Statistical Mechanics of Networks

TROISIEME CYCLE DE PHYSIQUE EN LA SUISSE ROMANDE

Guido Caldarelli

Together with

C.Caretta, M. Catanzaro, F. Colaiori, D. Garlaschelli, L. Pietronero, V. Servedio

INFM and Dipartimento di Fisica, Università “La Sapienza” Roma, Italy

L. Laura, S. Leonardi, S. Millozzi, A. Marchetti-Spaccamela

Dipartimento di Sistemistica e Elettronica, Università “La Sapienza” Roma Italy

P. De Los Rios[a], G. Bianconi[b], A. Capocci[b]

[a]Université de Lausanne, [b]Université de Fribourg, Switzerland

S. Battiston, A. Vespignani

Ecole Normale Supérieure and Université de Paris Sud, Paris France

•Contents

Part 1 *20-11-2003*

BASICS

A. **Networks as complex structures (9)**
B. **Fractals, Self-similarity (15)**
C. **Self-organization (8)**
D. **Data of scale-free networks (7)**
E. **Basic of Graphs (8)**

Part 2 *27-11-2003*

REAL GRAPHS

A. **Technological data: Internet, WWW**
B. **Social data: Finance and Board of Directors**
C. **Biological data: Proteins**

Part 3 *4-12-2003*

REAL TREES

A. **Geophysical data: the River Networks**
B. **Biological data: Taxonomy and Food Webs**
C. **Community Structures**

Part 4 *11-12-2003*

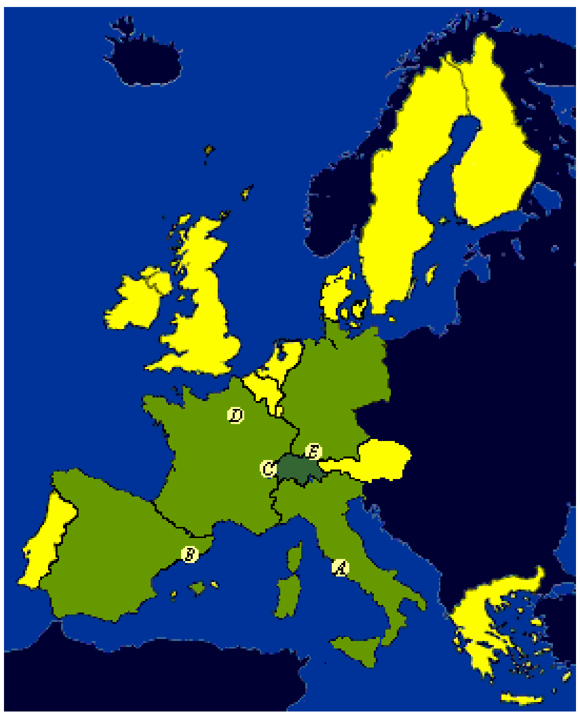
MODELS

A. **Random Graphs (Erdös-Renyi)**
B. **Small world**
C. **Preferential attachment**
D. **Fitness models**

COSIN
COevolution and Self-organisation In dynamical Networks

FET Open scheme RTD Shared Cost Contract IST-2001-33555

http://www.cosin.org



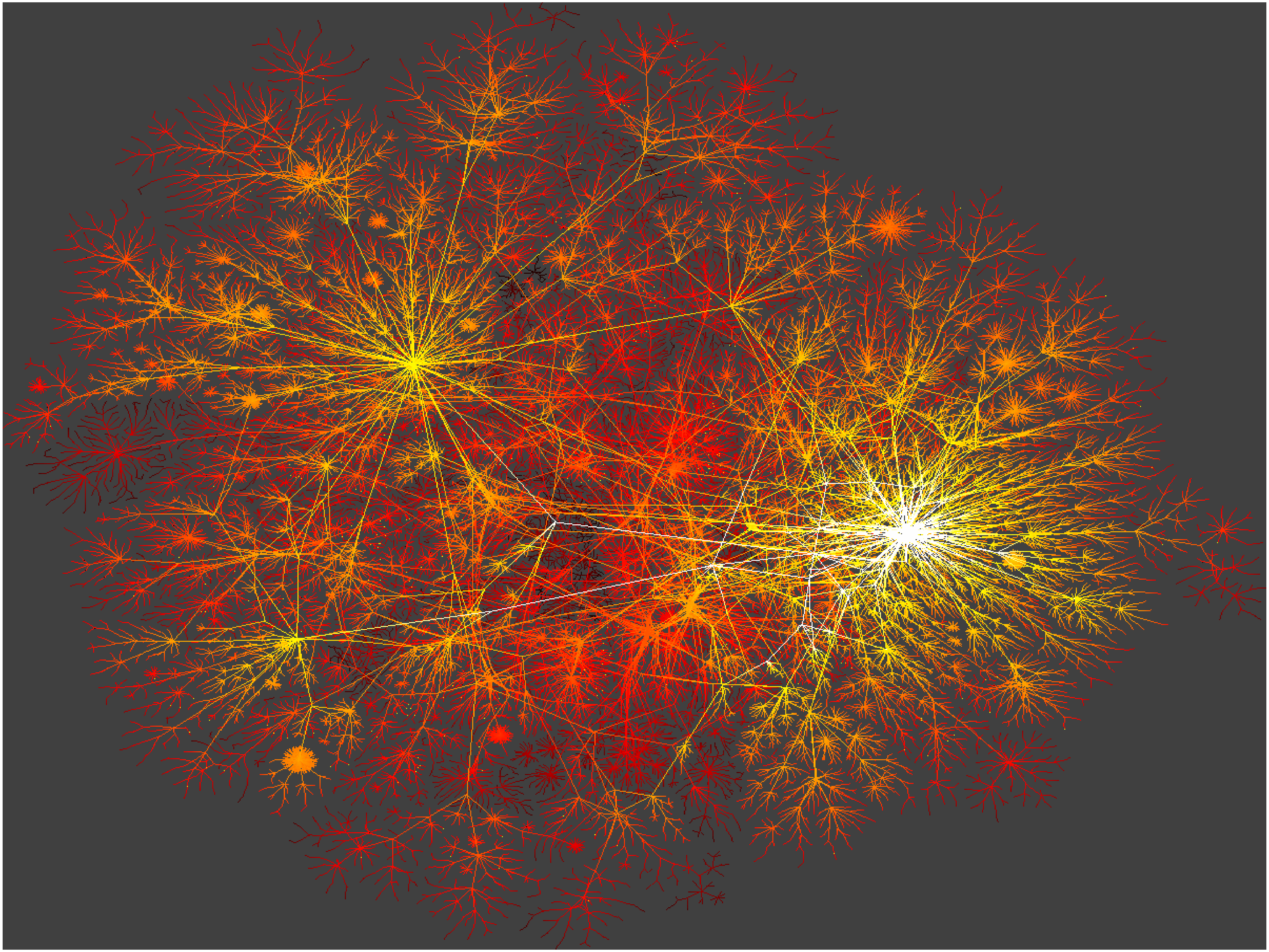
•	**Nodes**	6 in 5 countries
•	**Period of Activity**:	April 2002-April 2005
•	**Budget:**	1.256 M€
•	**Persons financed**:	8-10 researchers
•	**Human resources**:	371.5 Persons/months

EU countries

Non EU countries

EU COSIN participant

Non EU COSIN participant



•1A What is Complexity (for a physicist!)?

More is different !

P.W. Anderson, *Science* **177** 393-396 (1972)

•*quantitatively larger* systems are *qualitatively* different

Emergence of Complexity could be related to

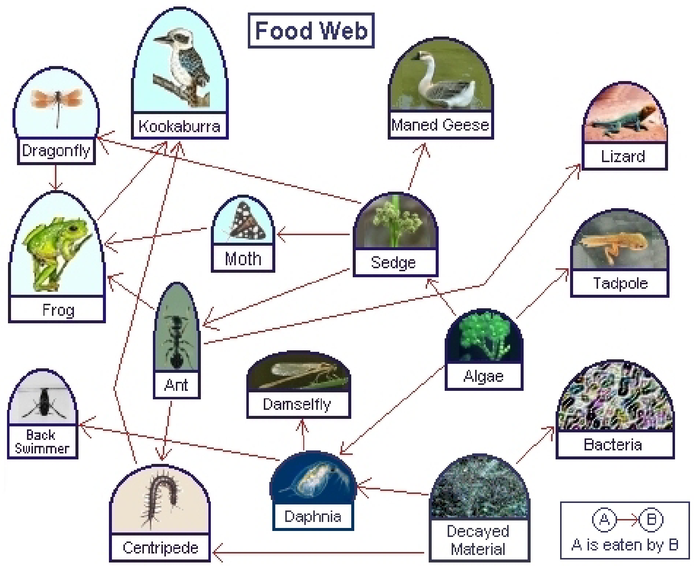
1) Microscopical interactions
2) Co-evolution
3) Self-Organisation

•1A (2) Complexity

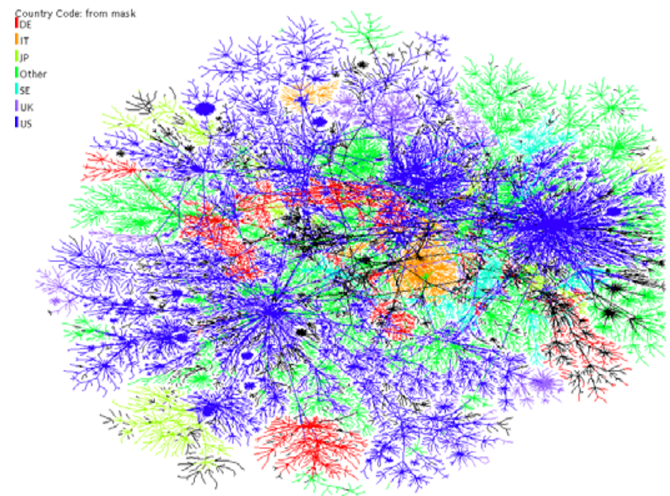
Networks represent an important example of **Complex Structures**
Through simple *microscopical interaction*
Complex Structures develop *long range correlations*.

Very different systems can be described through Graph Topology

River Networks

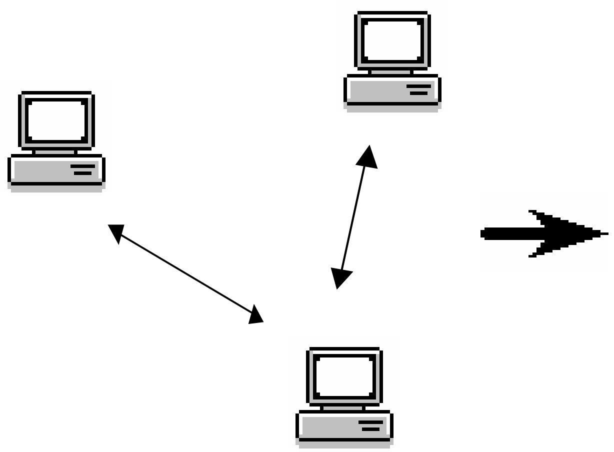


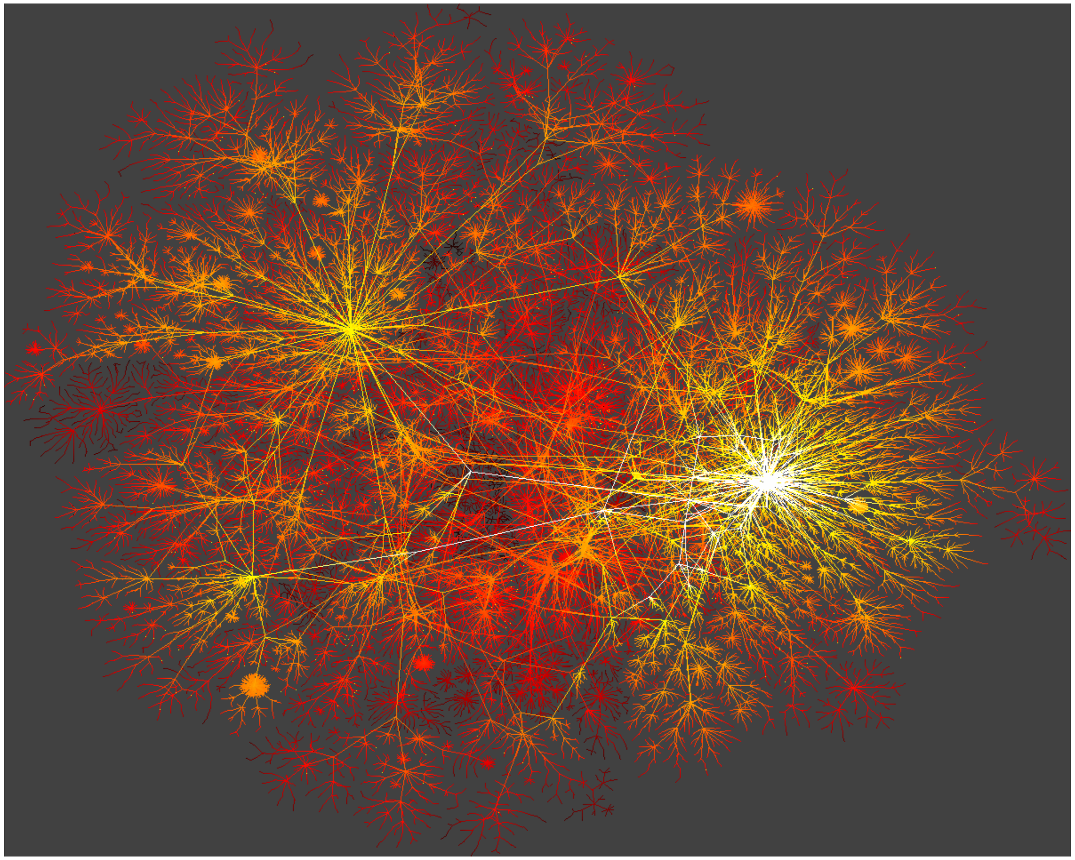
Food Webs



Internet

•1A (3) Complexity



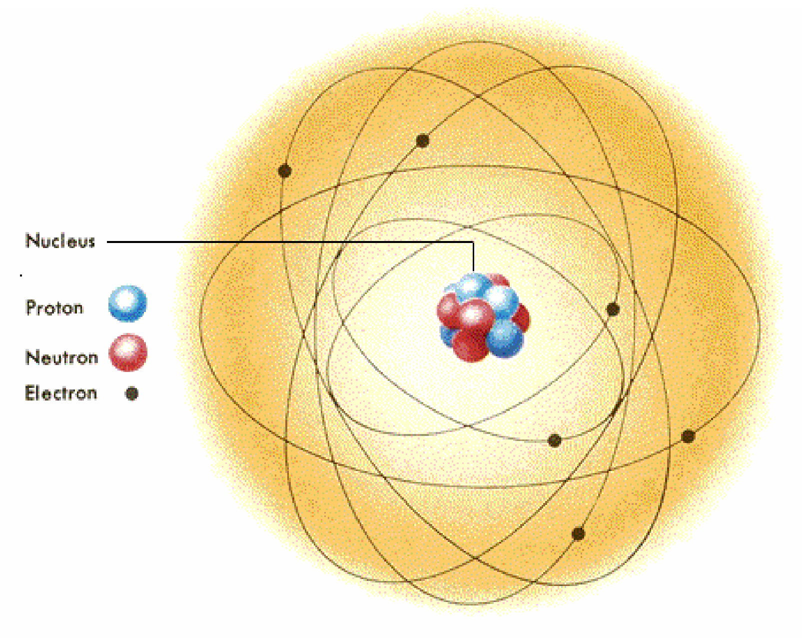


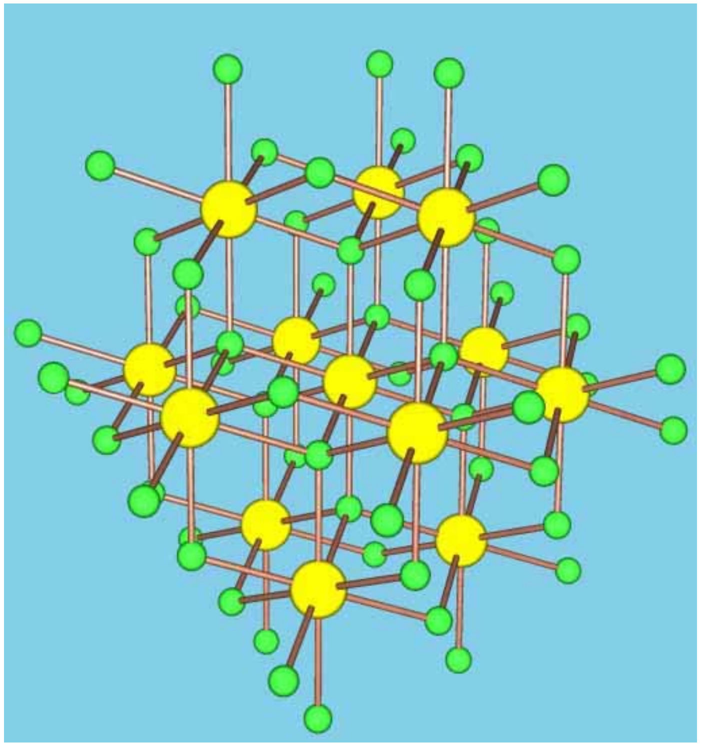
Router connections at small level produce a **complex** Internet structure.

•1A (4) Complexity

Atoms do not show the electrical features of macroscopic materials.

Complex rearrangement of electrons in cristals determine these new properties



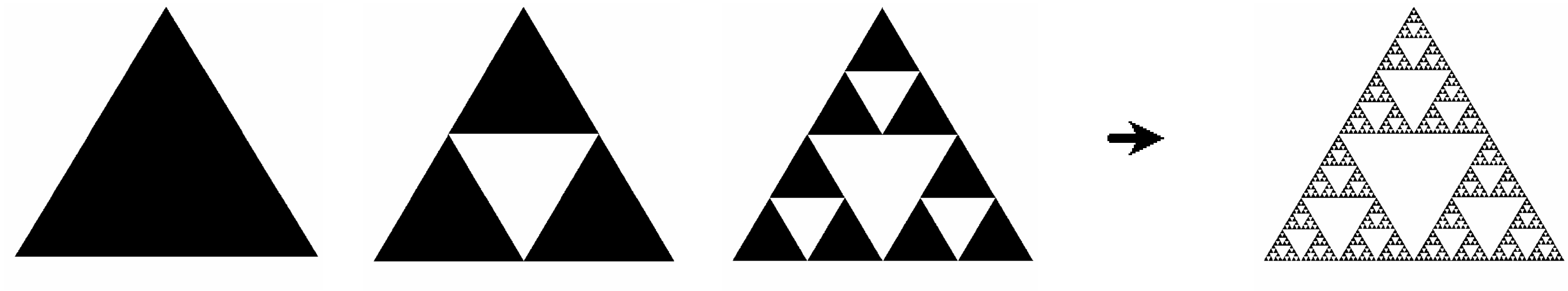


•1A (5) Complexity

Law of steepest descent produces the **complex** Network structure of rivers drainage basins.

•1B (1) The Fractal Geometry of Nature

Mathematicians provided the concept of ***Fractal Dimension***



The object obtained in the limit has "dimension" less than 2, in particular

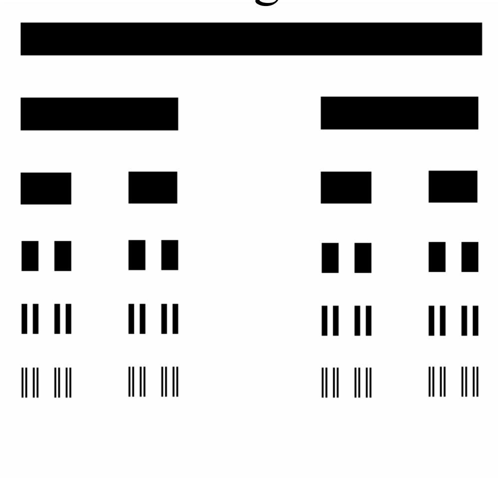
$$D = \lim_{\varepsilon \to 0} \frac{\ln N(\varepsilon)}{\ln(1/\varepsilon)} = \frac{\ln(3)}{\ln(2)} \approx 1.585$$

Where $N(\varepsilon)$ is the number of triangles of linear size ε needed to cover the structure

$$\boldsymbol{N(\varepsilon) = (1/\varepsilon)^D}$$

•1B (2) Deterministic Fractals

The Cantor Set is the dust of points obtained as the limit of this succession of segments

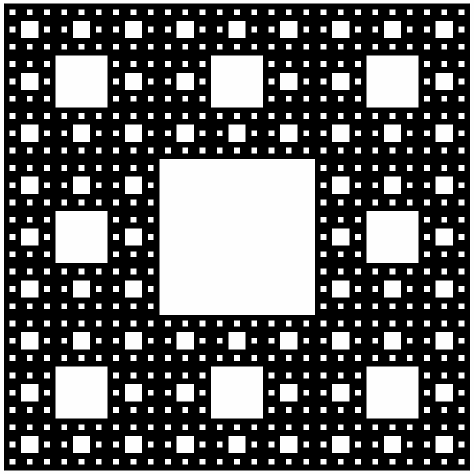


$N(\varepsilon) = 2^k$ where k is the iteration

And $\varepsilon = (1/3)^k$

D=ln(2)/ln(3) = 0.6309…

This is already the limit of succession of iterations



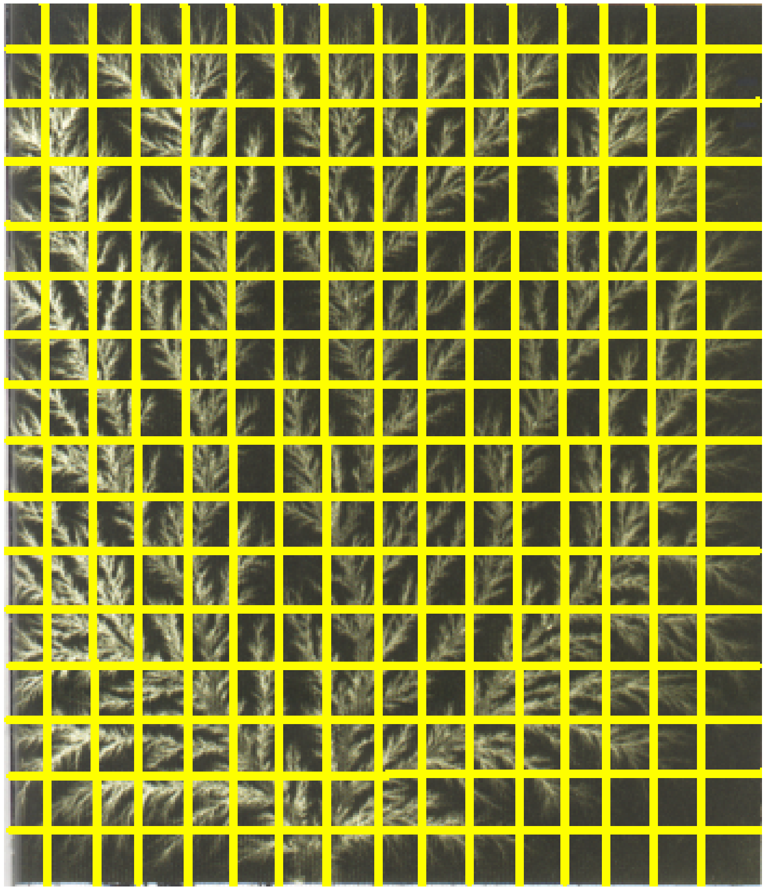
$N(\varepsilon) = 8^k$ where k is the iteration

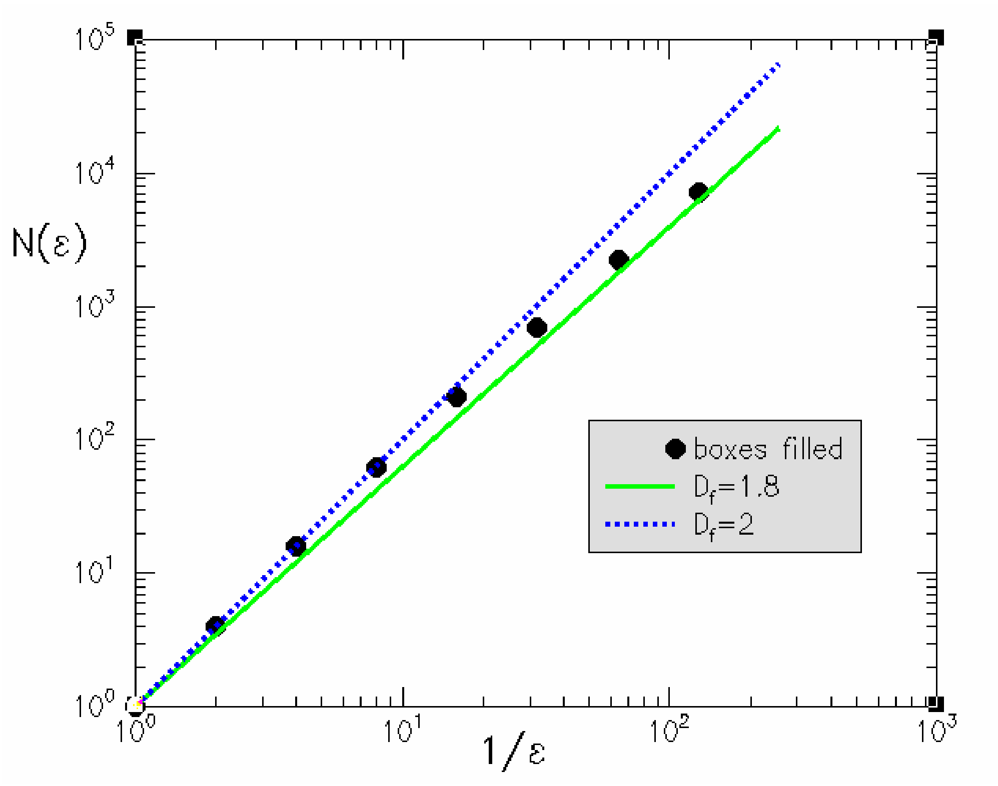
And $\varepsilon = (1/3)^k$

D=ln(8)/ln(3) = 1.8927…

•1B (3) Natural Fractals

More generally, Fractals are standard phenomena in Nature,
in this case their nature is intrinsically stochastic and not deterministic.

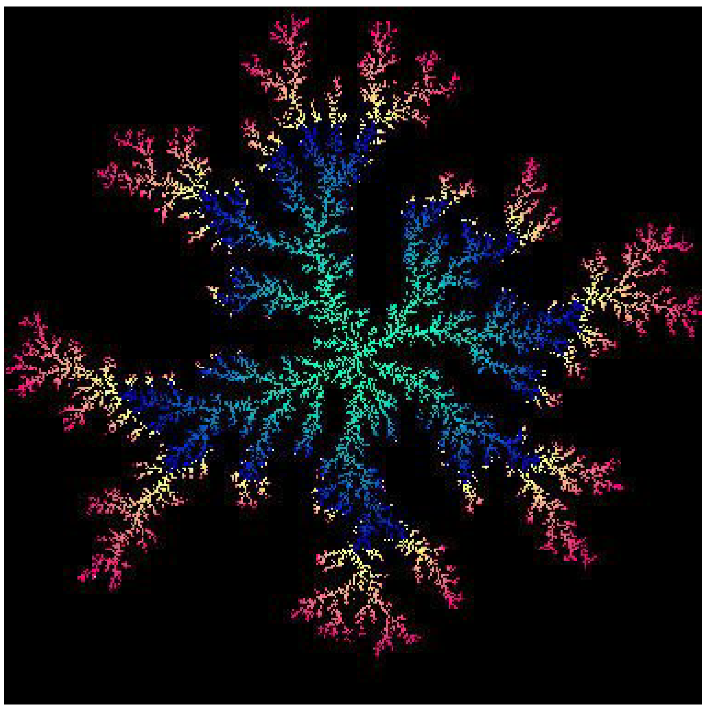


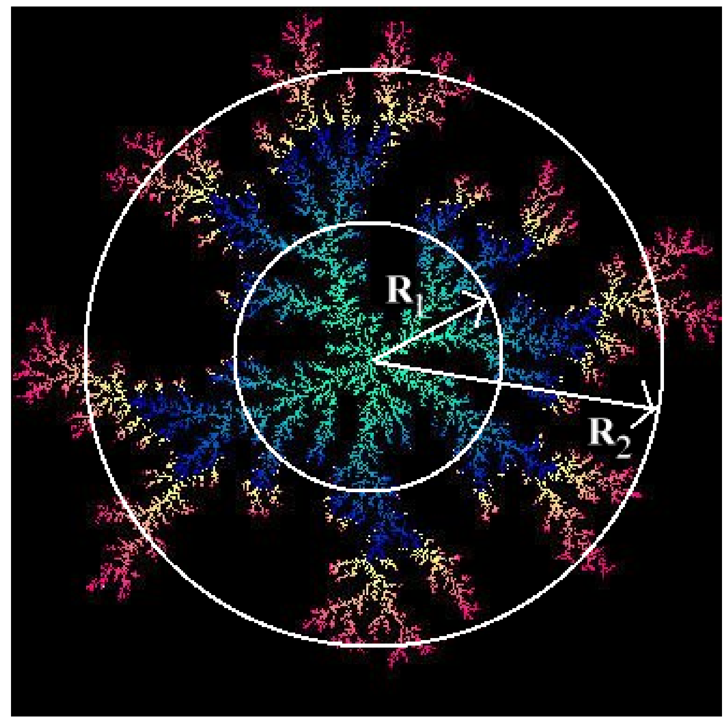


•1B (4) Natural Fractals

Another way to measure fractal dimension is through mass-length relation

$$M_1 \propto R_1^D \quad M_2 \propto R_2^D$$





•1B (5) Example

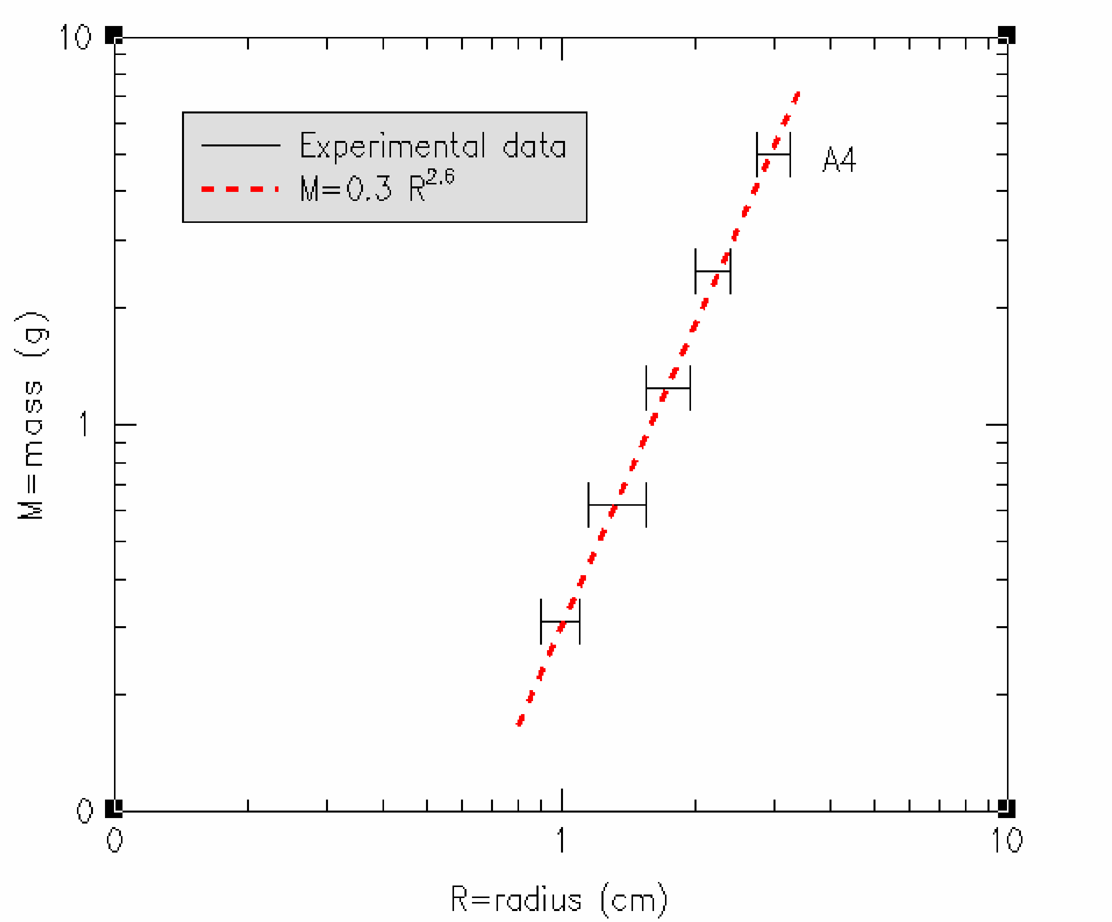
Let us consider an ordinary A4 sheet.
A4 format corresponds to 0.210 m X 0.297 m
Good quality printing paper weighs 80g/m^2
This means that one A4 weighs 0.297*0.21*80 g = 4.9896 g

Now fold an A4 sheet of 4.9896 g
Then fold

- one half of A4 (M=2.4948 g)
- one fourth of A4 (M=1.2474 g)
- …..

Measure the radius of the objects.

R (cm)	M (g)
3.0 ± 0.25	4.9896
2.2 ± 0.2	2.4948
1.75 ± 0.2	1.2474
1.35 ± 0.2	0.6237
1.0 ± 0.1	0.31185



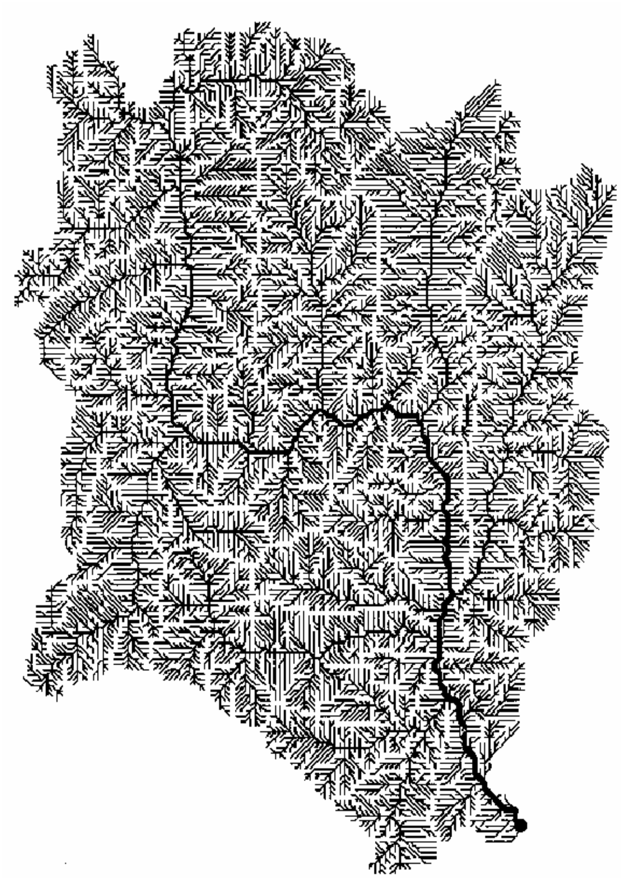
•1B (6) State of the art in Fractal Theory

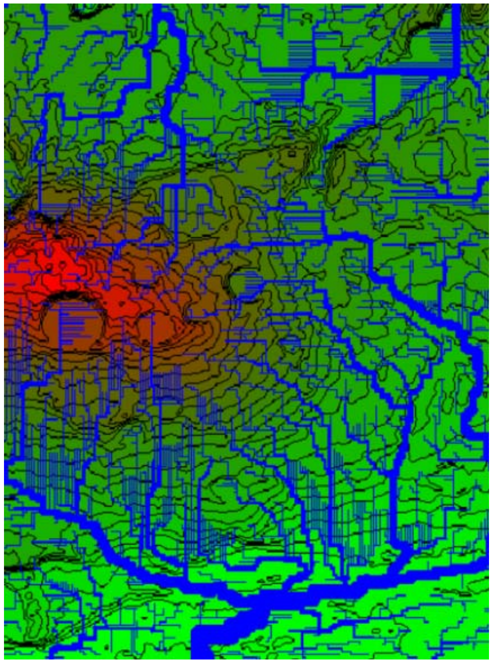
1) **Data collection (self-affinity in space and/or time)**
Coast lines, fractures, electrical breakdown, (invasion) percolation, price fluctuations, river networks, galaxy distribution, avalanches

2) **Modelling**
DLA (Diffusion Limited Aggregation)
DBM (Dieletric Breakdown Model)
IP (Invasion Percolation)
Sandpiles Models
BS (Bak and Sneppen model)

3) **General Theory**
Still lacking but good candidates:
Laplacian Fractals (Interplay between diffusion and disorder)
Self-Organised Criticality (Dynamics of the system keeps it in a self-similar state)

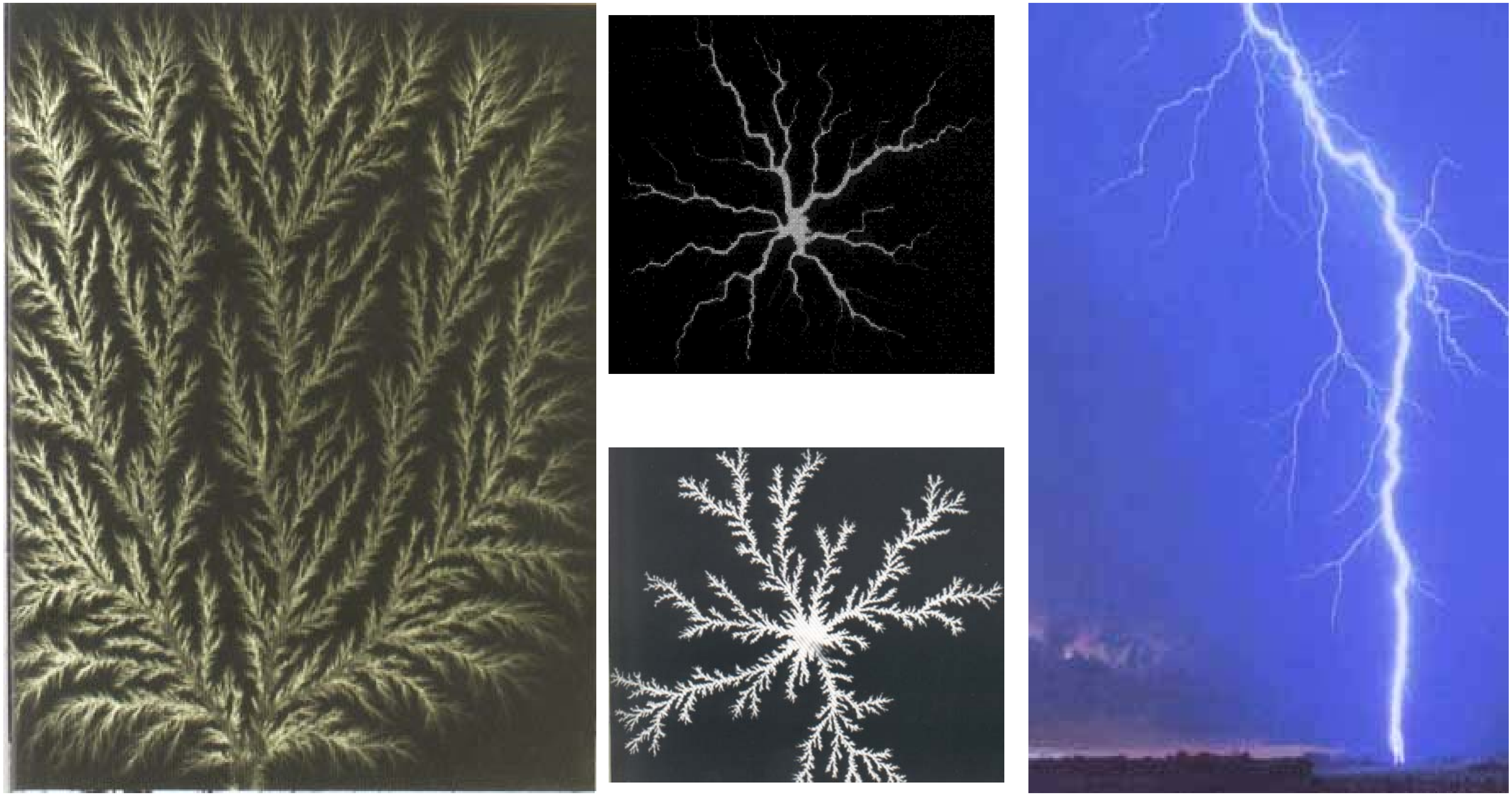
•1B (7) Fractal Structures in Nature





River Networks (Earth and Mars)

•1B (8) Fractal Structures in Nature



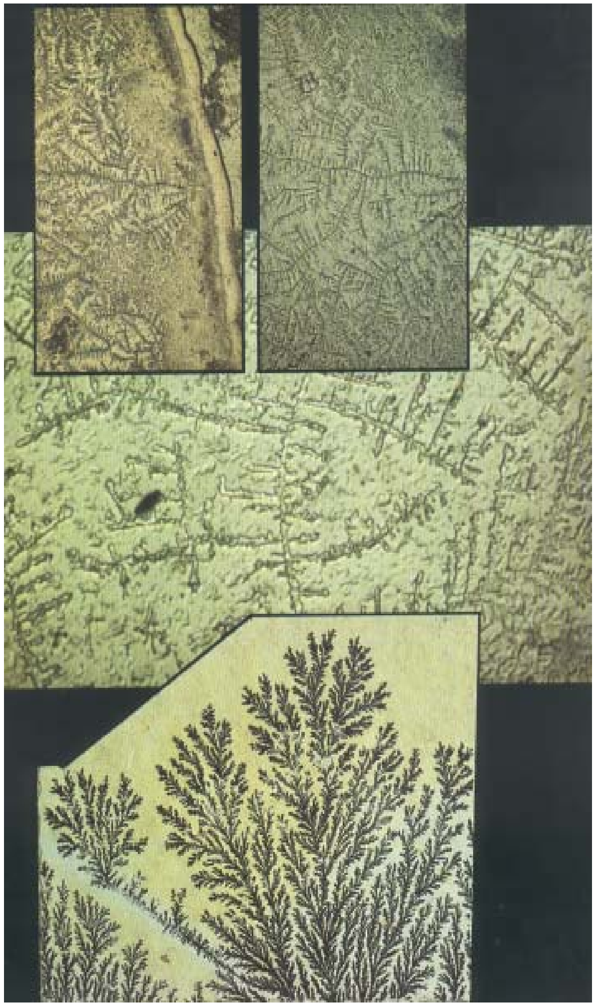
Electrical Discharge in dielectric, data and simulations

•1B (9) Fractal Structures in Nature

1104 Dome of Anagni (Italy)

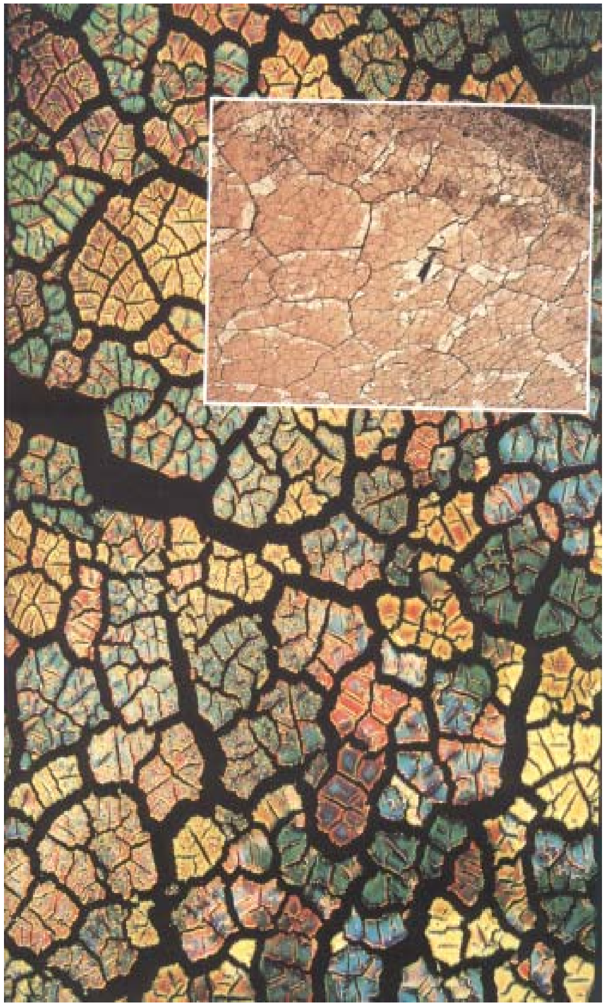
Viscous fingering (Lenormand)

•1B (10) Fractal Structures in Nature



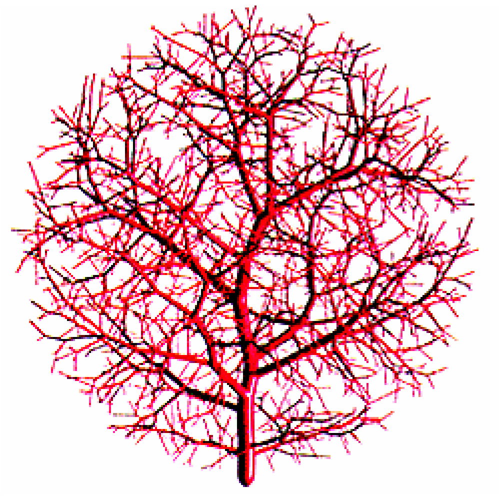
Olivine and MnO

Shock Waves (meteorite)



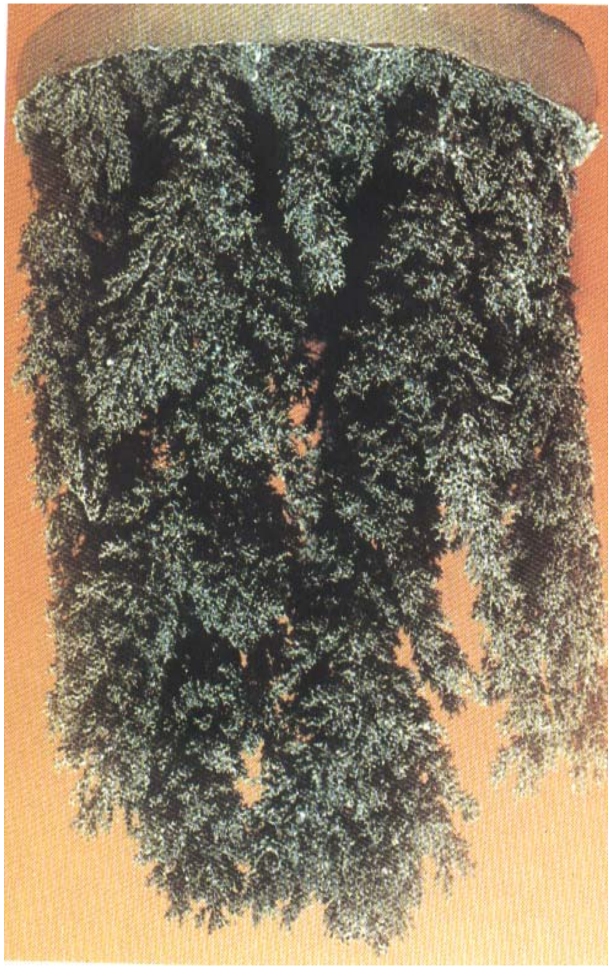
Cracks (speres and clay)

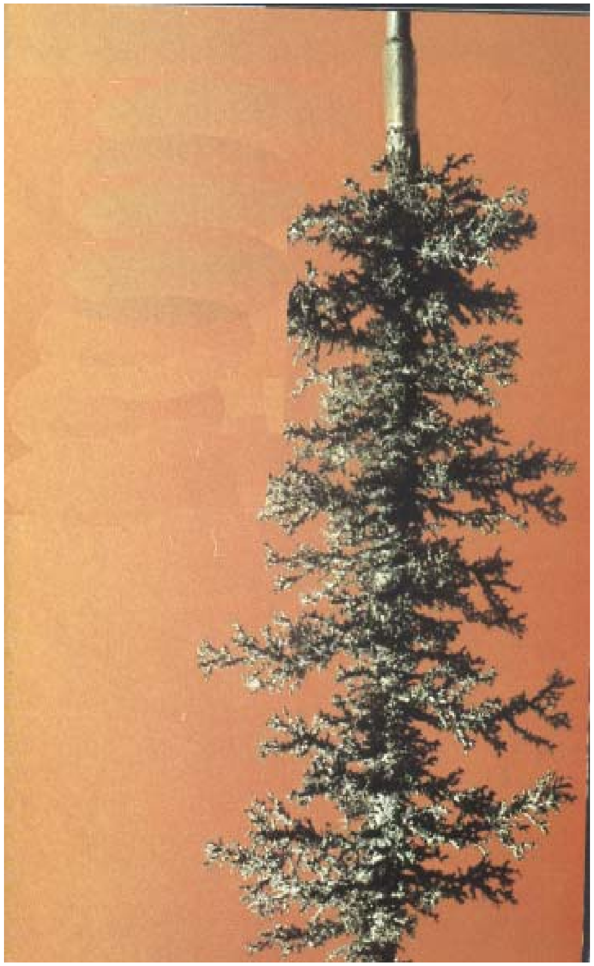
•1B (11) Fractal Structures in Nature

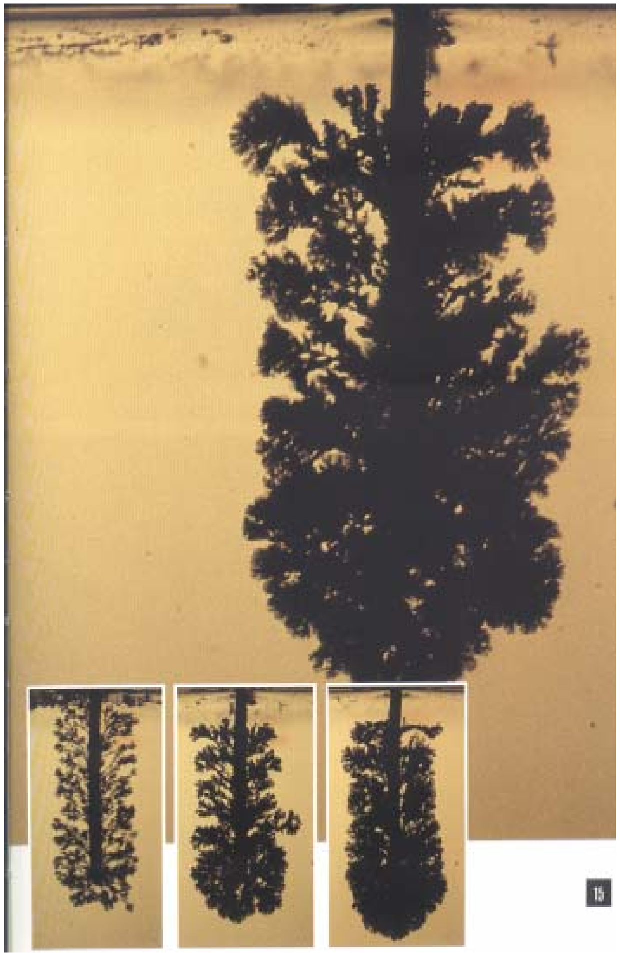


Cauliflowers, Blood vessels, Ferns

•1B (12) Fractal Structures in Nature

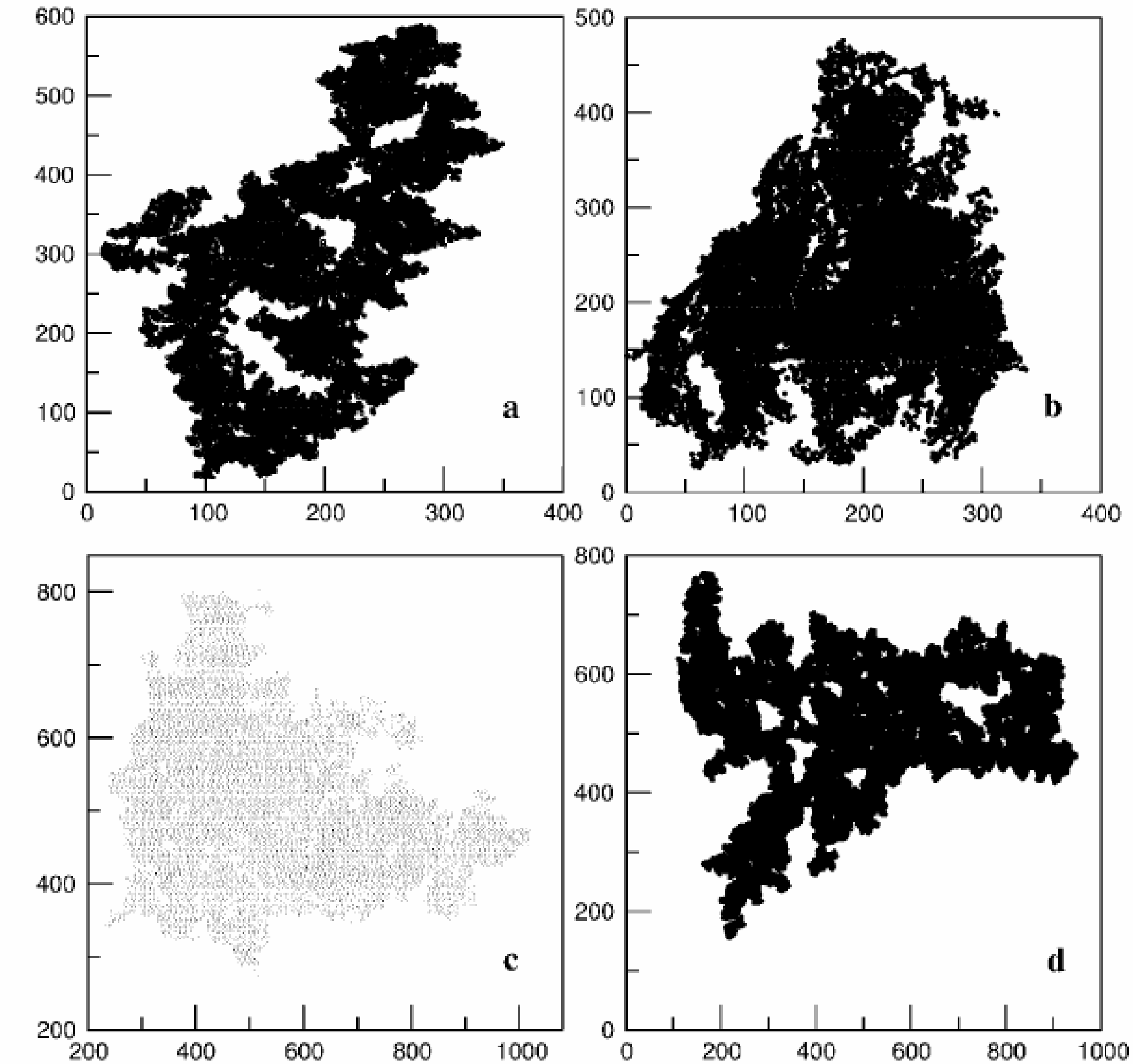






Electrochemical deposition

•1B (13) Fractal Structures in Nature

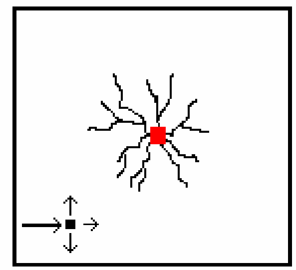


Different Wildfires

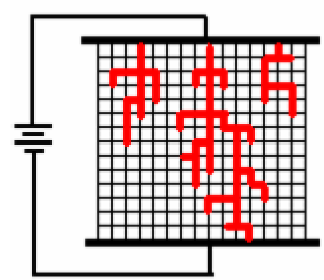
a) Valley of Biferno (I)
b) Penteli (Greece)
c) Cuenca (Spain)
d) A computer model (percolation)

•1B (14) Models of Fractal Growth

DLA Diffusion Limited Aggregation:
a random walker travels on a limited portion of the Euclidean space. In this region a "seed" is present, when the walker "touches" the seed the walker stops, sticks on the seed and a new walker starts on the region boundaries.
T.A. Witten, L. M. Sander PRL **47**, 1400 (1981).



DBM Dielectric Breakdown Model:
A dielectic material modelled by a regular lattice is kept under constant Electric Field. Step by step sites in the dielectric are removed with a probability proportional to the electrostatic difference of potential they see.
L. Niemeyer, L. Pietronero, H.J. Wiesman PRL **52**, 1033 (1984).



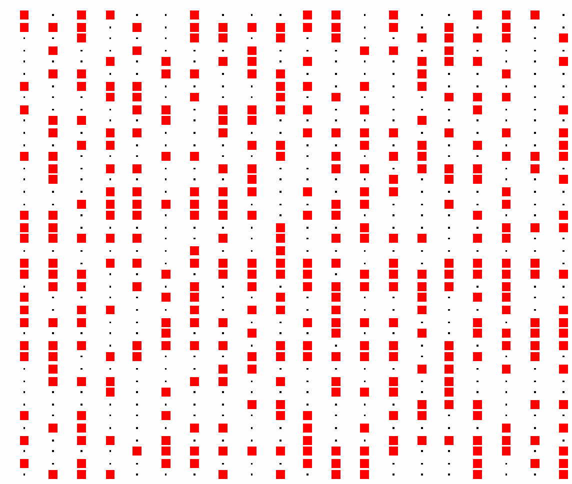
In both cases the physics of the problem is given by **Laplacian operator**.

DBM $\nabla^2 \phi = 0 \rightarrow \phi_i = \frac{1}{4} \Sigma_{<ji>} \phi_j$

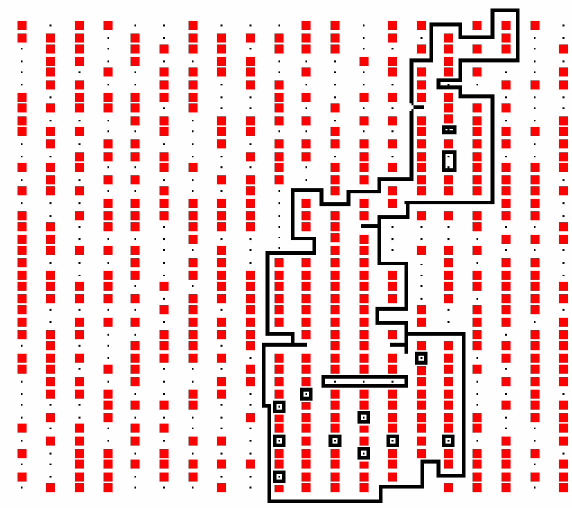
DLA $P_i = \frac{1}{4} \Sigma_{<ji>} P_j$

•1B (15) Models of Fractal Growth

Percolation:
In a lattice we switch on with probability p the various sites.
If $p=0$ no site is on.
If $p=1$ all the sites are on
Interesting situation for $0<p<1$

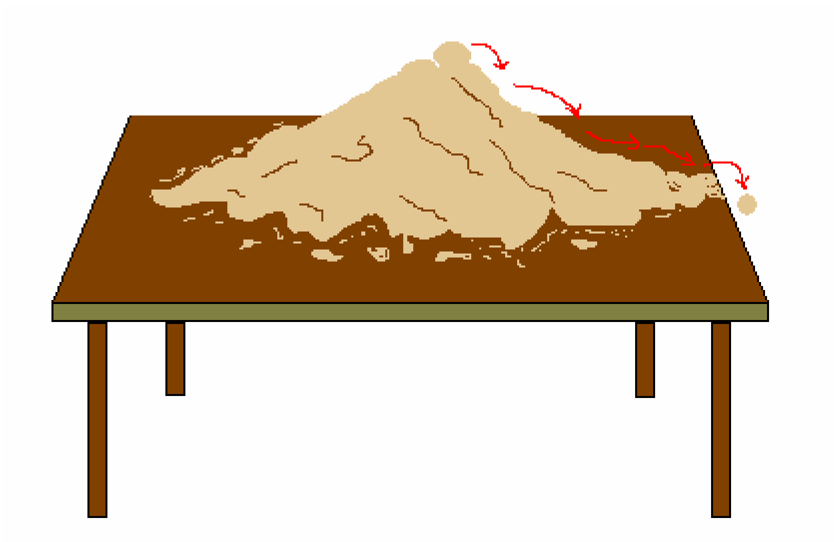


By tuning the value of p we pass from little isolated cluster to a large one that spans all the system.
It can be shown that when $p=p_c$ the cluster is fractal



D. Stauffer *Introduction to Percolation Theory* Springer (Berlin).

•1C (1) Self-Organization



Consider a pile of sand on a table

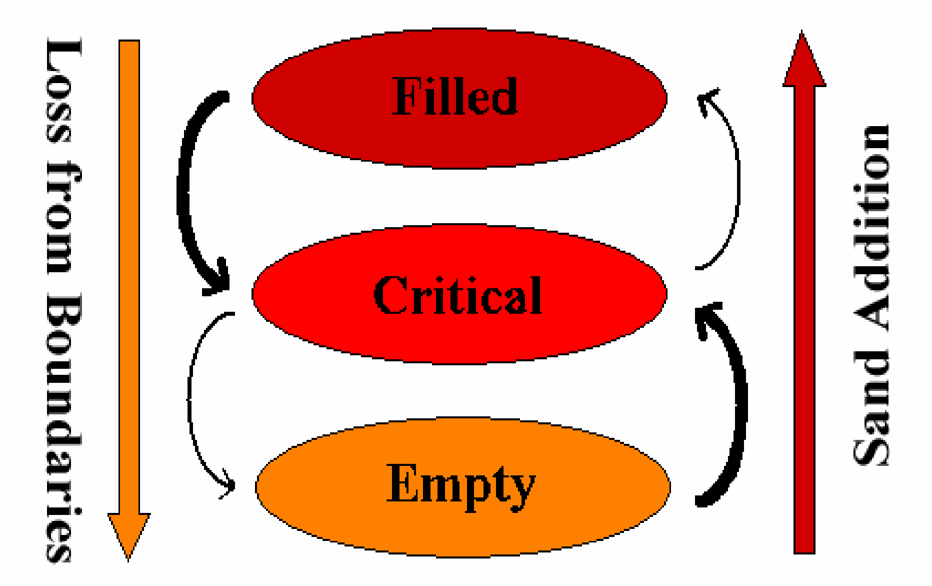
If the slope is very large a little perturbation matters

An avalanche starts towards the edges of the system

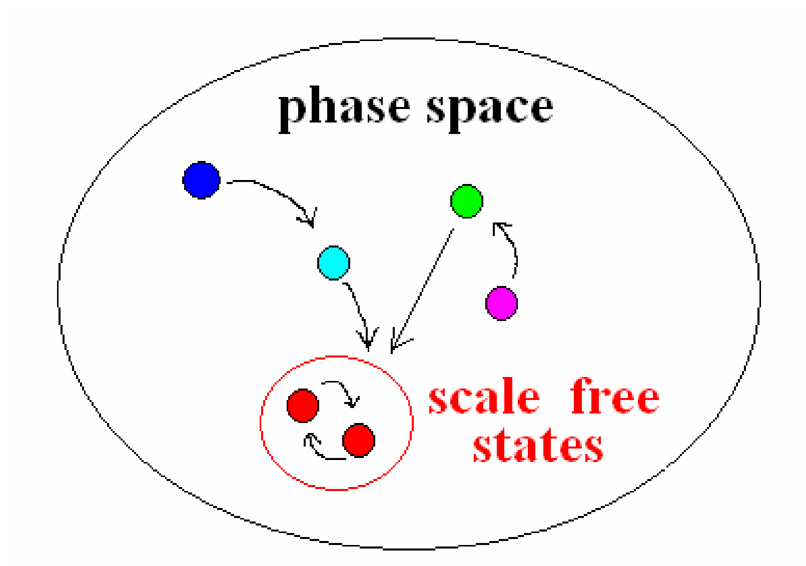
In most case there is no characteristic size for the avalanches

$$\mathbf{P(S) \propto S^{-\tau}}$$

•1C (2) Self-Organization



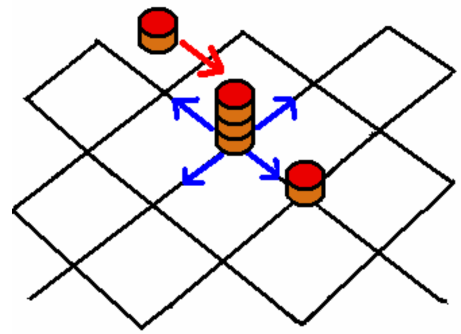
This process of self-organization
is based on this feedback.
The critical state attracts the dynamics



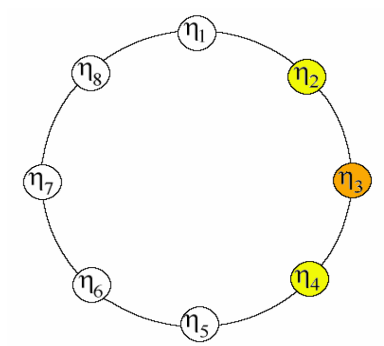
More generally one can justify the **ubiquity of fractals** whenever **the dynamics evolves in a scale-free state**

•1C (3) Models of Self-Organized Criticality

BTW Sandpile Model:
On a regular lattice, grains of sands are added on the sites of the lattice. When a critical threshold of grains is reached, the site topples on the neighbours triggering other topplings to form a macroscopic avalanche.
P. Bak, C. Tang, K. Weisenfeld PRL **52**, 1033 (1984).

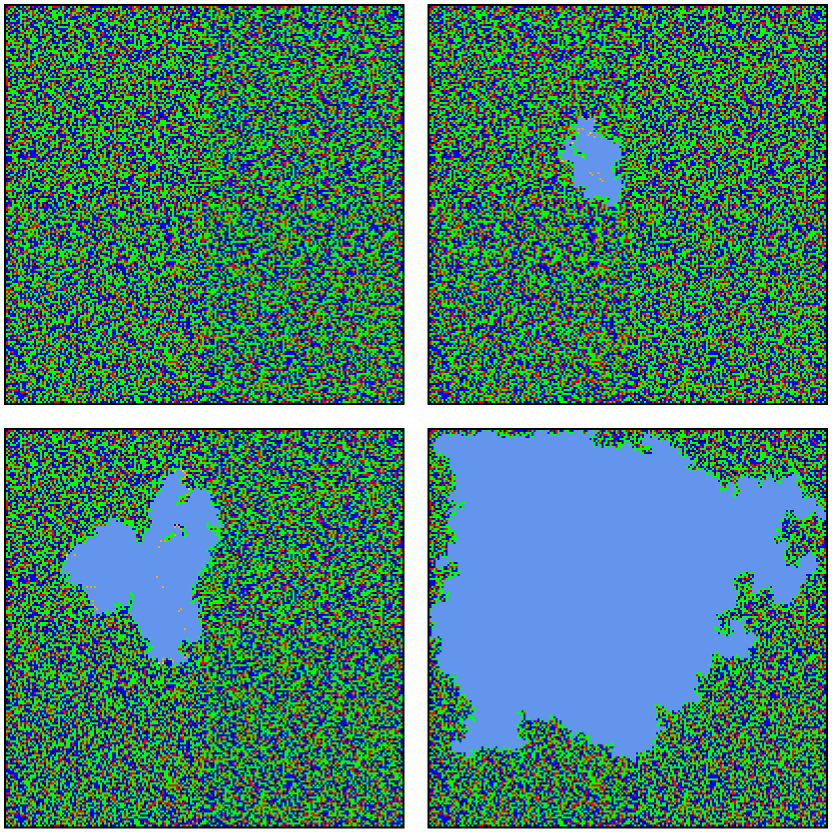


BS Bak and Sneppen Model:
An ecosystem of species is model through a system of species i characterized by a fitness η_i. Recursively the species with the minimum fitness and its neighbours are removed and changed with three new ones with random η_i
P. Bak, K. Sneppen PRL **71**, 4083 (1993).



IP Invasion Percolation:
A fluid (water) is injected in a porous medium to extract oil. Amongst the different channels on the boundaries the one with the minimum diameter is selected to be invaded.
D. Wilkinson and J. F. Willemsen, J. Phys. A London **16**, 3365 (1983).

•1C (4) Self-Organized Criticality



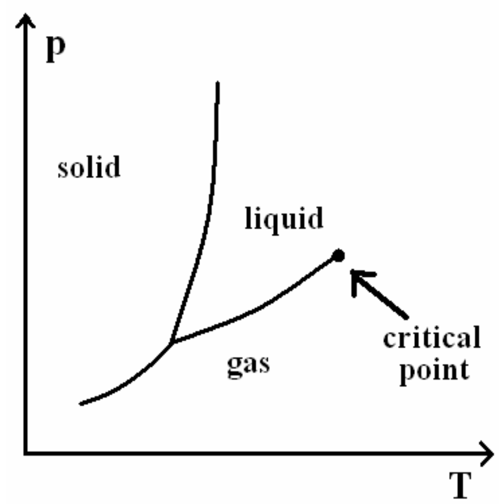
BTW and BS are critical in the sense of scale-free dynamics.
The spatial cluster are compact
But the size of the avalanches (measure of the burs of activity) are power-law

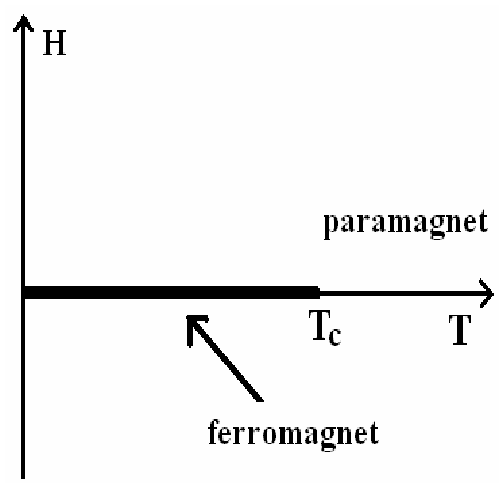
•1C (5) Self-similarity, Critical Phenomena, Renormalization Group

The presence of scale invariance was already noticed in the framework of **critical phenomena**

In certain conditions physical systems can abruptly change their macroscopic behaviour when Temperature or Pressure are smoothly varied

1. **Discontinuous transitions**
   When away from the critical point there is one phase whose properties are continuously connnected to one of the phase in the transition, generally the correlation length remains finite (Melting of 3d solid, condensation of gas in liquid)
2. **Continuous transitions**
   The two phases must become identical and generally the correlation lengths diverges (Curie Temperature in a Ferromagnet, liquid-gas critical point)





•1C (6) Self-similarity, Critical Phenomena, Renormalization Group

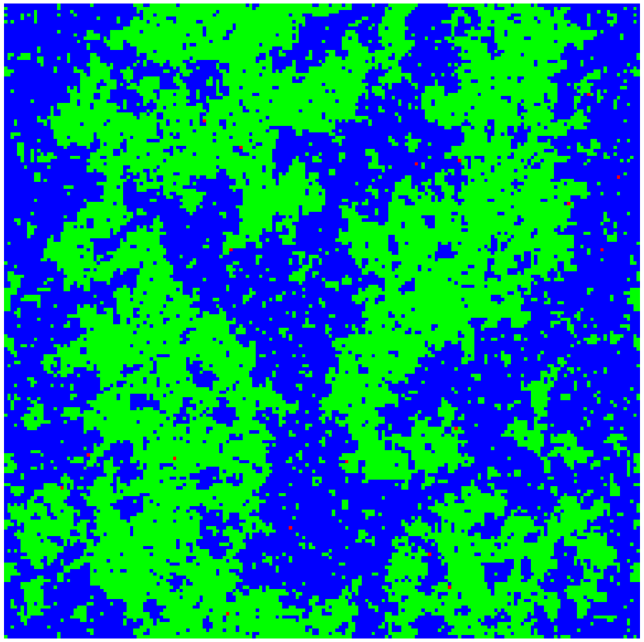
Typically close to a critical point, for a
continuous transition
most of the interesting quantities as specific heat, correlation length exhibit power-law scaling with respect to the distance from critical point ($T-T_c$)

- By defining $\mathbf{t \equiv (T-T_c)/T_c}$ and $\mathbf{h \equiv H/k_b T_c}$

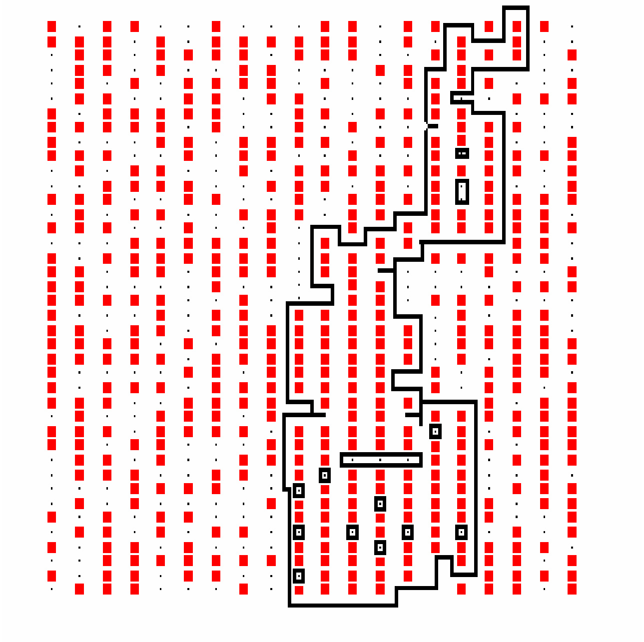
α The **specific heat** in zero field	$\mathbf{C \propto A\|t\|^{-\alpha}}$
β The **spontaneous magnetization**	$\mathbf{\lim_{H\to 0} M \propto -t^{\beta}}$
γ The zero field **susceptibility**	$\chi \equiv (\partial M/\partial H)\|_{H=0} \propto \mathbf{-t^{-\gamma}}$
δ At T=Tc the **magnetisation** versus h	$\mathbf{M \propto h^{1/\delta}}$
η The **Correlation Function G(r)**	$\mathbf{G(r) \propto 1/r^{d-2+\eta}}$
ν The **correlation length** ξ	$\xi \propto \mathbf{\|t\|^{-\nu}}$
z The typical **relaxation time** τ	$\tau \propto \xi^{\mathbf{z}}$

In a liquid $H=(p-p_c)$, $M=(\rho-\rho_c)$. The role χ is played by the *isothermal compressibility* κ_τ

•1C (7) Self-similarity, Critical Phenomena, Renormalization Group



Size of Domains in Ising model at critical state



Domains at percolation thresholds

•1C (8) Self-similarity, Critical Phenomena, Renormalization Group

The concept of scale invariance is used in the idea of the Renormalization Group

ISING MODEL $\mathcal{H}(\{s_i\}) = -J \sum_{<i,j>} s_i s_j$

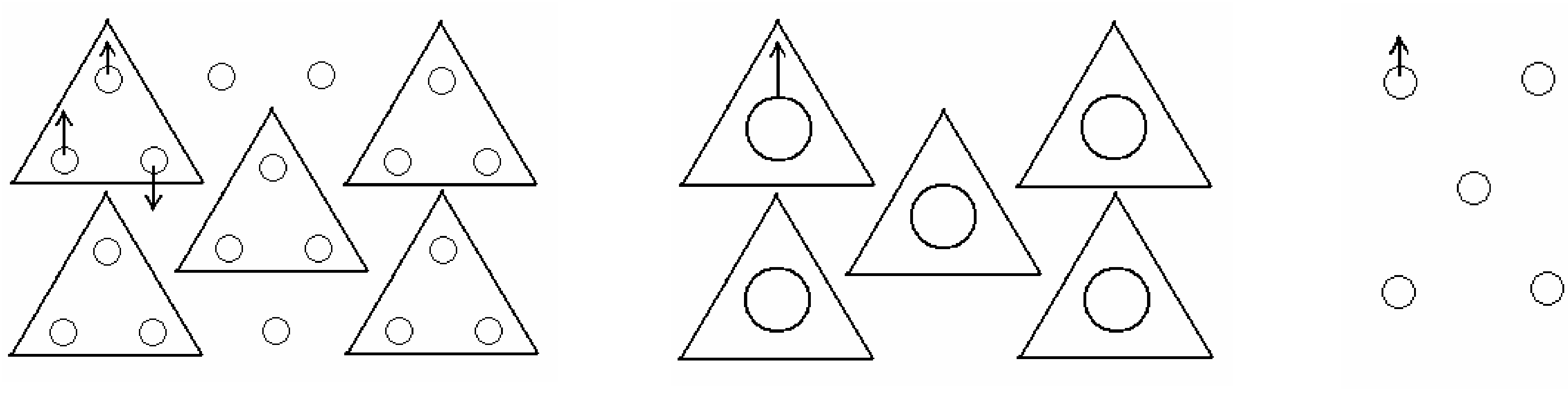
As in the example of the Ising model, we can start from a lattice of spins.

We replace (**decimation**) every three spin with a block spin

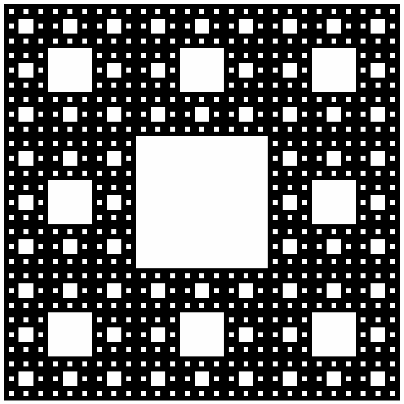
We then change (**rescaling**) the lattice constant.

The new system is now indistinguishable from the old.

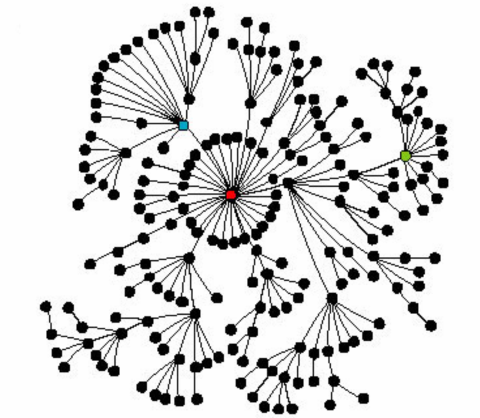
In order to have the same Hamiltonian we need to find the J* invariant under this transformation. From the condition on this parameter we can find an analytical way to compute the critical exponents.



•1D (1) Fractals and networks



These Complex structures extend the concepts of Self-similarity from *Metric Objects* (Fractals) to *Shape* (Networks).

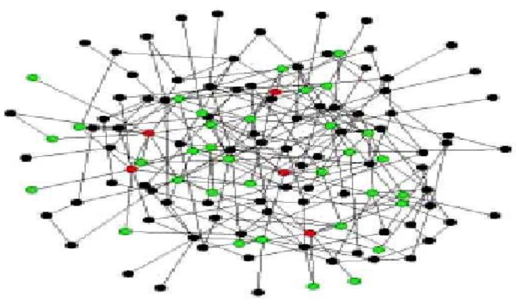


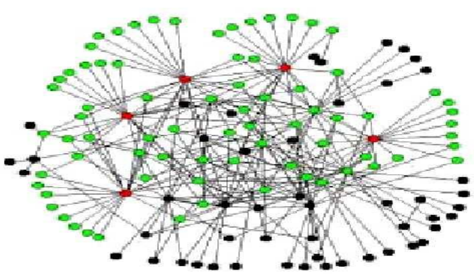
Fractal Dimension is defined

Fractal Dimension is **NOT** defined

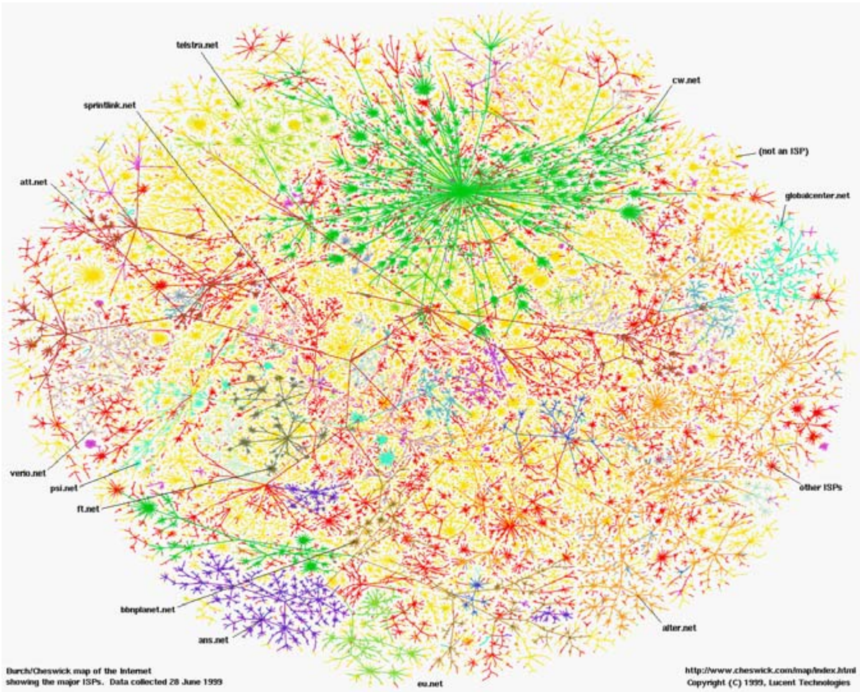
Similarly to Fractals, one finds self-similar properties in “some” distributions.

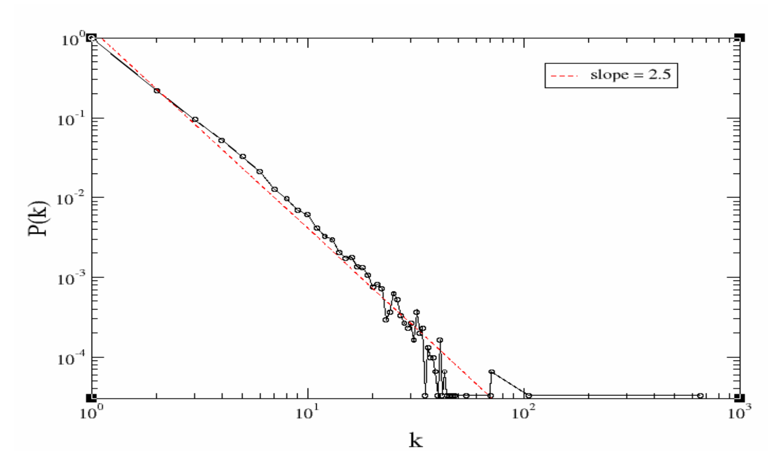
THE MOST NOTICEABLE IS THE NUMBER OF LINK PER SITE

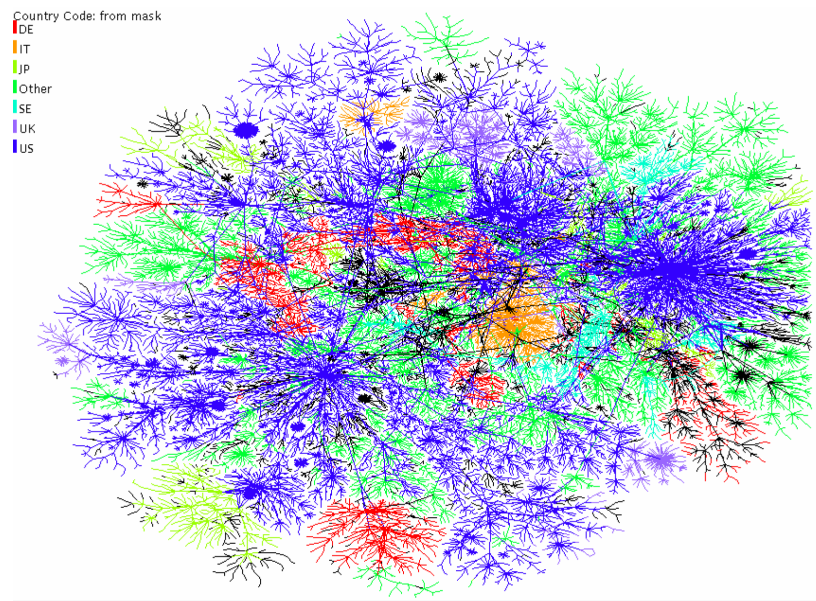




•1D (2) Evidence of scale freedom in networks

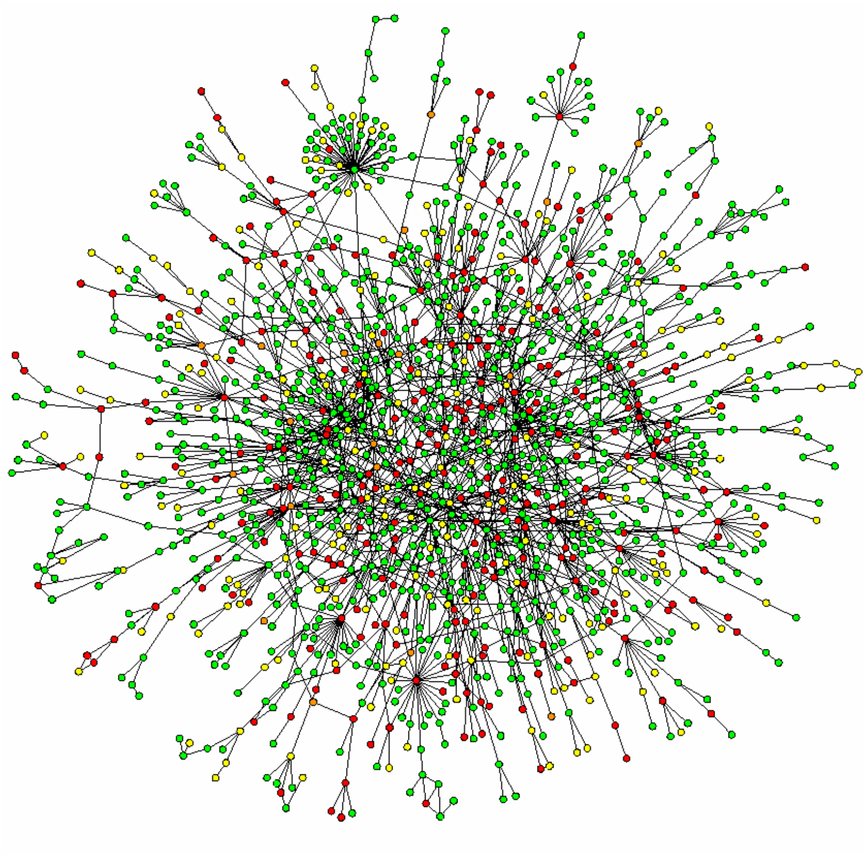




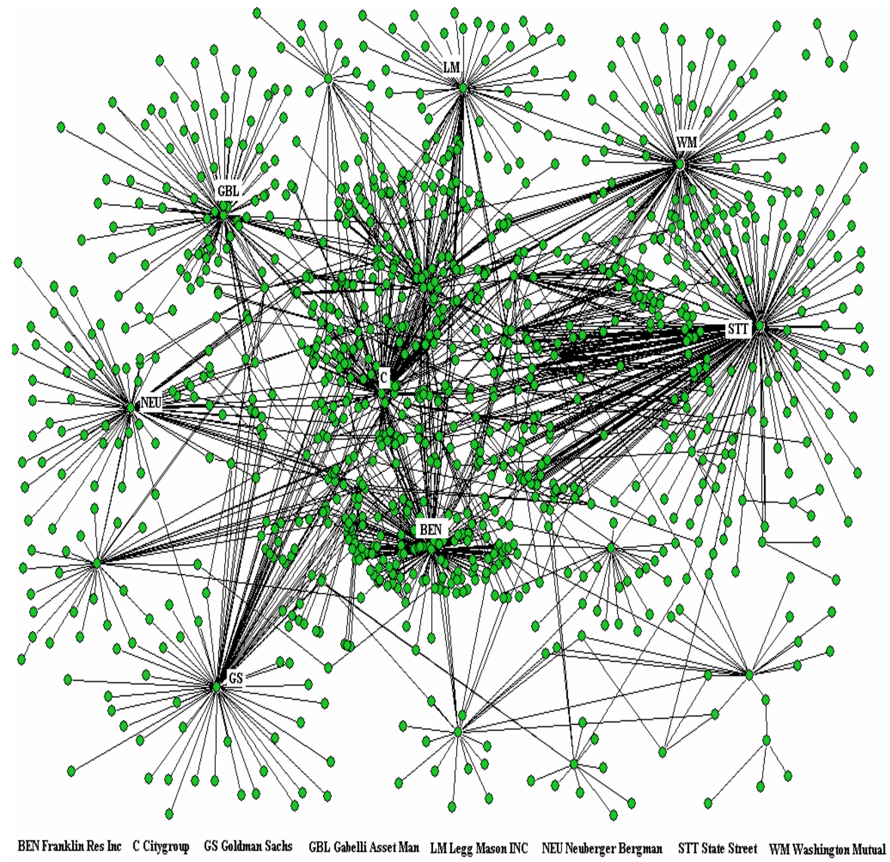


Different Traceroute maps
Of Internet

•1D (3) Evidence of scale freedom in networks

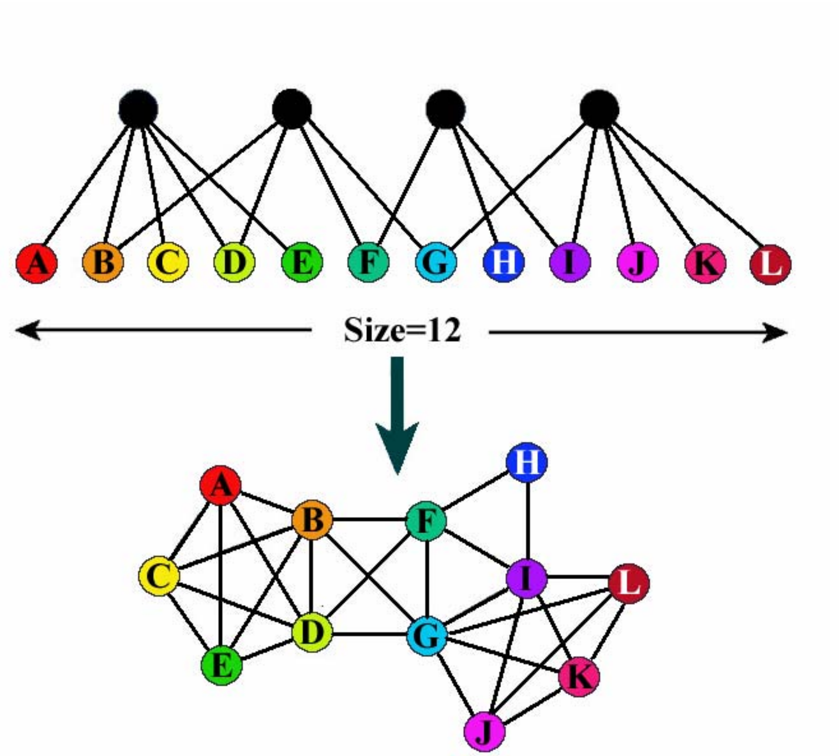


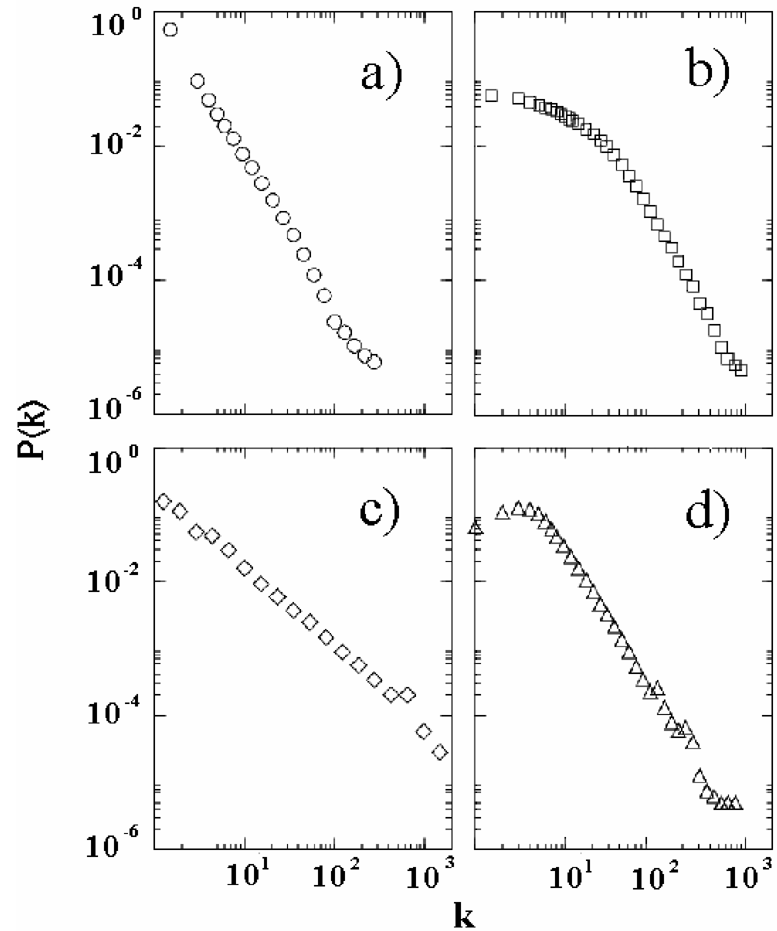
Protein Interaction Map
Saccaromyces Cerevisiae



Property network
In NYSE

•1D (4) Evidence of scale freedom in networks



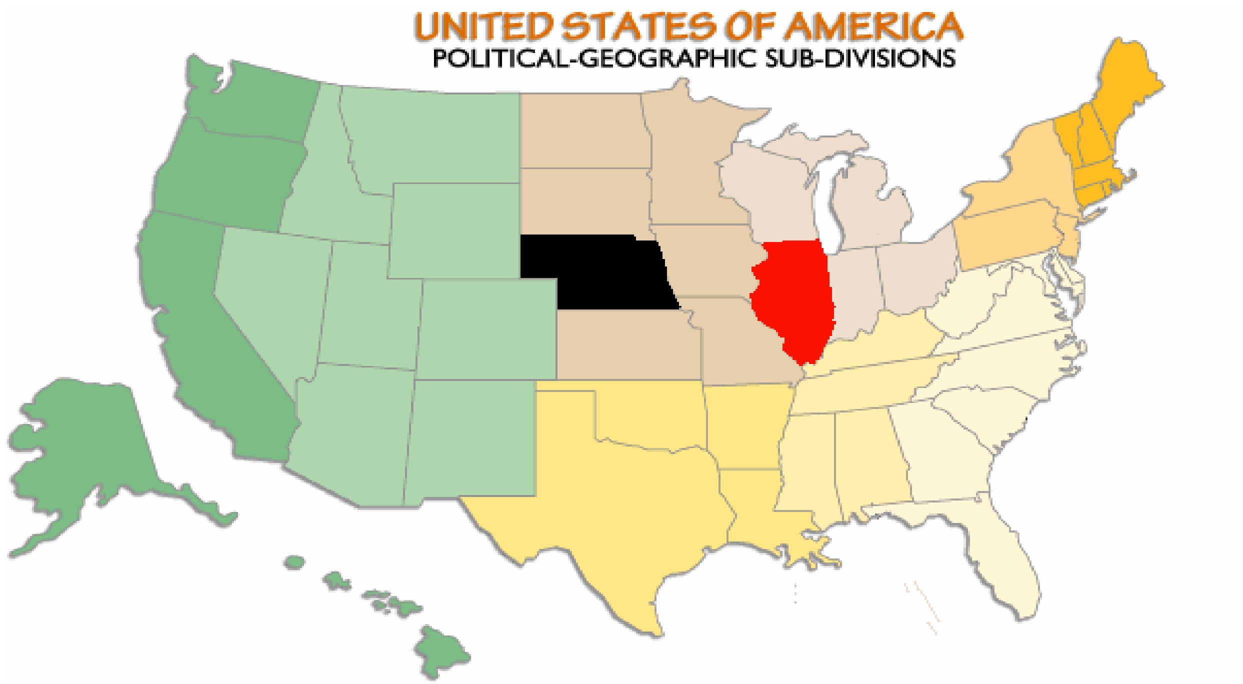


a) WWW b) Actors

c) Physicists d) Neuroscientists

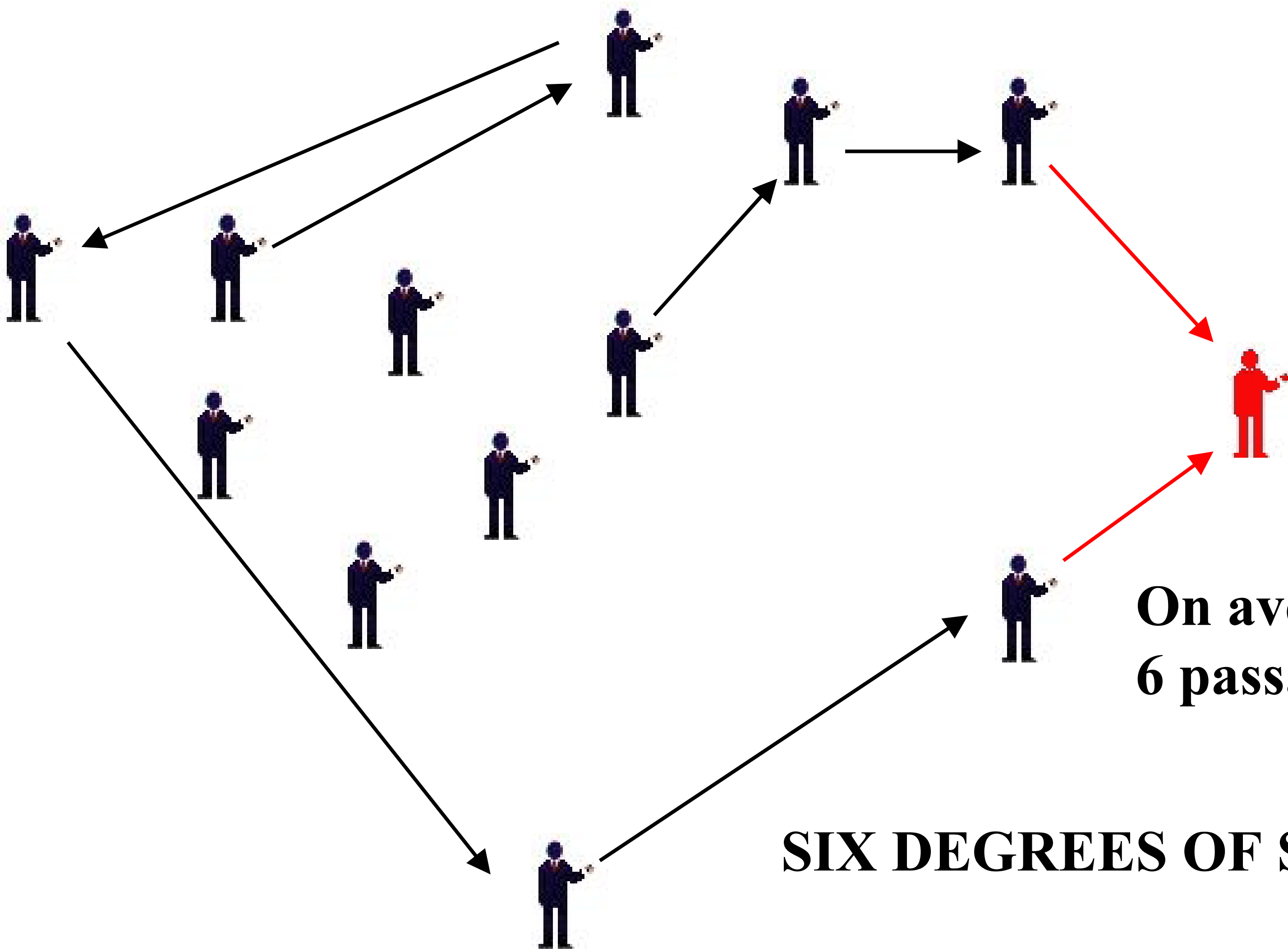
•1D (5) The Milgram Experiment (1967)

There is also another non trivial property owned by most of the networks: They are much more "connected" than expected. Let's see why:



Is it possible to deliver a message to a Stock dealer in Chicago starting from unrelated people in Nebraska?

•1D (6) The Small World Effect

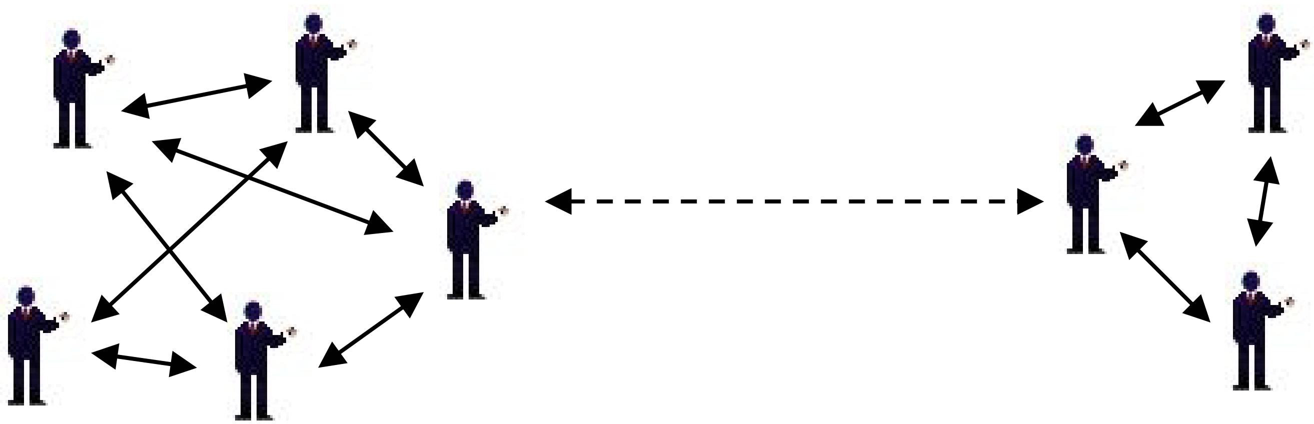


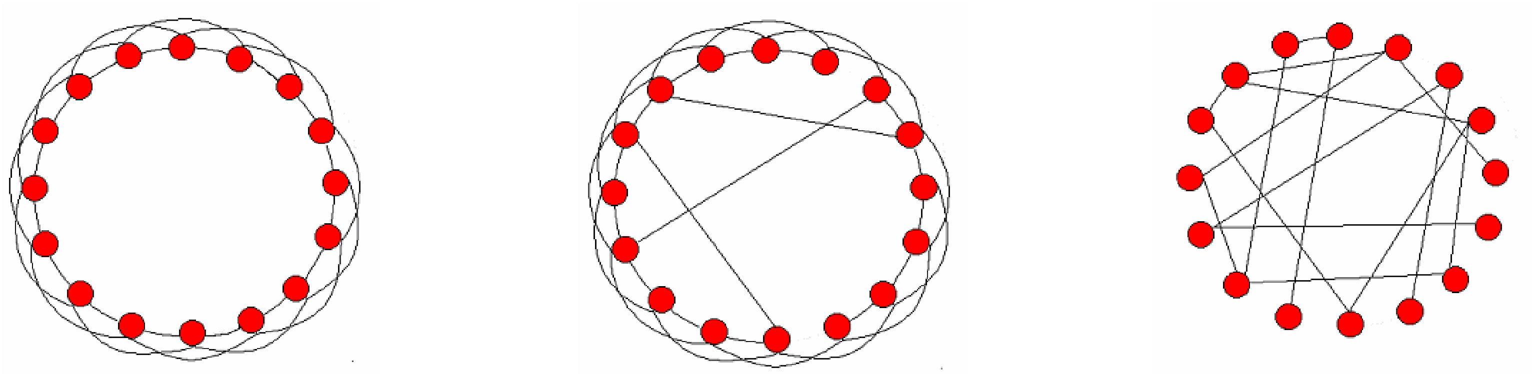
On average less than 6 passages!!

SIX DEGREES OF SEPARATION

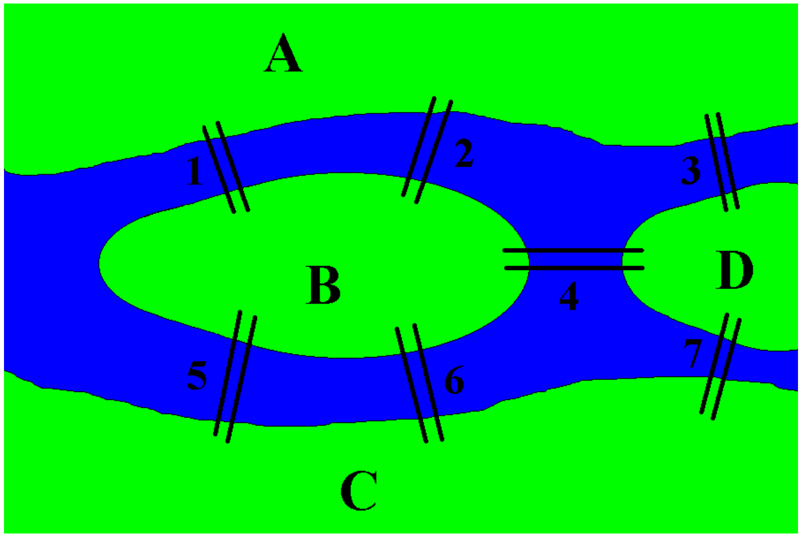
•1D (7) The structure of social networks

According to Mark Granovetter "weak links" work as shortcuts



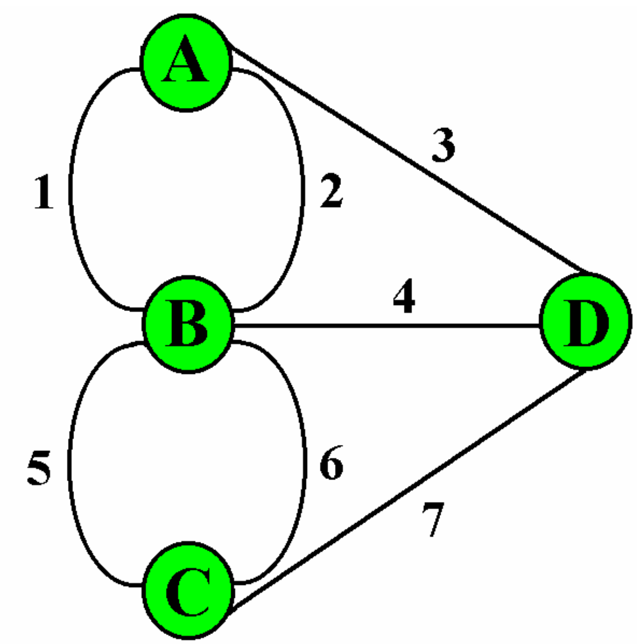


•1E (1) Basic Graph Theory



Is it possible to travel from one part of the city of Königsberg to any other PASSING ALL THE BRIDGES ON THE PREGEL ONLY ONCE ?

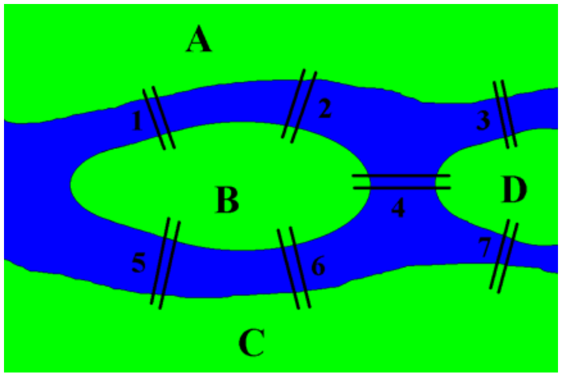
NO!

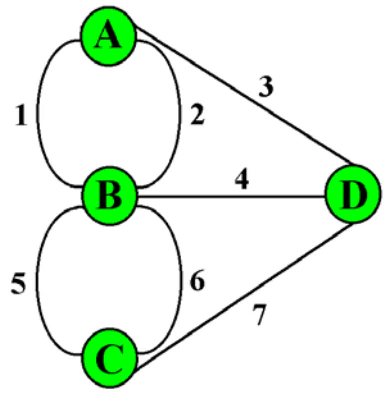


Euler (1736) pointed out that to be a "passage" point a vertex must have an even number of links. Only starting and ending points can have an odd number of links.

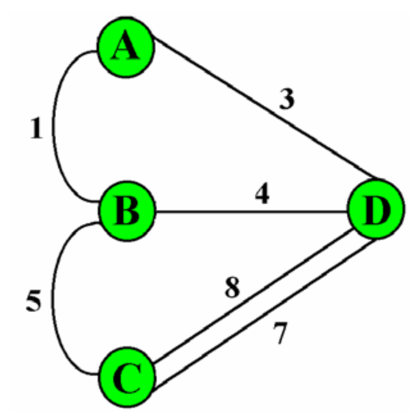
THIS IS NOT THE CASE FOR KÖNIGSBERG

•1E (2) Is the problem time dependent?



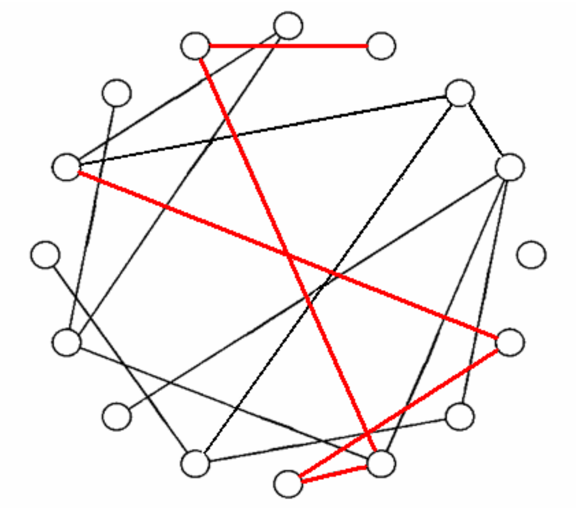


1736
All vertices have odd degree! → No way



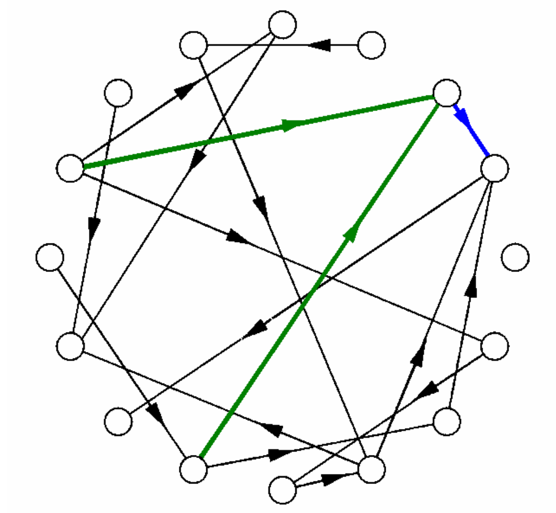
2003
Only B and C have odd degree! → we can do it!

•1E (3) Graph Definitions

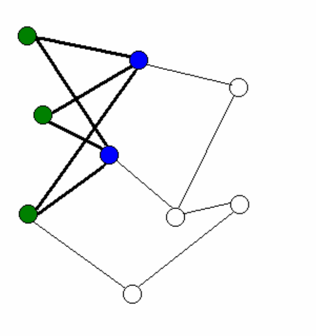


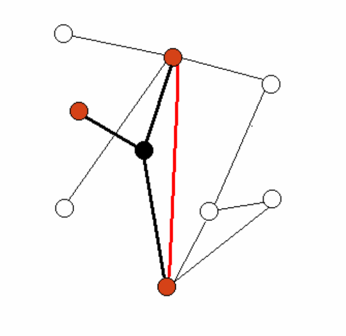
A **Graph *G(v,e)*** is an object composed by ***v*** *vertices* and ***e*** *edges*

Edges can be oriented →



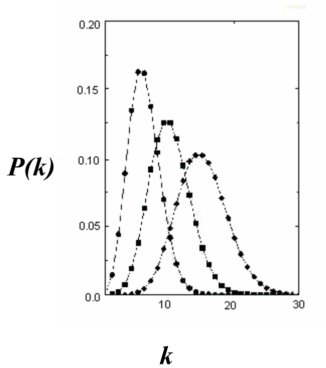
- Degree k (In-degree k_{in} and out-degree k_{out}) = number of edges (oriented) per vertex
- Distance d = number of edges amongst two vertices (in the connected region !)
- Diameter D = Maximum of the distances (in the connected region !)
- Clustering = cliques distribution, or clustering coefficient





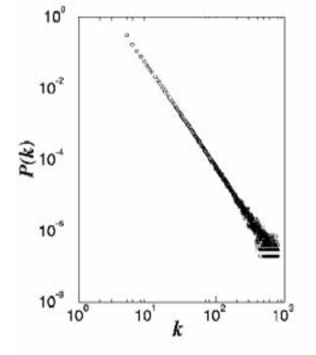
•1E (4) Statistical measures

·1 **Degree frequency density P(k)** = **how many times you find a vertex whose degree is k**

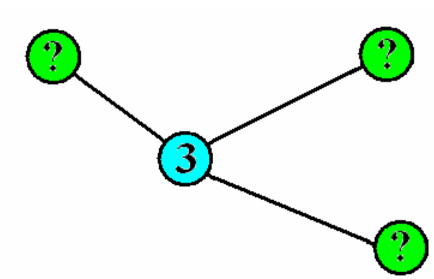


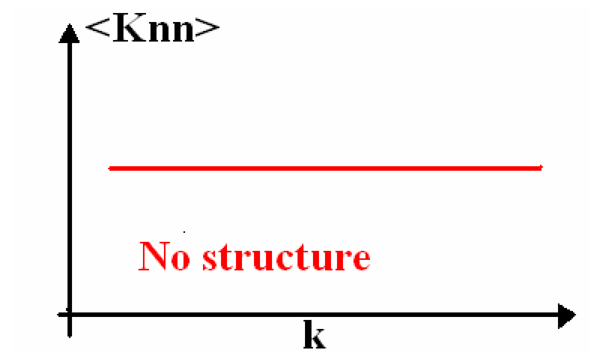
$$P(k) = e^{-pN}\frac{(pN)^k}{k!}$$

$$P(k) \propto k^{-\gamma}$$



·2 **Degree Correlation Knn (k)** = **average degree of a neighbour of a vertex with degree k**



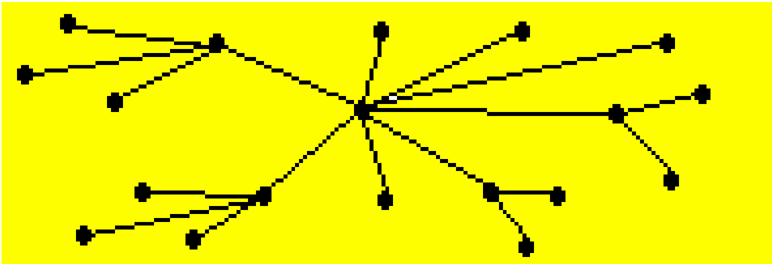


·3 **Clustering Coefficient (k)** = **the average value of c for a vertex whose degree is k**

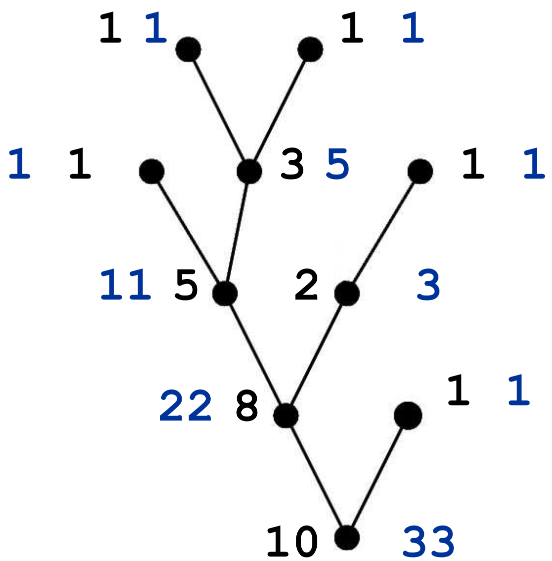
•1E (5) Statistical measures

·4 Centrality betweenness b(k) = The probability that a vertex whose degree is k has betweenness b

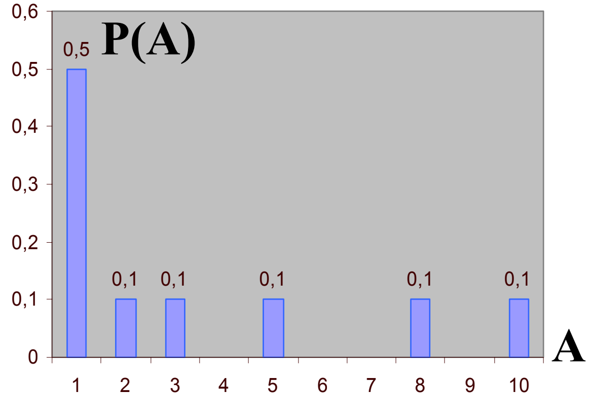
betweenness of I is the number of distances between any pair of vertices passing through I



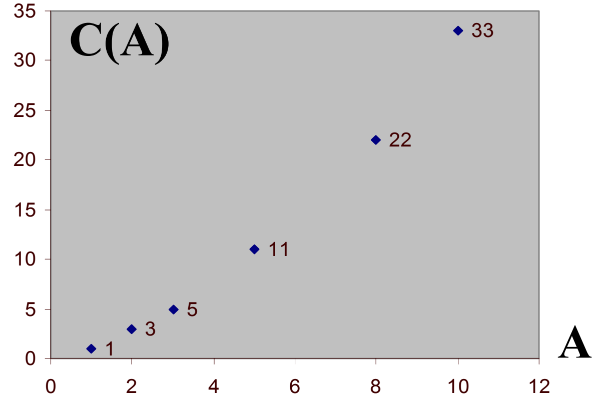
·5 TREES ONLY!!! P(A) = Probability Density for subbranches of size A



Size distribution:



Allometric relations:

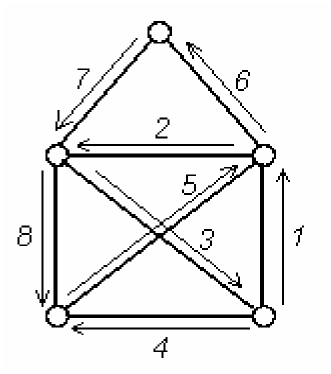


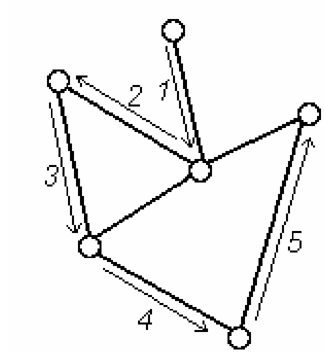
•1E (6) Boring stuff (1/3)

- The graph **size** is the number of its vertices.
- The graph **measure** is the number of its edges.
- The **degree** of a vertex in a graph is the number of edges that connects it to other vertices.
- In the case of an oriented graph the degree can be distinguished in **in-degree** and **out-degree**.
- Whenever all the vertices share the same degree the graph is called **regular**.
- A series of consecutive edges forms a **path**.
  - The number of edges in a path is called the length of the path.
  - A Hamiltonian path is a path that passes once through all the vertices (not necessarily through all the edges) in the graph.
  - A Hamiltonian cycle is a Hamiltonian path which begins and ends in the same vertex.
  - An Eulerian path is a path that passes once through all the edges (not necessarily once through all the vertices) in the graph.
  - An Eulerian cycle is an Eulerian path which begins and ends in the same edge.

•1E (6) Boring stuff (2/3)

- Whenever all the vertices share the same degree the graph is called **regular**.
- A series of consecutive edges forms a **path**.
  - The number of edges in a path is called the length of the path.
  - A Hamiltonian path is a path that passes once through all the vertices (not necessarily through all the edges) in the graph.
  - A Hamiltonian cycle is a Hamiltonian path which begins and ends in the same vertex.
  - An Eulerian path is a path that passes once through all the edges (not necessarily once through all the vertices) in the graph.
  - An Eulerian cycle is an Eulerian path which begins and ends in the same edge.





•1E (7) Boring stuff (3/3)

- A graph is **connected** if a path exists for any couple of vertices in the graph.
- A graph with no cycles is a **forest**. A **tree** is a connected forest.
- The **distance** between two vertices is the shortest number of edges one needs to travel to get from one vertex to the other.
- Therefore the **neighbours** of a vertex are all the vertices which are connected to that vertex by a single edge.
- A **dominating set** for a graph is a set of vertices whose neighbours, along with themselves, constitute all the vertices in the graph.
- A graph with size n cannot have a measure larger than $mmax = n(n-1)/2$. When all these possible edges are present the graph is **complete** and it is indicated with the symbol Kn.
- The opposite case happens when there are no edges at all. The measure is 0 and the graph is then **empty** and it is indicated by the symbol En.
- The **diameter** D of a graph is the longest distance you can find between two vertices in the graph.
- A complete **bipartite clique** Ki,j is a graph where every one of i nodes has an edge directed to each of the j nodes.
- The **clustering coefficient** C is a rougher characterization of clustering with respect to the clique distribution. C is given by the average fraction of pair of neighbours of a node that are also neighbours each other. For an empty graph En $C=0$ everywhere. For a complete graph Kn, $C=1$ everywhere.
- A **bipartite core** Ci,j is a graph on $i+j$ nodes that contains at least one Ki,j as a subgraph.