Complex (not only neural) networks

Péter Érdi
erdi@rmki.kfki.hu

Tamás Kiss
bognor@sunserv.kfki.hu

Department of Biophysics
KFKI Research Institute for Particle and Nuclear Physics of the
Hungaryan Academy of Sciences, Budapest, Hungary
http://www.rmki.kfki.hu/biofiz/cneuro

and

Center for Complex Systems Studies
Kalamazoo College
1200 Academy Street
Kalamazoo, Michigan, USA
http://www.kzoo.edu/physics/ccss/

Contents

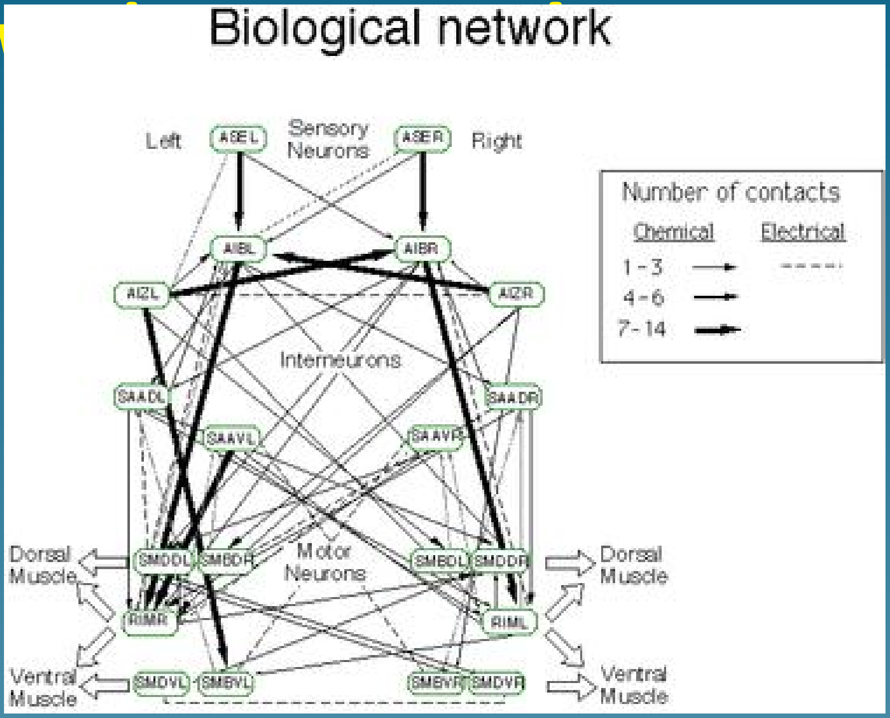
Contents

Contents

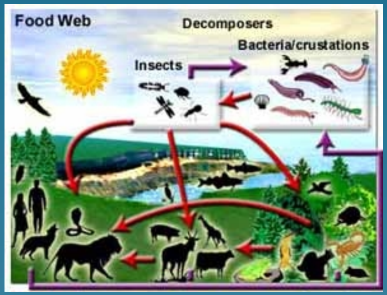
Contents

1. TOPOLOGY of REAL NETWORKS

Netw


Biological network
Left
ASEL
Sensory Neurons
ASER
Right
Number of contacts
Chemical
Electrical
1-3
4-6
7-14
AIBL
AIBR
AIZL
AIZR
Interneurons
SAADL
SAADR
SAAVL
SAAVR
Dorsal Muscle
SMDDL
SMBDR
Motor Neurons
SMBDL
SMDDR
Dorsal Muscle
RIMR
RIML
Ventral Muscle
SMDVL
SMBVL
SMBVR
SMDVR
Ventral Muscle

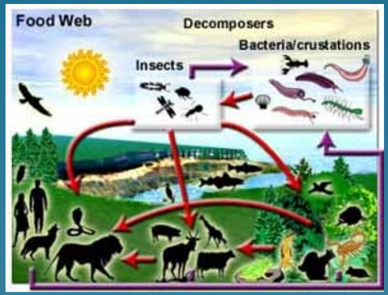
Networks everywhere



Food Web

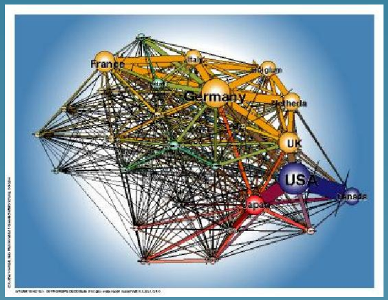
from bacteria to human beings

Networks everywhere



Food Web

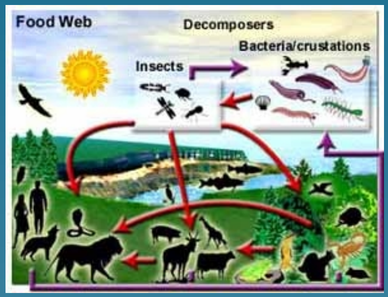
from bacteria to human beings



World Trade Networks

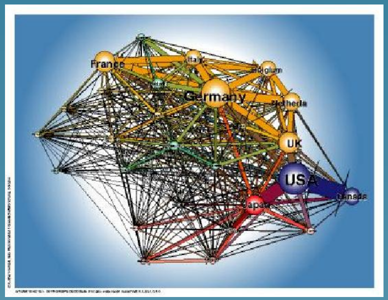
among 28 OECD countries

Networks everywhere



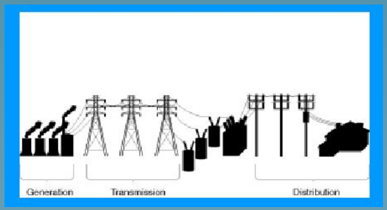
Food Web

from bacteria to human beings



World Trade Networks

among 28 OECD countries



Electric Power Network

nodes: generators, transformers

egdes: transformation lines

Social Networks

- *Acquaintanceship Networks* (Rapoport, Milgram) What is the probability that any two people selected arbitrarily from a large population will know each other?

Social Networks

- *Acquaintanceship Networks* (Rapoport, Milgram)
  What is the probability that any two people selected arbitrarily from a large population will know each other?

- *Hollywood Universe*
  From the shortest path from any actor to Kevin Bacon, using the association rule.
  Unlike society in general, film actor associations are well documented.

Social Networks

- *Acquaintanceship Networks* (Rapoport, Milgram)
  What is the probability that any two people selected arbitrarily from a large population will know each other?

- *Hollywood Universe*
  From the shortest path from any actor to Kevin Bacon, using the association rule.
  Unlike society in general, film actor associations are well documented.

- *Collaborative graphs*

Paul Erdős
(1913-1996)

Social Networks

Erdős Number

The Erdős Number Project
http://www.oakland.edu/grossman/erdosnp.html

E. n. 0	1	person
E. n. 1	502	people
E. n. 2	5713	people
E. n. 3	26422	people
E. n. 4	62136	people
E. n. 5	66158	people
E. n. 6	32280	people
E. n. 7	10431	people
E. n. 8	3214	people
E. n. 9	953	people
E. n.10	262	people
E. n.11	94	people
E. n.12	23	people
E. n.13	42	people
E. n.14	7	people
E. n.15	1	people
E. n.16	0	people

Average Erdős Number: 4.69

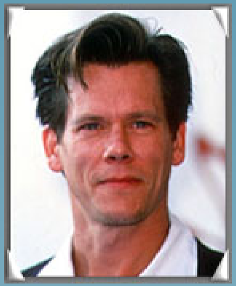
Social Networks

Hollywood Graph

429065 actors, 170479 films
Bacon numbers and their distributions

0	1
1	1465
2	109974
3	256183
4	56936
5	3932
6	477
7	69
8	27
9	1

average: 2.89
912 actors have smaller than average

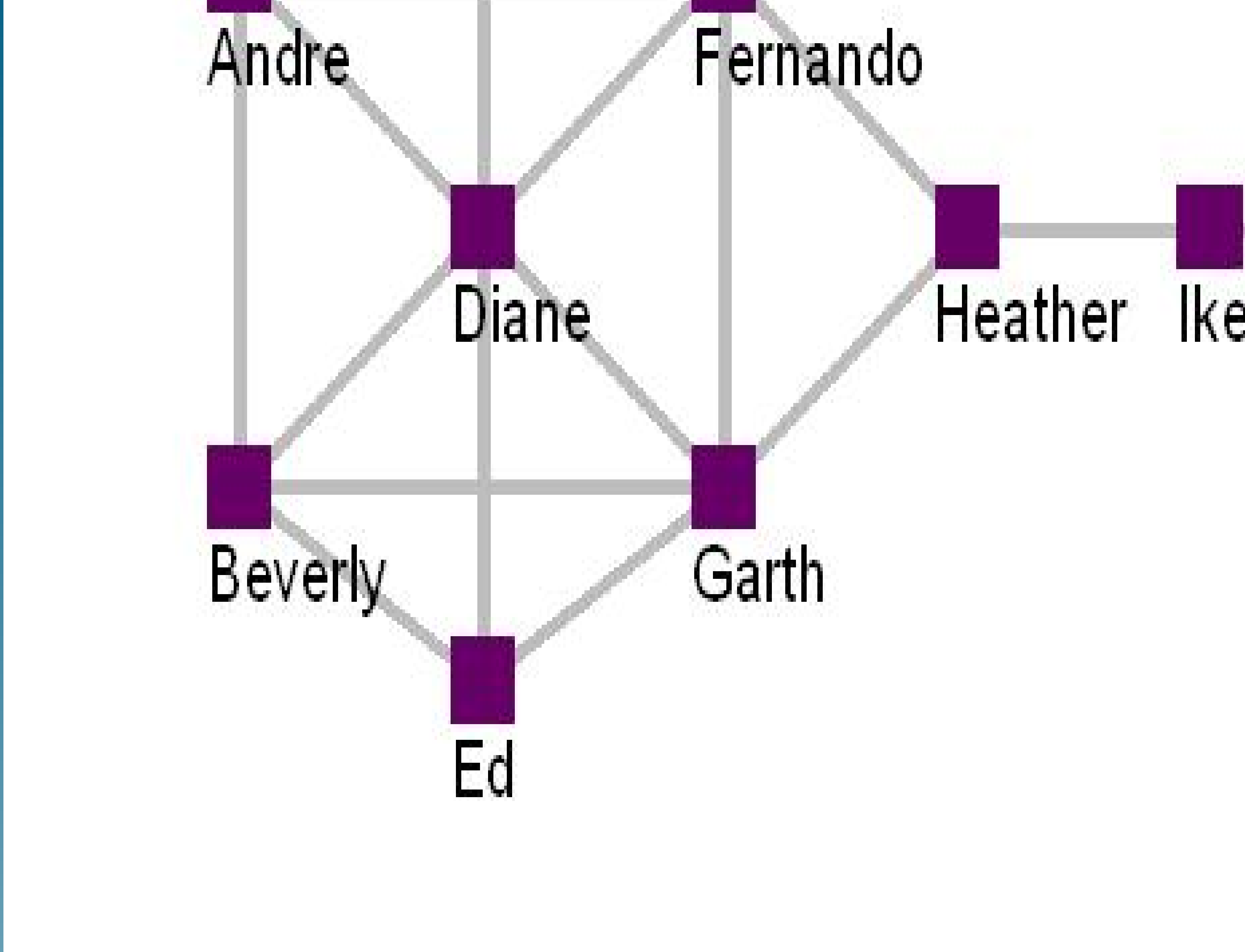


Kevin Bacon

Social Networks

Degree: the number of direct connections a node has. In the network above, Diane has the most direct connections in the network. She is a 'connector' or 'hub' in this network.

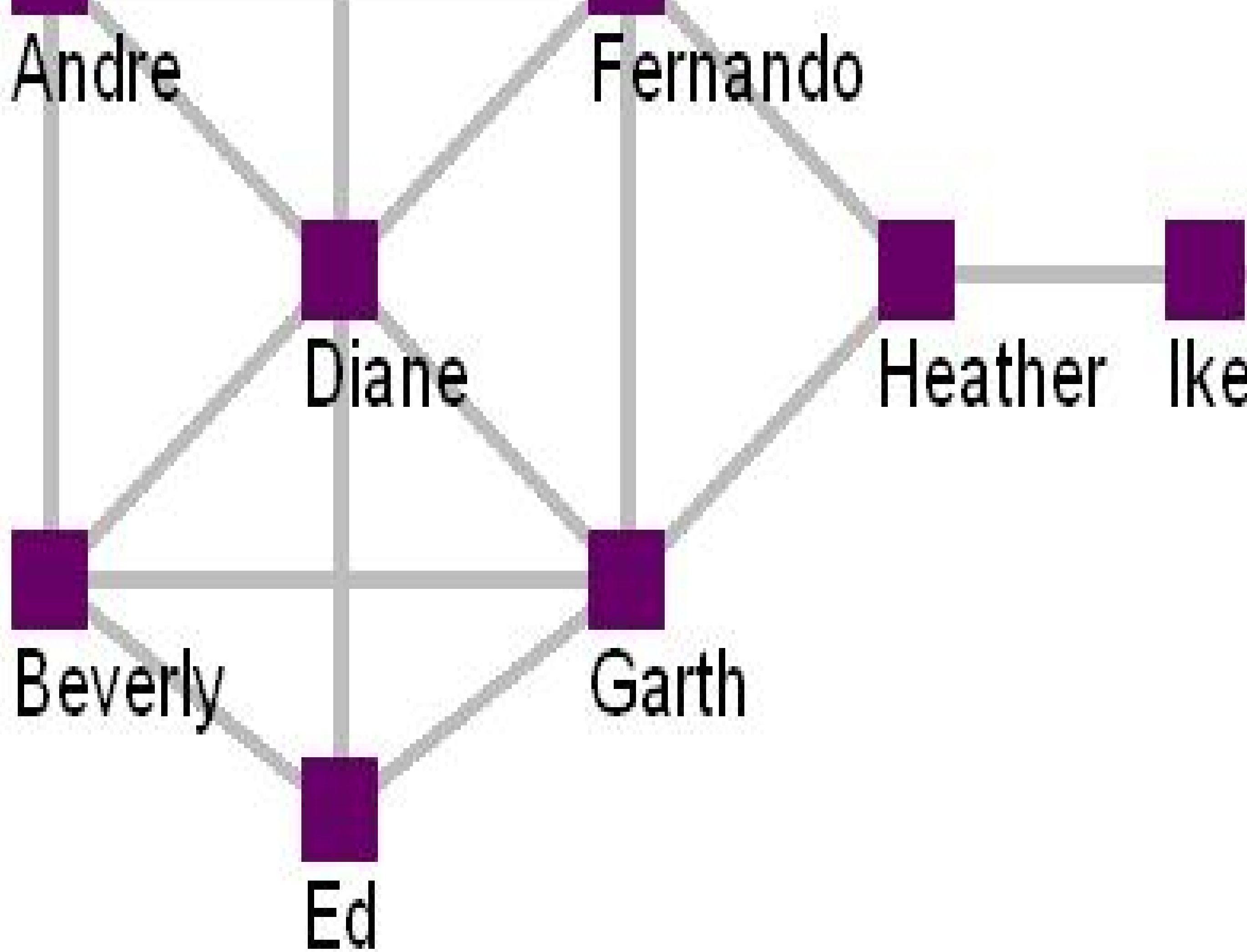
Social Networks


Andre
Fernando
Diane
Heather
Ike
Beverly
Garth
Ed

Betweenness: While Diane has many direct ties, Heather has few direct connections – fewer than the average in the network. Yet, in may ways, she has one of the best locations in the network – she is between two important constituencies. She plays a 'broker' role in the network.

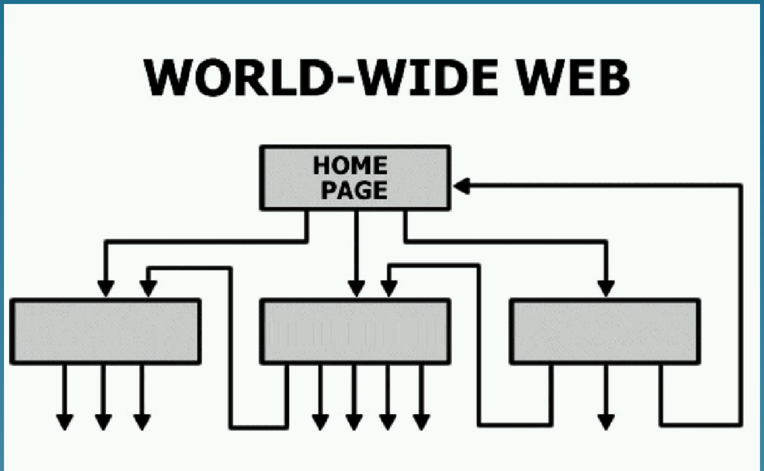
Social Networks

Social Network Analysis

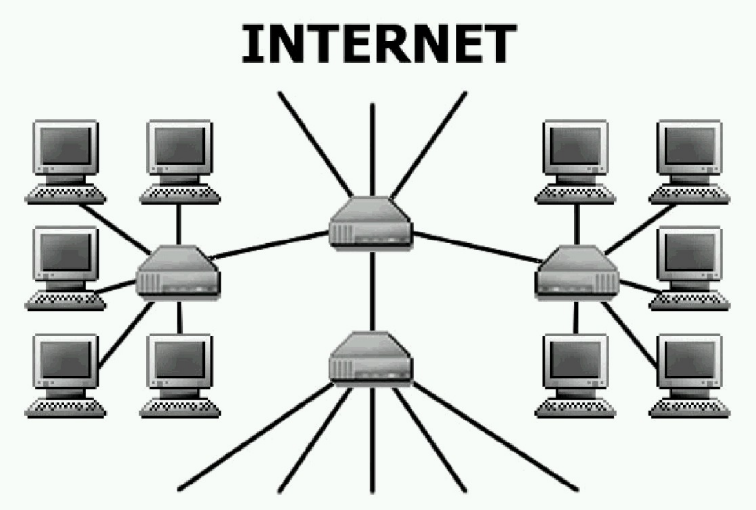

Andre
Fernando
Diane
Heather
Ike
Beverly
Garth
Ed

Closeness: Fernando and Garth have fewer connections than Diane, yet the pattern of their direct and indirect ties allow them to access all the nodes in the network more quickly than anyone else. They have the shortest paths to all others – they are close to everyone else.

World Wide Web and the Internet



nodes: web documents
edges: directed hyperlinks

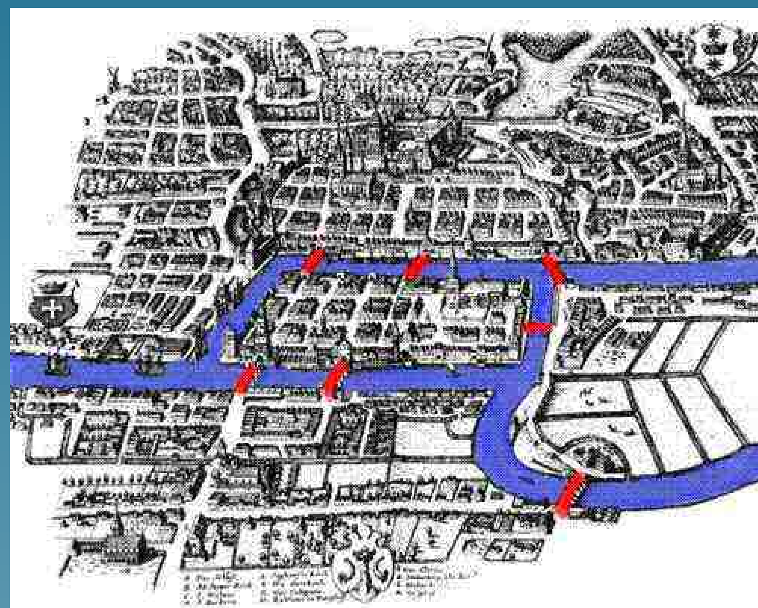


nodes: routers and computers
edges: wires and cables

2. STATISTICAL ANALYSIS of LARGE GRAPHS

Graphs

$<$nodes or vertices, edges$>$
Euler and the bridges of Königsberg (Kalinyingrad)

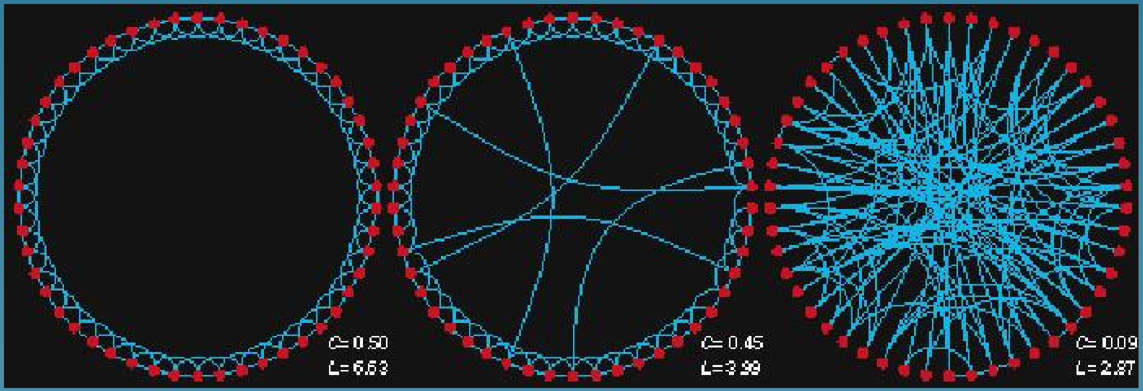


Statistical analysis of large graphs
Brain Hayes: Graph Theory in Practice I, II

American Scientist 88(1), 88(2) 2000

Graphs

Regular, Random and Real World Graphs



lattice like (several neighbours)	regular + random effects	random

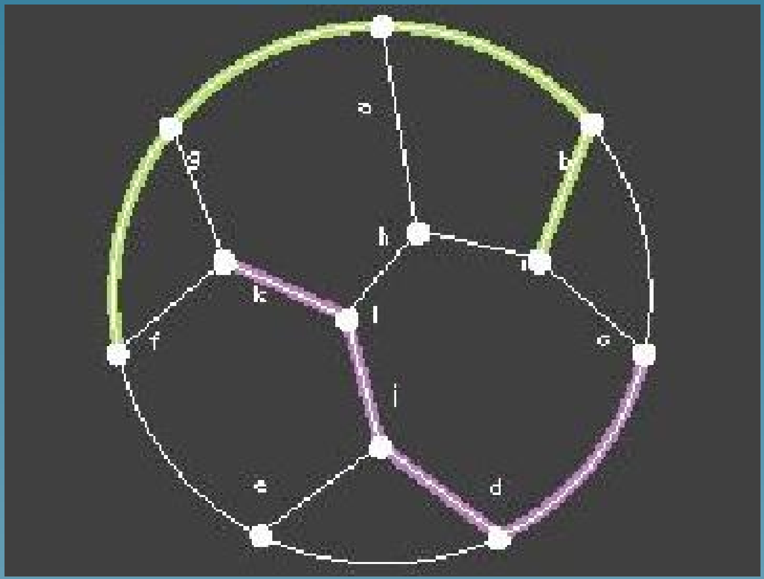
Graphs

Graph Characteristics

Diameter:
maximal minimal path

Characteristic path lenght:
minimal path lengths averaged for all pairs of vertices

(complete graph, clique)
many edges are necessary

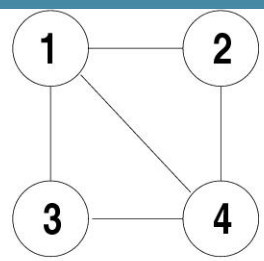


Graphs

Measure of Clustering (C)

(are my friends also friends of each other?)

List all the neighbours of a vertex, count the edges that link those neighbours, and divide by the maximum number of edges that could possibly be drawn among the neighbours; then repeat these operations for all the vertices, and take the average.



Node	1.	2.	3.	4.	
Neighbors	2, 3, 4	1, 4	1, 4	1, 2, 3	
Number of edges connecting the neighbor nodes	2	1	1	2	
Maximal number of edges connecting the neighbor nodes	3	1	1	3	
Ratio	2/3	1	1	2/3	5/6

Elements of Random Graph Theory

Illustration of the Erdős – Rényi Theorem

n nodes, no edges; p is the probability that we drag an edge to a pair of nodes, for each pair.

$p = 0 \longrightarrow$ no edge; $p = 1 \longrightarrow$ clique

generally the number of edges $pn(n-1)/2$; randomly and independently

for large graphs $(n \longrightarrow \infty, e \longrightarrow \infty, n/e = \text{const})$: $e > n/2 \longrightarrow$ there exists a 'giant component' i.e. a connected piece of the graph spanning most of the vertices,

the distance between two arbitrarily chosen points is 'rather small'

Elements of Random Graph Theory

Illustration of the Erdős – Rényi Theorem

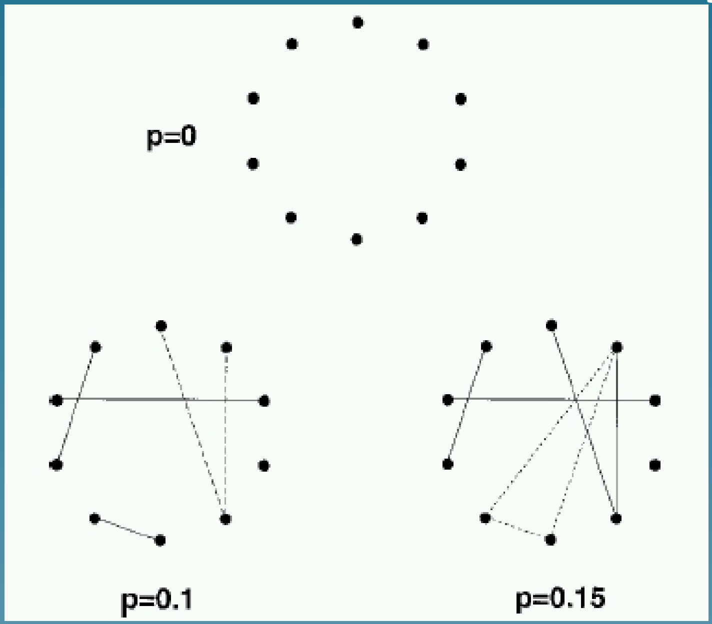
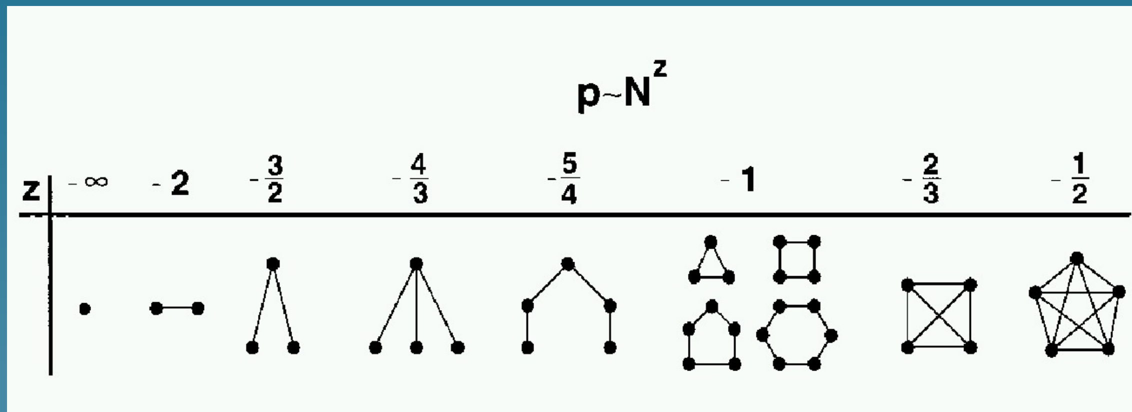


Illustration of the graph evolution process for the Erdős-Rényi model. We start with $N = 10$ isolated nodes, then connect every pair of nodes with probability p. The lower panel of the figure shows two different stages in the graph's development, corresponding to $p = 0.1$ and $p = 0.15$. We can notice the emergence of trees (a tree of order 3, drawn with long-dashed lines) and cycles (a cycle of order 3, drawn with short-dashed lines) in the graph, and a connected cluster that unites half of the nodes at $p = 0.15 = 1.5/N$.

Elements of Random Graph Theory

Illustration of the Erdős – Rényi Theorem

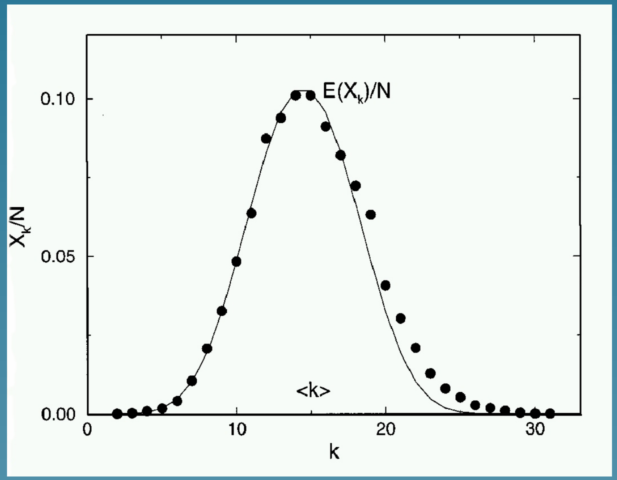


The threshold probabilities at which different subgraphs appear in a random graph. For $pN^{3/2} \longrightarrow 0$ the graph consists of isolated nodes and edges. For $p \sim N^{-3/2}$ trees of order 3 appear, while for $p \sim N^{-4/3}$ trees of order 4 appear. At $p \sim N^{-1}$ trees of all orders are present, and at the same time cycles of all orders appear. The probability $p \sim N^{-2/3}$ marks the appearance of complete subgraphs of order 4 and $p \sim N^{-1/2}$ corresponds to complete subgraphs of order 5. As z approaches 0, the graph contains complete subgraphs of increasing order.

Elements of Random Graph Theory

Illustration of the Erdős – Rényi Theorem

The Edge Distribution:



Comparison of the digree distribution (X_k/N) of a network of $N = 10000$ nodes and $p = 0.0015$ with the expectation value of the Poisson distribution below, $E\left(X_k\right)/N = P(k_i = k)$.

$$
\begin{aligned}
E(X_k) &= NP(k_i = k) = \lambda_k \\
\lambda_k &= NC^k_{N-1}p^k(1-p)^{N-1-k} \\
P(X_k = r) &= e^{-\lambda_k}\frac{\lambda_k^r}{r!}
\end{aligned}
$$

Elements of Random Graph Theory

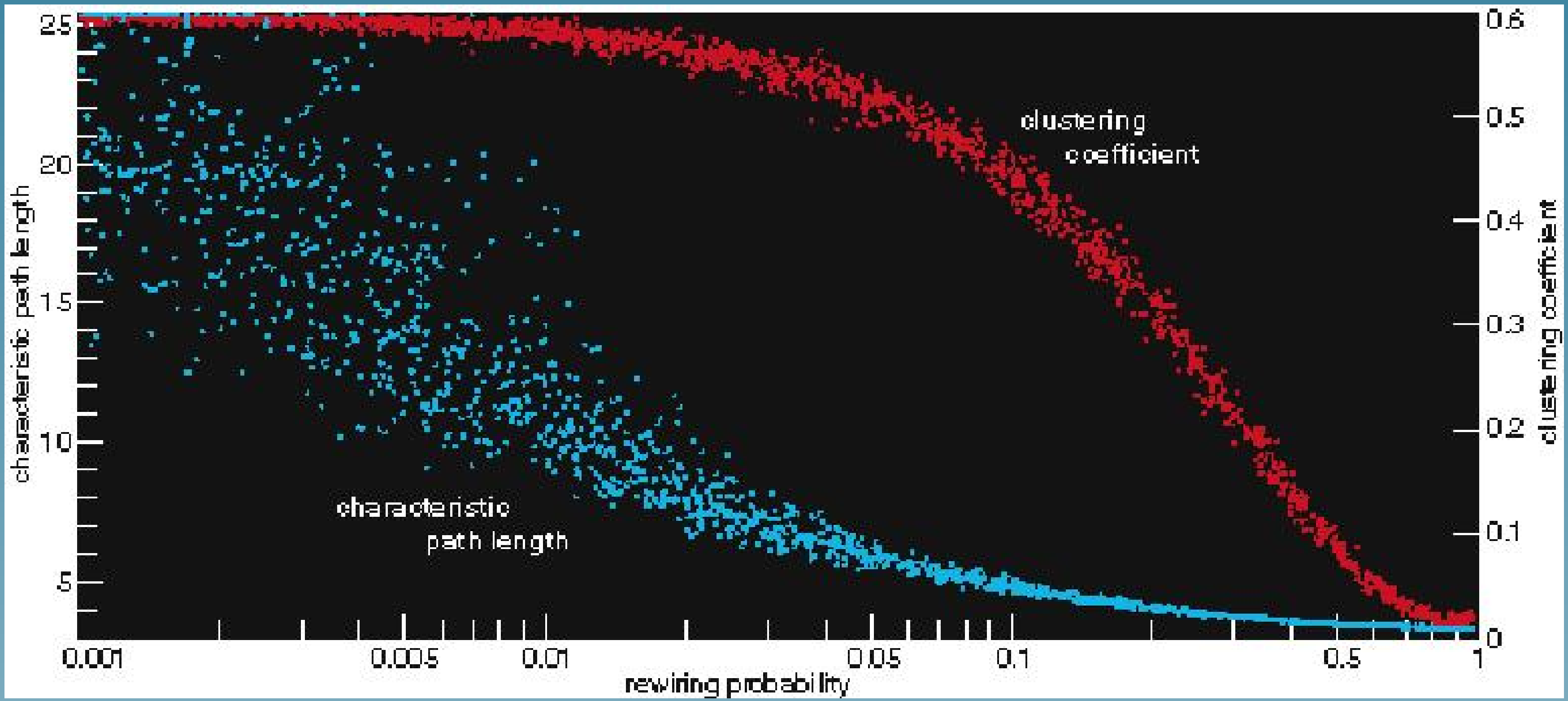
Real World Graphs: Small-World Networks

- It is not true that we know our neighbours only. It is also not true that our relationships are completely random.

Elements of Random Graph Theory

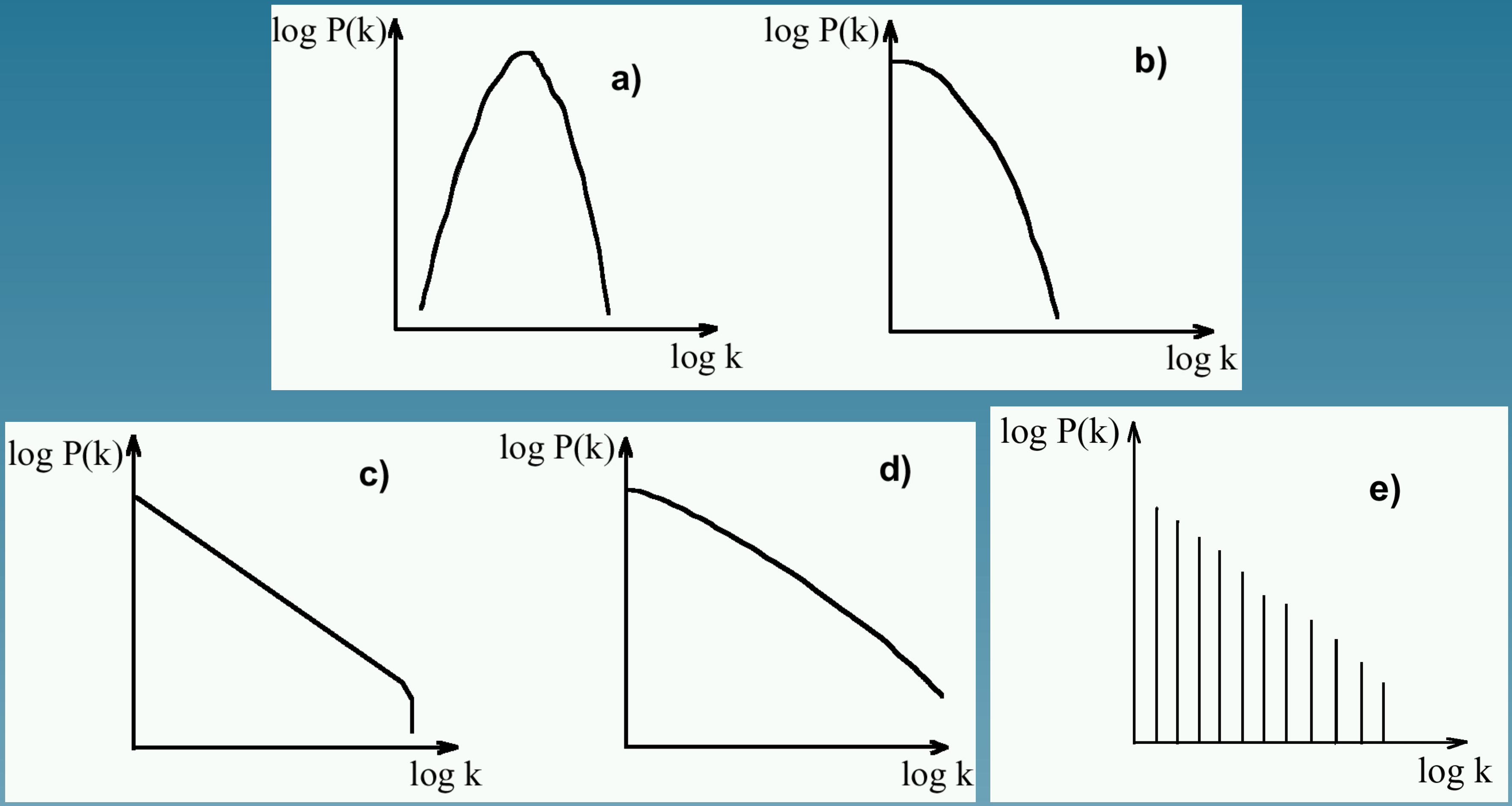
Real World Graphs: Small-World Networks

- It is not true that we know our neighbours only. It is also not true that our relationships are completely random.
- Generation of the Watts-Strogatz “Small World Graph”:
  1. initial configuration is a regular lattice
  2. each edge is examined, and is redirected with p probability to an other target node (chosen also randomly)



Elements of Random Graph Theory

Real World Graphs: Scale-free Model

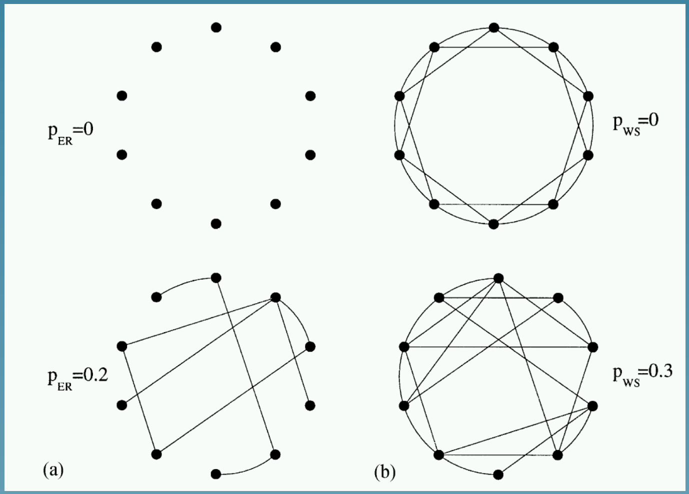


"Zoology" of degree distributions in networks. Main types of a degree distribution in log-log plots. Poisson (a), exponential (b), power-law (c), multifractal (d), and discrete (e) distributions.

3. DEVELOPMENT of NETWORKS

Formation of Scale-free Networks

Random Growth vs. Preferential Attachment



The **Erdős-Rényi** (ER) (a) and the **Watts-Strogatz** (WS) (b) models. A random network described by the ER model has N vertices connected with probability p_{ER}, the total number of edges in the system being $n = p_{ER}N(N-1)/2$. The example presents a network of $N = 10$ vertices for $p_{ER} = 0$ and $p_{ER} = 0.2$. The WS model starts with a regular one-dimensional lattice with edges between the nearest and next-nearest neighbors. Then a fraction p_{WS} of the edges is rewired randomly (their endpoint is changed to a randomly selected vertex). The example presents a network of $N = 10$ vertices. For $p_{WS} = 0$ the system is a regular lattice with $2N = 20$ edges. For $p_{WS} = 0.3$, $2p_{WS}N = 6$ edges have been rewired to randomly selected vertices.

Formation of Scale-free Networks

Random Growth vs. Preferential Attachment

The **Barabási-Albert** (BA) model.

(1) The ER and WS models assume that we start with a fixed number N of vertices that are then randomly connected or re-wired, without modifying N. In contrast, most real-world networks describe open systems that grow by the continuous addition of new nodes. In the BA model, after starting with a small, initial network, at every timestep a new node is added.

Formation of Scale-free Networks

Random Growth vs. Preferential Attachment

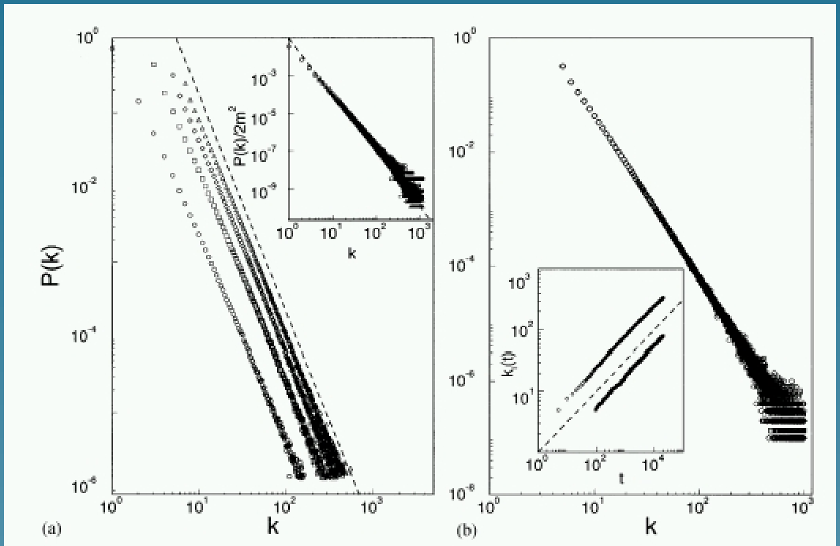
The **Barabási-Albert** (BA) model.
(1) The ER and WS models assume that we start with a fixed number N of vertices that are then randomly connected or re-wired, without modifying N. In contrast, most real-world networks describe open systems that grow by the continuous addition of new nodes. In the BA model, after starting with a small, initial network, at every timestep a new node is added.

(2) The ER and WS models assume that the probability that two nodes are connected (or their connection is rewired) is independent of the nodes degree, i.e., new edges are placed randomly. Most real networks however, exhibit preferential attachment, such that the likelihood of connecting to a node depends on the node s degree. The probability Π that a new node will be connected to node i depends on the degree k_i of node i , such that

$$\Pi\left(k_i\right) = \frac{k_i}{\sum_j k_j}$$

Formation of Scale-free Networks

The Barabási-Albert Model



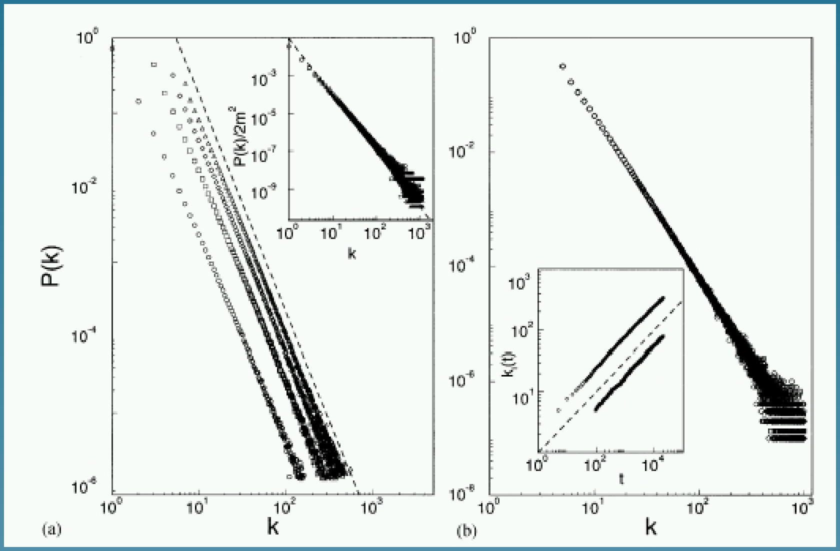
(a) Connectivity distribution of the BA model, with $N = m_0 + t = 300000$ and $m_0 = m = 1$ (circles), $m_0 = m = 3$ (squares), $m_0 = m = 5$ (diamonds) and $m_0 = m = 7$ (triangles). (b) P(k) for $m_0 = m = 5$ and system sizes $N = 100000$ (circles), $N = 150000$ (squares) and $N = 200000$ (diamonds). The inset shows the time-evolution for the connectivity of two vertices, added to the system at $t_1 = 5$ and $t_2 = 95$. Here $m_0 = m = 5$.

Given the assumption of growth and preferential attachment degree of nodes satisfies the dynamical equation:

$$\frac{\partial k_i}{\partial t} = m\Pi\left(k_i\right) = m\frac{k_i}{\sum_{j=1}^{N-1} k_j} \Longrightarrow k_i\left(t\right) = m\frac{t}{t_i}^{0.5}$$

Formation of Scale-free Networks

The Barabási-Albert Model



(a) Connectivity distribution of the BA model, with $N = m_0 + t = 300000$ and $m_0 = m = 1$ (circles), $m_0 = m = 3$ (squares), $m_0 = m = 5$ (diamonds) and $m_0 = m = 7$ (triangles). (b) P(k) for $m_0 = m = 5$ and system sizes $N = 100000$ (circles), $N = 150000$ (squares) and $N = 200000$ (diamonds). The inset shows the time-evolution for the connectivity of two vertices, added to the system at $t_1 = 5$ and $t_2 = 95$. Here $m_0 = m = 5$.

Given the assumption of growth and preferential attachment degree of nodes satisfies the dynamical equation:

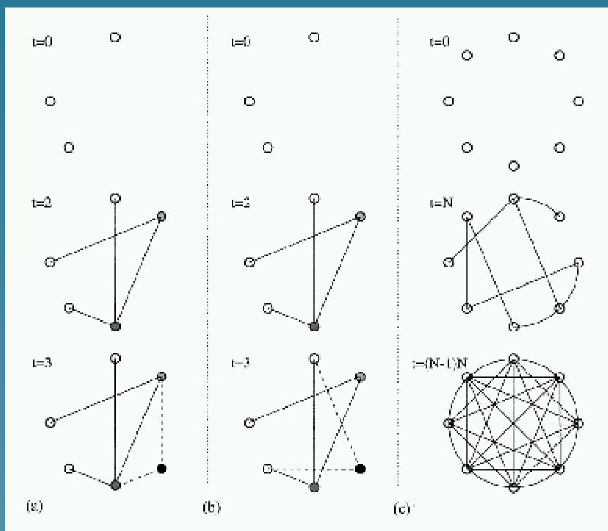
$$\frac{\partial k_i}{\partial t} = m\Pi\left(k_i\right) = m\frac{k_i}{\sum_{j=1}^{N-1} k_j} \Longrightarrow k_i\left(t\right) = m\frac{t}{t_i}^{0.5}$$

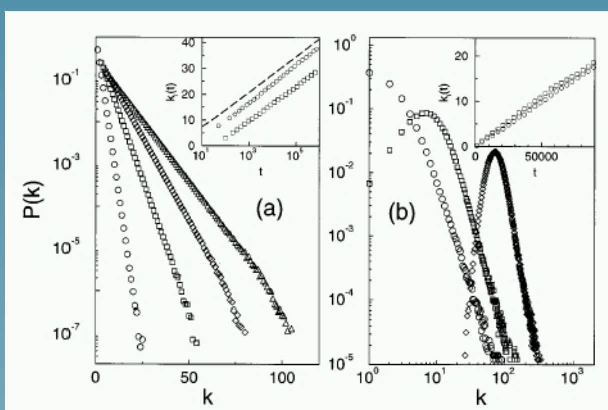
Probability density for $P(k)$ can be obtained:

$$P\left(k\right) = \frac{\partial P\left(k_i\left(t\right) < k\right)}{\partial k}$$

Formation of Scale-free Networks

The Barabási-Albert Model





Limited cases of the BA model:

Model A: growth without preferential attachment

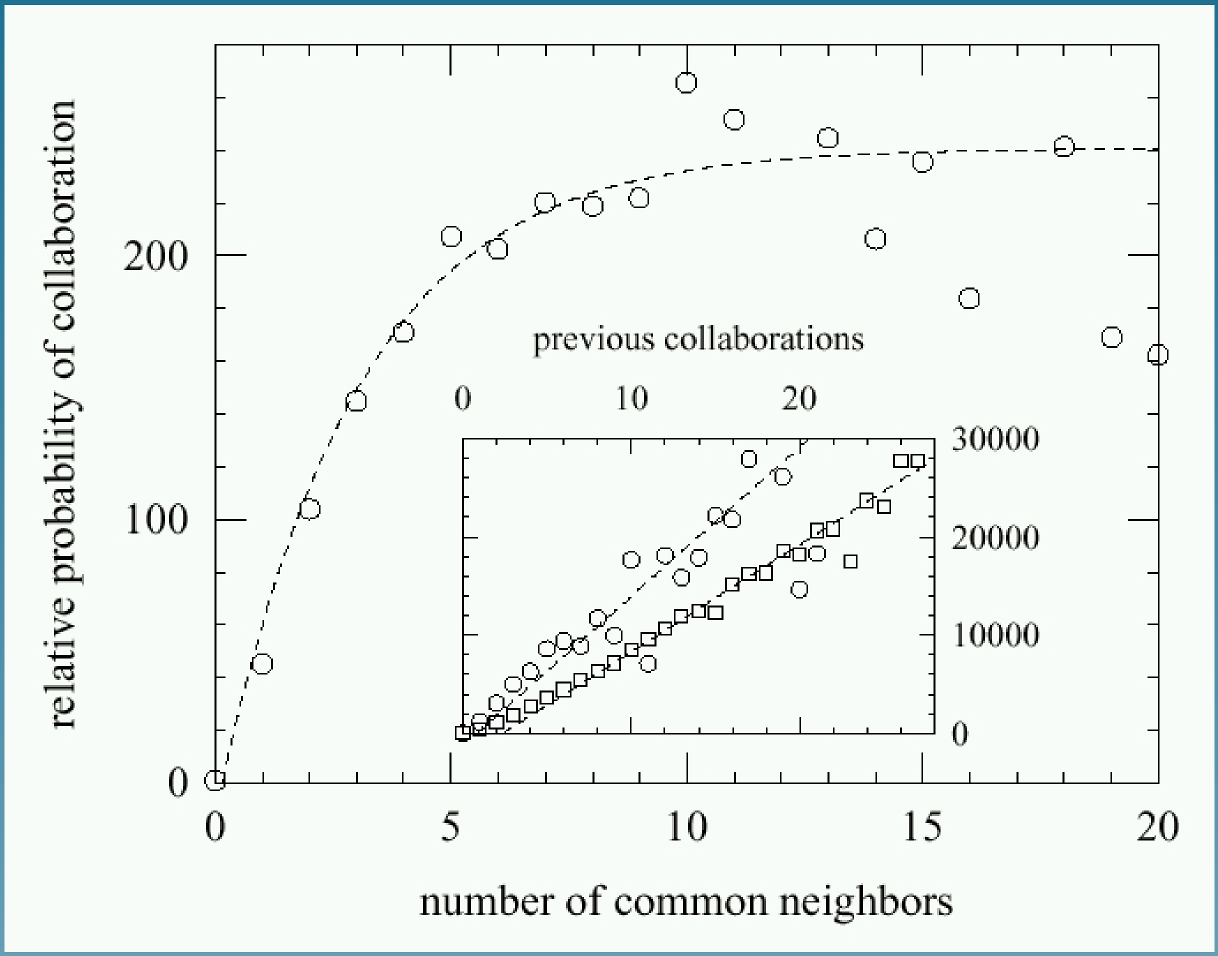
For $t \to \infty$ the degree distribution decays exponentially, $P(k) = \frac{e}{m} \exp\left(-\frac{k}{m}\right)$. The absence of preferential attachment eliminates the scale-free character.

Model B: preferential attachment without growing

After an initial transient when $P(k)$ follows a power-law scaling the degree distribution becomes a Gaussian.

Figures: upper panel show time evolution of (a) the BA model, (b) Model A and (c) Model B. Lower panel shows degree distributions for (a) Model A and (b) Model B.

Community Stuctures



Probability of collaboration between scientists in the Los Alamos Archive as a function of their number of mutual previous collaborators. Inset: the relative probability of collaboration as a function of number of previous collaborations of the same scientists, for the Los Alamos Archive (circles) and Medline (squares). The dotted lines are the best straight-line fits to the data. The data for Medline have been divided by a factor of 50 vertically to improve the clarity of the gure.

Sociopsychological Mechanisms and Algorithms

Rules of Connection Generation

- Trait vector: The trait is composed of different characteristics, such as psychological (temperament), physical, economical, social, aesthetic, intellectual etc.

Sociopsychological Mechanisms and Algorithms

Rules of Connection Generation

- Trait vector: The trait is composed of different characteristics, such as psychological (temperament), physical, economical, social, aesthetic, intellectual etc.
- Local: the establishment of an edge between the nodes of i and j depends only on the state/decisions of i and j.

Sociopsychological Mechanisms and Algorithms

Rules of Connection Generation

- Trait vector: The trait is composed of different characteristics, such as psychological (temperament), physical, economical, social, aesthetic, intellectual etc.
- Local: the establishment of an edge between the nodes of i and j depends only on the state/decisions of i and j.
- Global: there are other nodes, whose state/decision influence the edge formation.

Sociopsychological Mechanisms and Algorithms

Rules of Connection Generation

- Trait vector: The trait is composed of different characteristics, such as psychological (temperament), physical, economical, social, aesthetic, intellectual etc.
- Local: the establishment of an edge between the nodes of i and j depends only on the state/decisions of i and j.
- Global: there are other nodes, whose state/decision influence the edge formation.
- Passive: the formation of a new connection (or the deletion of an existing one) depends on the state of the "trait vectors" of the individuals.

Sociopsychological Mechanisms and Algorithms

Rules of Connection Generation

- Trait vector: The trait is composed of different characteristics, such as psychological (temperament), physical, economical, social, aesthetic, intellectual etc.
- Local: the establishment of an edge between the nodes of i and j depends only on the state/decisions of i and j.
- Global: there are other nodes, whose state/decision influence the edge formation.
- Passive: the formation of a new connection (or the deletion of an existing one) depends on the state of the "trait vectors" of the individuals.
- Active: The agents are active when they make decisions.

Sociopsychological Mechanisms and Algorithms

Rules of Connection Generation

- **Trait vector:** The trait is composed of different characteristics, such as psychological (temperament), physical, economical, social, aesthetic, intellectual etc.
- **Local:** the establishment of an edge between the nodes of i and j depends only on the state/decisions of i and j.
- **Global:** there are other nodes, whose state/decision influence the edge formation.
- **Passive:** the formation of a new connection (or the deletion of an existing one) depends on the state of the "trait vectors" of the individuals.
- **Active:** The agents are active when they make decisions.

Open problem: how to combine local and global, and passive and active rules, respectively.

The Small-world Network of the Human Language

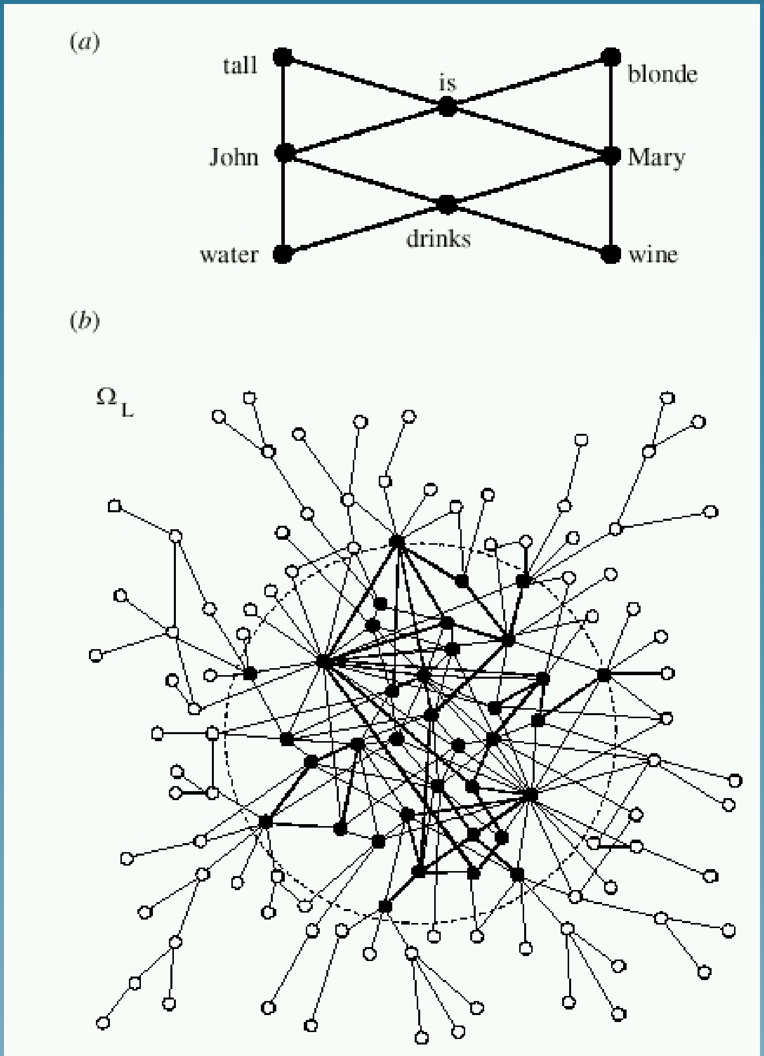
- sentence generation is rapid
- $1/f$ distribution in linguistics (Zipf's law)
- lexicons in human brain: $10^4 - 10^5$ words
- communication: kernel lexicon
- basic words and specialized words

$\Omega_L = (W_L, E_L)$, words: $W_L = \{w_i\}$, $i = 1, ..., N_L$, connections: $E_L = \{\{w_i, w_j\}\}$

connection between words: first- and second neighbours

[BUT: Chomsky: surface and dep structure $\longrightarrow$ distant neighbours]

The Small-world Network of the Human Language

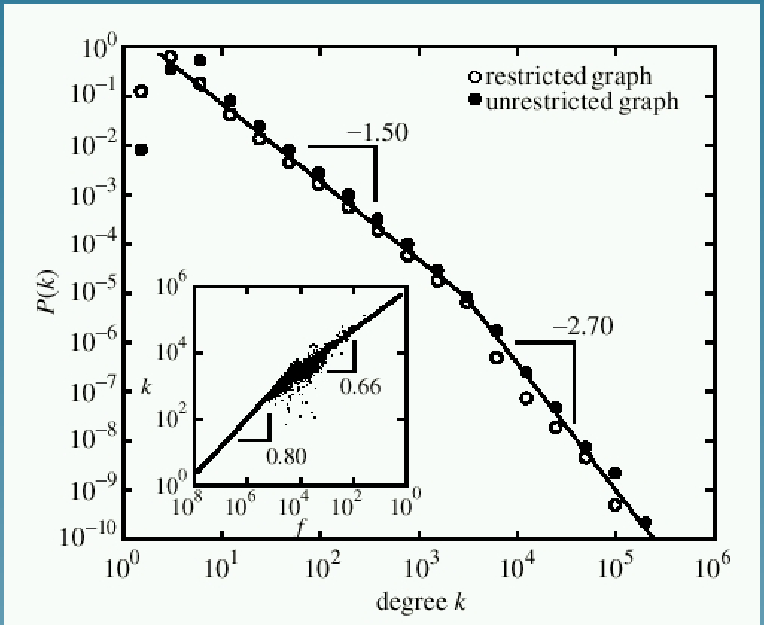


The figure shows an example of word networks. (a) A toy network constructed with four sentences: 'John is tall', 'John drinks water', 'Mary is blonde', 'Mary drinks wine'. (b) A possible pattern of wiring in Ω_L. Black nodes are common words and white nodes are rare words. Two words are linked if they co-occur significantly.

Assumption: from a sufficiently large text macroscopic properties of the network should emerge

Improvement: $p_{ij} > p_i p_j$: the presence of correlations beyond that expected from a random ordering of words (RWN: restricted word networks) UWN: else

The Small-world Network of the Human Language

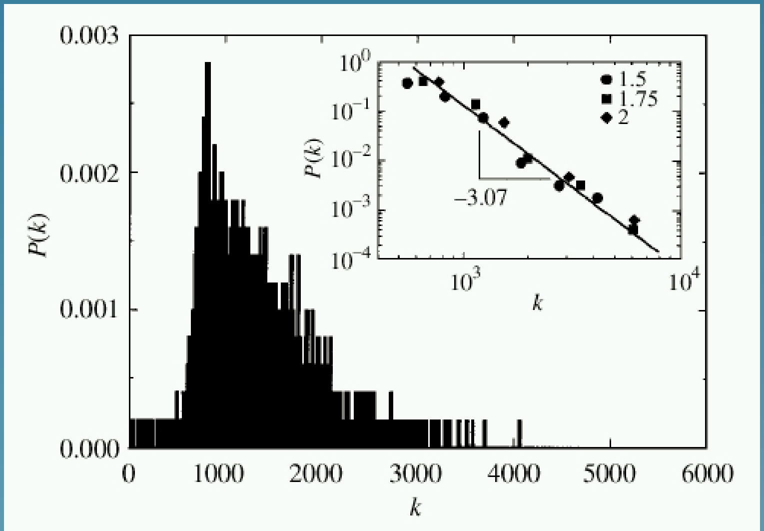


Degree distribution for the unrestricted word network (filled circles) and the restricted word network (open circles). The distribution function is obtained after processing about three- quarters of the 10^7 words of the British National Corpus (http://info.ox.ac.uk/bnc/). The obvious limitations of our methods are overcome by the use of a large amountPoints are grouped by powers of two. Inset: average degree as a function of frequency. Degree increases as a function with frequency, with exponent 0.80 for the first domain and 0.66 for the second one.

$\gamma_1 = -1.5$, $\gamma_2 = -2.7$

the network also has small world properites:
average minimum distance: ~ 2.6,
clustering coefficient: ~ 0.5

The Small-world Network of the Human Language



Connectivity distribution for the **kernel word network** (KWN), formed by the 5000 most connected vertices in RWN. Inset: power-law tail for $k > \overline{k}$ calculated by grouping in powers of 1.5, 1.75 and 2. The exponent of the power tail is $\gamma_{KWN} \approx -3$, indicating that preferential attachment is happening.

Every word on average is connected to 24% of the rest of the kernel words.

4. CONNECTIVITY of the BRAIN

Neural Networks

- The nervous system of the nematoda worm *Caeneorhabditis elegans* forms a small-world network.

Neural Networks

- The nervous system of the nematoda worm *Caeneorhabditis elegans* forms a small-world network.

- Mammalian cerebral cortex: its network is neither regular nor random

Neural Networks

- The nervous system of the nematoda worm *Caeneorhabditis elegans* forms a small-world network.
- Mammalian cerebral cortex: its network is neither regular nor random
- The distance of two arbitrarily choosen cortical neurons is 5 (John Szentágothai)

Neural Networks

Architectonics of the Cerebral Cortex,
edited by M. A. B. Brazier and H. Petsche.
Raven Press, New York © 1978

Specificity Versus (Quasi-) Randomness in Cortical Connectivity

J. Szentágothai

First Department of Anatomy, Semmelweis University Medical School, H-1450, Budapest, Hungary

NEURAL CONNECTIVITIES: BETWEEN DETERMINISM AND RANDOMNESS

P. Erdi[1,2] and J. Szentágothai[2]

[1] *Central Research Institute for Physics, Hungarian Academy of Sciences, Budapest, Hungary*

[2] *1st Department of Anatomy, Semmelweis University Medical School, Budapest, Hungary*

RÉNYI ALFRÉD—SZENTÁGOTHAI JÁNOS

AZ INGERÜLETÁTVITEL VALÓSZÍNŰSÉGE EGY EGYSZERŰ KONVERGENS KAPCSOLÁSÚ INTERNEURONÁLIS SYNAPSYS MODELLBEN

(Előzetes közlemény)

Theoretical Neuroanatomy "Small-world" Graphs

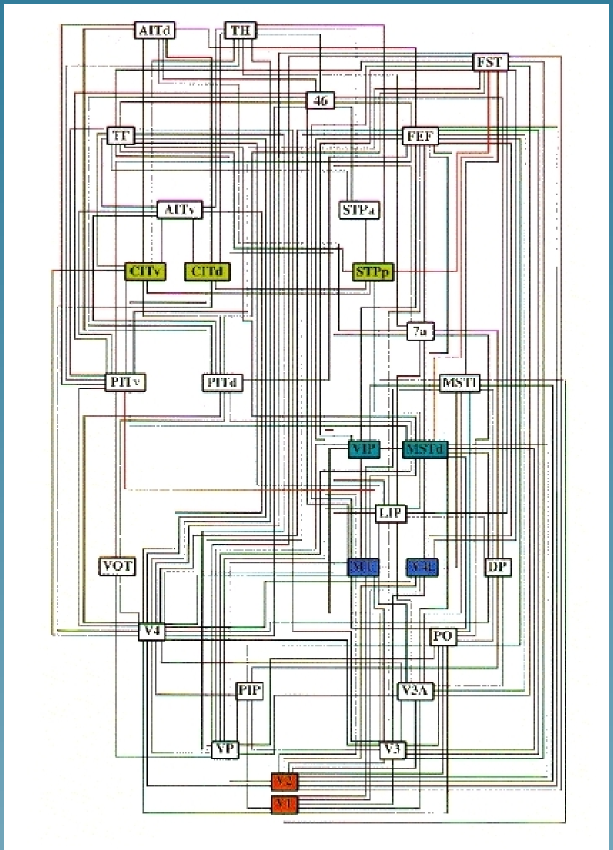
Neural Networks: Many Cells, Several Cell Types

Cell types of the hippocampus

T. Freund and colleagues

Theoretical Neuroanatomy “Small-world” Graphs

Hierarchy of Subsystems



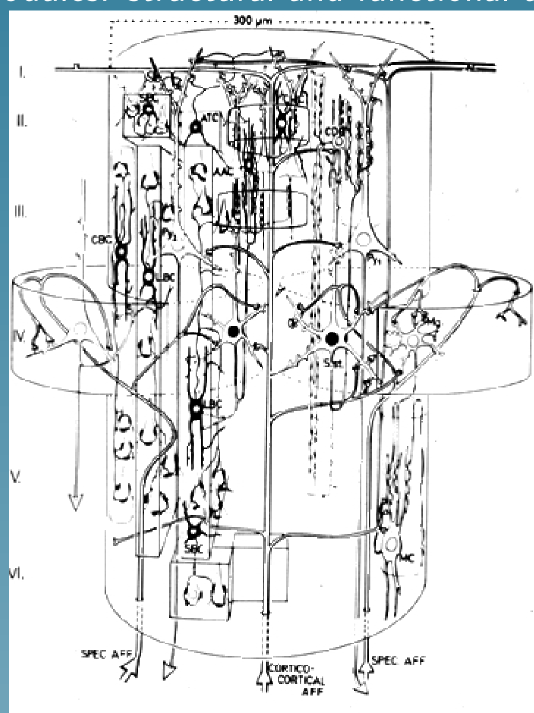
The Visual System

Van Essen and colleagues

Theoretical Neuroanatomy “Small-world” Graphs

Cortical Modules

- Wiring optimization in the brain - wiring economy principle?
- Limitations on the brain size require keeping the connection length as short as possible?
- Cortical modules: structural and functional units of information processing.



Cortical column as first visualized by J. Szentágothai

János Szentágothai (1912 – 1994)

Theoretical Neuroanatomy "Small-world" Graphs

Connectivity Databases & Neuroinformatics

- http://www.psychology.ncl.ac.uk/neuroinformatics.html (M. P. Young)
- The complexity of the brain, and the quantity and complexity of the data derived from it in the neuroscience, represents very substantial problems for brain science.
- Neuroinformatics: computer-based collation, management and analysis to neuroscience data, with the aims of making the complex data tractable, and of bringing mathematical and computational rigour to those areas that have not previousle benefitted from it.
- Cat cortico-cortical connectivity
- Parcellation of the cat thalamus
- Rat connectivity
- Connectivity data on the Macaque monkey

References

T. C. Ferree, B. A. Marcotte and S. R. Lockery: Neural network models of chemotaxis in the nematode Caenorhabditis elegans *Adv. in Neural Information Processing* 9:55-61, 1997

An Introduction to Social Network Analysis
`http://www.orgnet.com/sna.html`

Introduction to Social Network Methods A Working Bibliography on Social Network Analysis Method
`http://wizard.ucr.edu/~rhannema/networks/text/biblio.html`

A. Mrvar: Centrality and prestige
`http://mrvar.fdv.uni-lj.si/sola/info4/andrej/prpart4.htm`

R. Albert and A.-L. Barabási: Statistical mechanics of complex networks, *Reviews of Modern Physics* 74, 47 (2002)

A.-L. Barabási, R. Albert, and H. Jeong: Mean-field theory for scale-free random networks, *Physica A* 272 173-187 (1999)

B. Hayes: Graph Theory in Practice: Part I. *American Scientist* Vol. 88, No. 1, January-February 2000, pages 9-13.

B. Hayes: Graph Theory in Practice: Part II. *American Scientist* Vol. 88, No. 2, March-April 2000, pages 104-109

S. N. Dorogovtsev and J. F. F. Mendes: Accelerated growth of networks *arXiv:cond-mat/0204102* v1 4 Apr 2002

Evolution of Random Graphs
`http://www.business.auc.dk/evolution/esa/graphevol/graphevol.html`

P. Erdös and A. Rényi: On the Evolution of Random Graphs, *Publ. Math. Inst. Hung. Acad. Sci.*, Vol. 5, pp. 17-67, (1960)

M. Girvan and E. J. Newman: Community structure in social and biological networks, *Proc. Natl. Acad. Sci.* USA 99, 8271-8276 (2002)

References

R. F. i Cancho and R. V. Sole: The Small-World of Human Language, *Proceedings of the Royal Society of London Series, B-Biological Sciences* 268: (1482) 2261-2265, 2001

S. N. Dorogovtsev, J. F. F. Mendes, Language as an evolving word web, *Proceedings of the Royal Society of London Series, B-Biological Sciences* 268: (1485) 2603-2606, 2001

E. Motter, A. P. S. de Moura, Y.-C. Lai, and P. Dasgupta, Language, small-worlds, and cognition, *Physical Review E* 65, 065102(1-4) (2002)

J. Kleinberg. Navigation in a Small World. *Nature* 406(2000)

W. Aiello, F. Chung and L. Lu: A random graph model for massive graphs, *Proceedings of the Thirtysecond Annual ACM Symposium on Theory of Computing* (2000), 171-180

M. Steyvers, J .B. Tenenbaum: The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth `http://arxiv.org/abs/cond-mat/0110012`