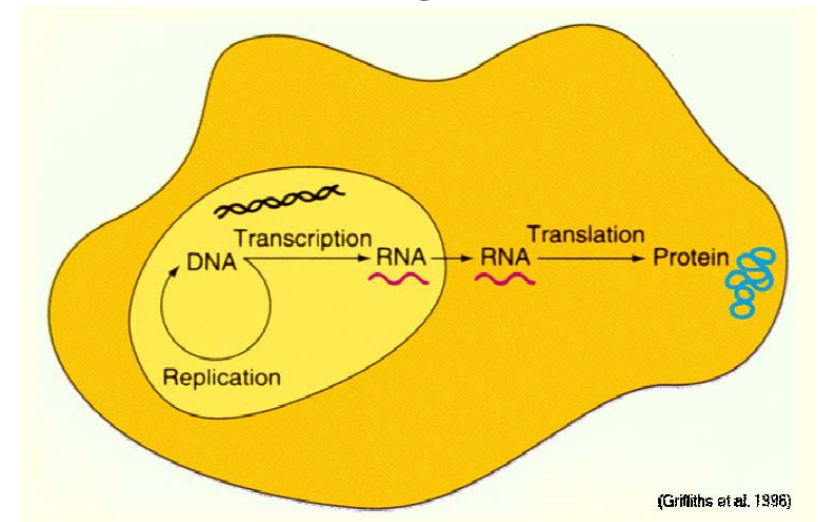


Introduction to Game Theory and Applications

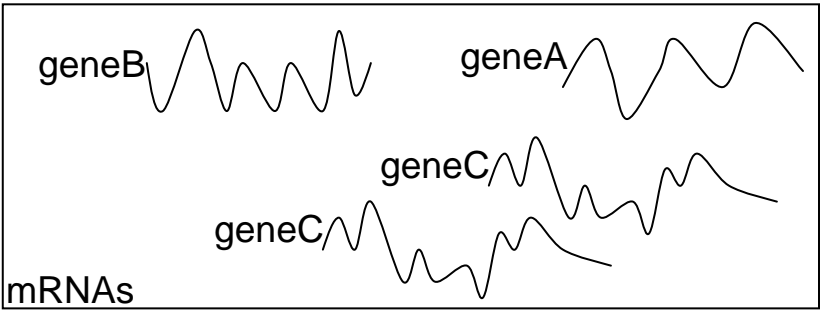
Stefano MORETTI and Fioravante PATRONE
LAMSADE (CNRS), Paris Dauphine
and
DIPTTEM, University of Genoa
Paris, Telecom ParisTech, 2010

Gene expression and microarray

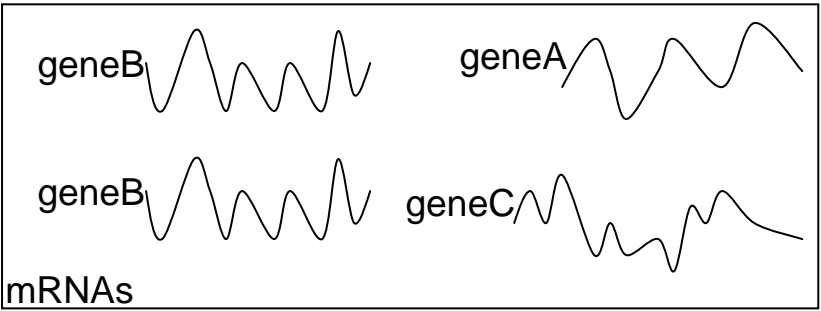
- **Gene expression** occurs when genetic information contained within DNA is *transcribed* into mRNA molecules and then *translated* into the **proteins**.
- Nowadays, **microarray** technology is available for taking “pictures” of gene expressions. Within a single experiment of this sophisticated technology, the level of expression of thousands of genes can be estimated in a sample of cells under a given condition.



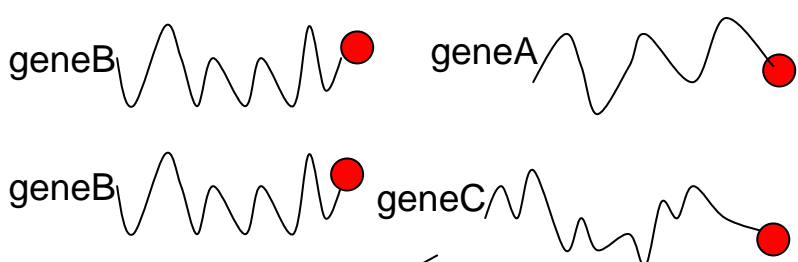
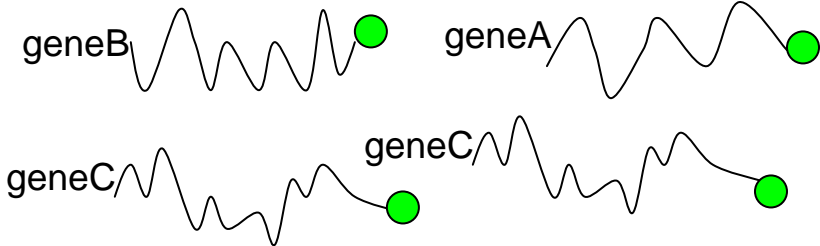
Normal Cell



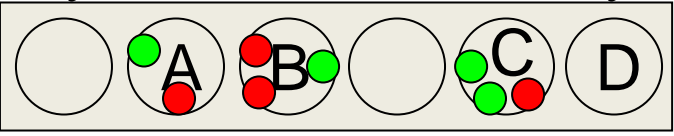
Tumor cell



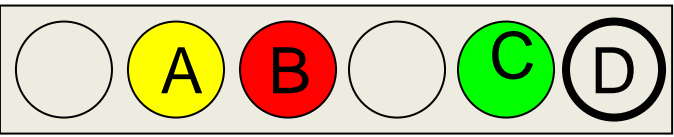
Fluorescent labelling reaction with reverse transcription



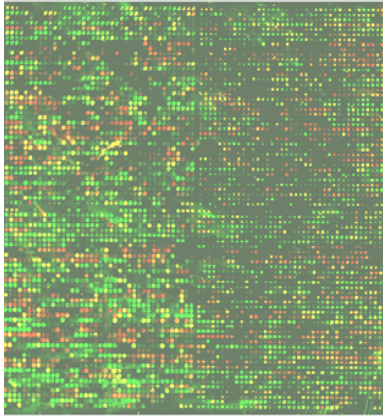
Hybridize to microarray



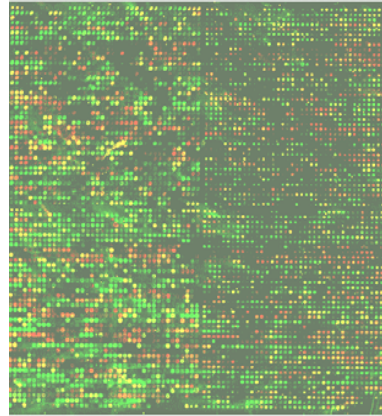
Scan image



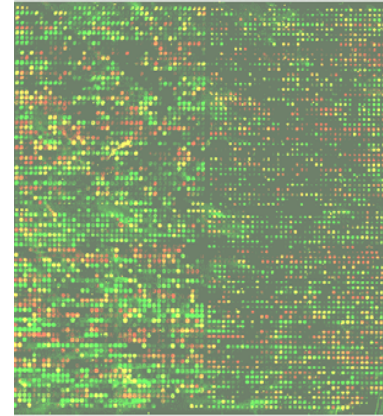
| | Normal | Tumor |
|--------|--------|-------|
| Gene A | 1 | 1 |
| Gene B | 1 | 2 |
| Gene C | 2 | 1 |
| Gene D | 0 | 0 |



Array1

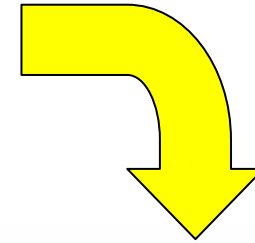


Array2



Array3

...



Arrays

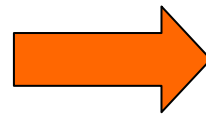
| | Array1 | Array2 | Array3 | Array4 | Array5 | ... |
|-------|--------|--------|--------|--------|--------|-----|
| Gene1 | 0.46 | 0.30 | 0.80 | 1.51 | 0.90 | ... |
| Gene2 | 0.10 | 0.49 | 0.24 | 0.06 | 0.46 | ... |
| Gene3 | 0.15 | 0.74 | 0.04 | 0.10 | 0.20 | ... |
| Gene4 | 0.45 | 1.03 | 0.79 | 0.56 | 0.32 | ... |
| Gene5 | 0.06 | 1.06 | 1.35 | 1.09 | 1.09 | ... |
| ... | ... | ... | ... | ... | ... | ... |

Expression level of gene 5 in array 4

From political and social science to genomics...

- Players are **genes**
- Who knows the **decision rule** in this context?
- IDEA: we can make a rule on microarray gene expression profiles.
- Example: we define a criterion to establish which genes have abnormal expressions on each array

| | array1 |
|-------|--------|
| gene1 | 1.121 |
| gene2 | 2.453 |
| gene3 | 3.586 |



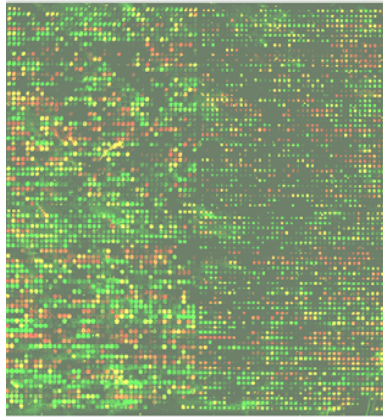
| | array1 |
|-------|--------|
| gene1 | 0 |
| gene2 | 1 |
| gene3 | 1 |

Decision rule

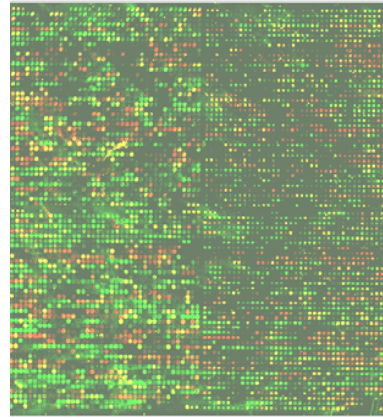
A group of genes is *winning* on a single array **if all genes** that have abnormal expressions **belong** to that group

| | array1 |
|-------|--------|
| gene1 | 0 |
| gene2 | 1 |
| gene3 | 1 |

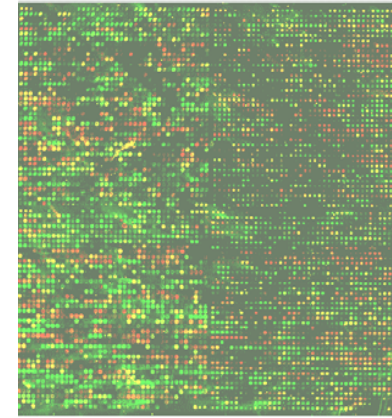
Both groups {gene2, gene3} and group {gene1, gene2, gene3} are winning.



Array1



Array2



Array3

...

| | array1 | array2 | array3 |
|-------|--------|--------|--------|
| gene1 | 0 | 1 | 0 |
| gene2 | 1 | 1 | 0 |
| gene3 | 1 | 0 | 1 |

- coalition {gene2, gene3} is winning two times out of three;
- coalition {gene1, gene2} is winning one time out of three;
- And so on for each coalition...

Example

| | Array1 | Array2 | Array3 |
|-------|--------|--------|--------|
| g_1 | 0 | 1 | 0 |
| g_2 | 1 | 1 | 0 |
| g_3 | 1 | 0 | 1 |

The corresponding *microarray game*

$\langle \{g_1, g_2, g_3\}, v \rangle$ tale che

$$v(\emptyset) = v(\{g_1\}) = v(\{g_2\}) = 0$$

$$v(\{g_1, g_3\}) = v(\{g_1, g_2\}) = v(\{g_3\}) = 1/3$$

$$v(\{g_2, g_3\}) = 2/3$$

$$v(\{g_1, g_2, g_3\}) = 1.$$

The Shapley value is

$$Sh_{g_1} = 1/6 \quad Sh_{g_2} = 1/3 \quad Sh_{g_3} = 1/2$$

Axioms for the Shapley value on microarray games

Property 1: Null Gene (NG)

A gene which does not contribute to change the worth of any coalition of genes, should receive zero power.

Prop.2: Equal Splitting (ES)

Each sample should receive the same level of reliability. So the power of a gene on two samples should be equal to the sum of the power on each sample divided by two.

| | | | | | | | | | | |
|----|----|----------|----|-----------|--|---|----|----|----|------------------------|
| | s1 | | s2 | + | | = | | s1 | s2 | |
| g1 | 0 | ψ_1 | 1 | ψ'_1 | | | g1 | 0 | 1 | $(\psi_1 + \psi'_1)/2$ |
| g2 | 0 | ψ_2 | 1 | ψ'_2 | | | g2 | 0 | 1 | $(\psi_2 + \psi'_2)/2$ |
| g3 | 1 | ψ_3 | 0 | ψ'_3 | | | g3 | 1 | 0 | $(\psi_3 + \psi'_3)/2$ |

Partnership of genes

A group of genes S such that does not exist a proper (\subset) subset of S which contributes in changing the worth of genes outside S .

Example

These two sets are partnerships of genes in the corresponding Microarray game

| | s1 | s2 | s3 |
|----|----|----|----|
| g1 | 0 | 1 | 1 |
| g2 | 0 | 1 | 1 |
| g3 | 1 | 0 | 1 |

Property 3: Partnership Monotonicity (PM)

(N, v) a microarray game. If two partnerships of genes S and T , with $|T| \geq |S|$ are such that they are

-*disjoint* ($S \cap T = \emptyset$),

-*equivalent* ($v(S) = v(T)$)

-*exhaustive* ($v(S \cup T) = v(N)$),

then genes in the smaller *partnership* S must receive more relevance than genes in T .

Example

| | s_1 | s_2 | |
|----------|-------|-------|---|
| ψ_1 | g1 | 0 | 1 |
| ψ_2 | g2 | 0 | 1 |
| ψ_3 | g3 | 1 | 0 |
| ψ_4 | g4 | 1 | 0 |
| ψ_5 | g5 | 1 | 0 |

$$\psi_i \geq \psi_k$$

For each

$$i \in \{1, 2\}$$

$$k \in \{3, 4, 5\}$$

Property 4: **Partnership Rationality (PR)**

The total amount of power index received from players of a partnership S should not be smaller than $v(S)$

Property 5: **Partnership Feasibility (PF)**

The total amount of power index received from players of a partnership S should not be greater than $v(N)$

Theorem (Moretti, Patrone, BOnassi (2007)):

The Shapley value is the unique solution which satisfies NG, ES, PM, PR, PF on the class of microarray games.

Networks and biology

- Network based methods have been found useful in biology,
 - protein interaction networks
 - gene regulatory networks
 - **gene co-expression networks**
 - ...
- The structure of a network can formally be represented by a graph $G = (V, E)$
 - The vertex set contains the genes: $V = \{x_{\text{gene}}, y_{\text{gene}}, z_{\text{gene}}, \dots\}$
 - The **edge set** contains *interactions*.

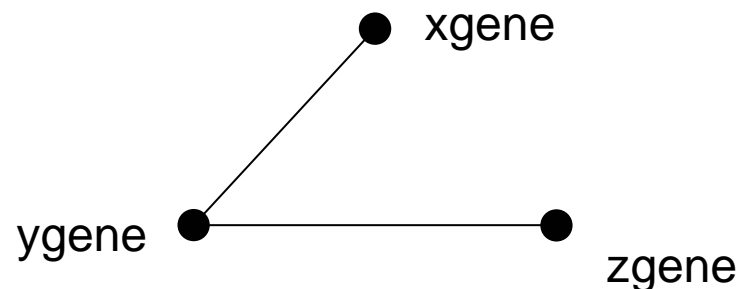
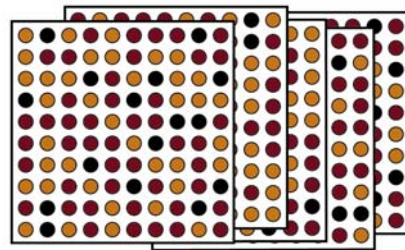


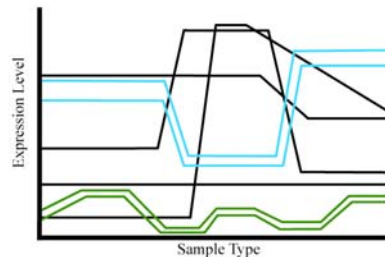
Figure 1

A Array Data



Data contains correlations

B Correlation Analysis



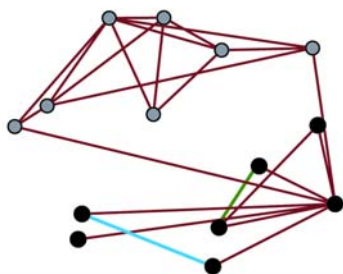
Correlation coefficients for all genes

C Correlation Matrix

| | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | G10 | G11 | G12 | G13 | G14 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| G1 | 1 | 0.9 | 0.9 | 0.9 | 0.9 | 0.8 | 0.9 | 0.1 | 0.9 | 0.1 | 0.1 | 0.8 | 0.2 | 0.2 |
| G2 | 0.9 | 1 | 0.9 | 0.3 | 0.3 | 0.7 | 0.0 | 0.5 | 0.3 | 0.1 | 0.1 | 0.2 | 0.4 | 0.3 |
| G3 | 0.9 | 0.9 | 1 | 0.9 | 0.0 | 0.2 | 0.5 | 0.7 | 0.6 | 0.5 | 0.2 | 0.6 | 0.1 | 0.0 |
| G4 | 0.9 | 0.3 | 0.9 | 1 | 0.5 | 0.3 | 0.6 | 0.3 | 0.0 | 0.5 | 0.1 | 0.2 | 0.2 | 0.6 |
| G5 | 0.9 | 0.3 | 0.0 | 0.5 | 1 | 0.1 | 0.6 | 0.1 | 0.3 | 0.3 | 0.3 | 0.5 | 0.2 | 0.5 |
| G6 | 0.8 | 0.7 | 0.2 | 0.3 | 0.1 | 1 | 0.9 | 0.2 | 0.1 | 0.1 | 0.5 | 0.3 | 0.1 | 0.1 |
| G7 | 0.9 | 0.0 | 0.5 | 0.6 | 0.6 | 0.9 | 1 | 0.3 | 0.1 | 0.5 | 0.1 | 0.3 | 0.5 | 0.2 |
| G8 | 0.1 | 0.5 | 0.7 | 0.3 | 0.1 | 0.2 | 0.3 | 1 | 0.9 | 0.9 | 0.9 | 0.8 | 0.8 | 0.9 |
| G9 | 0.9 | 0.3 | 0.6 | 0.0 | 0.3 | 0.1 | 0.1 | 0.9 | 1 | 0.8 | 0.1 | 0.3 | 0.5 | 0.3 |
| G10 | 0.1 | 0.1 | 0.5 | 0.5 | 0.3 | 0.1 | 0.5 | 0.9 | 0.8 | 1 | 0.8 | 1.0 | 0.2 | 0.3 |
| G11 | 0.1 | 0.1 | 0.2 | 0.1 | 0.3 | 0.5 | 0.1 | 0.9 | 0.1 | 0.8 | 1 | 0.5 | 0.8 | 0.9 |
| G12 | 0.8 | 0.2 | 0.6 | 0.2 | 0.5 | 0.3 | 0.3 | 0.8 | 0.3 | 1.0 | 0.5 | 1 | 0.8 | 0.1 |
| G13 | 0.2 | 0.4 | 0.1 | 0.2 | 0.2 | 0.1 | 0.5 | 0.8 | 0.5 | 0.2 | 0.8 | 0.8 | 1 | 0.9 |
| G14 | 0.2 | 0.3 | 0.0 | 0.6 | 0.5 | 0.1 | 0.2 | 0.9 | 0.3 | 0.3 | 0.9 | 0.1 | 0.9 | 1 |

Convert into Adjacency Matrix and Network

D Coexpression Network

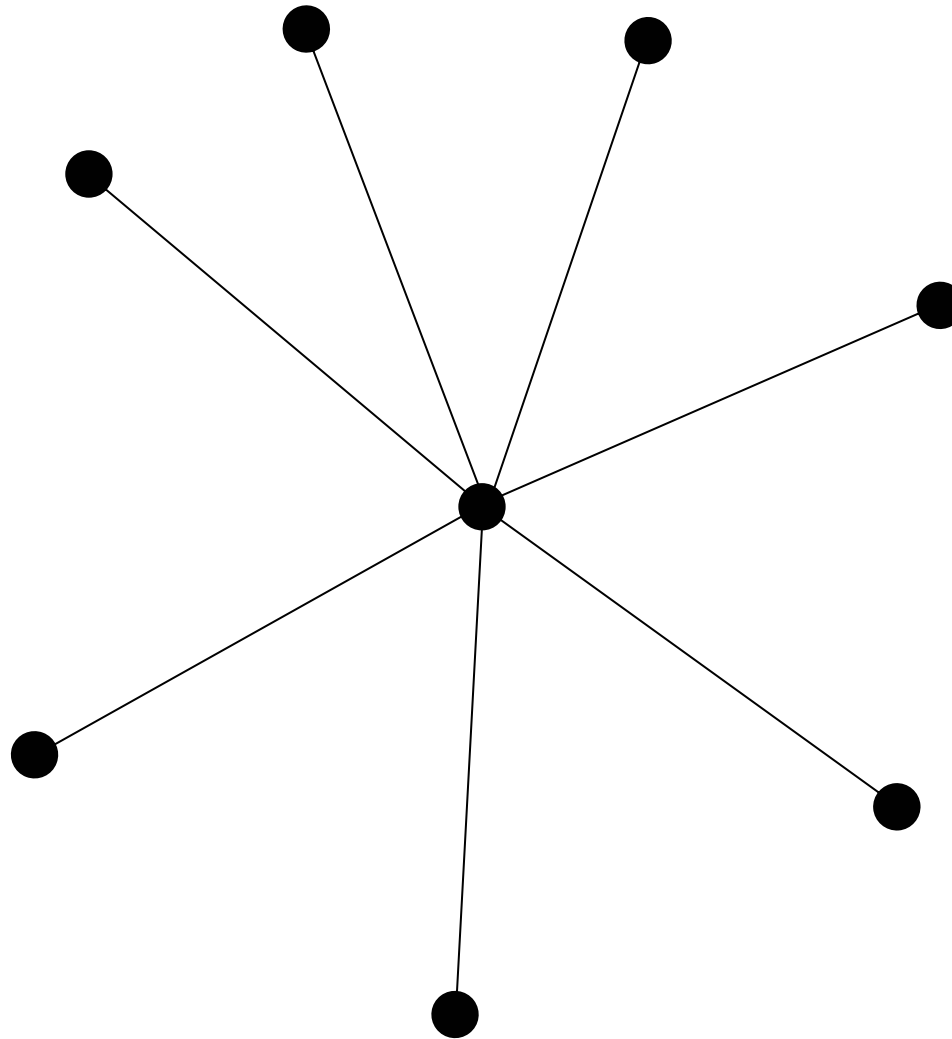


Steps for constructing a co-expression network

- A) Microarray gene expression data
- B) Measure concordance of gene expression with a Pearson correlation
- C) The Pearson correlation matrix is either dichotomized to arrive at an adjacency matrix → unweighted network
Or transformed continuously with the power adjacency function → weighted network

Centrality tries to measure the 'importance' of a vertex

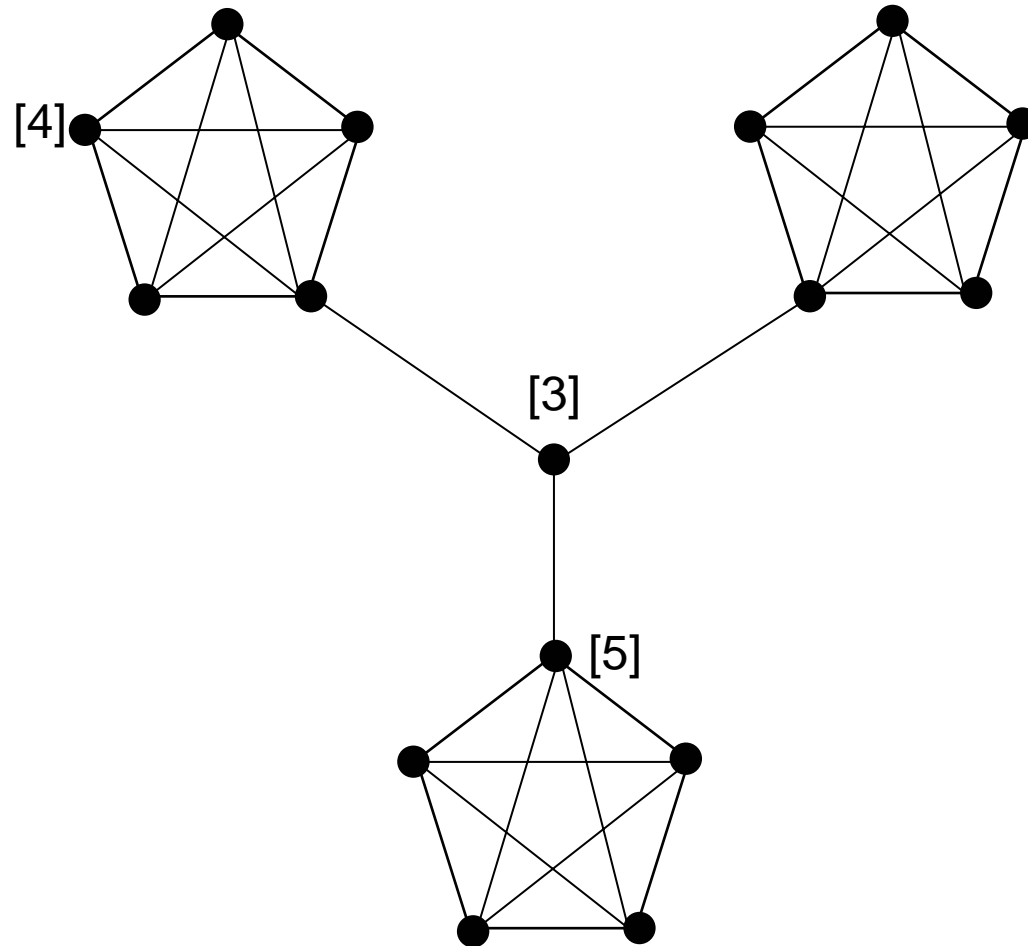
- **Degree centrality:** “How many nodes are connected to me?”
- **Closeness centrality:** “How close am I from all other nodes?”
- **Betweenness centrality:** “How important am I for shortest connections of any two other nodes?”



The maximal centrality is attained by the hub of a star, for all centrality measures.

(Shaw, 1954, and Nieminen, 1974)

Degree centrality

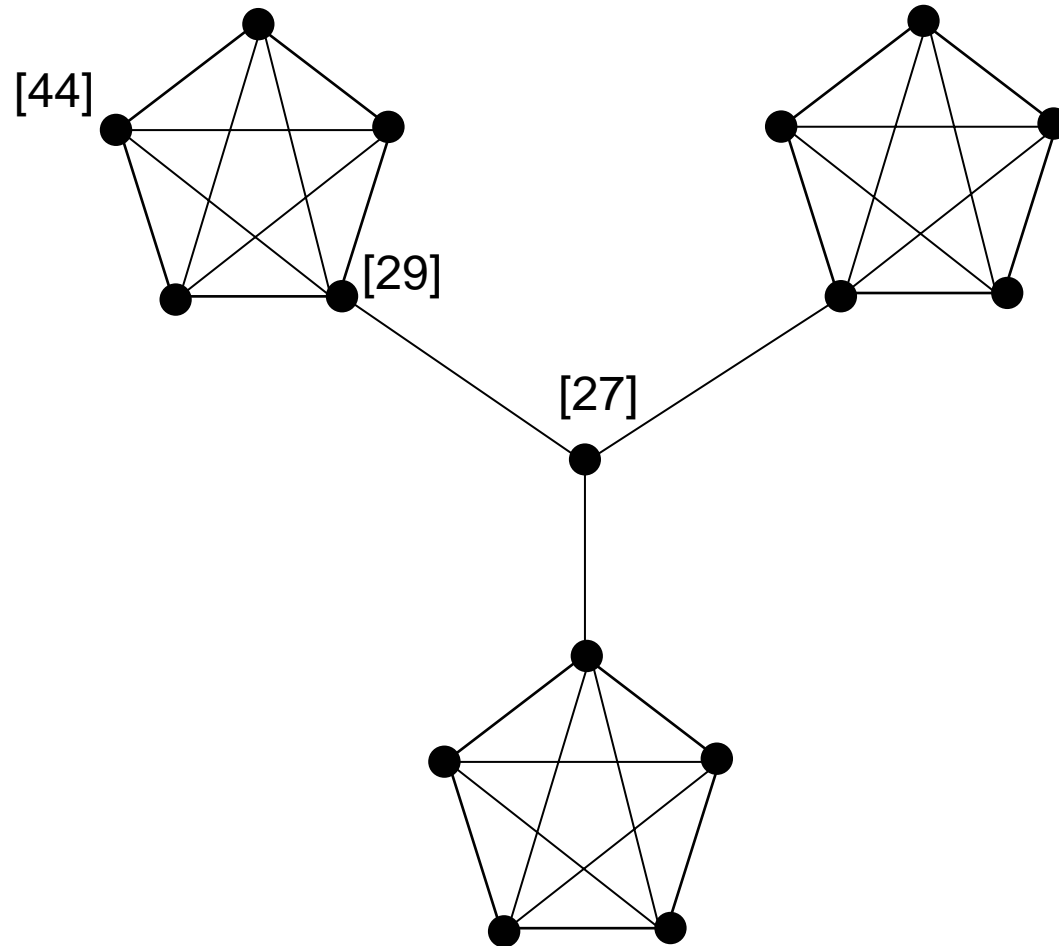


How many nodes are connected me?

(Beauchamp, 1965 and Sabidussi, 1966)

Closeness centrality

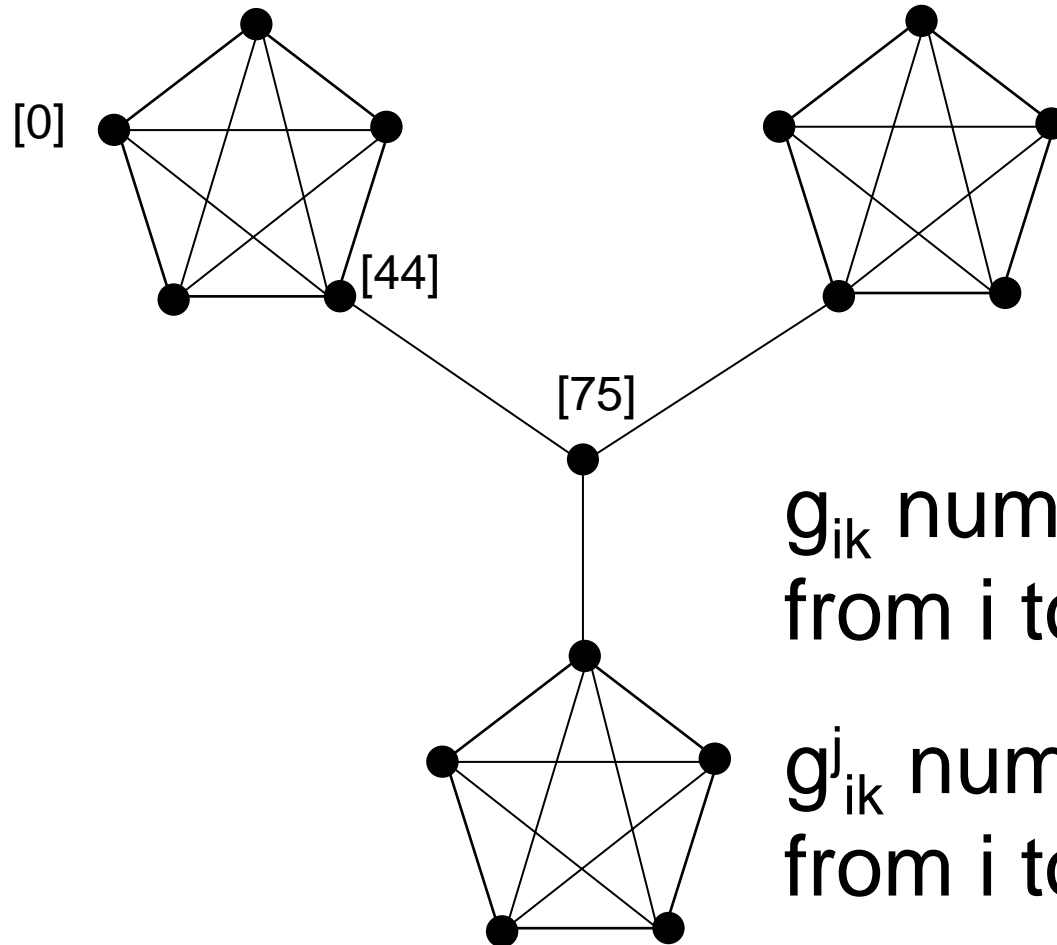
How close am I from all other nodes?



(Bavelas, 1948 and Freeman, 1977)

Betweenness centrality

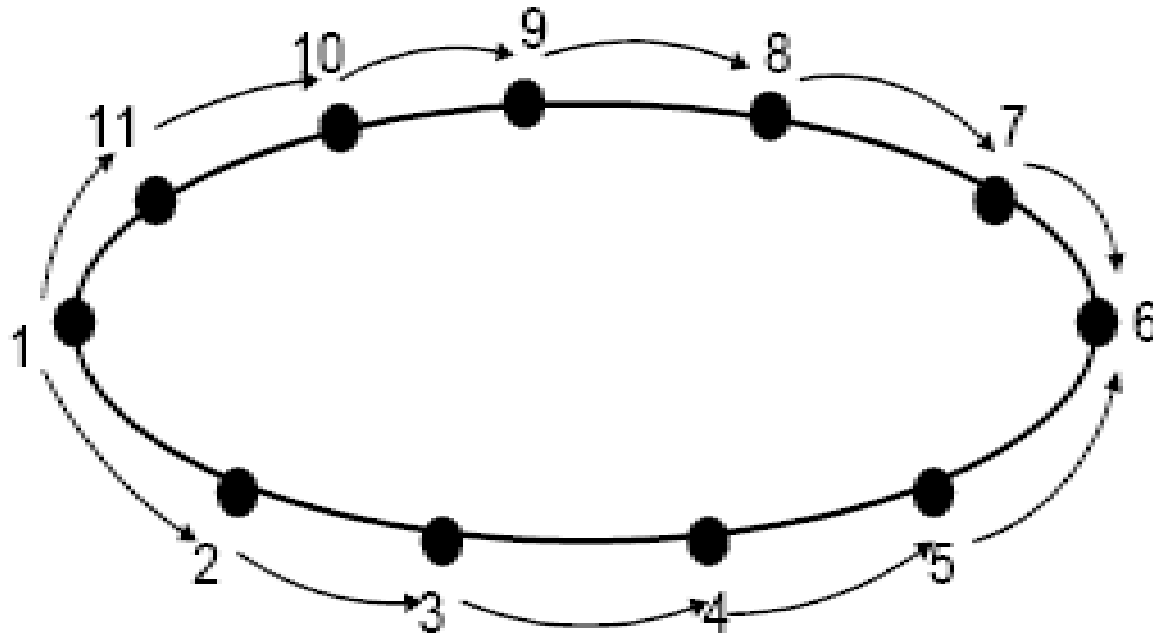
How important am I for the connection of every pair of nodes?



g_{ik} number of geodesics from i to k

g_{ik}^j number of geodesics from i to k which contain j

$$b^j = \sum_{k \neq i, i \neq j, k \neq j} g_{ik}^j / g_{ik}$$



it would seem foolish to discard the possibility of using a longer path, simply because a (slightly) shorter one exists.

Lethality and centrality

- Comprehensive efforts to determine the functional consequences of individual gene deletions in yeast (Giaever et al., *Nature*, 2002).
- Rank genes according to their degree centrality and correlate this with the phenotypic effect of their individual removal from the yeast genome and proteome.
 - the likelihood that removal of a protein will prove lethal correlates with centrality in protein networks (Jeong et al. *Nature* 2001, **411**).
 - analysis of unweighted gene co-expression networks have revealed a relationship between centrality and essentiality across all genes (Provero [arXiv:cond-mat/0207345], 2002).
 - strong positive correlations between gene connectivity within the whole weighted network and gene lethality (Carlson, *BMC Genomics*, 2006).

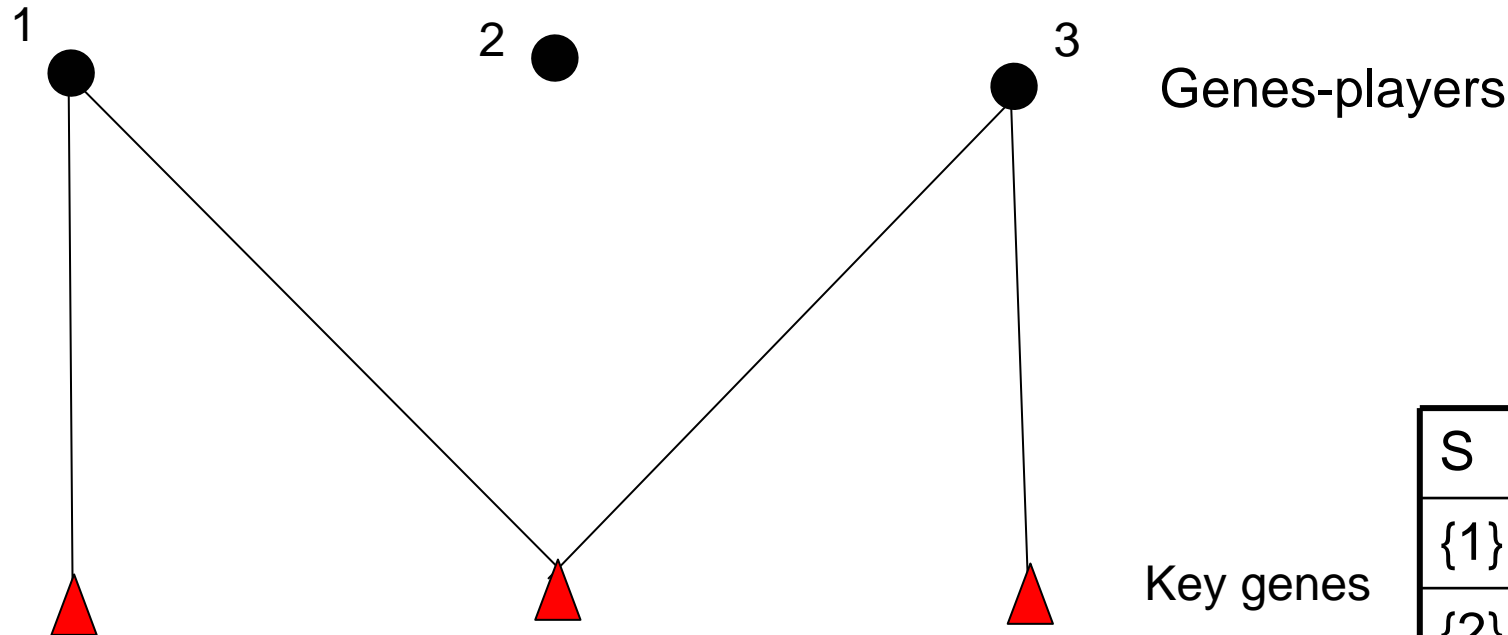
Co-expression network games

- Use a co-expression network (N, E) as a communication network
- The set N of players is the set of genes studied
- Links in E are co-expression relations
- What is an *a priori* game (N, v) ?

Co-expression network games

- a finer resolution of gene interaction investigated in the model, which is based on pair-wise relationships of genes in the network.
- integration of *a priori* knowledge concerning the co-expression with key genes, which may be obtained by previous studies.

a priori game (N, v) : the worth $v(S)$ of a coalition of genes in S is the number of key genes associated only to genes in S

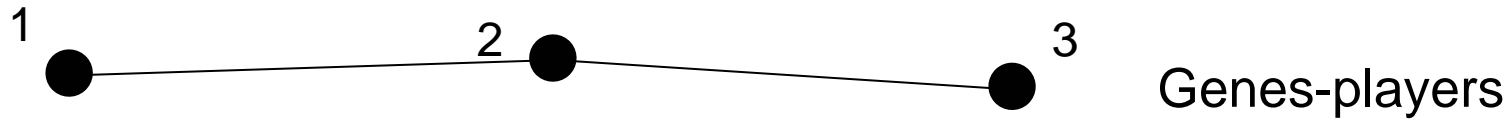


Key genes

| S | v | $\phi(v)$ |
|---------|---|-----------|
| {1} | 1 | 1.5 |
| {2} | 0 | 0 |
| {3} | 1 | 1.5 |
| {1,2} | 1 | |
| {1,3} | 3 | |
| {2,3} | 1 | |
| {1,2,3} | 3 | |

* Means that $v(S)$ is the number of key genes connected to S and not connected to gene-players out of S

Communication network: a co-expression network from experimental data ($\{1,2,3\},E$)



A priori game

| S | v | $\varphi(v)$ |
|---------|---|--------------|
| {1} | 1 | 1.5 |
| {2} | 0 | 0 |
| {3} | 1 | 1.5 |
| {1,2} | 1 | |
| {1,3} | 3 | |
| {2,3} | 1 | |
| {1,2,3} | 3 | |

Graph-restricted game

| S | w_E | $\varphi(w_E)$ |
|---------|-------|----------------|
| {1} | 1 | 4/3 |
| {2} | 0 | 1/3 |
| {3} | 1 | 4/3 |
| {1,2} | 1 | |
| {1,3} | 2 | |
| {2,3} | 1 | |
| {1,2,3} | 3 | |

| γ |
|----------|
| -1/6 |
| 1/3 |
| -1/6 |