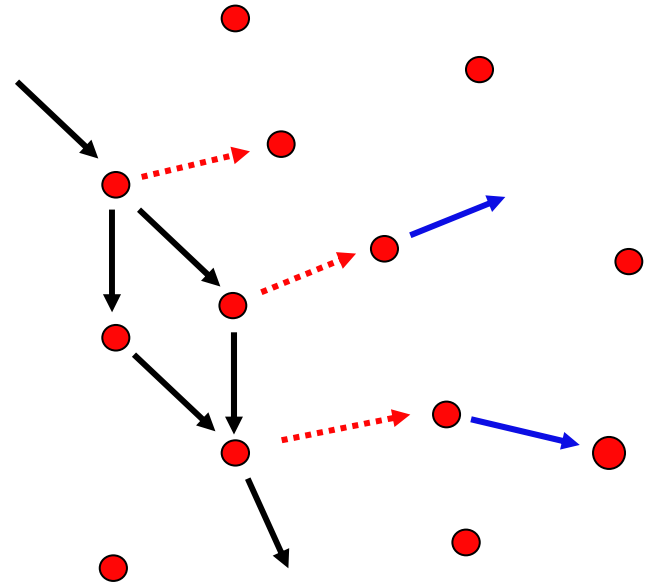
- A given set of metabolites: 
- A given set of possible reactions -
arrows not shown.
- A core metabolism: 
- A set of present reactions - **M**
black and **red** arrows

Restriction R:

A metabolism must define a connected graph

M + **R** defines

1. a set of deletable (dashed) edges **D(M)**: 
2. and a set of addable edges **A(M)**: 



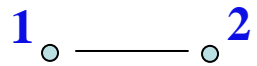
Let μ be the rate of deletion
 λ the rate of insertion

Then

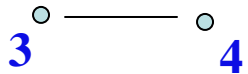
$$\frac{dP(M)}{dt} = \lambda \sum_{M' \in D(M)} P(M') + \mu \sum_{M'' \in A(M)} P(M'') - P(M)[\lambda|D(M)| + \mu|A(M)|]$$

Likelihood of Homologous Pathways

Number of Metabolisms:



+ 2 symmetrical versions



| n | Number of all graphs with n nodes | Number of states |
|-----|-------------------------------------|------------------|
| 1 | 1 | 1 |
| 2 | 2 | 2 |
| 3 | 8 | 8 |
| 4 | 64 | 61 |
| 5 | 1024 | 969 |
| 6 | 32768 | 31738 |
| 7 | 2097152 | 2069964 |
| 8 | 268435456 | 267270033 |
| 9 | 68719476736 | 68629753641 |
| 10 | 35184372088832 | 35171000942698 |

$$P_{\Theta}(\text{graph}_1, \text{graph}_2) = P_{\Theta}(\text{graph}_1) P_{\Theta}(\text{graph}_2 \mid \text{graph}_1)$$

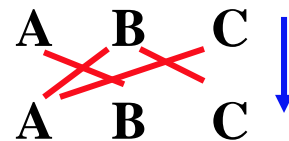
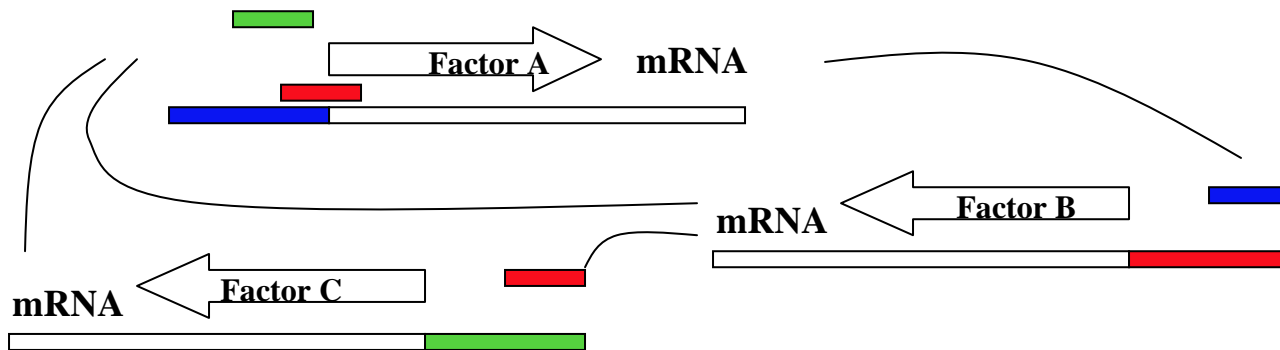
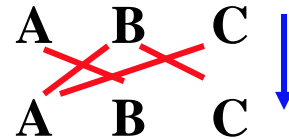
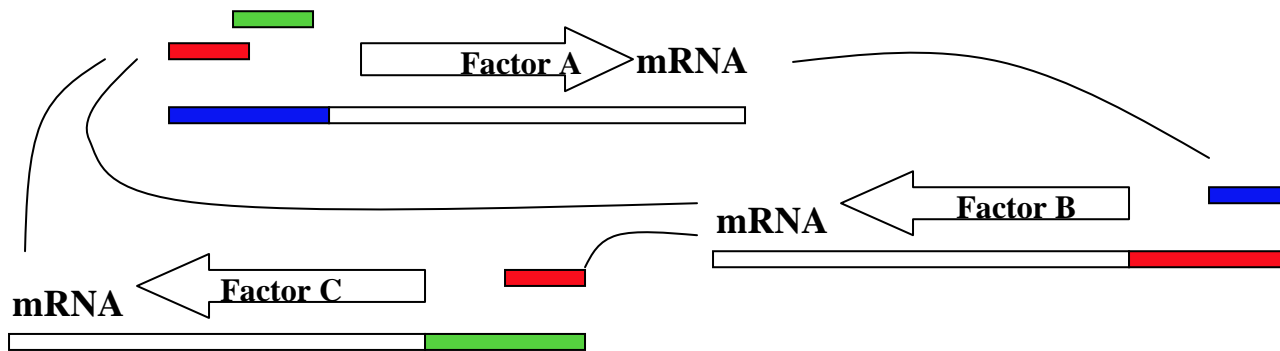
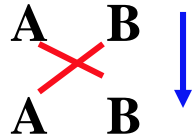
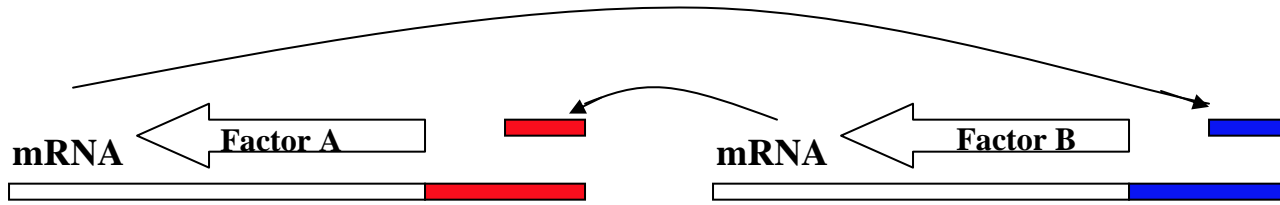
Approaches:

Continuous Time Markov Chains with computational tricks.

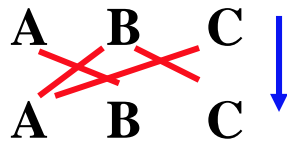
MCMC

Importance Sampling

B. Regulatory Networks



Boolean functions, Wiring Diagrams and Trajectories



| | | | |
|--------|---|---|---|
| Inputs | 2 | 1 | 1 |
| Rule | 4 | 2 | 2 |

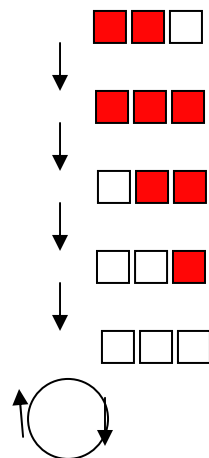
A activates B

B activates C

A is activated by B, inhibited by (B>C)

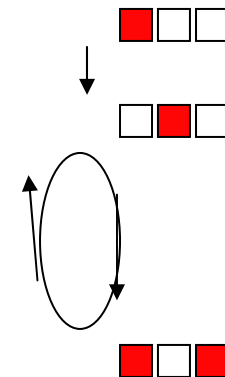
Point Attractor

| A | B | C |
|---|---|---|
| 1 | 1 | 0 |
| 1 | 1 | 1 |
| 0 | 1 | 1 |
| 0 | 0 | 1 |
| 0 | 0 | 0 |
| 0 | 0 | 0 |



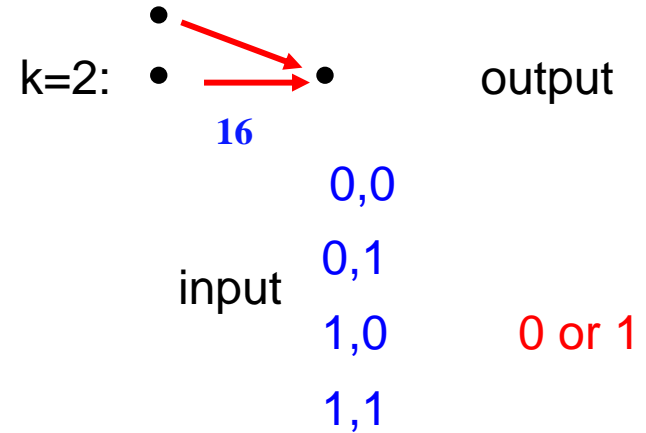
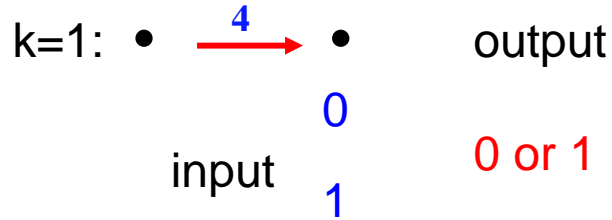
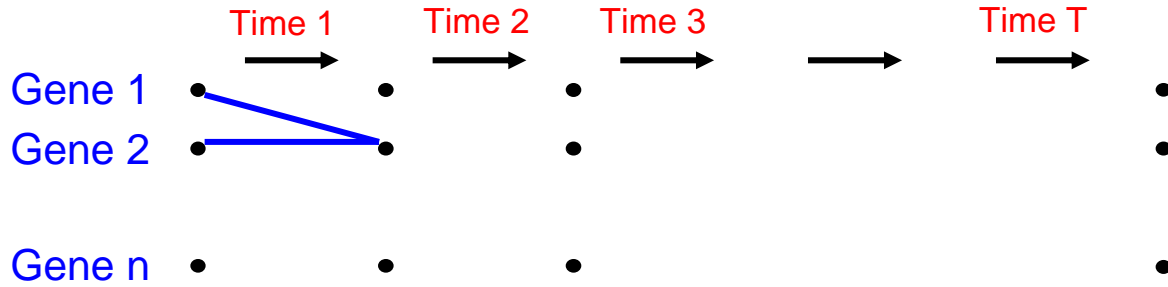
2 State Attractor

| A | B | C |
|---|---|---|
| 1 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 1 |
| 0 | 1 | 0 |



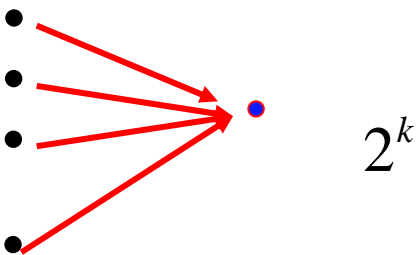
Boolean Networks

R.Somogyi & CA Sniegoski (1996) Modelling the Complexity of Genetic Networks Complexity 1.6.45-64.



A single function:

The whole set:



$$2^{2^k}$$

For each gene dependent on i genes: $\binom{k}{i}$ choices of dependent genes. Number of Boolean Rules $\left(\binom{k}{i} 2^i\right)^k$

Contradiction: Always turned off (biological meaningless) **Tautology:** Always turned on (household genes)

Reverse Engineering Algorithm-Reveal

D'haeseler et al.(2000) Genetic network Inference: from co-expression clustering to reverse engineering. Bioinformatics 16.8.707-

Assumptions:

Discrete known Generations

No Noise

| | | | | | |
|--------|--------|--------|--------|--------|---|
| 0 | 1 | 1 | 1 | 1 | 0 |
| 0 | 0 | 1 | 1 | 0 | 1 |
| 1 | 1 | 1 | 0 | 1 | 0 |
|→ |→ |→ |→ |→ | |

BOOL-1

Akutsu et al. (2000) Inferring qualitative relations in genetic networks and metabolic pathways. Bioinformatics 16.2.727-

Algorithm

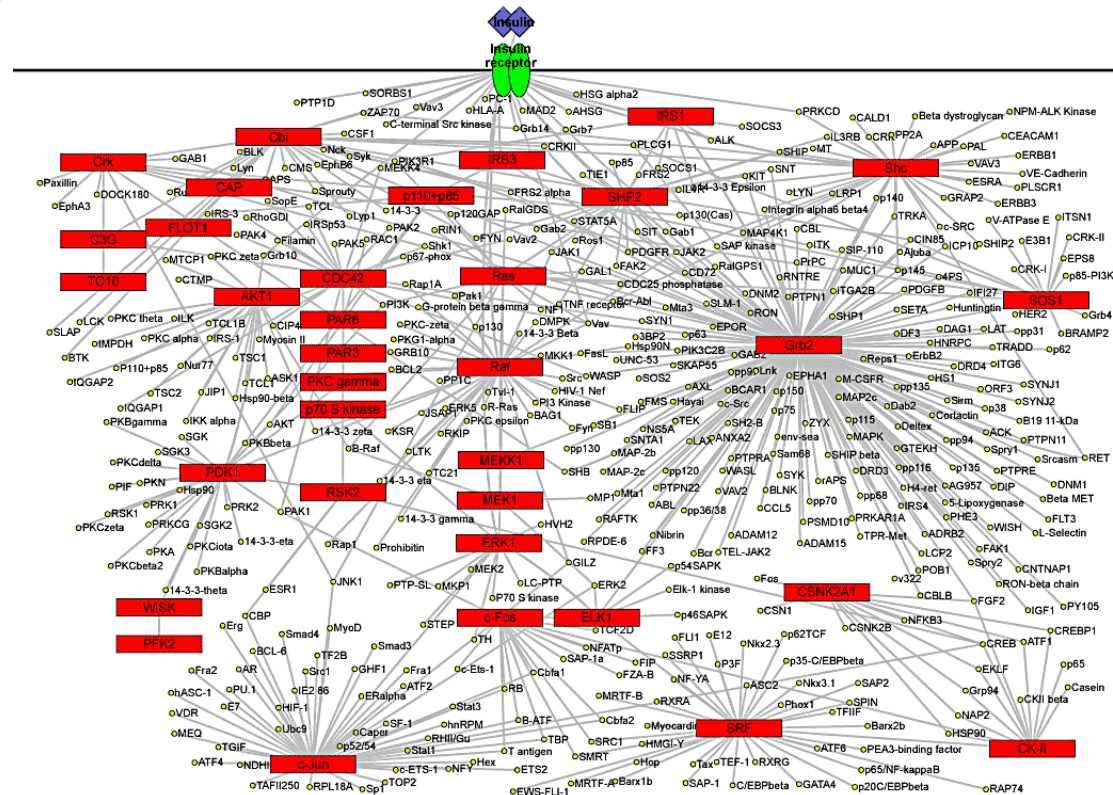
For each gene do (n)

For each boolean rule ($\leq k$ inputs) not violated, keep it.

If $O(2^{2k}[2k + \alpha]\log(n))$ INPUT patterns are given **uniformly randomly**, BOOL-1 correctly identifies the underlying network with probability $1-n^{-\alpha}$, where α is any fixed real number > 1 .

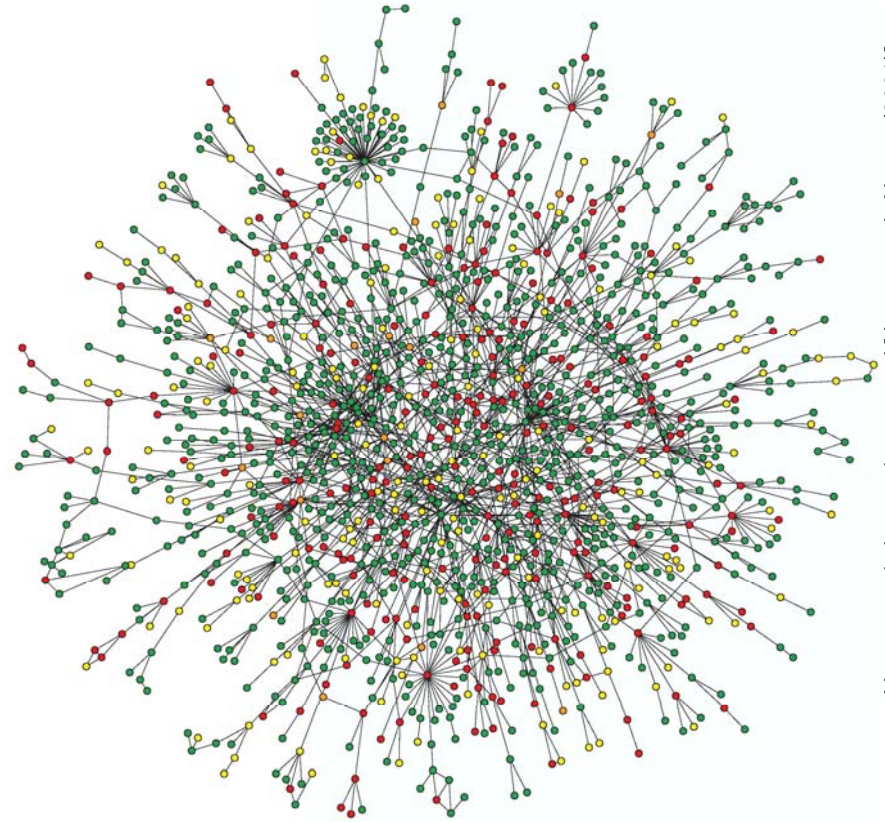
C. Signaling Pathways

- Transmits signals from membrane to gene regulation.
- Its function is enigmatic as some of the molecules involved are common to different functions and how cross-interaction is avoided is unknown.



D. Protein Interaction Network

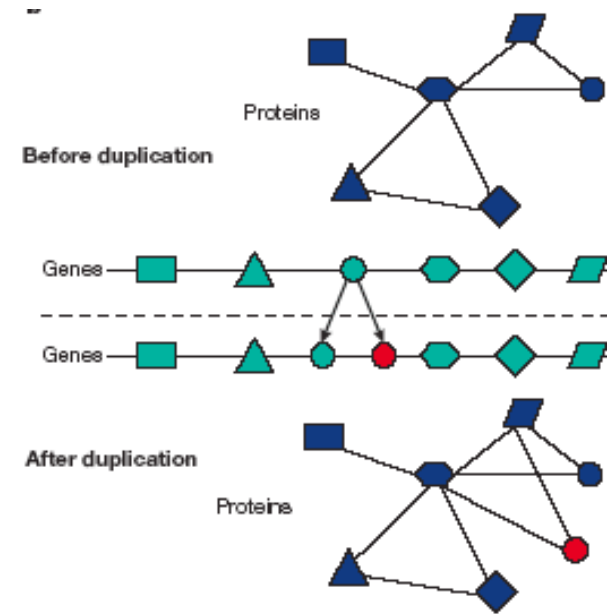
- The sticking together of different protein is measured by mass spectroscopy.
- The nodes will be all known proteins.
- Two nodes are connected if they stick together. This can be indicator of being part of a functional protein complex, but can also occur for other reasons.



PIN Network Evolution

Barabasi & Oltvai, 2004 & Berg et al. ,2004; Wiuf et al., 2006

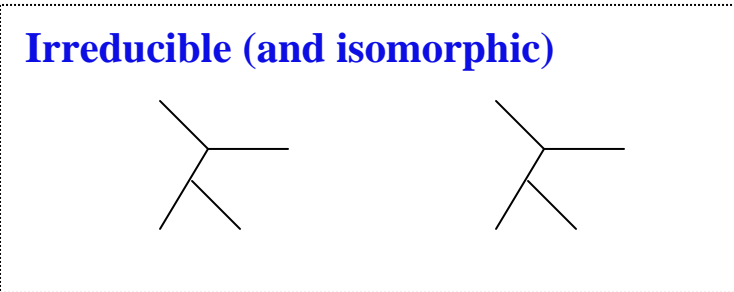
- A gene duplicates
- Inherits its connections
- The connections can change



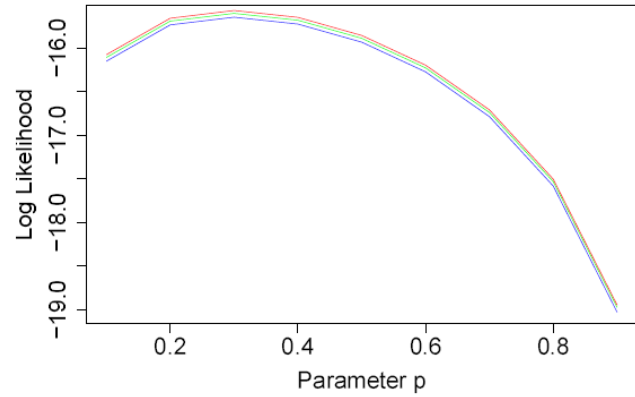
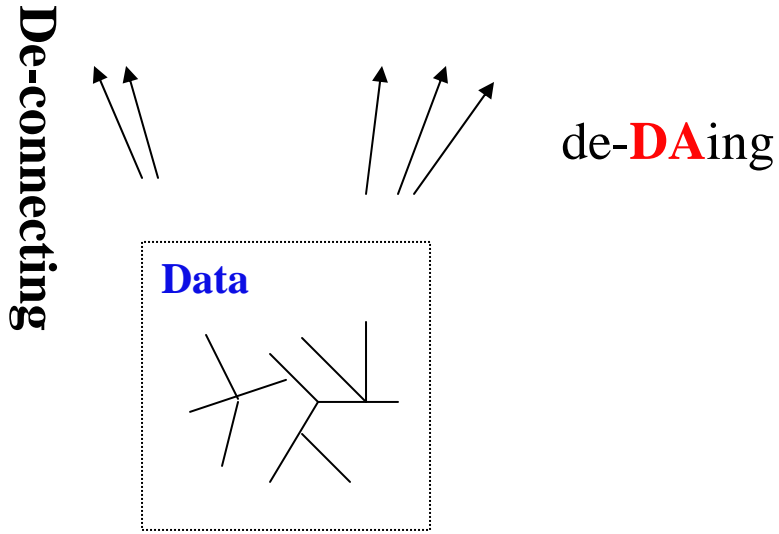
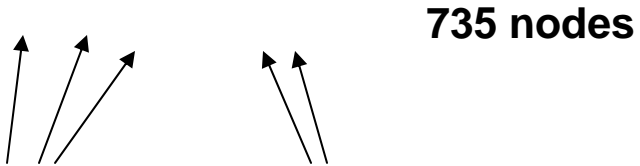
Berg et al. ,2004:

- Gene duplication slow $\sim 10^{-9}$ /year
- Connection evolution fast $\sim 10^{-6}$ /year
- Observed networks can be modeled as if node number was fixed.

Likelihood of PINs



- Can only handle 1 graph.
- Limited Evolution Model



2386 nodes and 7221 links

$$\theta_0 = (1, .66, .33, 0)$$