

M1 - BIM

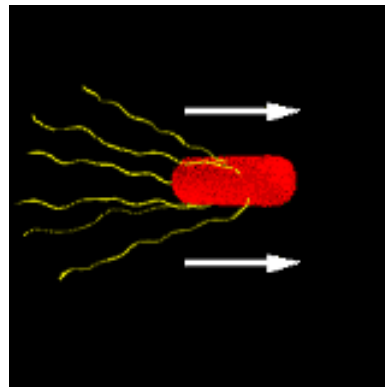
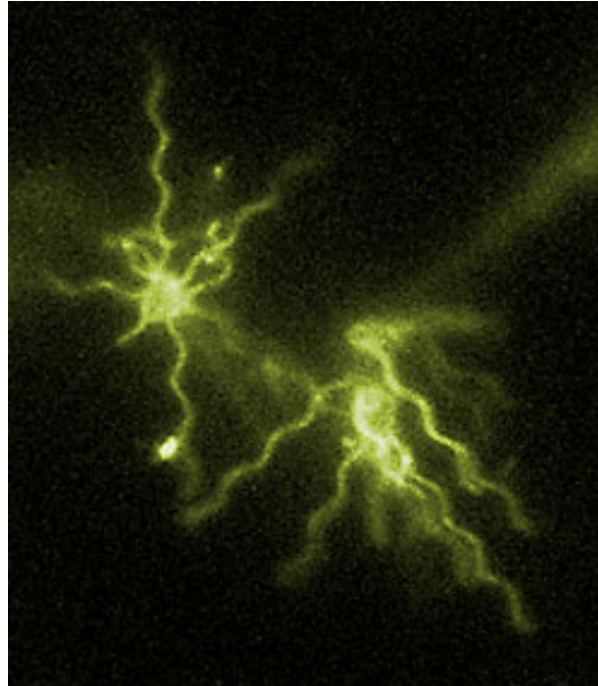
Réseaux biologiques et algorithmes pour leur
génération et analyse

Alessandra Carbone
Université Pierre et Marie Curie

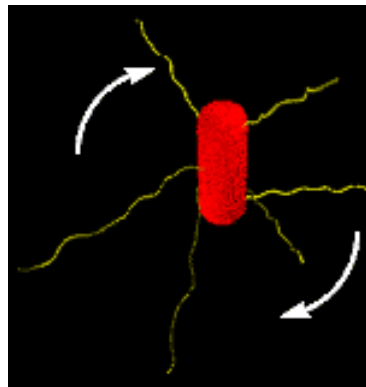
L'étude computationnel des systèmes biologiques est caractérisé par l'analyse des **relations entre** facteurs biologiques **plutôt que sur** les facteurs biologiques en soi même.

Les fonctions biologiques sont dérivées d'un **réseau d'interactions** entre molécules qui est **complexe, redondant** et **robuste**. Il traite l'information, prend des décisions sophistiquées et survit dans un environnement dynamique.

Exemple: la chemotaxis

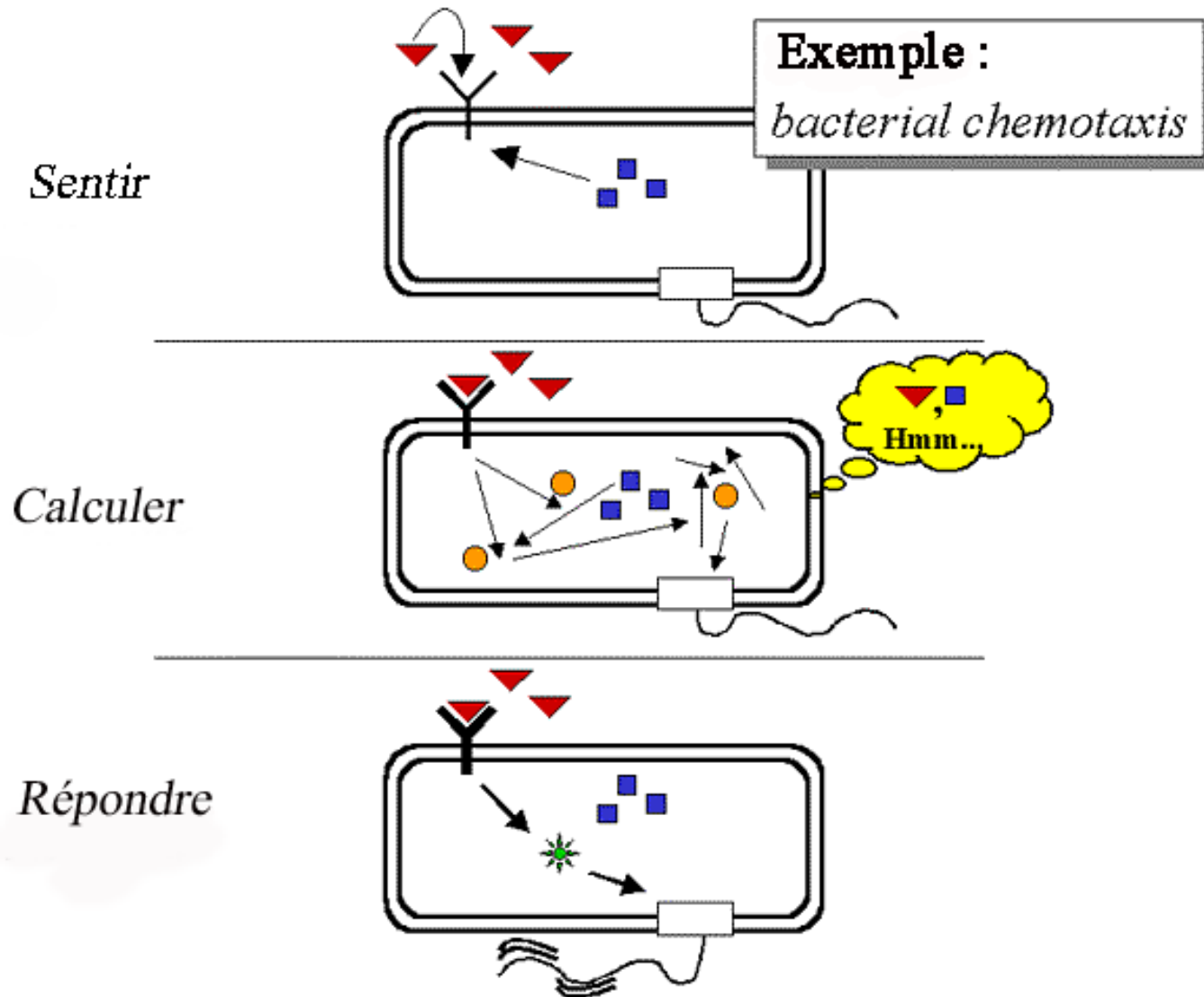


nager...



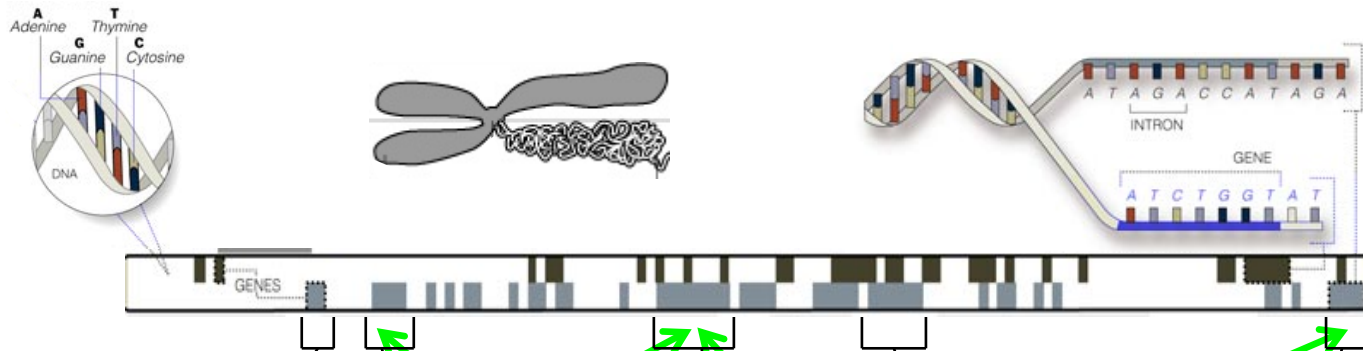
tâter...

Cellules utilisent des réseaux de protéines pour contrôler leurs comportement



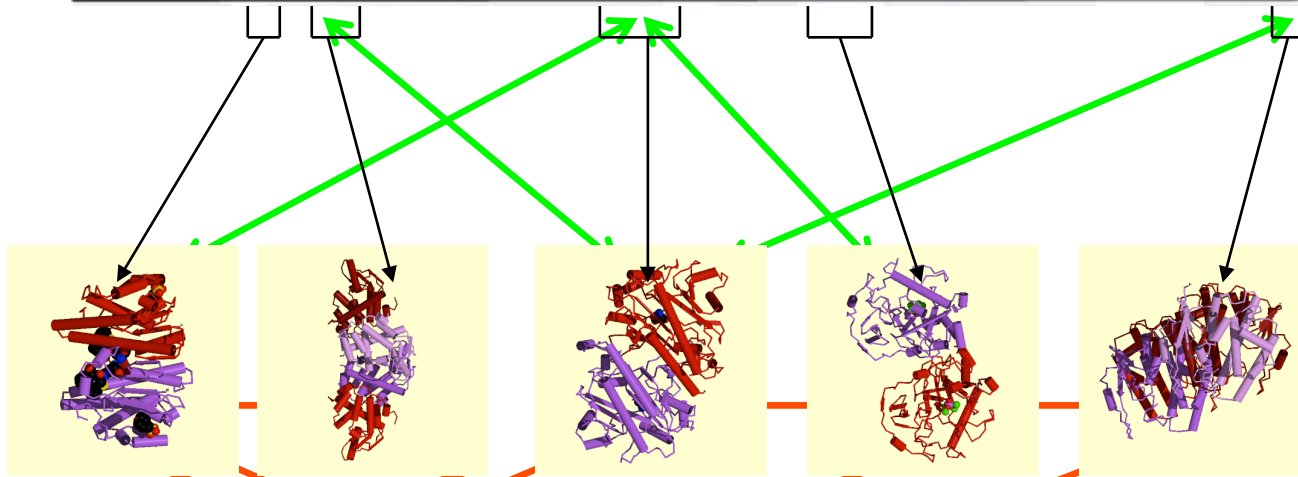
Ils existent plusieurs types d'interaction biologique:

1. Control de la transcription – relation entre gènes et protéines
2. Protéine-protéine
3. Métabolites et protéines



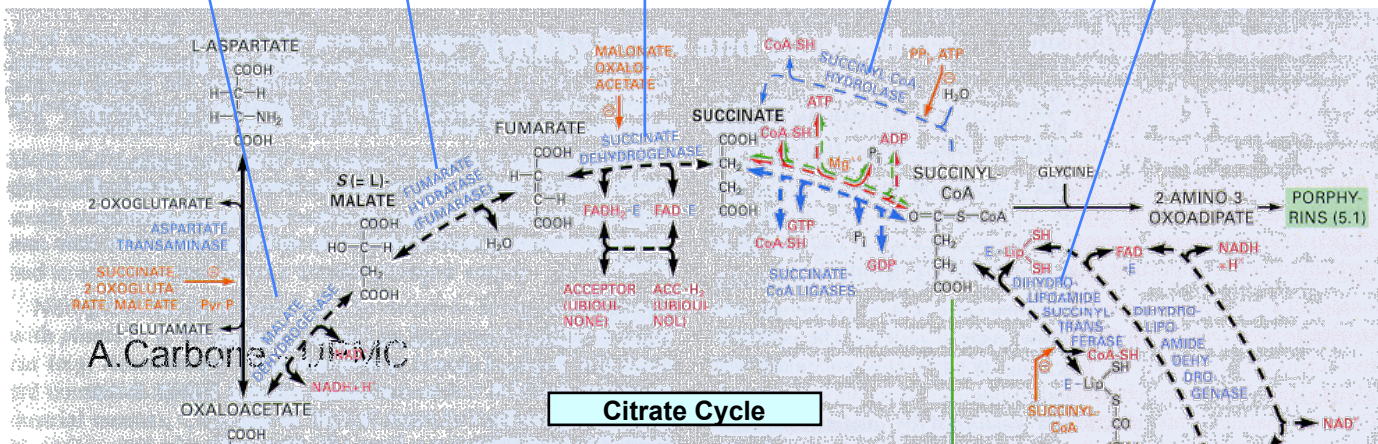
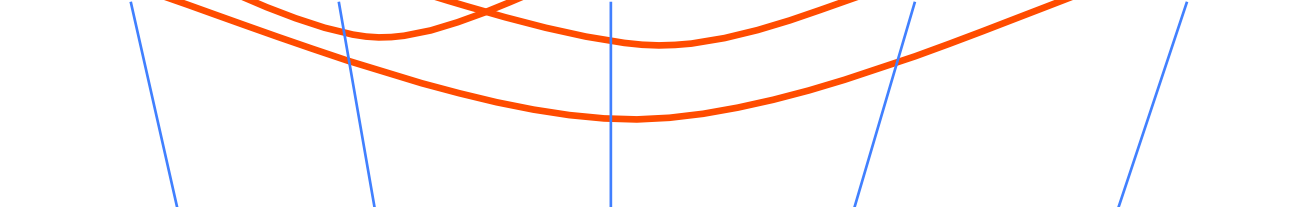
GENOME

Interaction
gènes -
protéines



PROTEOME

Interactions
protéine -
protéine



METABOLISME

Interactions
protéines
métabolites

Réactions bio-⁶
chimiques

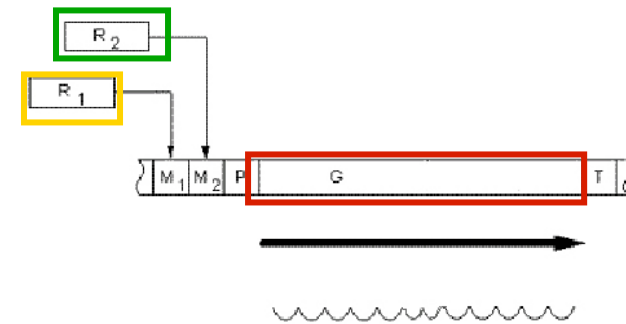
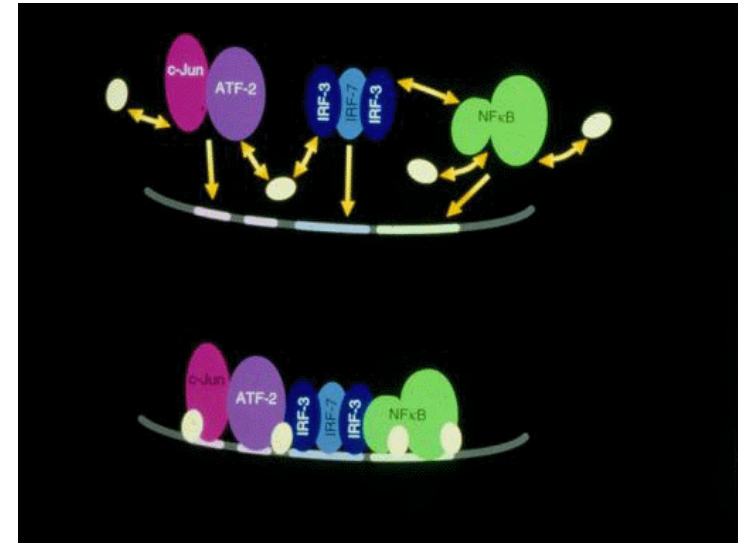
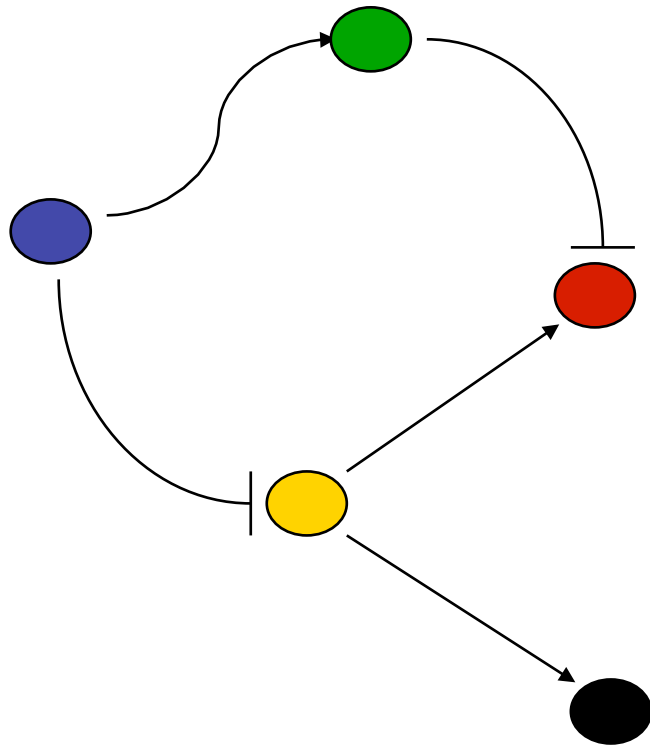
1. Contrôle de la transcription – réseau transcriptionnel

Chaque cellule contient une copie de l'ADN de l'organisme, mais les gènes ne sont pas tout le temps produits ni générés au même taux de production.

Des protéines spécifiques, appelées **facteurs de transcription**, rentrent en contact avec l'ADN à des différents endroits proches du début du gène (entre 50 et 20000bp) et catalysent la réaction qui démarre la transcription.

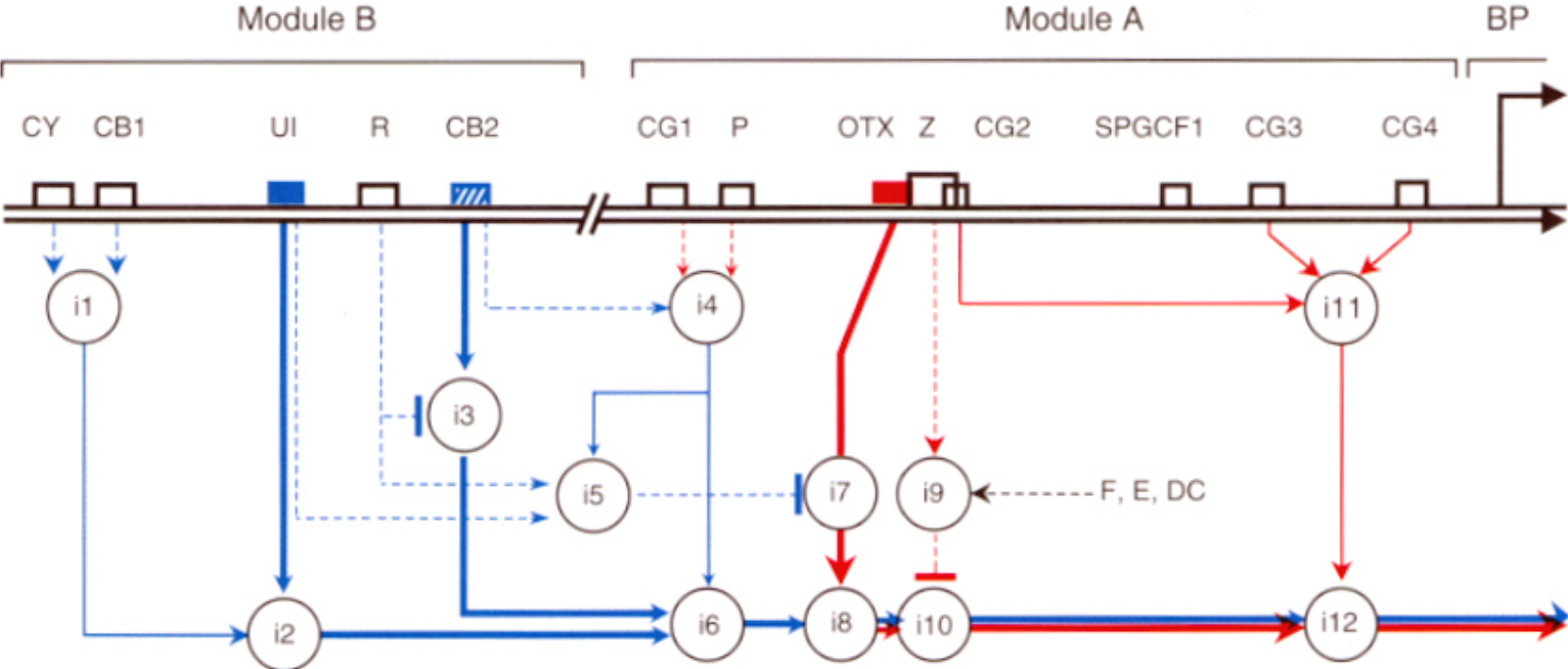
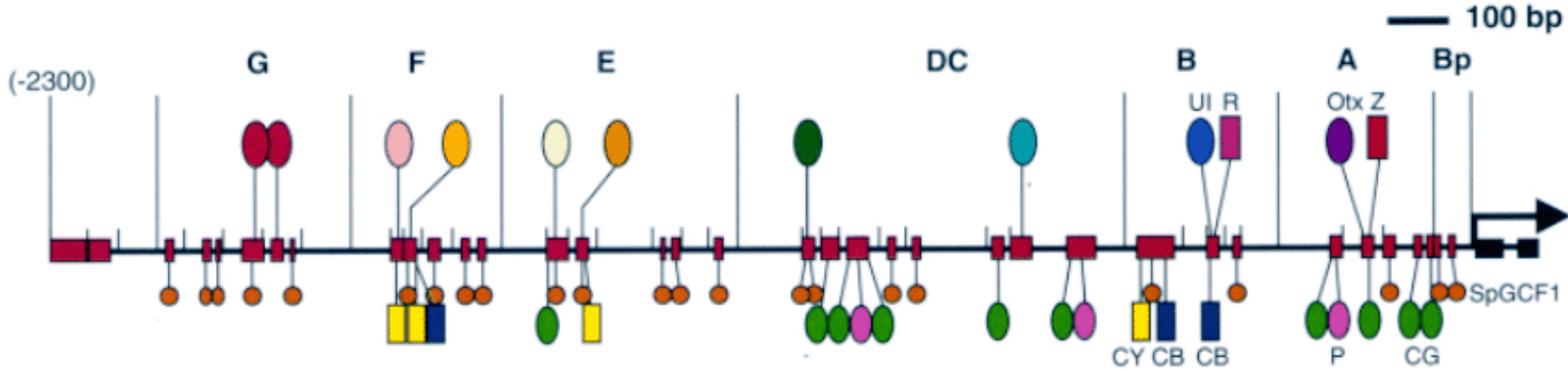
Ces protéines, sont eux même des produits des gènes, et elles sont sousjetttes au contrôle transcriptionnel. Pour cette raison on peut penser de modéliser ce système comme un réseau de gènes affectant la transcription l'un de l'autre de façon dynamique.

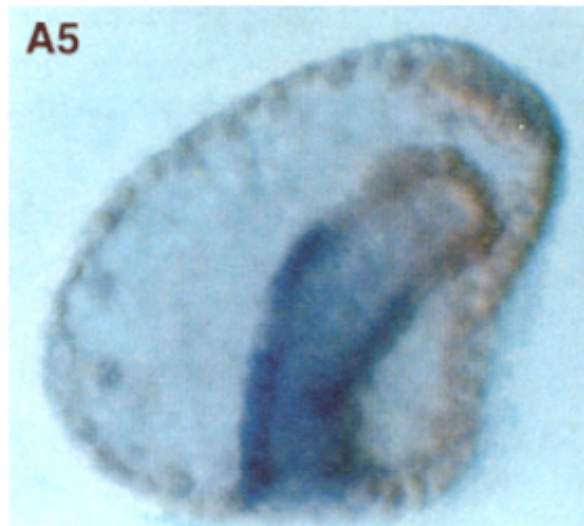
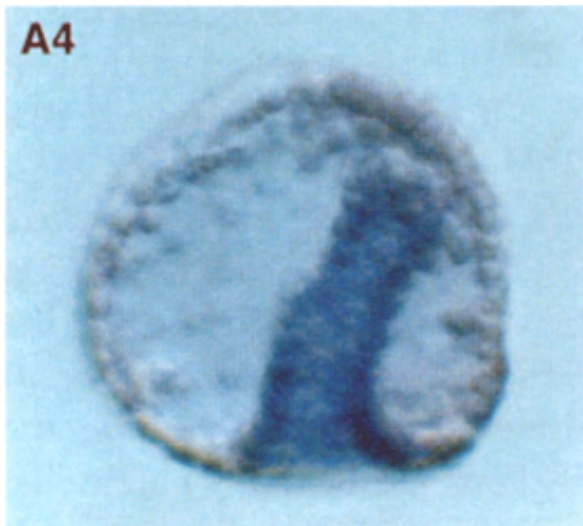
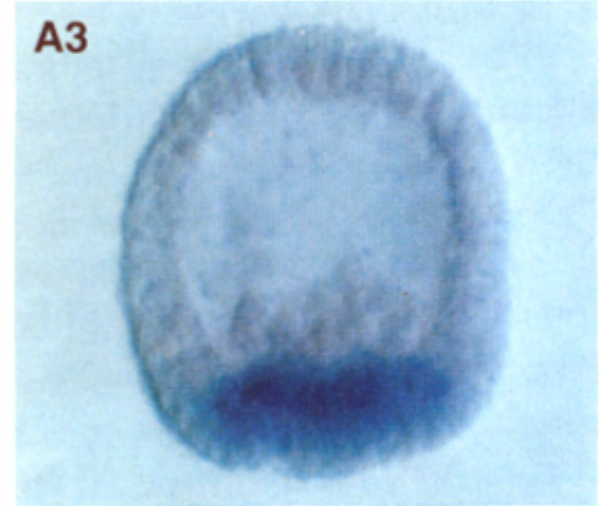
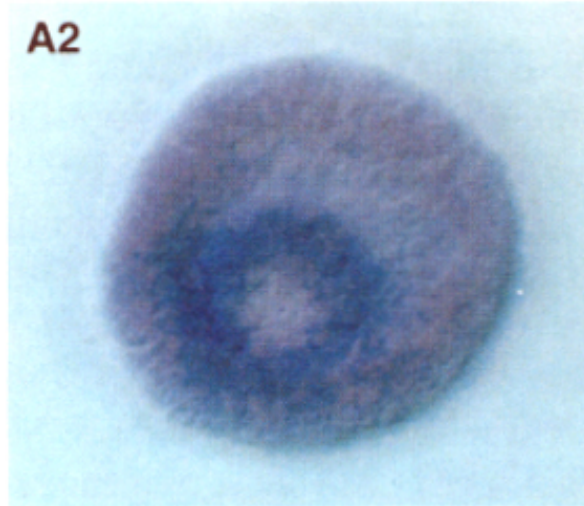
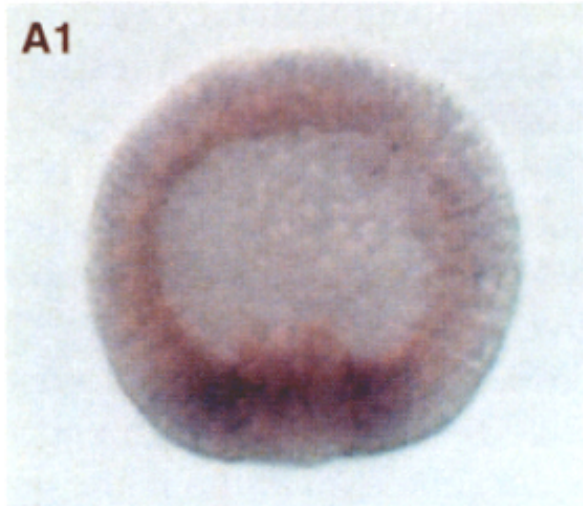
Eléments positifs et négatifs



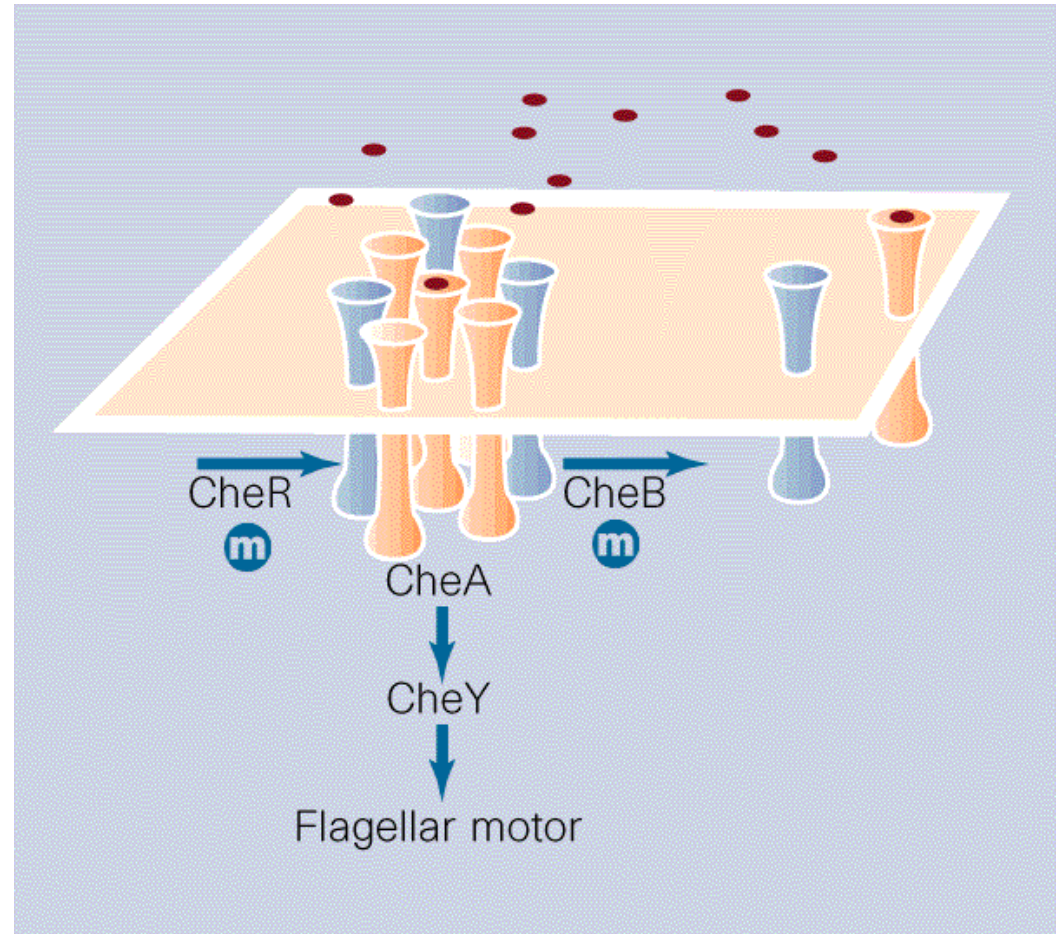
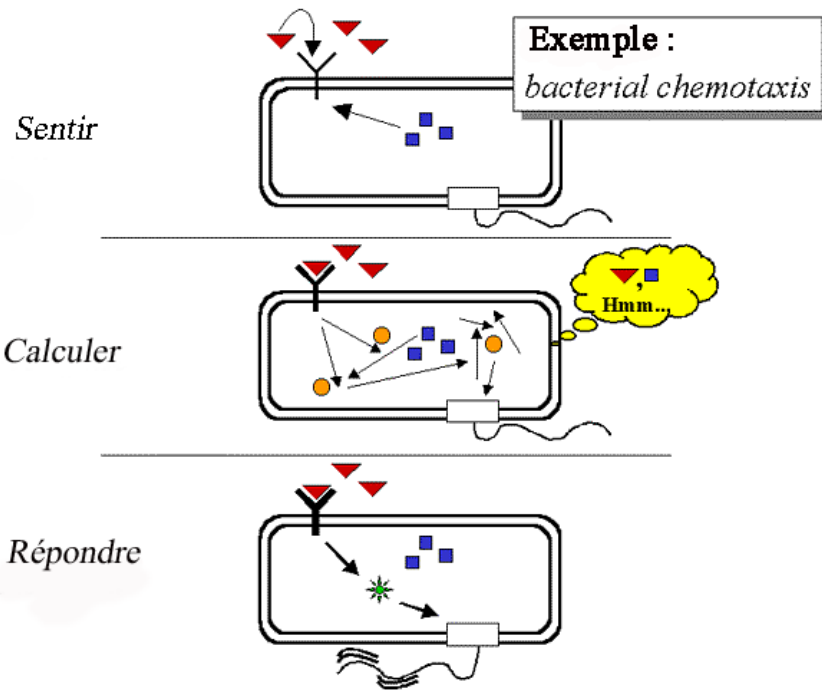
Presence		Expression
R ₁	R ₂	AND
Yes	Yes	OFF
Yes	No	ON
No	Yes	OFF
No	No	OFF

Exemple : le gène endo16





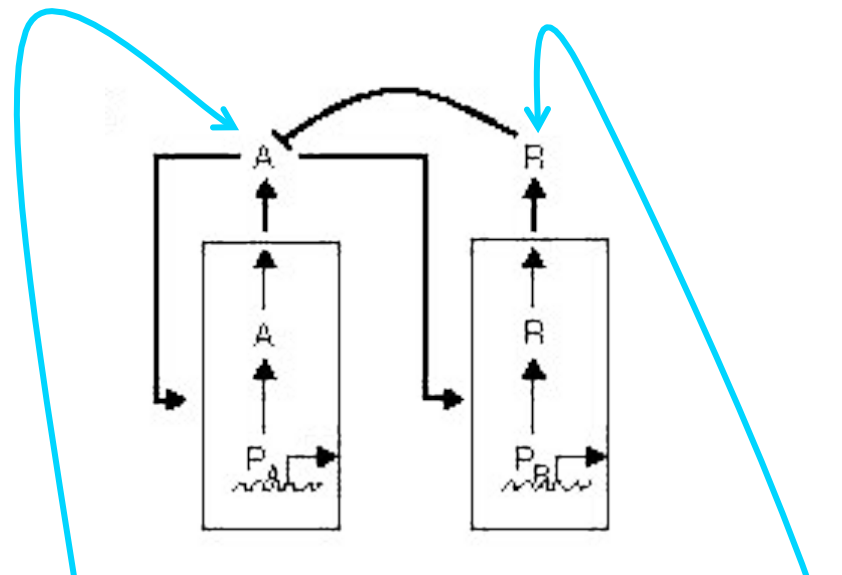
Chemotaxis: exemple d'homologie des gènes



L'homologie entre gènes montre des circuits topologiquement équivalents

A.Carbone - UPMC quand on change d'espèce.

Horloges Circadiens : exemple d'homologie des circuits



Eléments positifs

KaiA	<i>Synachococcus</i>
Clc and Cyc	<i>Drosophila</i>
Clock and Bmal	<i>Mice</i>

Eléments négatifs

KaiB KaiC	<i>Synachococcus</i>
Tim and Per	<i>Drosophila</i>
Tim and Per1,2	<i>Mice</i>

Interactions protéine-protéine

Il s'agit de graphes construits à partir de données des interactions **par paires de protéines** (par exemple, par des systèmes double-hybrides) ou **par groupes de protéines** (par spectrométrie de masse).

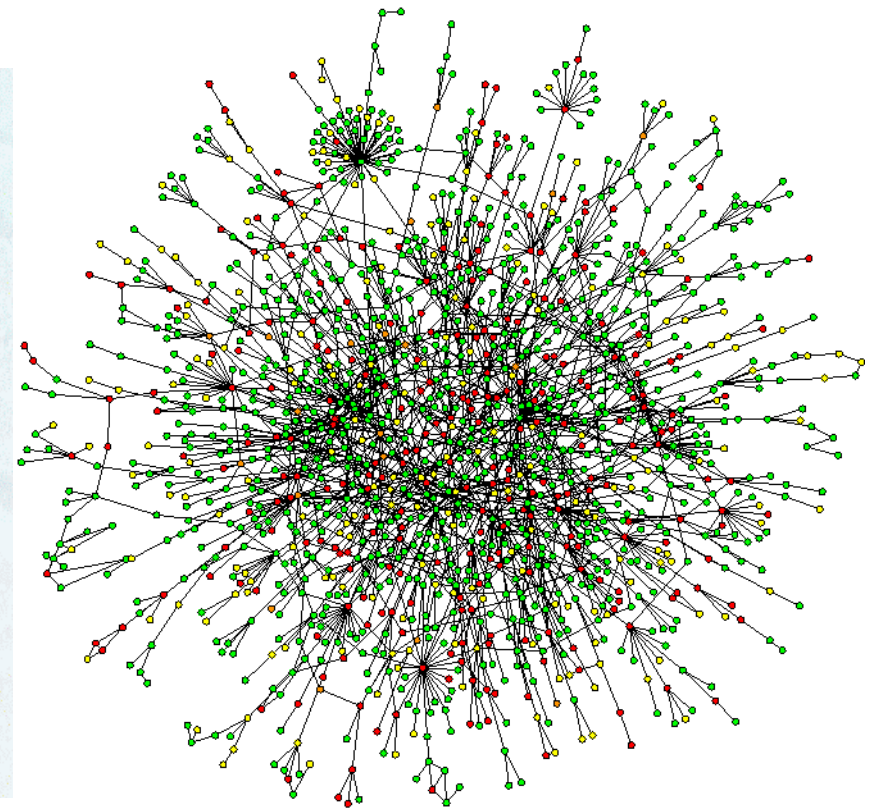
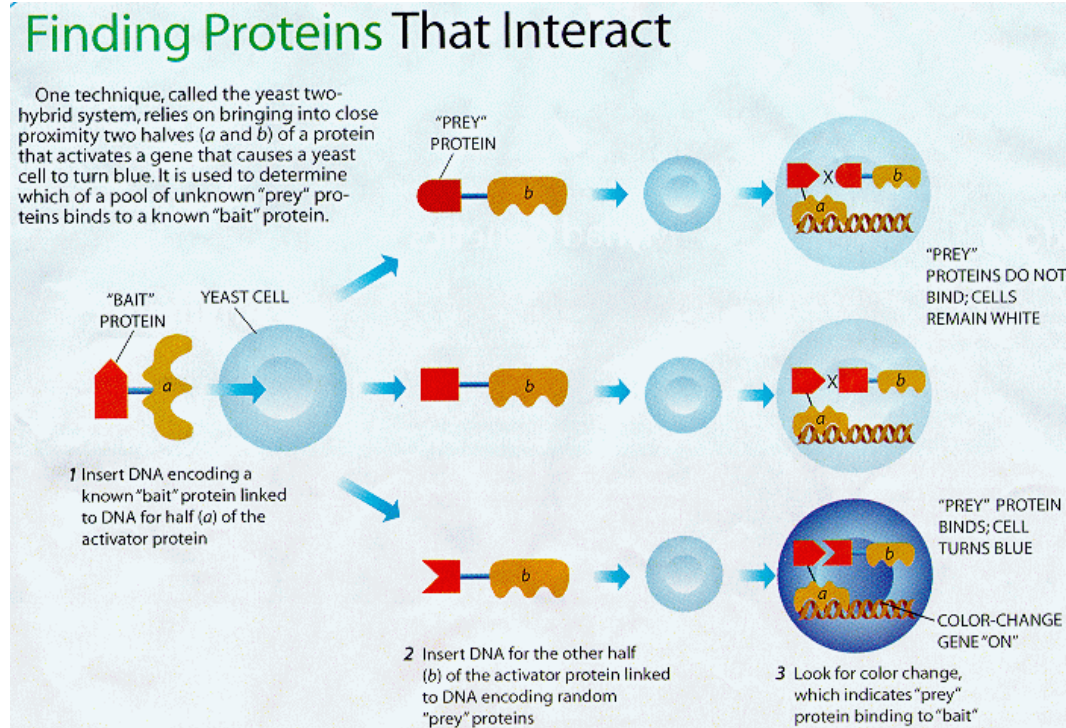
Les réseaux sont d'habitude assez différents :

- les **techniques expérimentales différentes** recouvrent des fractions différentes de l'ensemble complet des protéines d'un génome donné
- les données sont fortement dépendantes des protocoles expérimentaux.

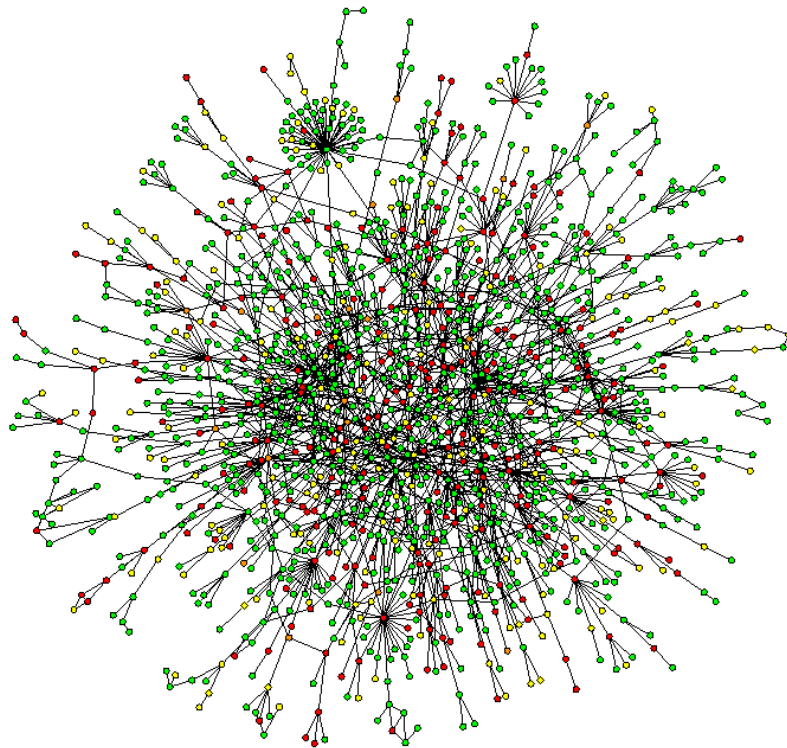
Réseau d'interaction de la levure

Noeuds: protéines

Arcs: interactions physiques (binding)

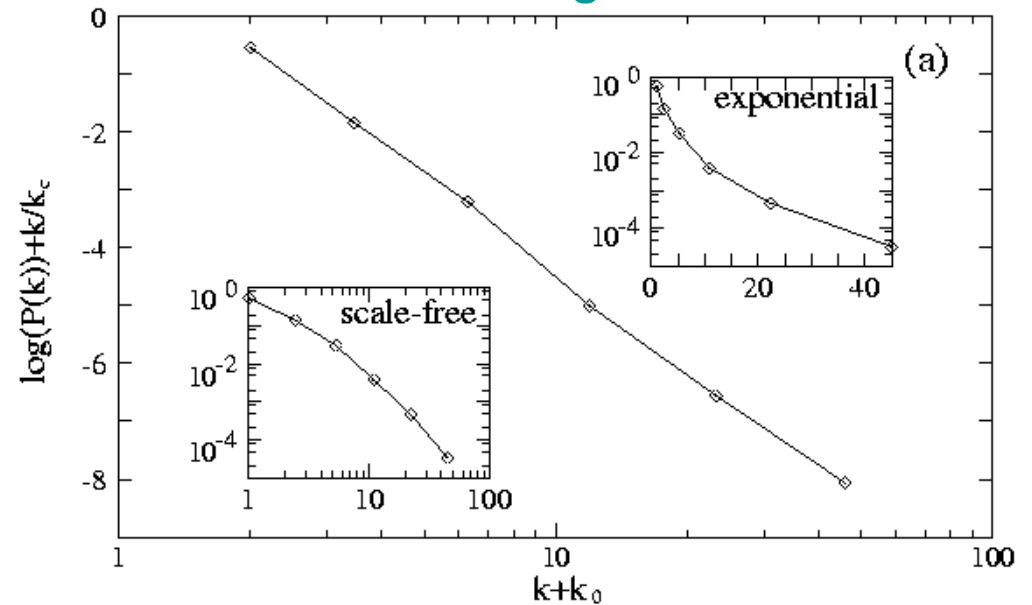


Topologie du réseau d'interaction protéine-protéine



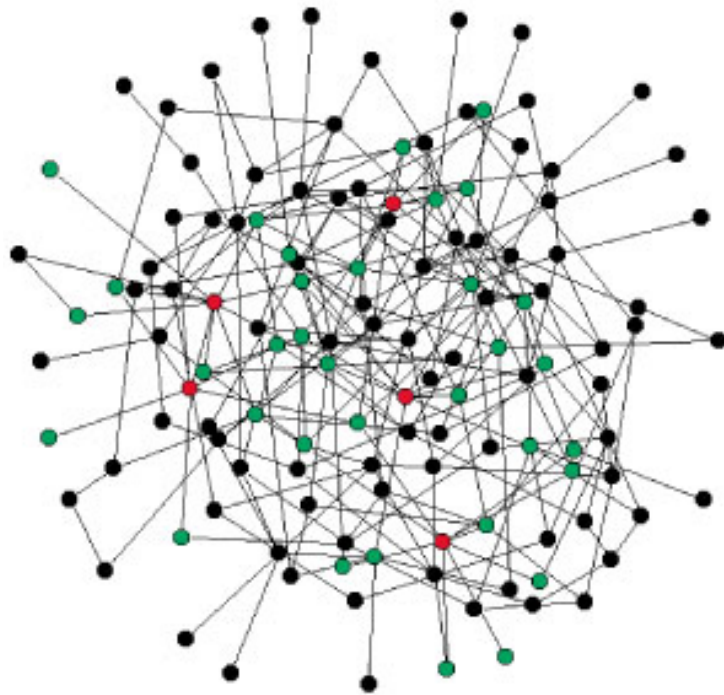
Protéines de la levure

Distribution des degrés des noeuds

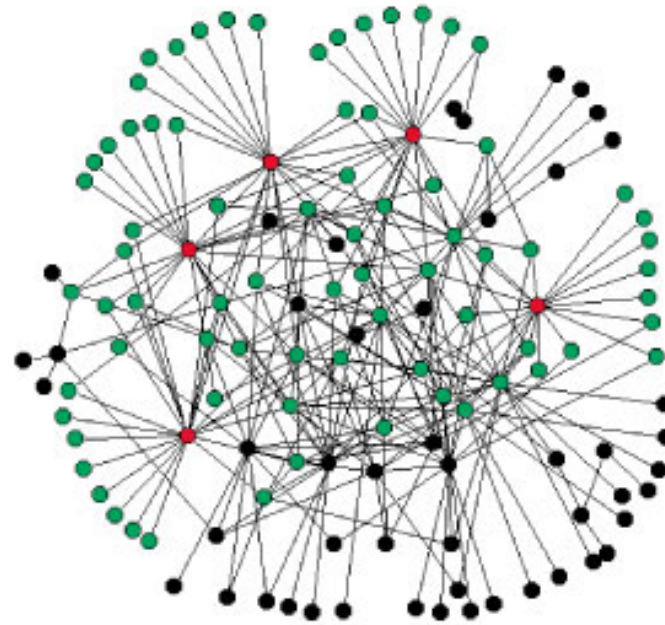


$$P(k) \sim (k + k_0)^{-\gamma} \exp\left(-\frac{k + k_0}{k_\tau}\right)$$

H. Jeong, S.P. Mason, A.-L. Barabasi & Z.N. Oltvai, **Nature**, 2001



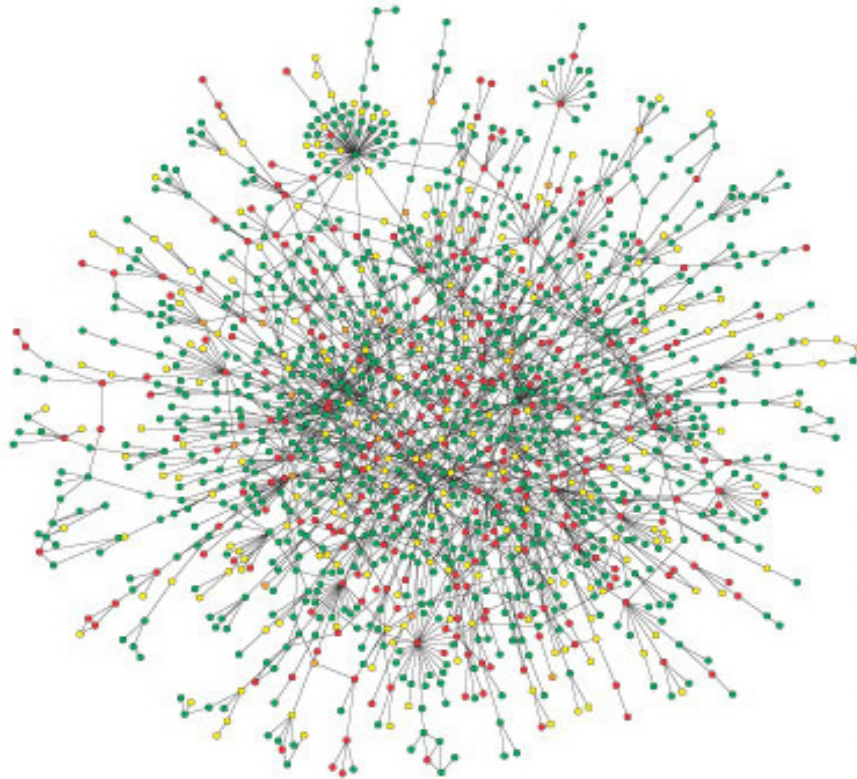
Exponential



Scale-free

Réseau exponentiel : il existe une « connectivité moyenne » des nœuds. La plupart des nœuds est connecté par un nombre donné de liens, et seulement très peu de nœuds ont un degré différent de la moyenne.

Réseau sans échelle : il n'y a **pas de moyenne significative de liens**. Le nombre de nœuds ayant un nombre donné de connections simplement décroît quand le nombre de nœuds augmente. **Plusieurs nœuds sont reliés au réseau par un seul arc**; moins encore ont 2 arcs, moins en ont 3 etc. Au contraire que pour les réseaux exponentiels, il y a un nombre limité mais significatif de nœuds ayant plusieurs connections.



1870 nœuds = protéines
2240 arcs = interactions

le graphe est épars (sparse graph)

Composante géante de 466 protéines : 78% du nombre totale des nœuds connectés.

Les autres 163 sous-graphes connectés sont petits.

Dans la grande composante :

Longueur moyenne du plus petit chemin est 6.8, cad le 4/5 de la longueur moyenne du plus petit chemin pour un graphe aléatoire.

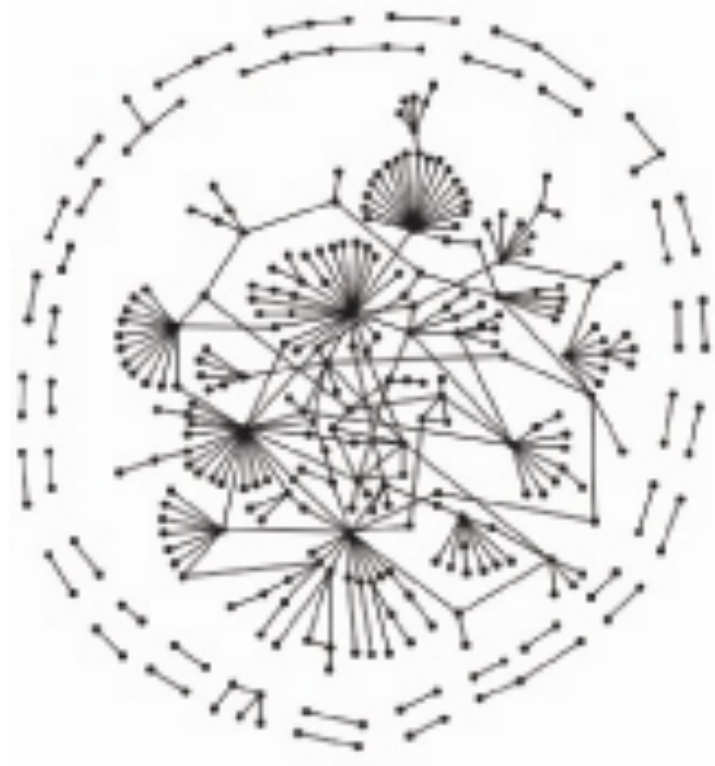
La distribution des degrés des nœuds décroît plus lentement que un exponentiel :

$$(k+k_0)^{-\gamma} e^{-k/k_{cut}}$$

où $\gamma \sim 2.4$, $k_0 = 1$ et $k_{cut} = 20$

← Valeurs dépendantes de la taille finie du réseau

Pour récapituler :



Partie d'un réseau typique d'interaction protéine-protéine (Maslov-Sneppen 2002)

Pourquoi étudier les interactions entre protéines:

Les protéines sont les “chevaux de trait” de la cellule :

- réalisent les réactions catalytiques, de transport, etc
- forment les capsides virales
- transmettent l'information de l'ADN via l'ARN, en traversant les membranes
- forment des canaux de communication et règlent leur comportement,
- rendent possible la synthèse de nouvelles protéines
- sont responsables de la dégradation de protéines qui ne sont plus nécessaires à la cellule
- permettent les réponses immunitaires

....

L'identification de la fonction d'une protéine est basée sur l'identification des ses partenaires. Si la fonction d'au moins un des partenaires est connue, alors on essaiera d'associer sa fonction et le chemin fonctionnel dans lequel la protéine est sensée participer. **Il est donc, à travers le réseaux intriqué de ces interactions que l'on peut espérer de réaliser une carte des chemins fonctionnels de la cellule, leur interconnectivités et leur régulation dynamique.**

Pourquoi étudier la structure des réseaux PPI?

- On recherche les propriétés partagées par les différents réseaux d'organismes
- Ces propriétés peuvent aider à connecter la structure à la fonction
- Position **relative** d'une protéine dans le réseau
- Corrélation de **modules fonctionnels conservés** avec les complexes protéiques.

Comment est-ce que l'on sait que les protéines interagissent?

- ***In vivo* (dans la cellule)**

- 1. Système double hybride

- A comprehensive two-hybrid analysis to explore yeast protein interactome – Ito,et.al

- 2. Spectrométrie de masse

- Functional organization of the yeast proteome by systematic analysis of protein complexes – Gavin et.al.

- 3. Corrélation des données d'expression (mRNA)

- Functional Discovery via a compendium of expression profiles – Hughes et.al.

- 4. Interactions génétiques

- MIPS: A database for genomes and protein sequences – Mewes et.al.

- ***In silico* (prédictions computationnelles)**

- Profiles phylogénétiques

- Rosetta stone

- Gene neighbors

- Co-evolution

- Gene clusters

Une présentation de toutes ces méthodes et une comparaison :

Comparative assessment of large-scale data sets of protein-protein interactions – von Mering

Problème des faux positifs et des faux négatifs

- Les cartes d'interaction pour la levure, *C.elegans* et la *Drosophile* ont été réalisées avec des **couvertures très basses**.
- Cette couverture très basse explique le chevauchement limité observé entre différents ensembles de données.
- **Faux négatifs** peuvent apparaître à partir
 - d'une couverture biaisée ou limitée
- **Faux positifs** peuvent apparaître à partir de
 - limites dans la procédure expérimentale
 - absence du milieu cellulaire

Nombre de protéines
(nœuds du réseau)

Nombre d'interactions
(arêtes du réseau)

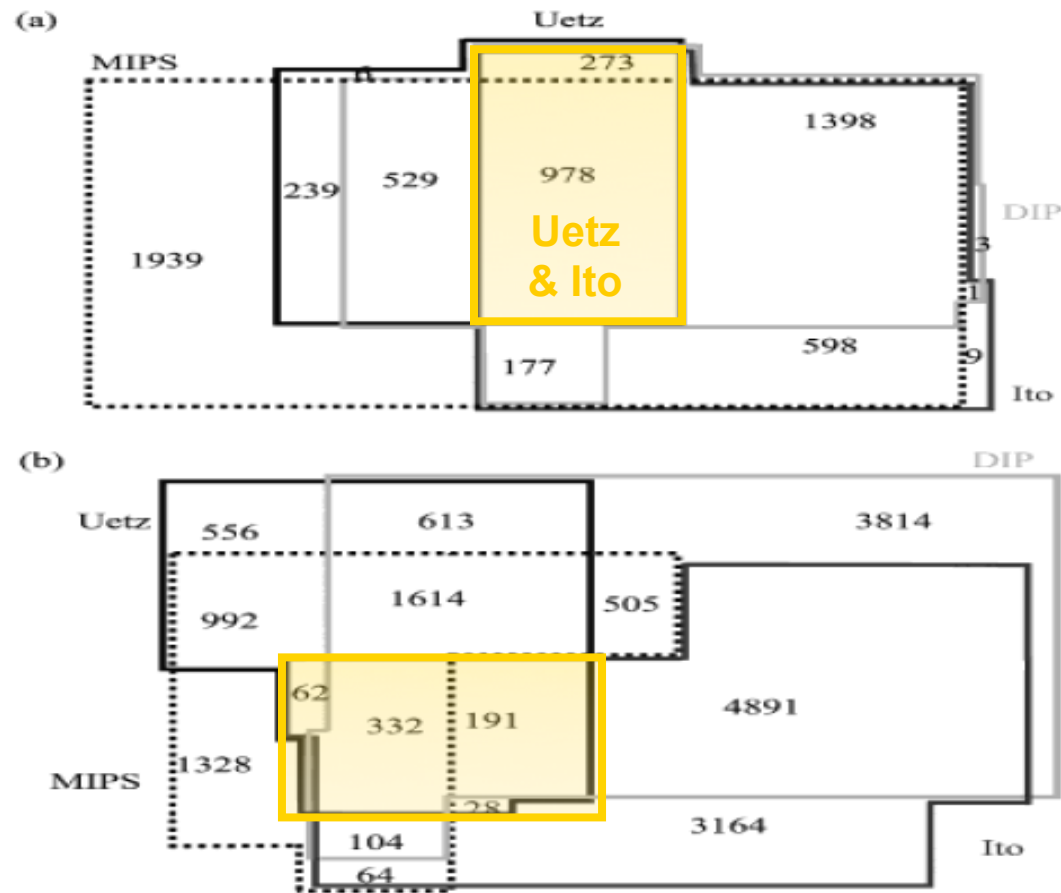


Figure 1. Relationship between the four studied databases (a) at the protein and at (b) the interaction level. The sum of the numbers within each color boundary denotes (a) the total number of proteins or (b) the total number of interactions found in the corresponding database. For example, (a) indicates that the MIPS database has altogether 6745 proteins, of which 1939 proteins with at least one interaction do not show interactions in any other databases. Note that while the overlap between the Uetz and Ito maps, (a) at the protein level is as high as 30%, (b) at the interaction level it is much smaller.

Métabolites et protéines : réseaux métaboliques

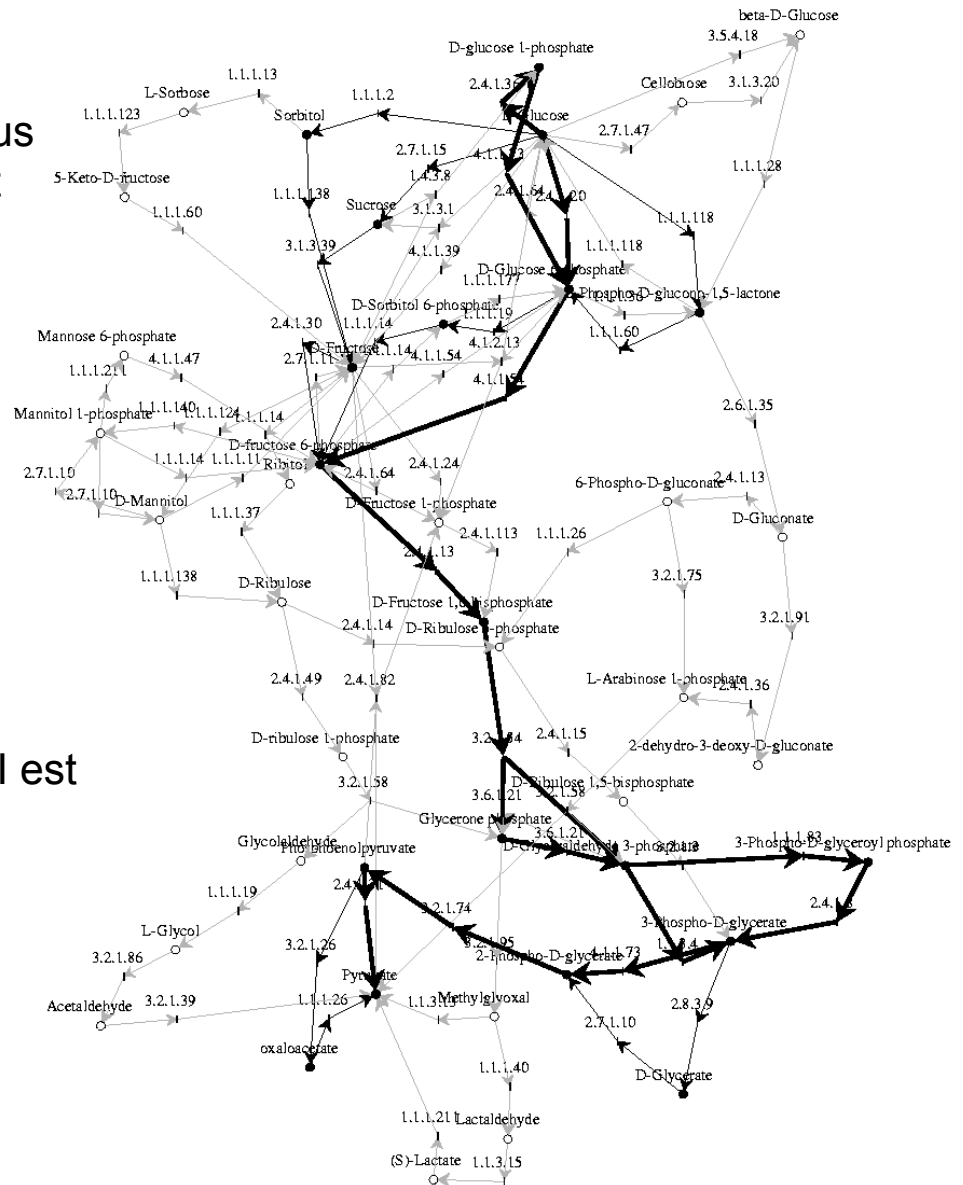
Les réseaux métaboliques représentent la bio-chimie d'un organisme, cad les processus qui génèrent la masse, l'énergie, le transfert d'information et la mort d'un organisme.

Graphes orientés:

arcs: métabolites/protéines (soustraits)

noeuds: réactions biochimiques

Une différence de base avec les réseaux PPI est que les réseaux PPI sont constitués par des **clusters disjoints** que dans les réseaux métaboliques sont connectés.



Biochemical Pathways

1

2

3

4

5

6

7

8

9

10

A

B

C

D

E

F

G

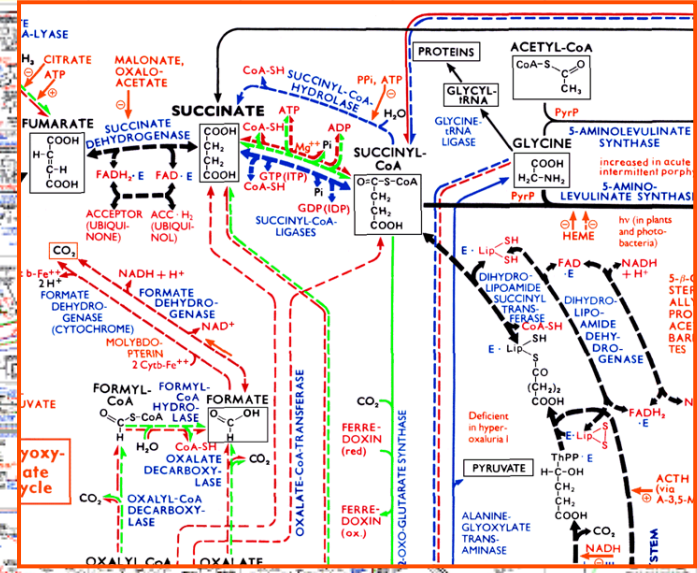
H

I

J

K

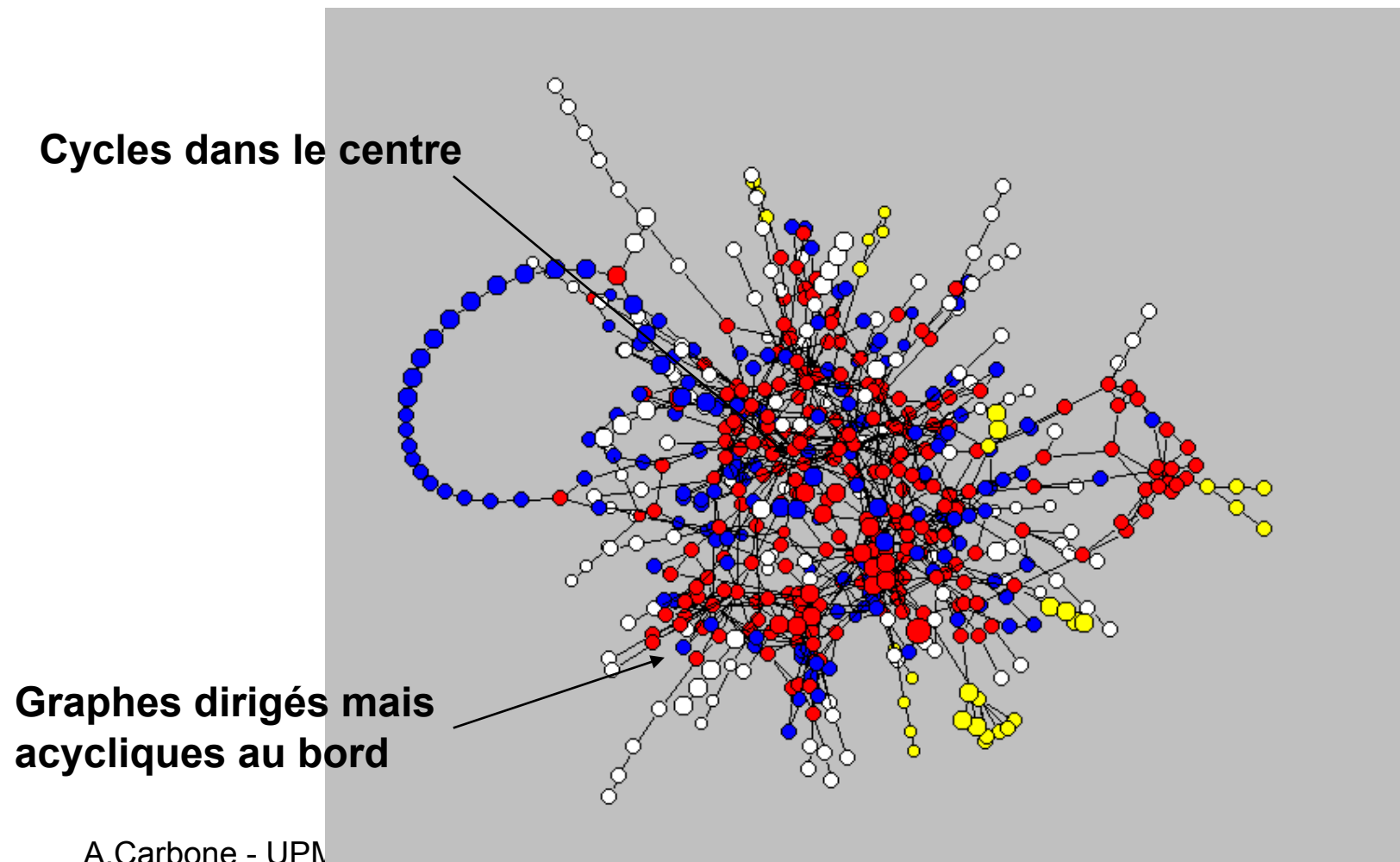
L



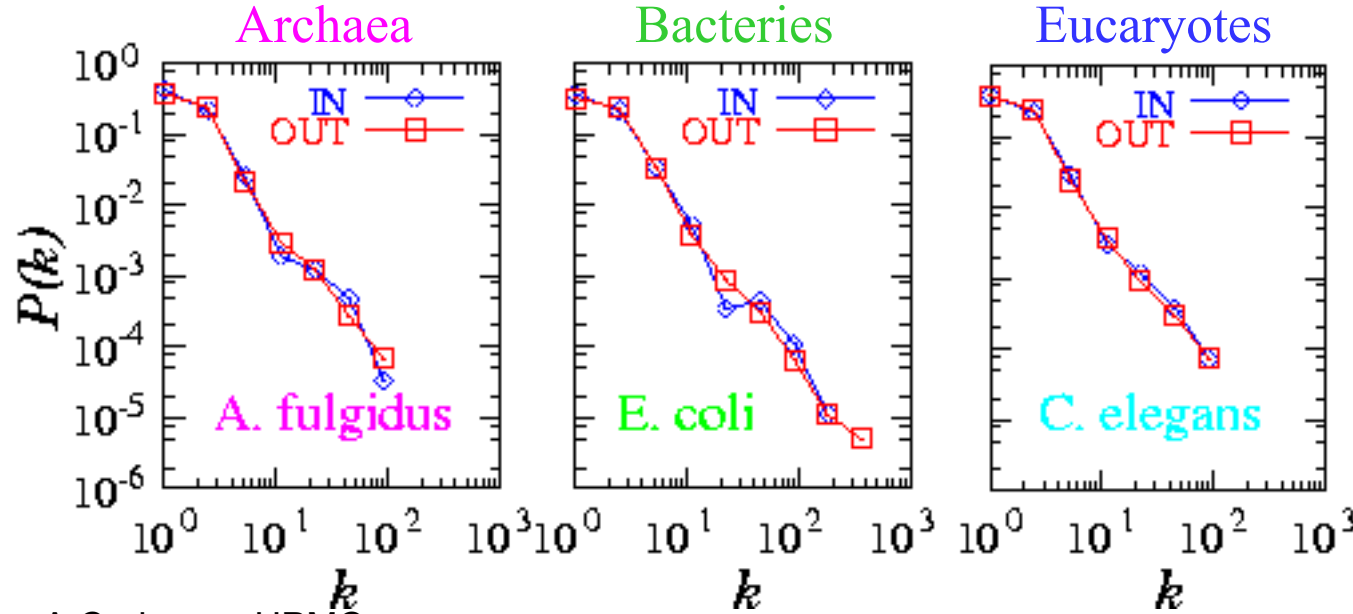
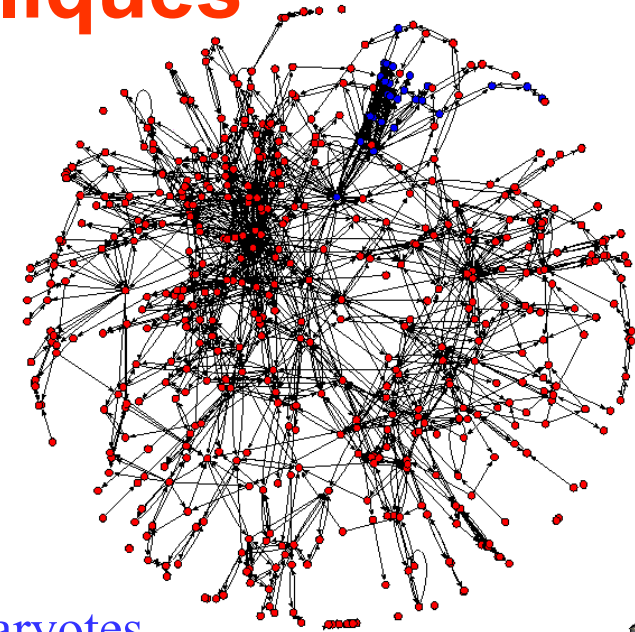
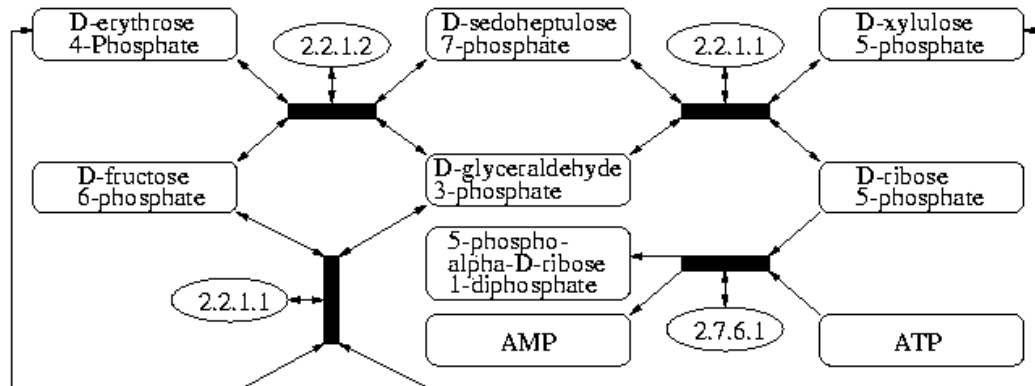
Réseaux métaboliques

Noeuds: métabolites/protéines
(soustraits)

Arcs: réactions biochimiques



Réseaux métaboliques



Les organismes de tous les trois domaines de la vie ont des réseaux **sans échelle!**

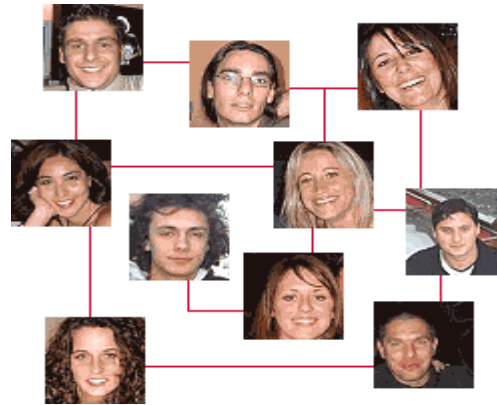
But de la théorie des réseaux

1. Trouver des propriétés statistiques qui caractérisent la structure et le comportement des systèmes définis par les réseaux
2. Créer des modèles des réseaux qui peuvent aider à comprendre la signification des propriétés des réseaux réels.
3. Prédire le comportement des systèmes ayant un réseaux sous-jacent.

Types de réseaux, autres que les réseaux biologiques

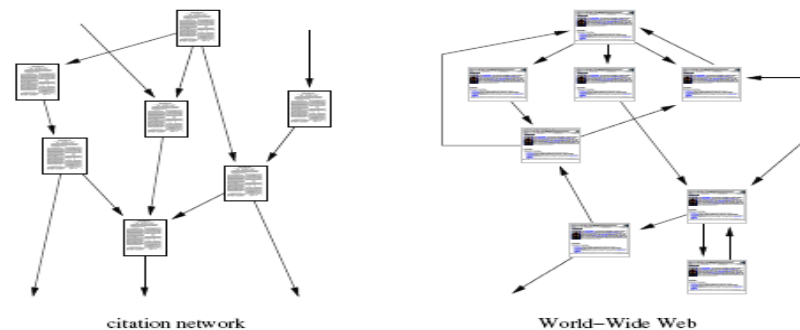
A. Réseaux sociaux

Exemples: la relation d'amitié entre individus, les relation d'affaires entre compagnies et les mariages entre familles.



B. Réseaux d'information

Exemples: réseaux des citations, World Wide Web



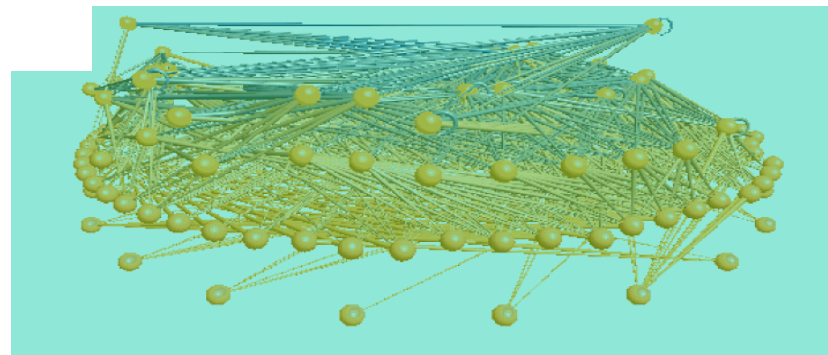
C. Réseaux Technologiques

Exemples: réseaux des centrales électriques, réseaux des compagnies aériennes, routières et ferroviaires, réseaux des rivières



D. Réseaux Biologiques

Réseaux des interactions entre protéines, des chemins métaboliques, de régulation des gènes, **des chemins de signalisation, réseaux des neurones**



Il s'agit de réseaux « sans échelle » (scale free)

- La plupart des réseaux biologiques est **sans échelle**, cad que la distribution des degrés est approximée par une loi de puissance, $P(k) \sim k^{-\gamma}$, ou γ est le degré de l'exposant
- pour $2 < \gamma < 3$ il y a une hiérarchie de hubs
- pour $\gamma = 2$ on a la formation émergente d'un hub-and-spoke du réseau
- Pour $\gamma \geq 3$ la plupart des caractéristiques non habituelles est absente

Propriété observée pour plusieurs réseaux : réseaux des citations, World Wide Web, Internet, réseaux métabolique, graphe des appels téléphoniques.

Propriétés des réseaux

- Effet “small world”
- Transitivité/Clustering
- Effet “sans échelle” (Scale Free)
- Degré maximum
- Résilience et robustesse d'un réseau
- Mélange de motifs et hétérogénéité
- Corrélations entre degrés
- Structure des communautés
- Origine évolutionnaire
- Centralité “betweenness” entre noeuds

Effet “small world” : les chemins courts

- La plupart des paires de noeuds dans un réseaux « small world » semblent être connectées par un court chemin

$$\ell = \frac{1}{\frac{1}{2}n(n-1)} \sum_{i \geq j} d_{ij},$$

ℓ est la distance minimale moyenne (mean geodesic distance)

d_{ij} est la distance minimale (geodesic distance) entre les noeuds i et j

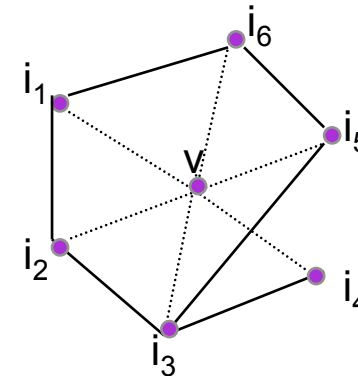
- effet sur la dynamique des processus occurrent dans un réseaux

Transitivité/Clustering

- Un réseaux montre du clustering si la probabilité qu'une paire de noeuds soient adjacents est plus grande quand les deux noeuds ont un voisin commun.
- Le coefficient de clustering C est défini comme la probabilité moyenne que deux voisins d'un noeud soient adjacents.

$$C_v = \frac{2E_v}{d_v(d_v - 1)}$$

$$C_v \in [0, 1]$$

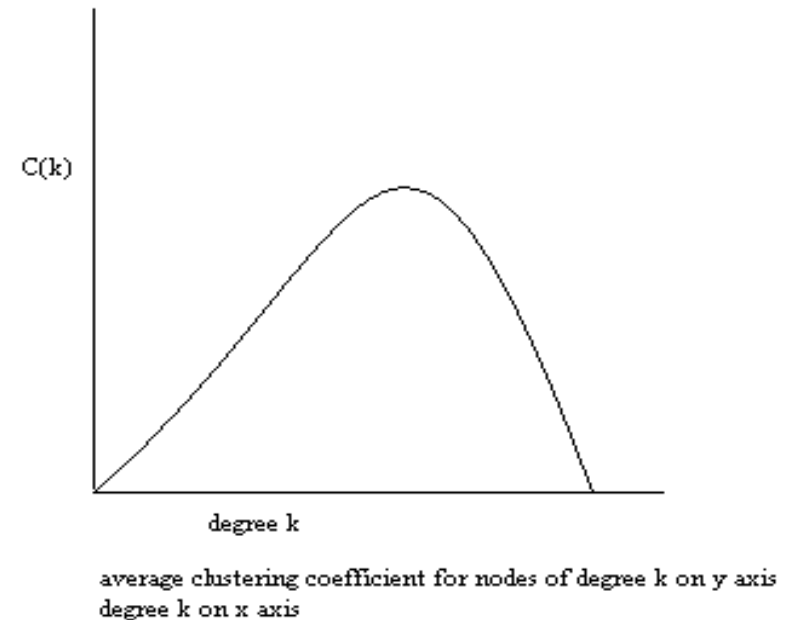
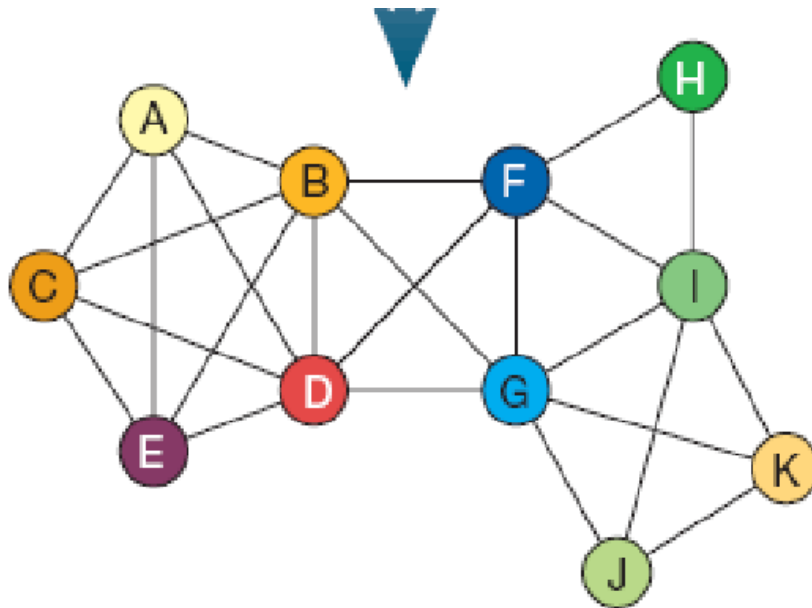


E_v est le nombre d'arêtes entre voisins de v .
 d_v est le nombre de voisins du nœud v .
 v nœud du réseau

C (le coefficient de clustering du réseau entier) = la moyenne des C_v , pour tout noeud v du réseau.

Réseaux complexes du monde réel montrent un degré de clustering élevé, cad que le coefficient de clustering est beaucoup plus grand que pour les réseaux aléatoires.

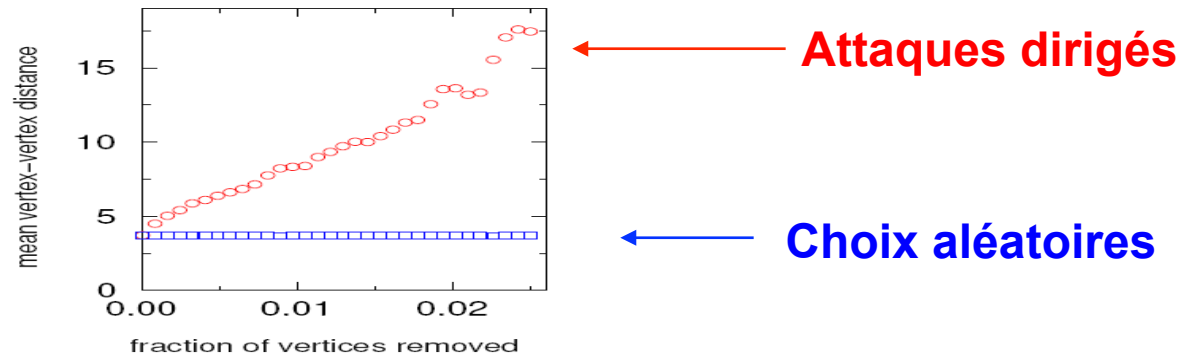
Exemple: graphe avec coefficient de clustering C élevé



Une mesure importante de la structure des réseaux est la fonction C_k , cad le coefficient moyen de clustering de tous les noeuds ayant k liens.

Résilience et robustesse des réseaux : robustesse topologique

- la plupart des réseaux que nous avons considéré se basent, pour fonctionner, sur leur connectivité.
- Effet de la suppression de noeuds sur **la longueur du chemin le plus court**, pour le graphe de l'internet :



- L'internet est fortement résilient contre les échecs aléatoires de ses noeuds mais fortement vulnérable à des attaques dirigés aux noeuds de plus fort degré.

Les réseaux sans échelle n'ont pas de seuil critique de désintégration — ils sont particulièrement robustes contre les accidents aléatoires.

Résilience et robustesse des réseaux : robustesse fonctionnelle et dynamique

- L'effet d'une perturbation ne peut pas seulement dépendre du degré des noeuds ! Mais aussi de la **fonctionnalité** des nœuds.
- Complexes de protéines identifiées expérimentalement tendent à être composés de molécules **seulement essentielles** ou **seulement non-essentiels** de façon systématique.

Une analyse de délétion systématique de protéines dans la levure *S.cerevisiae* a indiqué que pour :

~10% ayant < 5 interactions sont essentielles

>60% ayant >15 interactions sont essentielles

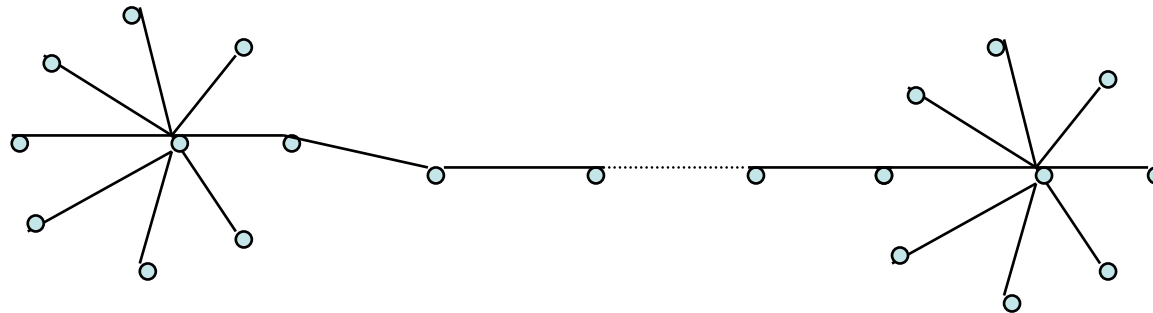
Les protéines ayant des fortes interactions dans *S.cerevisiae* sont plus conservées:

- elles ont une distance évolutive plus petite avec leurs orthologues chez *C.elegans*

-elles ont, avec forte probabilité, des orthologues dans d'autres organismes.

Homogénéité des connexions dans les réseaux biologiques

Dans les réseaux d'interaction protéine-protéine les noeuds fortement connectés (hubs) ne se relient pas directement à d'autres hubs, et à la place ils se connectent à des protéines ayant très peu d'interactions.



Structure communautaire

- Une “structure communautaire” est un **groupe de noeuds qui ont une densité élevée d’arcs entre eux**, et une densité basse d’arcs avec d’autres groupes.

Exemple: réseau de liens d’amitié d’enfants à l’école.

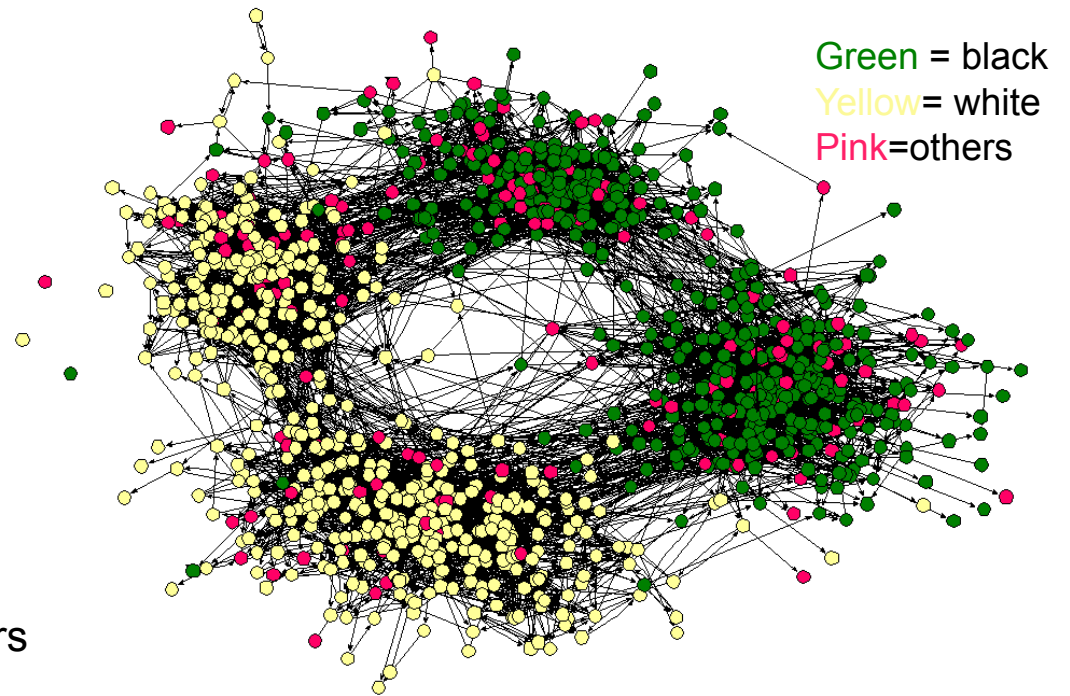
Autres exemples :

Réseaux de citations: domaines particuliers de recherche.

World Wide Web: contenu des pages

Réseaux métaboliques : unités fonctionnelles

- clustering hiérarchique



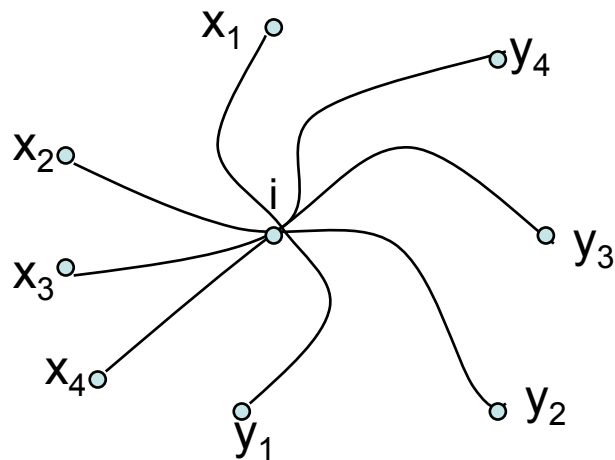
Division :

-raciale

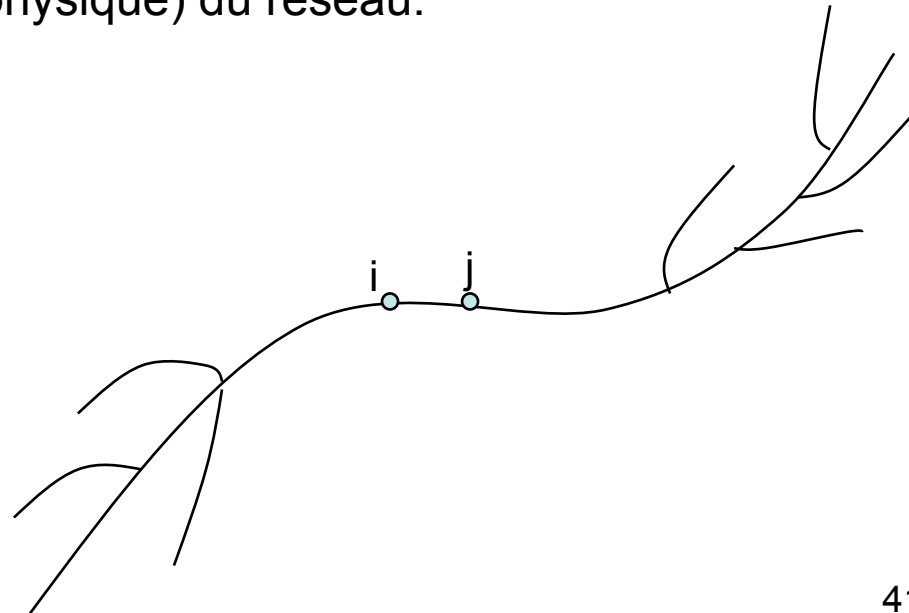
-selon l’age: école élémentaire/collège et lycée

“Betweenness centrality” des noeuds

- La “betweenness centrality” d’un noeud i est le nombre de chemins géodésiques entre deux noeuds qui passent par i .
- “Betweenness centrality” semble suivre une loi de puissance pour plusieurs réseaux !
- “Betweenness centrality” peut aussi être vue comme une mesure de la **résilience** (cad de la résistance physique) du réseau.



A.Carbone - UPMC



Stratégie pour détecter la structure des communautés basée sur la « betweenness » **des arcs**

- Pour des graphes faiblement connectés : tous les plus courts chemins doivent passer par un même arc.
- Donc, les arcs qui relient les communautés auront une forte betweenness des arcs, et en supprimant ces arcs on séparera les groupes et on trouvera les communautés.

Algorithme de détection des communautés

n^2 paires de noeuds à tester



1. Calculer la betweenness de tous les arcs dans le réseau
2. Supprimer l'arc avec betweenness la plus forte
3. Recalculer la betweenness entre tous les nœuds affectés par la suppression
4. Répéter l'étape 2 jusqu'à quand plus d'arcs restent

Complexité : $O(mn)$, $m = \# \text{ arcs}$ et $n = \# \text{ nœuds}$

Pire cas : $O(m^2n)$

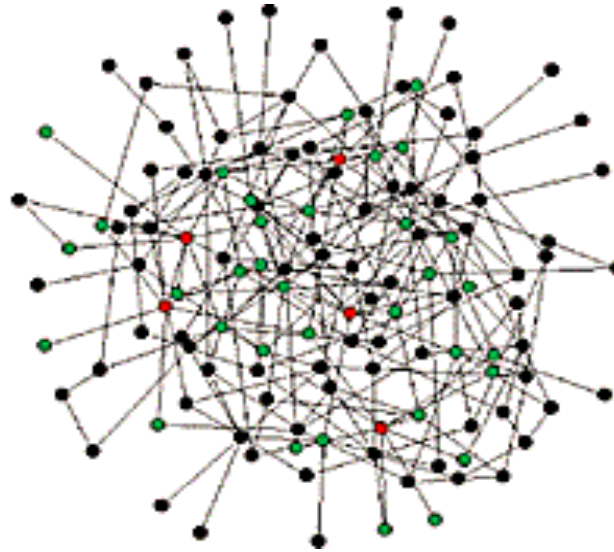
Note sur l'étape 3 de l'algorithme:

- On peut penser de calculer la betweenness seulement une fois pour efficacité.
- Si deux communautés sont connectées par plus que un arc alors il n'est pas garanti que tous ces arcs ont une betweenness élevée.
- En recalculant on est sûr que au moins seulement un des arcs restant entre deux communautés aura une valeur élevée.

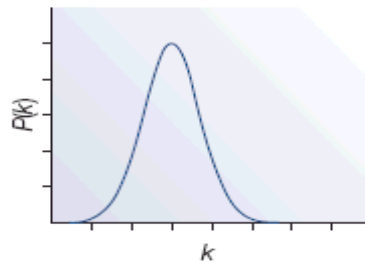
Modèles de réseaux

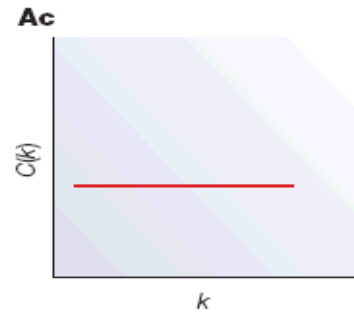
- Réseaux aléatoires
- Réseaux “sans échelle” (scale free)
- Réseaux hiérarchiques

Réseaux aléatoires



- Le modèle de Erdős–Rényi (ER) d'un réseau aléatoire commence par N noeuds et relie chaque paire de noeuds avec une probabilité p , et cela génère un graphe de, approximativement, $pN(N-1)/2$ liens placés aléatoirement.
- Les degrés des noeuds suivent une distribution de Poisson





Coefficient de clustering constant

Table 1: Clustering coefficients, C , for a number of different networks; n is the number of node, z is the mean degree. Taken from [146].

Network	n	z	C measured	C for random graph
Internet [153]	6,374	3.8	0.24	0.00060
World Wide Web (sites) [2]	153,127	35.2	0.11	0.00023
power grid [192]	4,941	2.7	0.080	0.00054
biology collaborations [140]	1,520,251	15.5	0.081	0.000010
mathematics collaborations [141]	253,339	3.9	0.15	0.000015
film actor collaborations [149]	449,913	113.4	0.20	0.00025
company directors [149]	7,673	14.4	0.59	0.0019
word co-occurrence [90]	460,902	70.1	0.44	0.00015
neural network [192]	282	14.0	0.28	0.049
metabolic network [69]	315	28.3	0.59	0.090
food web [138]	134	8.7	0.22	0.065

les graphes aléatoires ont été utilisés comme des modèles idéalisés de certaines réseaux génétiques, d'écosystèmes, de la prolifération de maladies infectieuses et des virus d'ordinateurs.

Modèle de Erdős-Renyi : la taille des composantes

Erdős and Renyi ont étudié comment la topologie d'un graphe aléatoire de n noeuds change en fonction du **nombre d'arcs m** .

Quand m est petit, le graphe est fragmenté (avec forte probabilité) dans plusieurs petites composantes connexes ayant un ensemble de noeuds de taille au plus $O(\log n)$.

Si m augment, les composantes croissent en ajoutant des noeuds isolés, et ainsi de suite en fusionnant avec d'autres composantes.

Une **transition** se passe a $m = n/2$, avec le collapse spontané de plusieurs clusters et la formation d'une composante unique et grande appelée **composante géante** (giant component). Son ensemble de noeuds est beaucoup plus grand que l'ensemble de noeuds des autres composantes. Elle contient $O(n)$ noeuds, et la deuxième composante la plus large contient $O(\log n)$ noeuds.

En physique statistique, ce phénomène est appelé **percolation**.

Modèle de Erdős-Renyi : la longueur du chemin le plus court

La **longueur du chemin le plus court** entre chaque paire de noeuds dans la composante géante croit comme $\log n$. Pour cette raison, ces graphes sont appelés „**small worlds**“.

Les propriétés des graphes aléatoires ont été particulièrement étudiés ; par exemple, in B. Bollobas. *Random Graphs*. Academic, London, 1985, 2004

Les graphes aléatoires ne sont pas des modèles adéquats pour les réseaux réel parce que

- (1) les réseaux réels ont une distribution des degrés qui suit une loi de puissance (mais les graphes aléatoires suivent une distribution de Poisson)
- (2) les réseaux réels montrent des clustering très forts, et le coefficient de clustering pour un graphe aléatoire au contraire est $C = \text{constante}$, indépendamment si deux noeuds ont un voisin en commun.

Graphes aléatoires généralisés

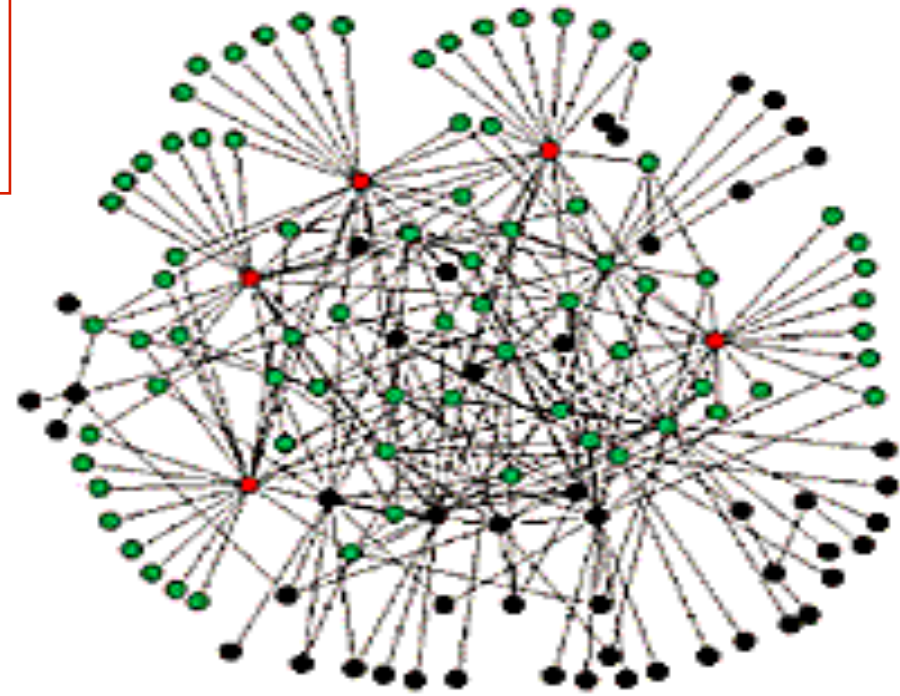
But: permettre une distribution de degrés dans le graphe qui suit une loi de puissance, et préserver toutes autres propriétés d'un modèle de réseau aléatoire.

Algorithme: étant donné une séquence de degrés (par exemple une distribution respectant une loi de puissance) on peut générer un graphe aléatoire en assignant à un noeud i un degré k_i appartenant à la séquence de degrés données. Alors, choisir des paires de noeuds uniformément et aléatoirement pour construire les arcs dans telle façon que les degrés assignés soient préservés.

Une fois que tous les degrés ont été utilisés pour créer les arcs, le graphe résultant est un membre aléatoire de l'ensemble des graphes ayant la distribution des degrés souhaitée.

Problème: cette méthode ne permet pas la spécification de coefficients de clustering. En fait, cette propriété rend possible la détermination de plusieurs autres propriétés de ces graphes pour des n larges. Par exemple, presque tous les graphes aléatoires ayant une distribution de degrés fixée et n'ayant pas de noeuds isolés (de degré < 2) ont une composante géante unique.

Réseaux sans échelle (scale free)



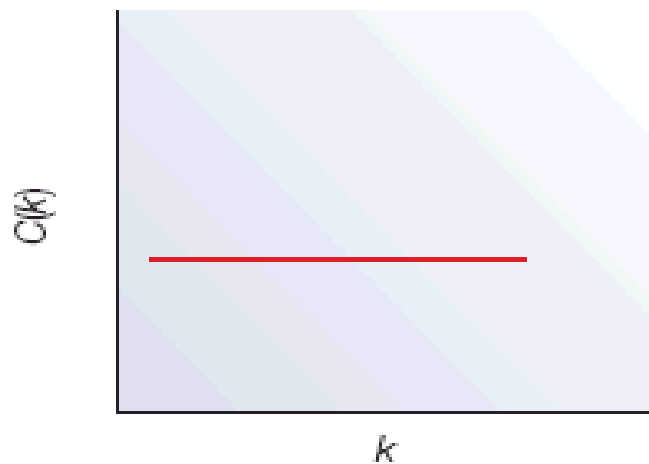
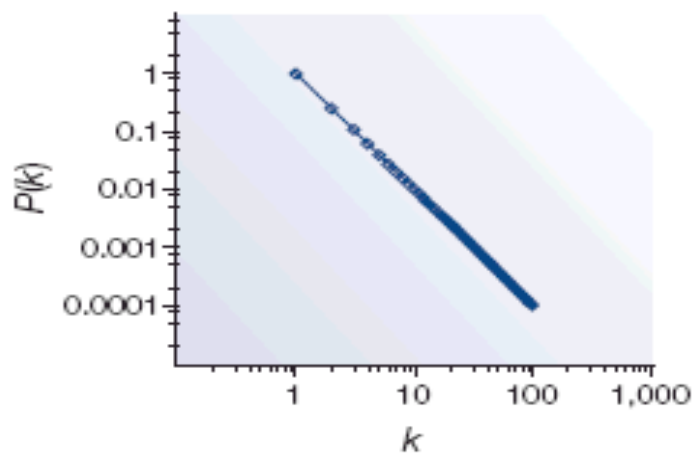
$P(k) \sim k^{-\gamma}$, où γ est le degré de l'exposant.

Les propriétés du réseau sont déterminés par les **hubs**

Modèle proposé pour expliquer le réseau : **processus de croissance** appelé “modèle de Barabási–Albert”

Caractéristiques des réseaux sans échelle qui « devraient » être garanties

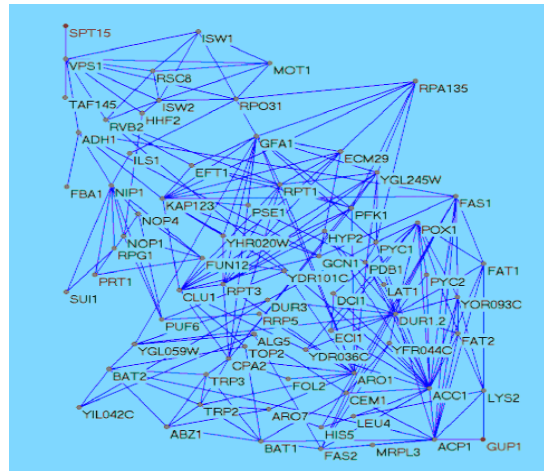
- Les réseaux sans échelle ont un exposants de degré $2 < \gamma < 3$, un intervalle qui est souvent observé dans les réseaux biologiques et non-biologiques comme l'Internet, le World Wide Web, le réseau des réactions métaboliques et le graphe des appels téléphoniques.



- La longueur moyenne du plus court chemin est proportionnelle à $\log(n)/\log(\log(n))$

D'après une analyse des réseaux sans échelle appliquée aux réseaux biologiques : un exemple

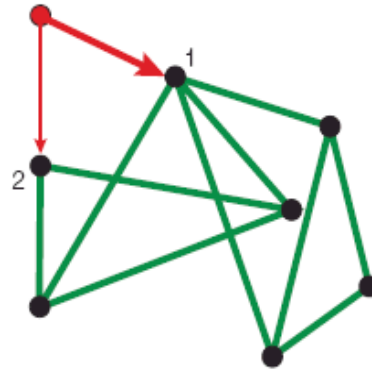
Exemple d'un réseau prédit, ayant topologie sans échelle



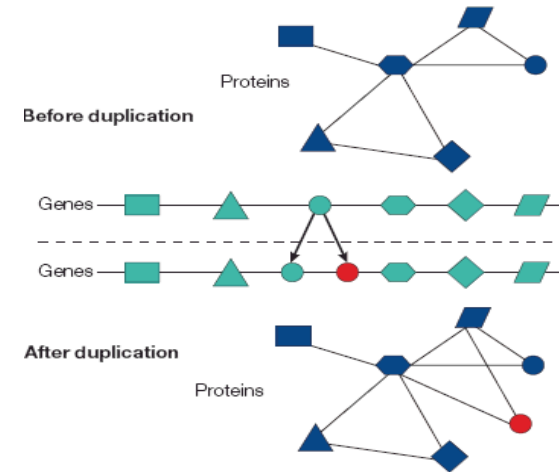
- L'analyse des réseaux métaboliques de **43 organismes** du WIT database a montré une topologie scale-free avec $P(k) \sim k^{-2.2}$ pour les degrés internes et les degrés externes.
- Le diamètre des réseaux métaboliques était le même
- Quelques hubs dominaient ces réseaux et après une suppression séquentielle des noeuds les plus connectés le diamètre a augmenté brusquement
- seulement à peu près le 4% des noeuds sont présents dans toutes les espèces et ces noeuds étaient aussi les plus connectés dans tous les organismes. ← **conservations**
- en supprimant aléatoirement les noeuds de ces réseaux, la longueur moyenne du plus court chemin n'a pas changé, et cela est un indicateur de l'insensibilité aux erreurs aléatoires dans les réseaux.

Modèle : origine évolutionnaire des réseaux sans échelle

- **Croissance** et **liaison préférentielle** sont responsables conjointement de l'émergence des propriétés scale-free des réseaux complexes.



- Réseaux du à la duplication de gènes et sous-réseaux des protéines.



- **Evidence empirique pour l'attachement préférentiel :**
 - Une inspection des hubs métaboliques indique que **ce qui reste du monde à ARN**, comme les co-enzymes A-NAD et GTP, constituent les soustrats les plus connectés du réseau métabolique, ainsi que les éléments des plus anciens chemins métaboliques, comme la glycolyse et le cycle acide-tricarboxylique.
 - dans le contexte des réseaux d'interaction des protéines, la comparaison des génomes a permis de trouver que, en moyenne, les protéines qui sont phylogénétiquement les **plus anciennes** ont plus de liens à d'autres protéines que leurs contrepartie plus jeune.

Modèle sans échelle de Barabasi-Albert

Entrée où

n_0 est le nombre initiale de noeuds,
 m ($m \leq n_0$) est le nombre d'arcs ajoutés à chaque fois que un nouveau noeuds est ajouté au graphe
 t est le nombre d'itérations.

Algorithme

- a) Commencer par n_0 noeuds isolés.
- b) A chaque fois que un nouveau noeud v est ajouté, m arcs seront liés aux noeuds existants avec une **probabilité de liaison préférentielle**

$$\Pi_i(k_i) = \frac{k_i}{\sum_{i=1}^{N-1} k_i}$$

où k_i est le nombre de liens au i -ème noeud.

A la fin, le graphe aura n_0+t noeuds et mt arcs au plus.

Barabasi-Albert est un modèle sans échelle

$P(k) \propto k^{-\gamma}$ avec $\gamma = 3$.

Réseaux réel montrent souvent $\gamma \approx 2.1 - 2.4$

Observation : si soit la condition sur la croissance soit la condition sur l'attachement préférentiel est éliminé, le réseau résultant ne montre pas des propriétés sans échelle.

La **longueur moyenne des chemins** dans le modèle BA est proportionnelle à $\ln(n)/\ln \ln(n)$, cad une longueur estimée plus petite que pour les graphes aléatoires. Dans ce sens, les réseaux sans échelle BA sont des **ultrasmall worlds**.

Observation: on peut observer une corrélation (= clustering) non triviale entre les degrés des noeuds connectés.

Résultats numériques pour le modèle BA : $C \approx n^{-0.75}$. Pas de prédiction analytiques sur C jusqu'à maintenant.

Propriétés des modèles sans échelle

- Graphes épars (sparse graphs)
- Ayant plusieurs “cliques” (cliquish)
- Chemins entre deux noeuds quelconques sont courts
- Une loi de puissance gouverne la distribution du degré des noeuds
 $P(k) \approx k^{-\gamma}$ ou $k \gg K$
- Très peu de noeuds fortement connectés
- Si on supprime un noeud choisi aléatoirement, la distance moyenne ne change pas: **insensibilité aux erreurs aléatoires, robustesse**. (Cela est dû au fait que les hubs de degré élevé ne sont pas nombreux.)
- **Vulnérables aux attaques** sur les hubs.
- **Applications biologiques**: diamètre et soustraités connectés sont les mêmes pour des organismes différents.

Ces propriétés ont été confirmées numériquement et en étudiant analytiquement la longueur moyenne des chemins et la taille de la composante géante.

Modifications du modèle sans échelle de Barabasi-Albert

Le modèle BA est un modèle minimale que capture les mécanismes responsables de la distribution des degrés suivant une loi de puissance observée dans les réseaux réels.

Une divergence est l'exposant $\gamma = 3$ de la distribution prédite.

→ est-ce que le modèle BA décrit l'évolution des réseaux biologiques „vrais“?

Efforts récents :

- étude de variantes avec des propriétés mathématiques plus claires (Bollobas, modèle LCD)
- inclusion d'effets obtenus en ajoutant ou en re-connectant les arcs, permettent aux noeuds de vieillir de telle façon qu'ils puissent refuser d'accepter des nouveaux noeuds ou qu'ils acceptent des variations de formes de liaison préférentielle.

Ces modèles prédisent aussi des distributions exponentielles et des lois de puissance tronquées à certaines régimes des paramètres.

Mais **ils ne tiennent pas en compte les duplications massives ni partielles des génomes.**

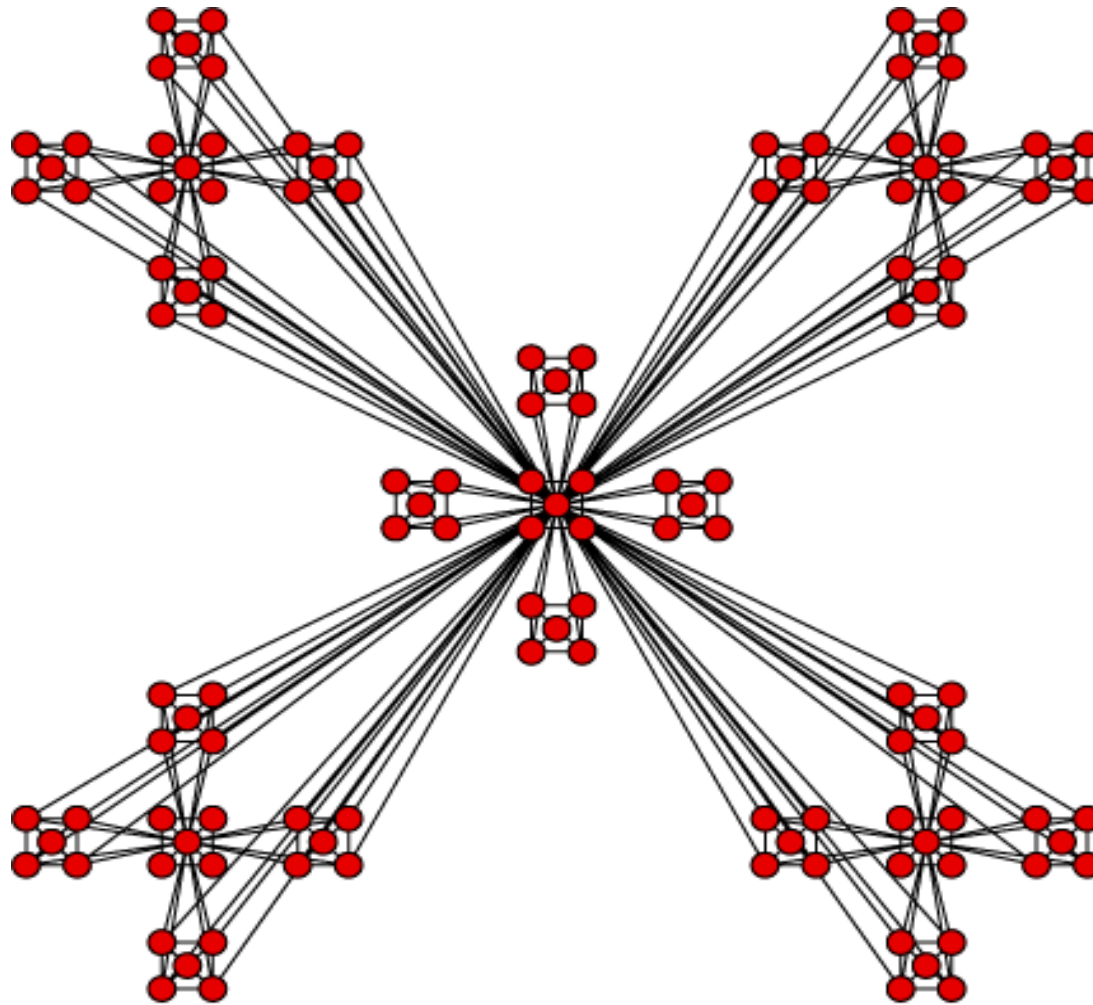
Evlampiev K, Isambert H. Conservation and topology of protein interaction networks under duplication-divergence evolution. *Proc Natl Acad Sci U S A*. 2008 Jul 22;105(29):9863-8.

Evlampiev K, Isambert H. Modeling protein network evolution under genome duplication and domain shuffling. *BMC Syst Biol*. 2007 Nov 13;1:49.

Modèle d'évolution de duplication-divergence. Etude des propriétés topologiques et de conservativité.

Les réseaux conservés sparses, qui sont les seuls d'intérêt biologique potentiel, sont forcément scale-free, indépendamment des variations évolutives des paramètres du modèle.

Réseaux Hiérarchiques



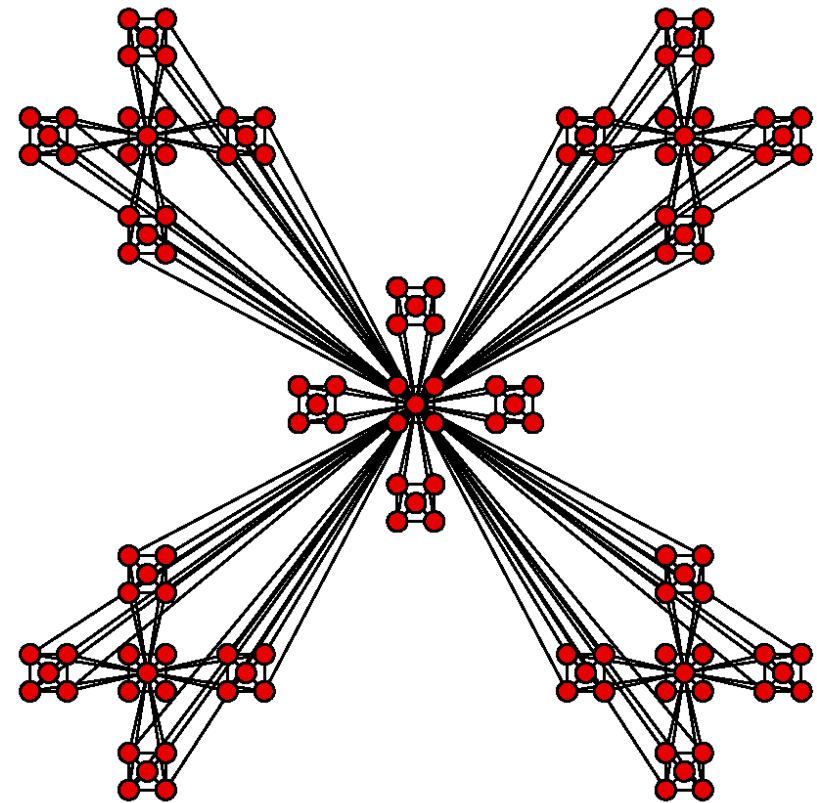
Les réseaux réels ont une topologie hiérarchique

Qu'est-ce que cela signifie ? Pour tenir en compte de la coexistence de la modularité, le clustering locale et la topologie scale-free dans plusieurs systèmes réels, cette méthode fait l'hypothèse que les clusters se combinent de manière itérative, en générant un réseau hiérarchique.

Pleins de petits clusters très connectés
se combinent dans
des grands clusters, peu nombreux et
moins connectés
qui se combinent dans
des clusters même plus grands, et même
moins connectés

➤ Le coefficient de clusterisation d'un nœud avec k liens suit la loi d'échelle:

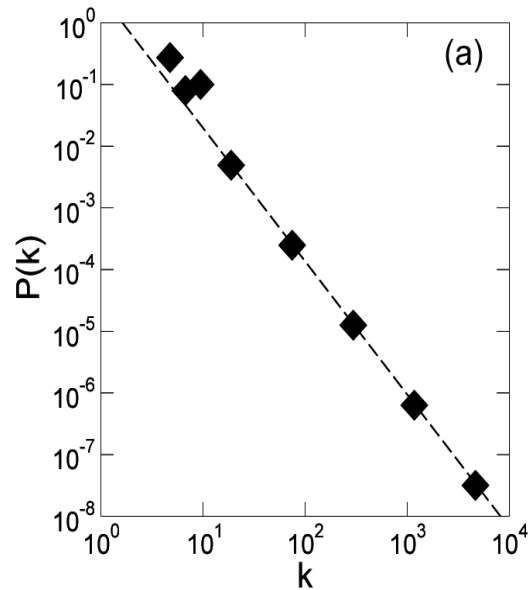
$$C(k) \sim k^{-\beta}$$



Propriétés des réseaux hiérarchiques

1. Sans échelle

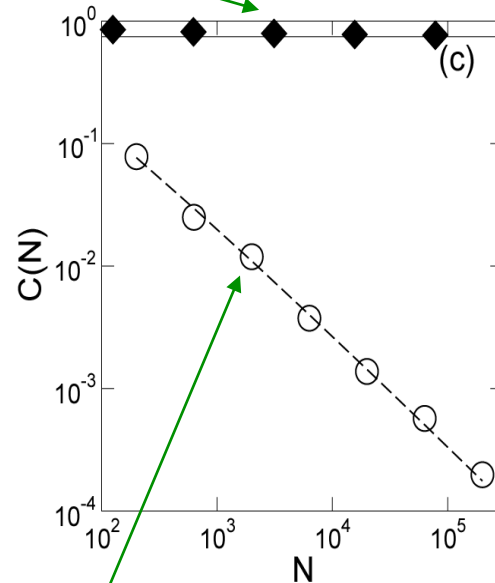
$$\gamma = 1 + \frac{\ln 5}{\ln 4} \\ = 2.161$$



2. Coefficient de clustering (moyenne) indépendant de N

$$C(N) = \text{const.}$$

Modèle hiérarchique



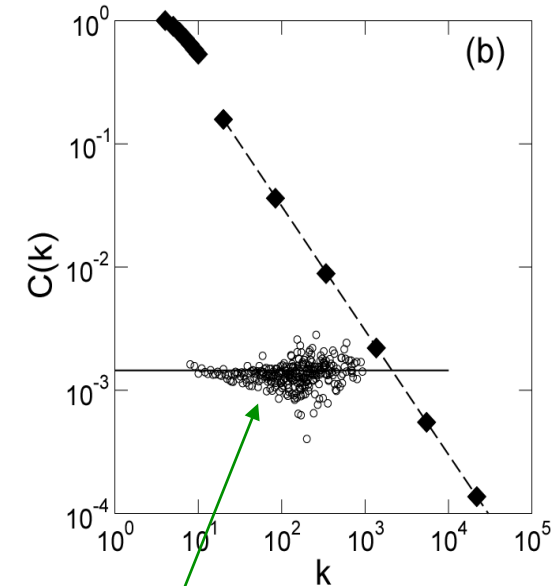
Modèle sans échelle

N = taille du réseau

3. Echelles de coefficients de clustering

$$C(k) \sim k^{-1}$$

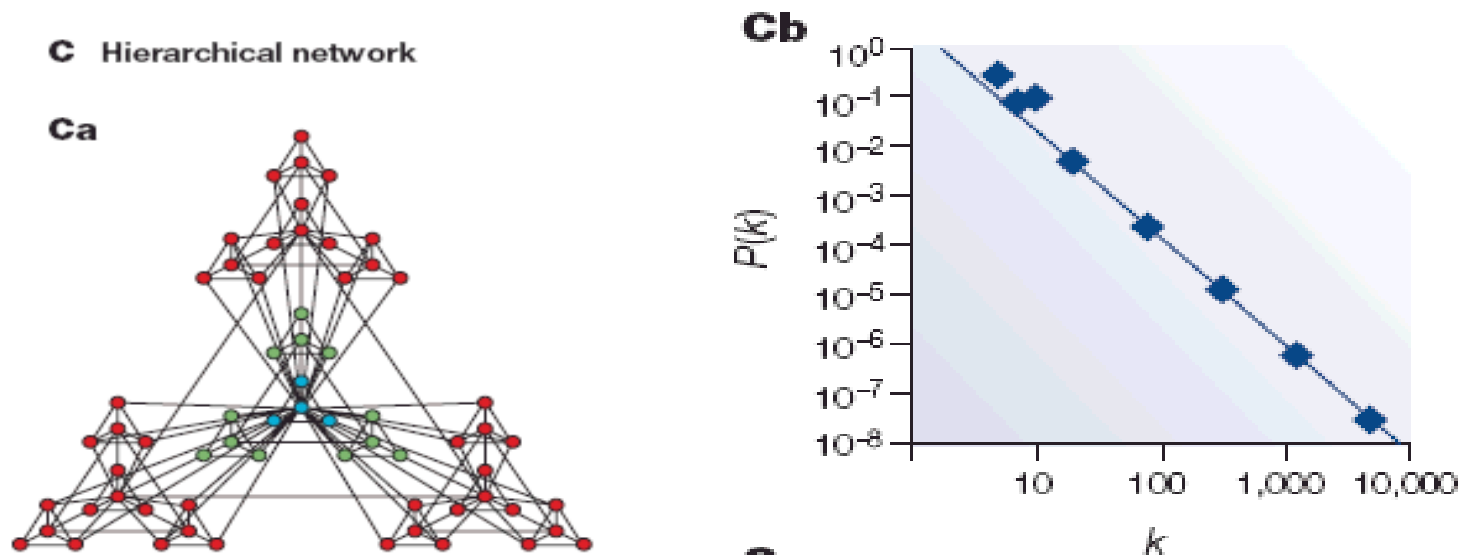
pour un nœud de k liens



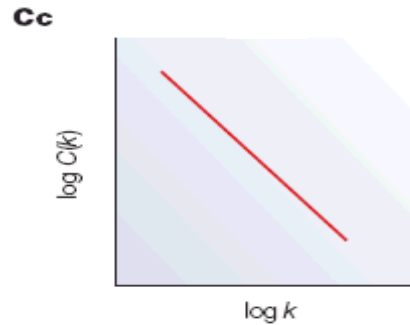
$C(k)$ pour un modèle sans échelle de la même taille

Réseaux hiérarchiques, exemple

Le modèle de réseau hiérarchique intègre une topologie scale-free avec structure modulaire inhérente, et génère un réseau qui a une loi de puissance pour distribution des degrés, avec un exposant de degré $\gamma = 1 + \ln 4 / \ln 3 = 2.26$



- Il a un coefficient de clustering moyen qui est indépendant de la grande taille du système : $\langle C \rangle \sim 0.6$. La caractéristique la plus importante de la modularité hiérarchique est la **mise en échelle du coefficient de clustering**, qui suit $C(k) \sim k^{-1}$ une droite de pente -1 sur un plot log-log



- Une architecture hiérarchique implique que **les noeuds peu connectés font partie de zones très clustérisées**, avec une communication entre les différents voisins très clustérisés qui est maintenues par un nombre très limité de hubs. Quelques exemples de réseaux hiérarchique scale-free :

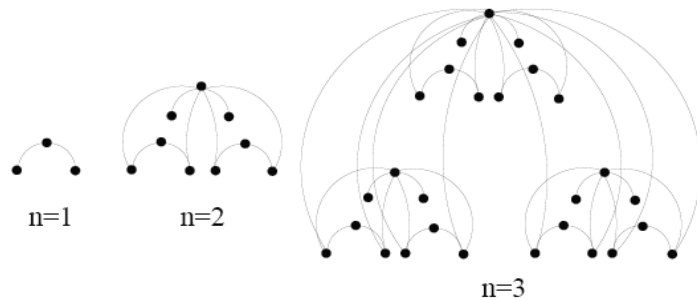


Figure 4: Scheme of the growth of a scale-free deterministic hierarchical graph. Adapted from [23].

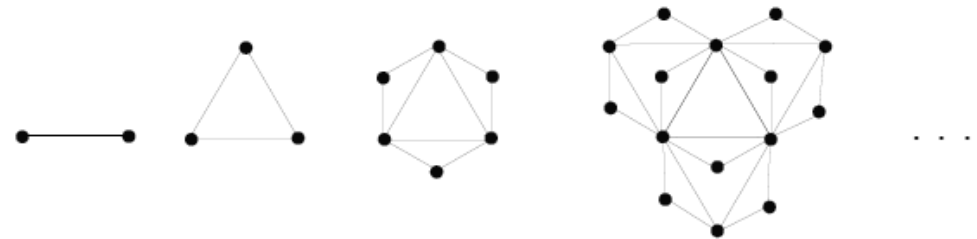


Figure 5: Scheme of the growth of the scale-free pseudofractal graph. The graph at time step $t + 1$ can be constructed by “connecting together” three graphs from step t . Adapted from [55].

Réseaux hiérarchiques et réseaux métaboliques

Modèle hiérarchique appliqué aux réseaux biologiques

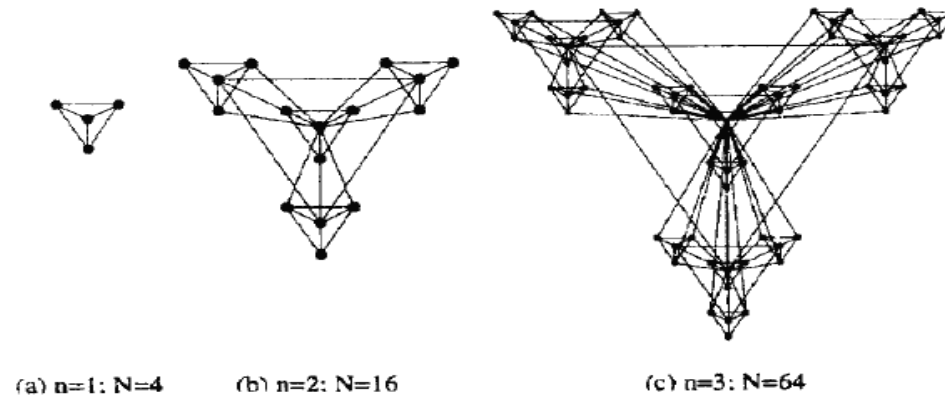


Figure 10: Three steps in the construction of a hierarchical model network.
Taken from [162].

- il a été établi que le modèle chevauche assez bien le réseau métabolique de *E.coli*.
- Les réseaux hiérarchiques ont eu leur succès dans la modélisation du réseau métabolique de *E.coli* ainsi que des réseaux de tous les 43 organismes précédemment modélisés par des réseaux scale-free.

Données online de PPI

- 1.>The Munich Information Center for Protein Sequences (MIPS)
<http://mips.gsf.de>
- 2.>Yeast Proteomics Database (YPD)
<http://www.incyte.com/sequence/proteome/databases/YPD.html>
- 3.>Human Reference Protein Database (HRPD)
<http://www.hrpd.org>
- 4.>The Biomolecular Interaction Network Database
<http://www.biond.org/>
- 5.>The General Repository for Interaction Datasets (GRID)
<http://biodata.mshri.on.ca/grid/>
- 6.>The Molecular INTeraction database (MINT)
mint.bio.uniroma2.it/mint/
- 7.>Online Predicted Human Interaction Database (OPHID)
<http://ophid.utoronto.ca>

Références sur le réseaux

Nadav Kashtan and Uri Alon, *From the Cover: Spontaneous evolution of modularity and network motifs* PNAS, 102: 13773-13778 (2005).

R Milo, S Itzkovitz, N Kashtan, R Levitt, S Shen-Orr, I Ayzenshtat, M Sheffer & U Alon, *Superfamilies of designed and evolved networks*, Science, 303:1538-42 (2004).

R Milo, S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii & U Alon, *Network Motifs: Simple Building Blocks of Complex Networks*, Science, 298:824-827 (2002).

S Shen-Orr, R Milo, S Mangan & U Alon, *Network motifs in the transcriptional regulation network of Escherichia coli*. Nature Genetics, 31:64-68 (2002).

Abrahamson, E. & Rosenkopf, L. Social network effects on the extent of innovation diffusion: A computer simulation. *Organization Science* 8(3), 289-309 (1997).

Adamic L. A. The small world web. *Proceedings of the ECDL'99 Conference*, pp.443-452 (Springer, 1999).

Adamic, L. A. and Huberman, B. A. Technical comment to 'Emergence of scaling in random networks', Science 286, 509-512 (1999). Available online at <http://www.parc.xerox.com/istl/groups/iea/www/bacomment.html> (1999).

Aguirre, B. E., Quarantelli, E. L. and Mendoza, J. L. The collective behavior of fads: the characteristics, effects, and career of streaking. *American Sociological Review* 53, 569-584 (1988).

Aiello, W., Chung, F. and Lu, L. A random graph model for massive graphs. *Proceedings of the 32nd Annual ACM Symposium on the Theory of Computing* (2000). 66

Albert, Reka and Albert-Laszlo Barabasi. Dynamics of Complex Systems: Calculating the Period of Boolean Networks, *Physical Review Letters* 84 5660-5663 (2000). (See <http://www.nd.edu/~networks/papers.htm> for electronic copies.)

Albert, Reka and Albert-Laszlo Barabasi. Topology of Evolving Networks: Local Events and Universality, *cond-mat/0005085*. (See <http://www.nd.edu/~networks/papers.htm> for electronic copies.)

Albert, Reka, Hawoong Jeong and Albert-Laszlo Barabasi. Attack and Error Tolerance of Complex Networks, *Nature* 406, 378-382 (2000). (See <http://www.nd.edu/~networks/papers.htm> for electronic copies.)

Albert, Reka, Hawoong Jeong and Albert-Laszlo Barabasi. Diameter of the World Wide Web , *Nature* 401 130-131 (1999). (See <http://www.nd.edu/~networks/papers.htm> for electronic copies.)

Amaral, L. A. N., Scala, A., Barthelemy, M., and Stanley, H. E. Classes of behavior of small-world networks. <http://xxx.lanl.gov/abs/cond-mat/0001458> (2000).

Andreeva, N. Do the math — it is a small world. *Business Week*, August 17, 54-55 (1998).

Arthur W. B. and Lane D. A. Information contagion. *Structural Change and Economic Dynamics*. 4(1), 81-103 (1993).

Axelrod, Robert. "On Six Advances in Cooperation Theory." In a Special Issue of *Analyse and Kritik* on The Evolution of Cooperation, forthcoming in 2000.

Axelrod, Robert "Advancing the Art of Simulation in the Social Sciences." Published in Rosario Conte, Rainer Hegselmann and Pietro Terna (eds.), *Simulating Social Phenomena*, (Berlin: Springer, 1997), pp21-40.

Axelrod, Robert, Rick L. Riolo, and Michael D. Cohen. "Beyond Geography: Cooperation with Persistent Links in the Absence of Clustered Neighborhoods." To appear in *Adaptation and Evolution: Essays in Honor of John H. Holland*, L. Booker, S. Forrest, M. Mitchell and R. Riolo (eds), forthcoming in 2000. 67

Axtel, R. Effects of interaction topology and activation regime in several multi-agent systems. Brookings Institute working paper, Available from raxtell@brook.edu (2000).

Ball, F., Mollison, D., and Scalia-Tomba, G. Epidemics with two levels of mixing. *Annals of Applied Probability* 7(1), 46-89 (1997).

Banerjee, A. V. A simple model of herd behavior. *Quarterly Journal of Economics* 107, 797-817 (1992).

Barabasi, A. and Albert, R. Emergence of scaling in random networks. *Science* 286, 509-512 (1999).

Barabasi, Albert-Laszlo, Reka Albert and Hawoong Jeong. Mean-Field Theory for Scale-Free Random Networks, *Physica A* 272 173-187 (1999). (See <http://www.nd.edu/~networks/papers.htm> for electronic copies.)

Barrat, A. and Weigt, M. On the properties of small-world network models. *European Physical Journal B* (in press, 2000).

Barthelemy, M and Amaral, L. A. N. Small-world networks: Evidence for a crossover picture. *Physical Review Letters* 82, 3180-3183 (1999).

Bikhchandani, S., Hirshleifer, D. and Welch. I. A theory of fads, fashion, custom and cultural change as informational cascades. *Journal of Political Economy* 100(5), 992-1026 (1992).

Blakeslee, S. Mathematicians prove that it's a small world. *New York Times* (June 16), F3 (1998).

Blythe, S. P., Castillo-Chavez, C. and Palmer, J. S. Toward a unified theory of sexual mixing and pair formation. *Mathematical Biosciences* 107, 379-405 (1991).

Bolton, P. and Dewatripont, M. The firm as a communication network. *Quarterly Journal of Economics* 109(4), 809-839 (1994).

Boots, M. and A. Sasaki. 'Small worlds' and the evolution of virulence: Infection occurs locally and at a distance. *Proceedings of the Royal Society, Series B* 266, 1933-1938 (1999).

Bornholdt, Stefan and Thimo Rohlf, *Topological Evolution of Dynamical Networks: Global Criticality from Local Dynamics*, Phys. Rev. Lett. 84 (2000) 6114.

Bowles, S. and H. Gintis *Optimal Parochialism: The Dynamics of Trust and Exclusion in Networks*, Santa Fe Institute working paper, 2000.

Callaway, D. S., Newman, M.E.J., Strogatz, S. H. and Watts, D. J. Exact solution of percolation on random graphs with arbitrary degree distributions. <http://xxx.lanl.gov/abs/cond-mat/0007300> (2000)

Carlson, J. M. and Doyle, J. Highly optimized tolerance: a mechanism for power laws in designed systems. *Physical Review E*. 60(2), 1412-1427 (1999).

S. Chakrabarti, B. Dom, D. Gibson, J. Kleinberg, S.R. Kumar, P. Raghavan, S. Rajagopalan, A. Tomkins, Hypersearching the Web. Scientific American, June 1999. [[html](#)]

Coase, R. The Nature of the Firm: Meaning. *Journal of Law, Economics and Organization*. 4(1) (1988)

Cohen, J. E. and Newman, C. M. A stochastic theory of community food webs I. Models and aggregated data. *Proceedings of the Royal Society of London, Series B* 224, 421-448 (1985).

Cohen, M. D., Riolo, R., and Axelrod, R. *The emergence of social organization in the Prisoner's Dilemma: how context preservation and other factors promote cooperation*. Santa Fe Institute working paper 99-01-002 (1999).

Cohen, Michael D., Rick L. Riolo, and Robert Axelrod. *The Role of Social Structure in the Maintenance of Cooperative Regimes*. Note: Due to pre-publications restrictions by the journal to which this paper is being submitted, only one-half of the paper is available on the Web. The entire paper is available in hardcopy on request to axe@umich.edu.

Coleman, J. S., Menzel, H., and Katz, E. The diffusion of an innovation among physicians. *Sociometry* 20, 253-270 (1957).

Davis, G. F. The significance of board interlocks for corporate governance. *Corporate Governance*, 4(3), 154-159 (1996).

Davis, G. F., and Greve, H. R. Corporate elite networks and governance changes in the 1980s. *American Journal of Sociology* 103(1), 1-37 (1997).

Demers, A., Greene, D., Hauser, C., Irish, W., Larson, J., Shenker, S., Sturgis, H., Swinehart, D., and Terry, D. Epidemic algorithms for replicated database maintenance. *ACM Operations Systems Review* 22, 8-32 (1988).

De Vany, A. S. and Walls, W. D. Bose-Einstein dynamics and adaptive contracting in the motion picture industry. *The Economic Journal* **106**, 1493-1514 (1996).

Dodds, P.S. and Rothman, D. Unified view of scaling laws for river networks. *Physical Review E* 59, 4865-4877 (1999).

Erdős, P. and Rényi, A. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5, 17-61 (1960).

Fararo, T.J. Reflections on mathematical sociology. *Sociological Forum* 12(1), 73-101 (1997).

Freeman, L. C. A set of measures of centrality based on betweenness. *Sociometry* 40, 35-41 (1977).

Freeman, L. C. Centered graphs and the structure of ego networks. *Mathematical Social Sciences* 3, 291-304 (1982).

Friedkin, N. E. Theoretical foundations for centrality measures. *American Journal of Sociology* 96(6),⁷⁰ 1478-1504 (1991).

Glance, N. S. and Huberman, B. A. The outbreak of cooperation. *Journal of Mathematical Sociology*. **17(4)**, 281-302 (1993).

Granovetter, Mark. A Theoretical Agenda for Economic Sociology. To appear in *Economic Sociology at the Millenium*, edited by Mauro F. Guillen, Randall Collins, Paula England, and Marshall Meyer (New York: Russell Sage Foundation, 2001).

Granovetter M. Threshold models of collective behavior. *American Journal of Sociology*. **83(6)**, 1420-1443 (1978).

Granovetter, M., and Soong, R. Threshold models of diffusion and collective behavior. *Journal of Mathematical Sociology* 9, 165-179 (1983).

Hauer, J. F. and Dagel. J. E. Consortium for electric reliability technology solutions, grid of the future: White paper on review of recent reliability issues and system events. U.S. Dept. of Energy. (1999).

Hecht, R., R. Happel, P. Schuster, and [P. Stadler](#), *Autocatalytic Networks with Intermediates I: Irreversible Reactions*, Santa Fe Institute working paper, 1996.

Hess, G. Disease in metapopulation models: implications for conservation. *Ecology* **77**, 1617-1632 (1996).

Hopcroft, J. E., and Tarjan, R. E. Dividing a graph into triconnected components. *SIAM Journal of Computing* 2(1), 135-145 (1973).

Jain, S. and S. Krishna, Autocatalytic sets and the growth of complexity in a simple model, *Phys. Rev. Lett.* 81 (1998) 5684-5687. <http://arXiv.org/abs/nlin.AO/9809003>

Jain, S. and S. Krishna, Emergence and growth of complex networks in adaptive systems, *Computer Phys. Comm.* 121-122 (1999) 116-121. <http://arXiv.org/abs/nlin.AO/9810005>

Jain, S. and S. Krishna, A model for the emergence of cooperation, interdependence and structure in evolving networks, SFI preprint (2000). <http://arXiv.org/abs/nlin.AO/0005039>

Kareiva, P. Population dynamics in spatially complex environments: Theory and data. *Philosophical Transactions of the Royal Society of London, Series B* 330, 175-190 (1988).

Keeling, M. J. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London, Series B* 266(1421), 859-867 (1999).

Kleinberg, J. *Authoritative Sources in a Hyperlinked Environment*. Proc. 9th ACM-SIAM Symposium on Discrete Algorithms, 1998. Extended version in *Journal of the ACM* 46(1999). Also appears as IBM Research Report RJ 10076, May 1997.

Kleinberg, J. The small-world phenomenon: An algorithmic perspective. Cornell Computer Science Technical Report 99-1776, (1999).

Kleinberg, J., Kumar, S. R., Raghavan, P., Rajagopalan, S., and Tomkins, A. The Web as a graph: Measurements, models and methods. Invited survey at the International Conference on Combinatorics and Computing (1999). Available online at <http://www.cs.cornell.edu/home/kleinber/kleinber.html>

Kogut B. and G. Walker. The small world of firm ownership in Germany: social capital and structural holes in large firm acquisitions--1993-1997. *Working paper*, (Reginald H. Jones Center, Wharton School 1999, <http://www.hbs.edu/units/cs/smallworldsr.PDF>).

Kosterev, D. N., Taylor, C. W. and Mittelstadt, W. A. Model validation for the August 10, 1996 WSCC System Outage. *IEEE Trans. on Power Systems*. 14(3), 967-979 (1999).

Kretschmar, M. and Morris, M. Measures of concurrency in networks and the spread of infectious disease. *Mathematical Biosciences* 133, 165-195 (1996).

Kuramoto, Y. Self-entrainment of a population of coupled nonlinear oscillators. In: Araki, H. (ed.) *International Symposium on Mathematical Problems in Theoretical Physics*. (Lecture Notes in Physics, vol. 39.) Springer, New York. (1975), pp.420-422.

Lago-Fernandez, L. F., R. Huerta, F. Corbacho, and J.A. Siguenza. Fast response and temporal coding on coherent oscillations in small-world networks. *Physical Review Letters*, 84, 2758 (2000).

Leighton, T. Methods for message routing in parallel machines. *Proceedings of 24th Annual ACM Symposium on the theory of computing*, 77-96 (1992).

Lohmann S. The dynamics of informational cascades: the Monday demonstrations in Leipzig, East Germany, 1989-91. *World Politics*. 47, 42-101 (1994).

Longini, I. M., Jr. A mathematical model for predicting the geographic spread of new infectious agents. *Mathematical Biosciences* 90, 367-383 (1988).

Matthews, P. C., Mirollo, R. E., and Strogatz, S. H. Dynamics of a large system of coupled nonlinear oscillators. *Physica D* 52, 293 (1991).

Matthews, R. Get connected. *New Scientist* 164, 24-28 (1999).

Milgram, S. The small world problem. *Psychology Today* 2, 60-67 (1967).

Mirollo, R. E. and Strogatz, S. H. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics* 50, 1645-1662 (1990).

Mitchell, M., Crutchfield, J. C., and Hraber, P. T. Evolving cellular automata to perform computations: mechanisms and impediments. *Physica D* 75, 361-391 (1994).

Mizruchi, M. S. and Potts, B. B. Centrality and power revisited: actor success in group decision making. *Social Networks* 20, 353-387 (1998).

Monasson, R. Diffusion, localization and dispersion relations on 'small-world' lattices. *European Physical Journal B* 12(4), 555-567 (1999).

Moody, J. and D.R. White. Social Cohesion and Embeddedness: A Hierarchical Conception of Social Groups. *American Sociological Review* 68(1):1-25.

Moukarzel, C. F. Spreading and shortest path in systems with sparse long-range connections. *Physical Review E*. 60(6), 6263-6266 (1999).

Newman, M. E. J. Models of the Small World: A Review. <http://xxx.lanl.gov/abs/cond-mat/0001118> (2000).

Newman, M. E. J. The structure of scientific collaboration networks. *Proceedings of the National Academy of Science* (submitted). Available online at <http://xxx.lanl.gov/abs/cond-mat/0007214> (2000).

Newman, M. E. J., Moore, C. and Watts, D. J. Mean-field solution of the small-world network model. *Physical Review Letters* 84, 3201-3204 (2000).

Newman, M. E. J. and D.J. Watts. a Scaling and percolation in the small-world network model. *Physical Review E* 60, 7332-7342 (1999).

Newman, M.E.J. and Watts, D. J. b Renormalization group analysis of the small-world network model. *Physics Letters A* **263**, 341-346 (1999).

Newman, M.E.J., Watts, D. J. and Strogatz, S. H. Collaboration networks as random bipartite graphs. *Nature* (submitted, July 2000).

Newman, M.E.J., Strogatz, S. H. and Watts, D. J. Random graphs with arbitrary degree distributions and their applications. <http://xxx.lanl.gov/abs/cond-mat/0007235> (2000).

Nowak, M. A. and May, R. M. Evolutionary games and spatial chaos. *Nature* 359, 826-829 (1992).

Nowak, M. A. and May, R. M. The spatial dilemmas of evolution. *International Journal of Bifurcations and Chaos* 3 (1), 35-78 (1993).

Radner, R. The organization of Decentralized information processing. *Econometrica* 61(5), 1109-1146 (1993).

Radner, R. Bounded rationality, indeterminacy, and the theory of the firm. *The Economic Journal*. 106, 1360-1373 (1996).

Rapoport, A. Spread of information through a population with socio-structural bias. I. Assumption of transitivity. *Bulletin of Mathematical Biophysics* 15, 523-533 (1953 a.).

Rapoport, A. Spread of information through a population with socio-structural bias. II. Various models of partial transitivity. *Bulletin of Mathematical Biophysics* 15, 535-546 (1953 b.).

Rapoport, A. A contribution to the theory of random and biased nets. *Bulletin of Mathematical Biophysics* 19, 257-271 (1957).

Sachtjen, M. L., Carreras, B. A., and Lynch, V. E. Disturbances in a power transmission system. *Physical Review E*. 61(5), 4877-4882 (2000).

Sakaguchi, H., Shinomoto, S., and Kuramoto, Y. Local and global self-entrainment in oscillator lattices. *Progress of Theoretical Physics* 77(5), 1005-1010 (1987).

Satoh, K. Computer experiment on the cooperative behavior of a network of interacting nonlinear oscillators. *Journal of the Physical Society of Japan* 58(6), 2010-2021 (1989).

Sattenspiel, L. and Simon, C. P. The spread and persistence of infectious diseases in structured populations. *Mathematical Biosciences* 90, 341-366 (1988).

Schelling, T. C. A study of binary choices with externalities. *Journal of Conflict Resolution*. 17(3), 381-428 (1973).

- Shulman, P. From Muhammad Ali to Grandma Rose. *Discover* 19 (12), 84-89 (1998).
- Shrager, J., Hogg, T., and Huberman, B. A. Observation of phase transitions in spreading activation networks. *Science* 236, 1092-1094 (1987).
- Skvoretz, J. Random and biased networks: simulations and approximations. *Social Networks* 7, 225-261 (1985).
- Skvoretz, J. Theoretical and methodological models of networks and relations. *Social Networks* 13(3), 275-300 (1991).
- Solomonoff, R. and Rapoport, A. Connectivity of random nets. *Bulletin of Mathematical Biophysics* 13, 107-117 (1951).
- Sporns O., Tononi G. and Edelman G. M. Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex* (in press).
- Stark, D. C. Recombinant Property in East European Capitalism. *American Journal of Sociology*. 101(4), 993-1027 (1996).
- Stauffer, D. Simulation of Griffeath majority rule on large square lattice. *International Journal of Modern Physics C*. 8(5), 1141-1144 (1997).
- Strang, G. and H. Eriksson, Preprint, Mathematics Department, MIT (1999).
- Strogatz, S. H. and Stewart, I. Coupled oscillators and biological synchronization. *Scientific American* 269 (6), December, 102-109 (1993).
- Valente T. W. Social network thresholds in the diffusion of innovations. *Social Networks*, 18, 69-89 (1996).
- Watts, D. J. Networks, dynamics and the small-world phenomenon. *American Journal of Sociology* 105(2):⁷⁶ 493-527 (1999 a.).

Watts, D. J. Global cascades on binary switching networks. Santa Fe Institute working paper, available on request (2000).

Watts D. J. and Strogatz S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440-442 (1998).

West, G. B., Brown, J. H., and Enquist, B. J. A General Model for the Origin of Allometric Scaling Laws in Biology. *Science* 276, 122-126 (1997).

White, D. R., and F. Harary. The Cohesiveness of Blocks in Social Networks: Connectivity and Conditional Density. *Sociological Methodology*(2001), 31:1, 305-359 (2001). Blackwell Publishers, Inc., Boston, USA and Oxford, UK.

White, D. R., and Jorion, P. Kinship Networks and Discrete Structure Theory: Applications and Implications. *Social Networks* 18:267-314 (1996).

White, D. R., and Jorion, P. Representing and Computing Kinship: A New Approach *Current Anthropology*, 33:4, 454-463 (1992).

White, H. C. What is the center of the small world? Paper presented at AAAS annual symposium, Washinton D.C. (2000).

Wiesenfeld, K., Colet, P., and Strogatz, S. H. Synchronization transitions in a disordered Josephson series array. *Physical Review Letters* 76, 404-407 (1996).

Wilf, H. S. *Generatingfunctionology* (Academic Press, San Diego, 1990).

Winfrey, A. T. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of Theoretical Biology* 16, 15-42 (1967).

Wuensche, A., *Discrete Dynamical Networks and Their Attractor Basins*, Santa Fe Institute working paper, 1998.