Core concepts

COMPLEX NETWORKS: A PRIMER

Shakti N. Menon

The Institute of Mathematical Sciences, Chennai

May 22nd 2025

Brains, Dynamics & Computation: A Workshop on Network Neuroscience

HOW CLOSE ARE REPRESENTATIONS TO REALITY?

René Magritte, *La Trahison des images* (“*The Treachery of Images*”, 1929)

“…What do you consider the largest map that would be really useful?”

“About six inches to the mile.”

“Only six inches!” exclaimed Mein Herr. “We very soon got to six yards to the mile. Then we tried a hundred yards to the mile. And then came the grandest idea of all! We actually made a map of the country, on the scale of a mile to the mile!”

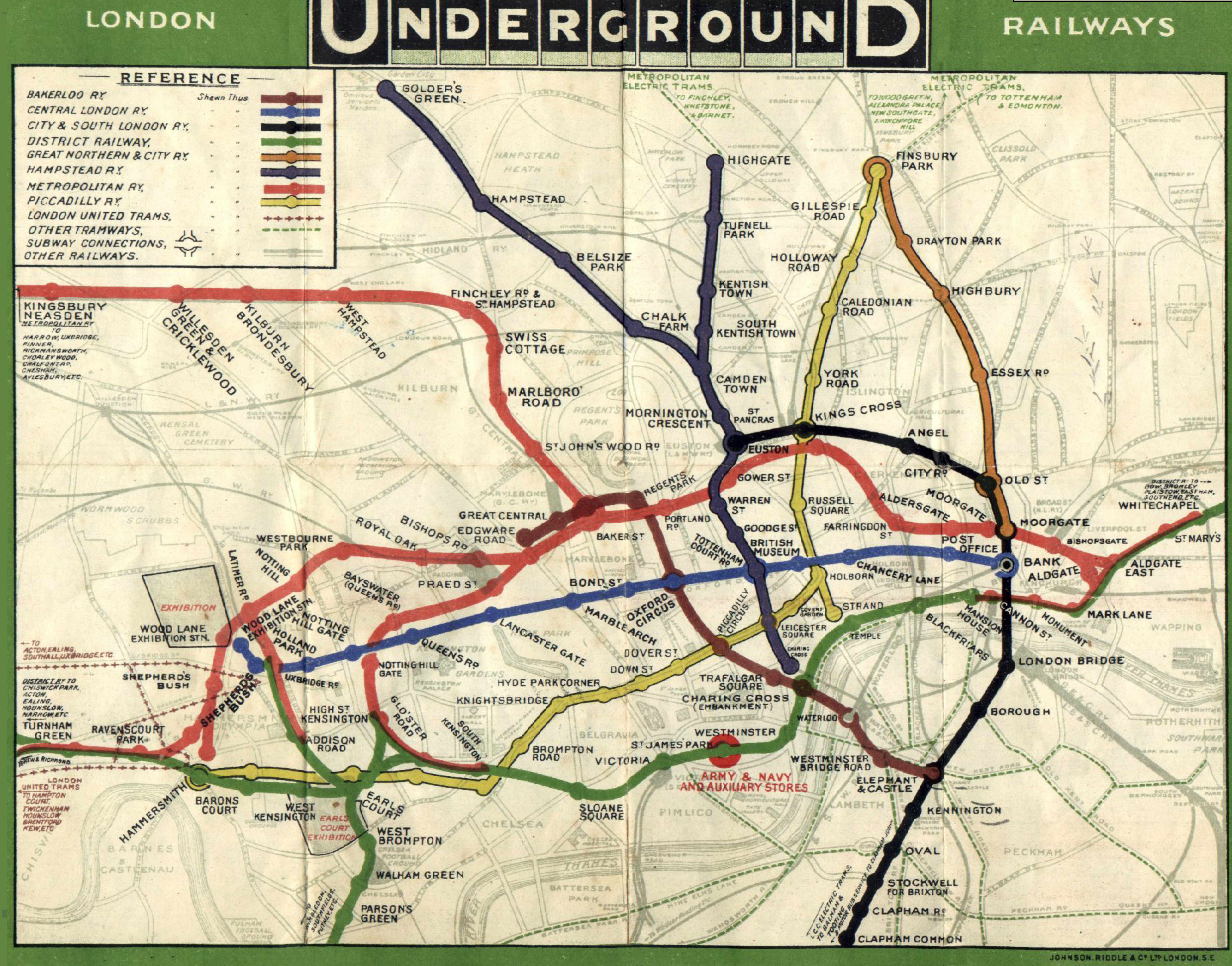
“Have you used it much?" I enquired.

“It has never been spread out, yet,” said Mein Herr: “the farmers objected: they said it would cover the whole country, and shut out the sunlight! So we now use the country itself, as its own map, and I assure you it does nearly as well.”

Lewis Carroll, *“Sylvie and Bruno Concluded” (1893)*

Representations cannot be a substitute for reality… or can they?

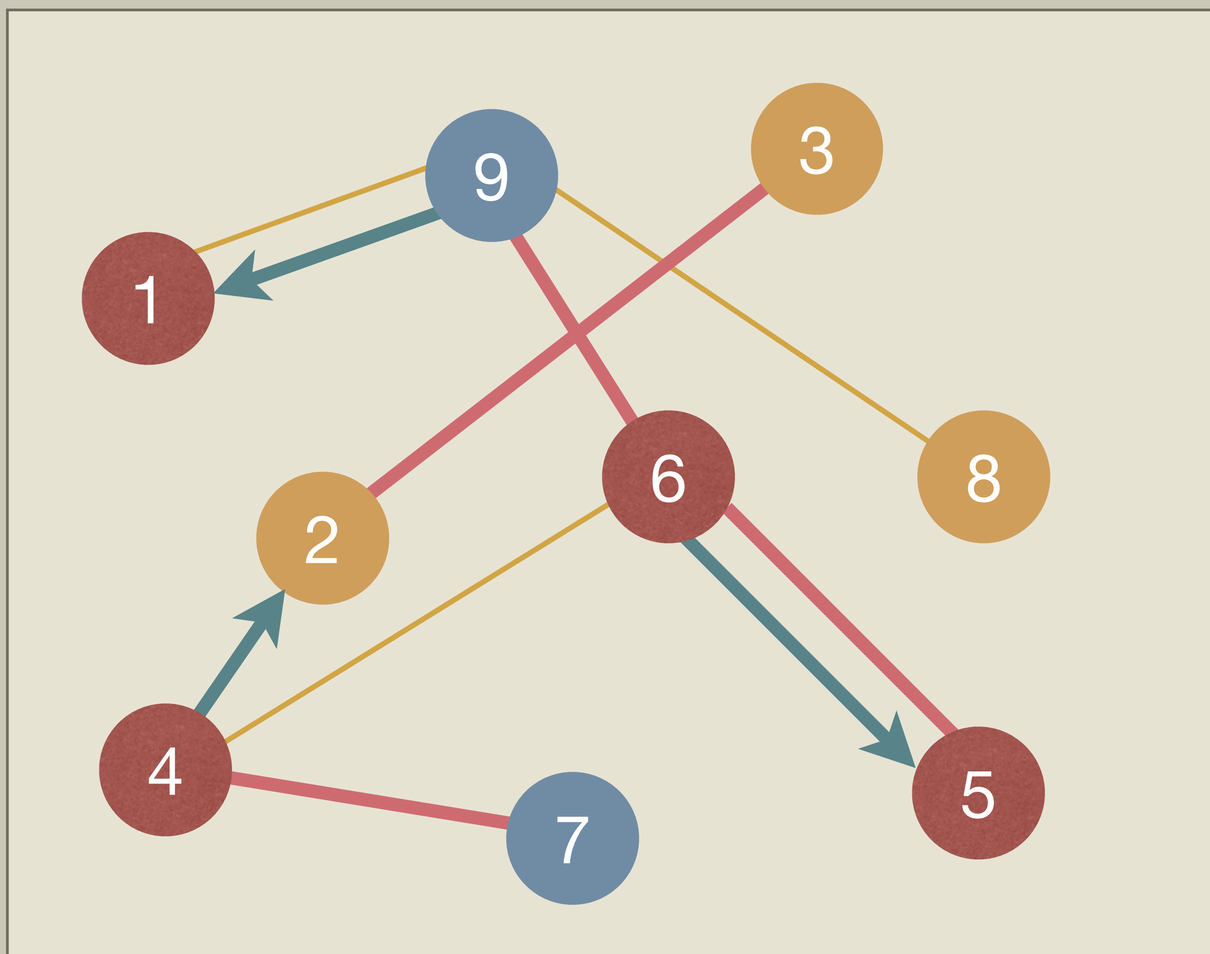
image: public domain (1908)

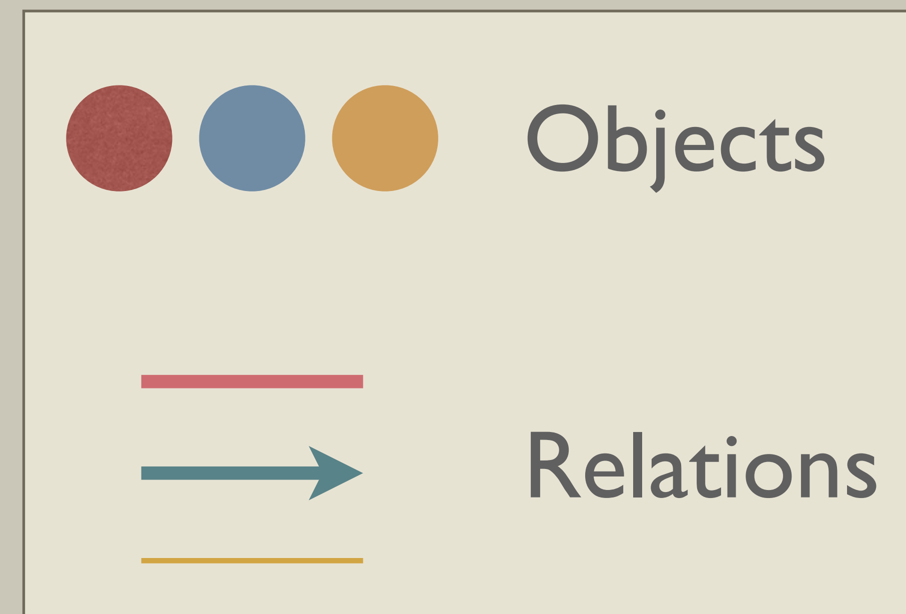


Harry Beck's Tube map (1933)

image: Rex Features

NETWORKS AS AN ABSTRACTION OF REALITY





In general, this scenario can change over time

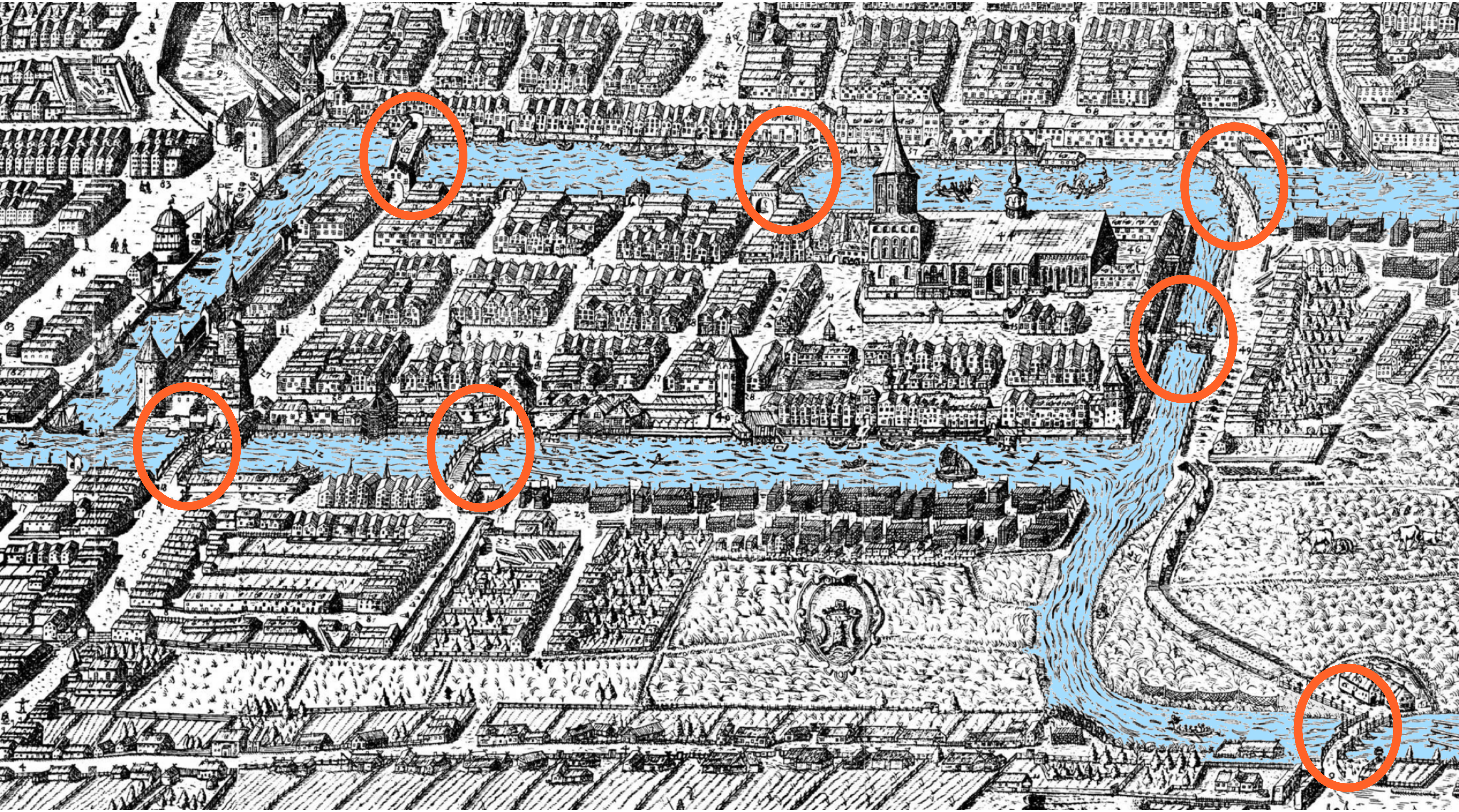
Every object has an associated set of attributes. Objects can also be classified based on certain common attributes.

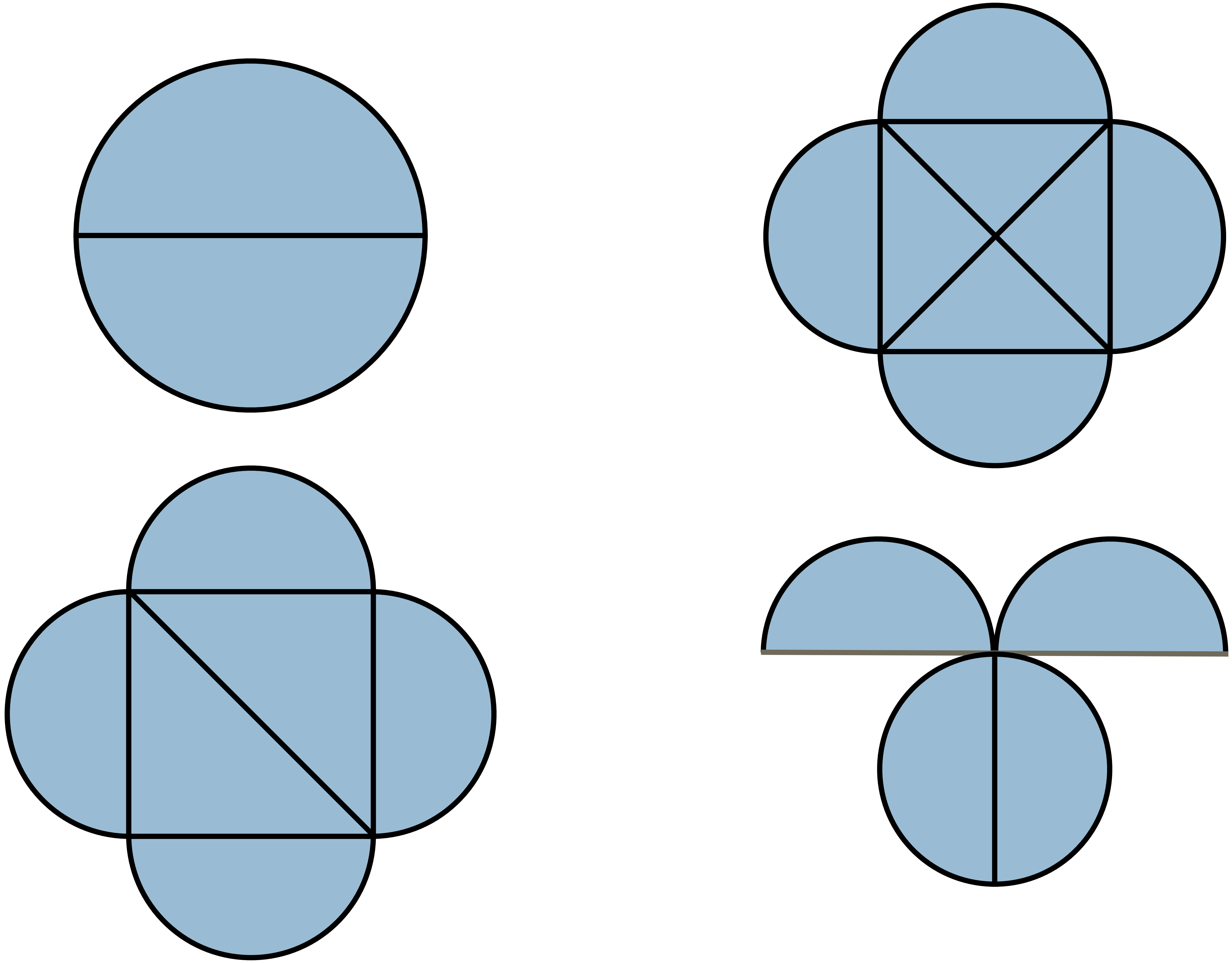
Relations are classified based on type, direction, intensity, etc.

Wiring diagram of a human brain

image: Laboratory of Neuro Imaging and Martinos Center for Biomedical Imaging, Consortium of the Human Connectome Project

THE SEVEN BRIDGES OF KÖNIGSBERG

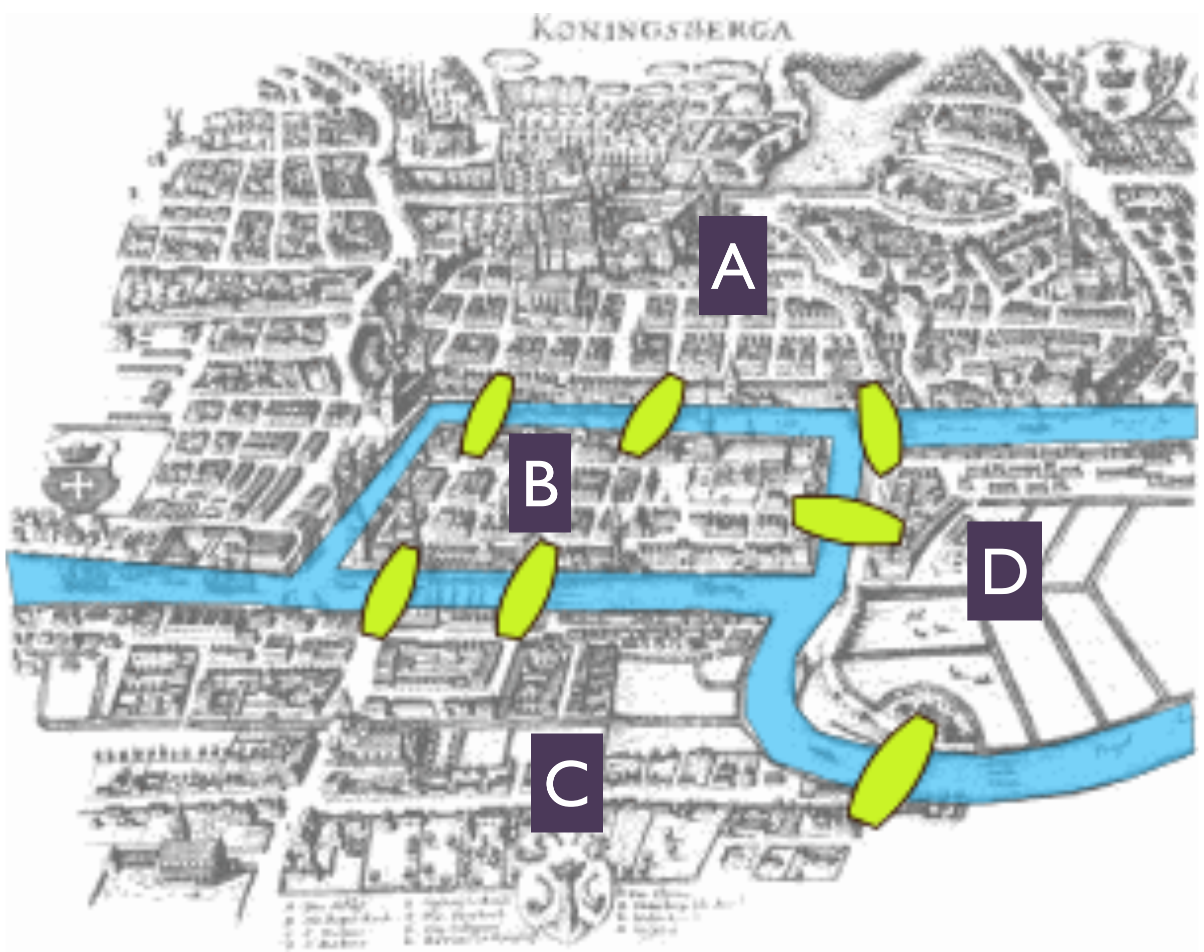


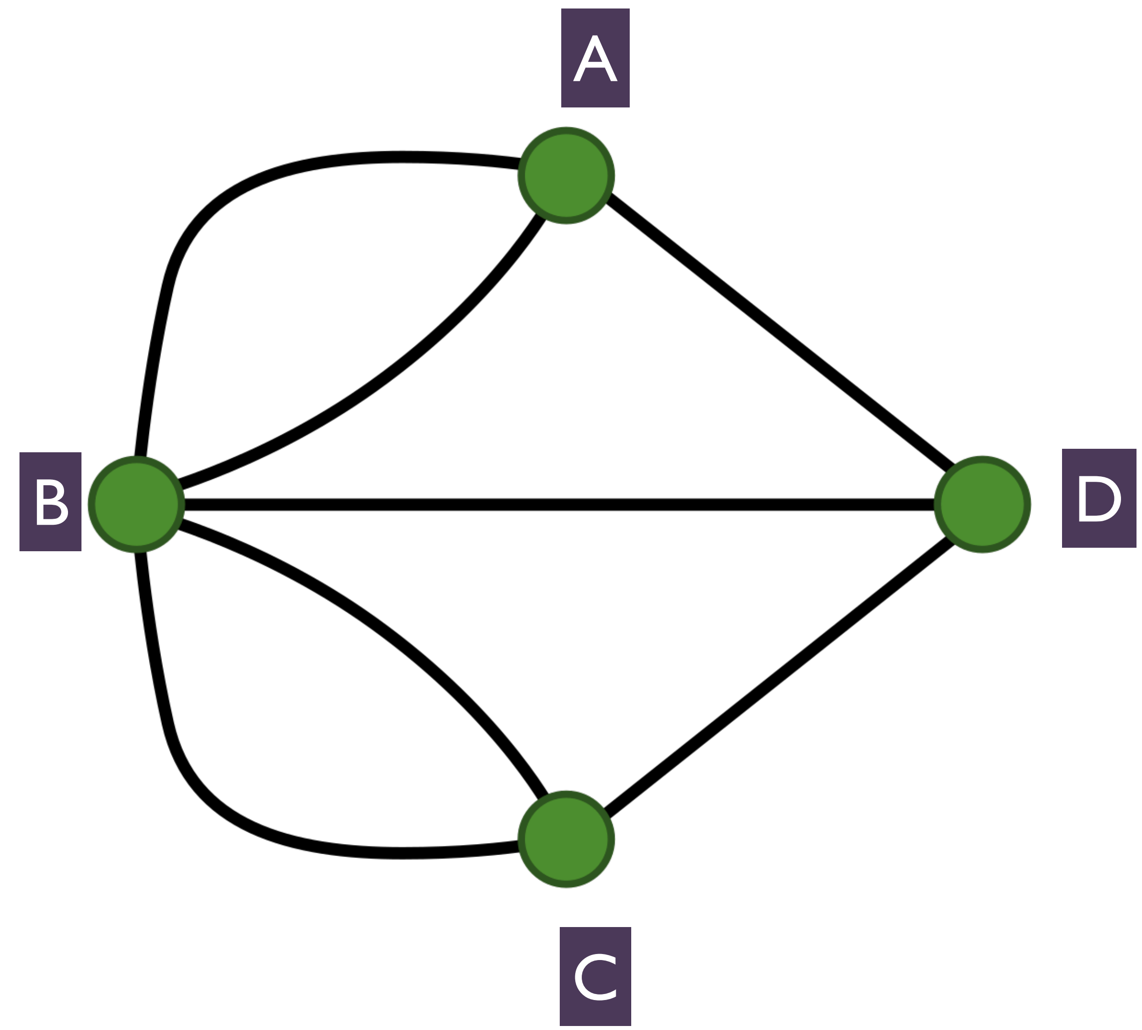


Can you draw these patterns

- without taking your pen off the paper, and
- without crossing any path twice?

EULER'S SOLUTION IN 1736





Each land mass can be viewed as a "vertex" and each bridge as a "link".

Only terminal vertices can have an odd number of links.

GRAPHS AND NETWORKS

Euler's work laid the foundation for the field of **graph theory**.

Any network of connections between entities can be analysed by viewing it as a graph that describes the manner in which a set of objects are connected.

Conversely, a network can simply be thought of as a graph where the objects and relations can be mapped to some real world setting.

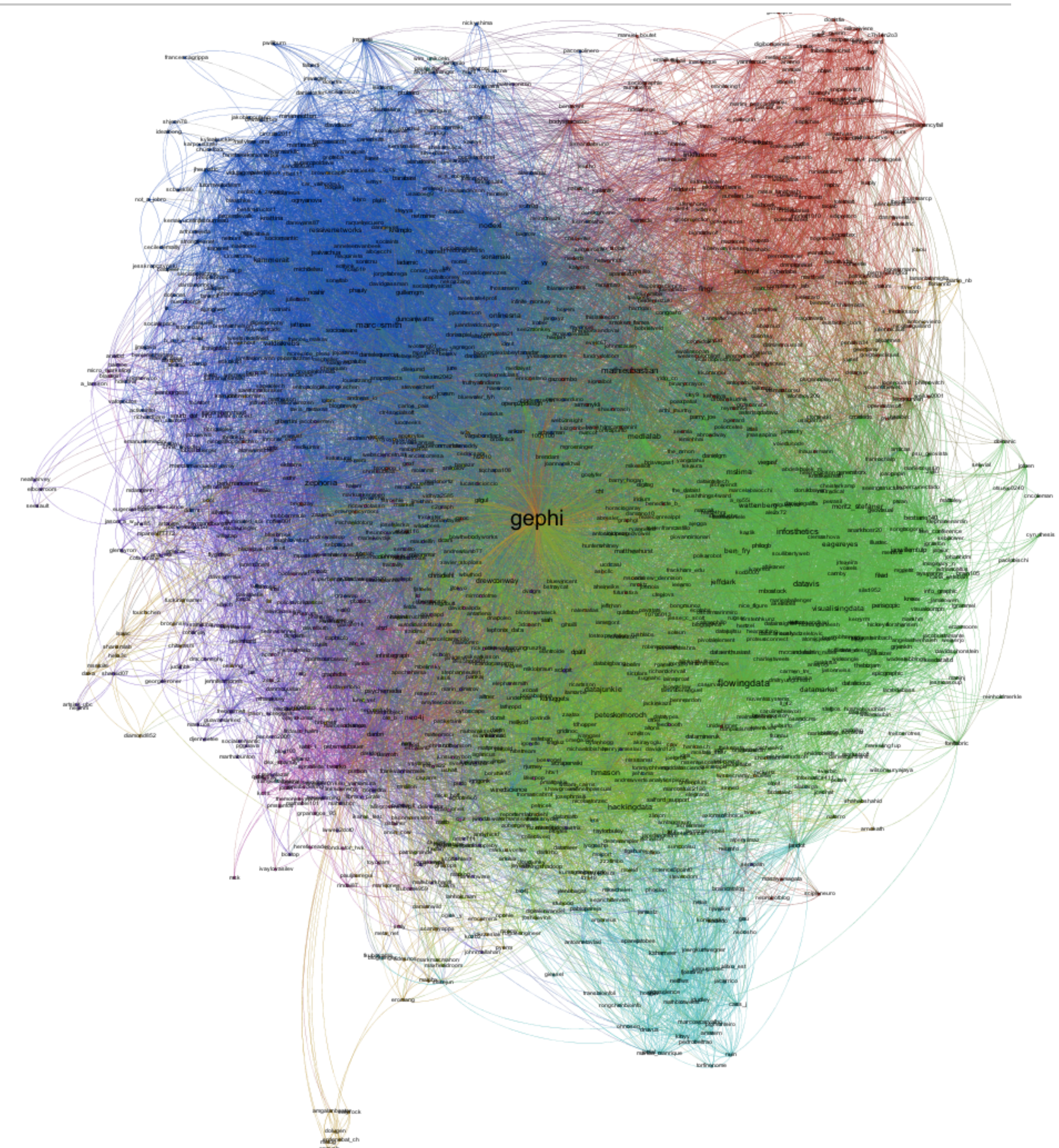
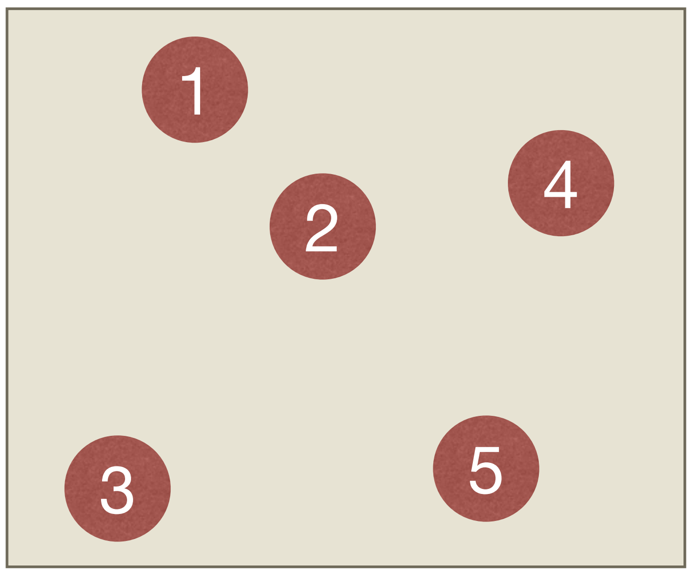
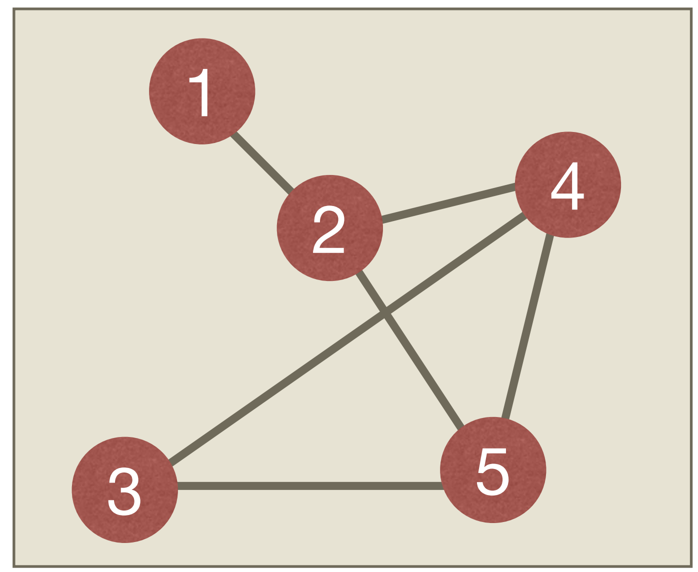


image: http://social-dynamics.org/a-gephi-visualization-of-gephi-on-twitter/

FUNDAMENTAL CONCEPTS: NODES AND LINKS



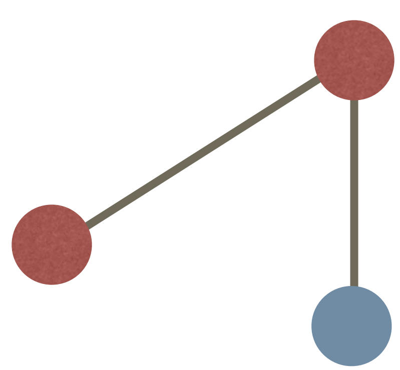
Nodes (or "Vertices")



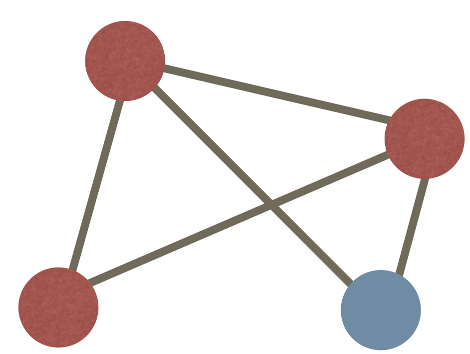
Links (or "Edges")

NODE	DEGREE
1	1
2	3
3	2
4	3
5	3

The total number of links associated with a node is its **degree** (k).

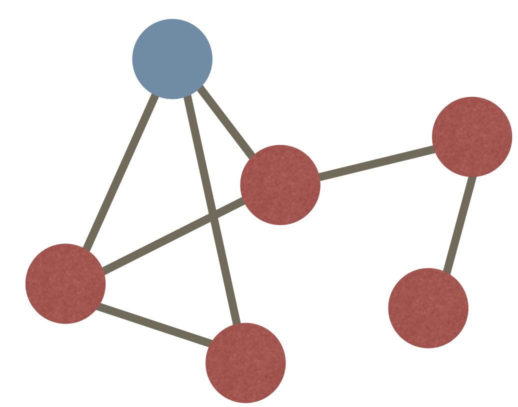


$k = 1$



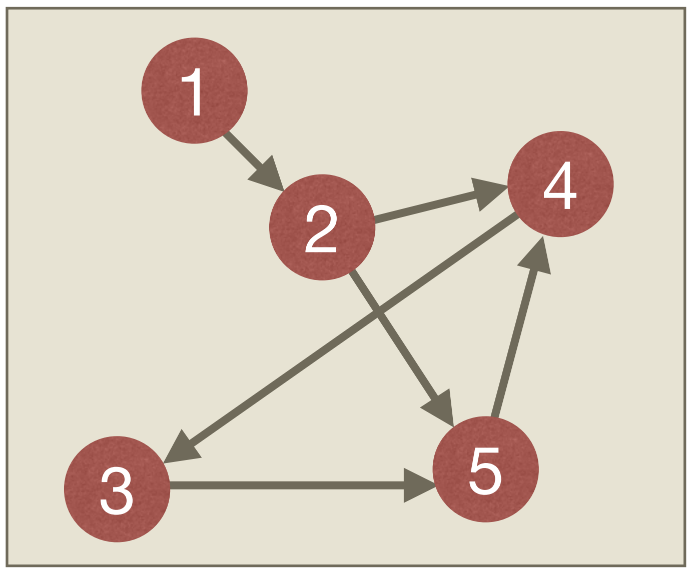
$k = 2$

$k = 3$

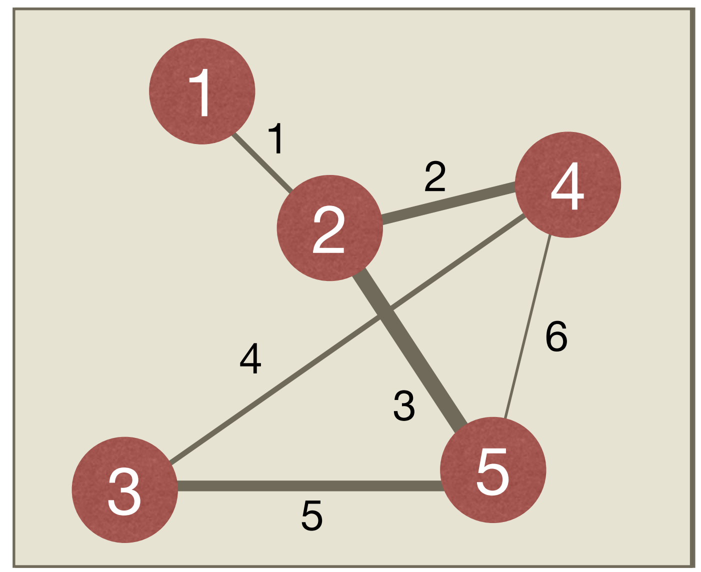


FUNDAMENTAL CONCEPTS:
DIRECTED AND WEIGHTED GRAPHS

Directed graph



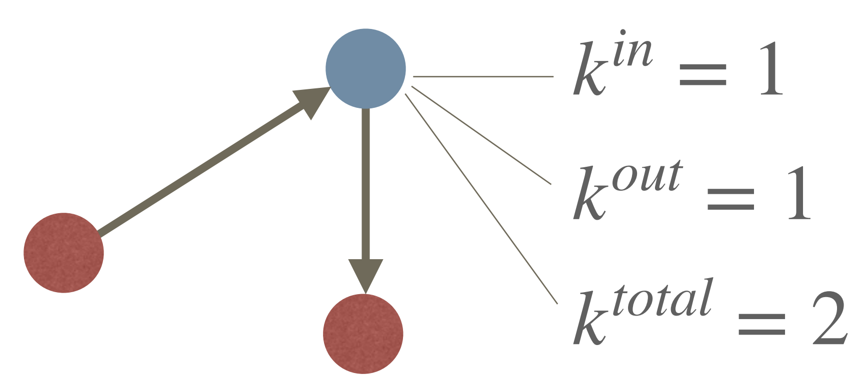
Weighted graph



EDGE	WEIGHT
1	1
2	2
3	3
4	1
5	2
6	0.5

NODE	IN-DEGREE	OUT-DEGREE	TOTAL DEGREE
1	0	1	1
2	1	2	3
3	1	1	2
4	2	1	3
5	2	1	3

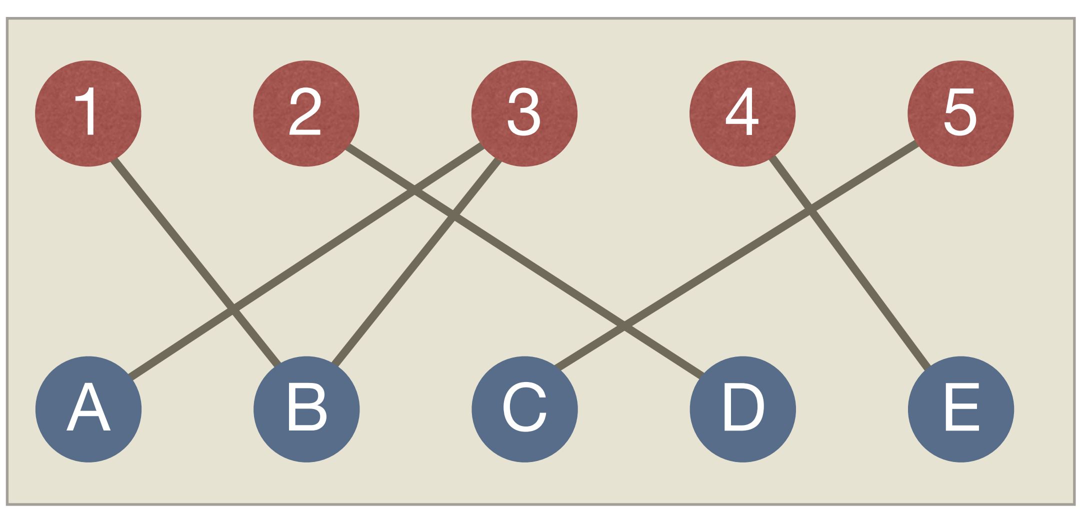
In a directed graph a node's in-degree can be different to its out-degree



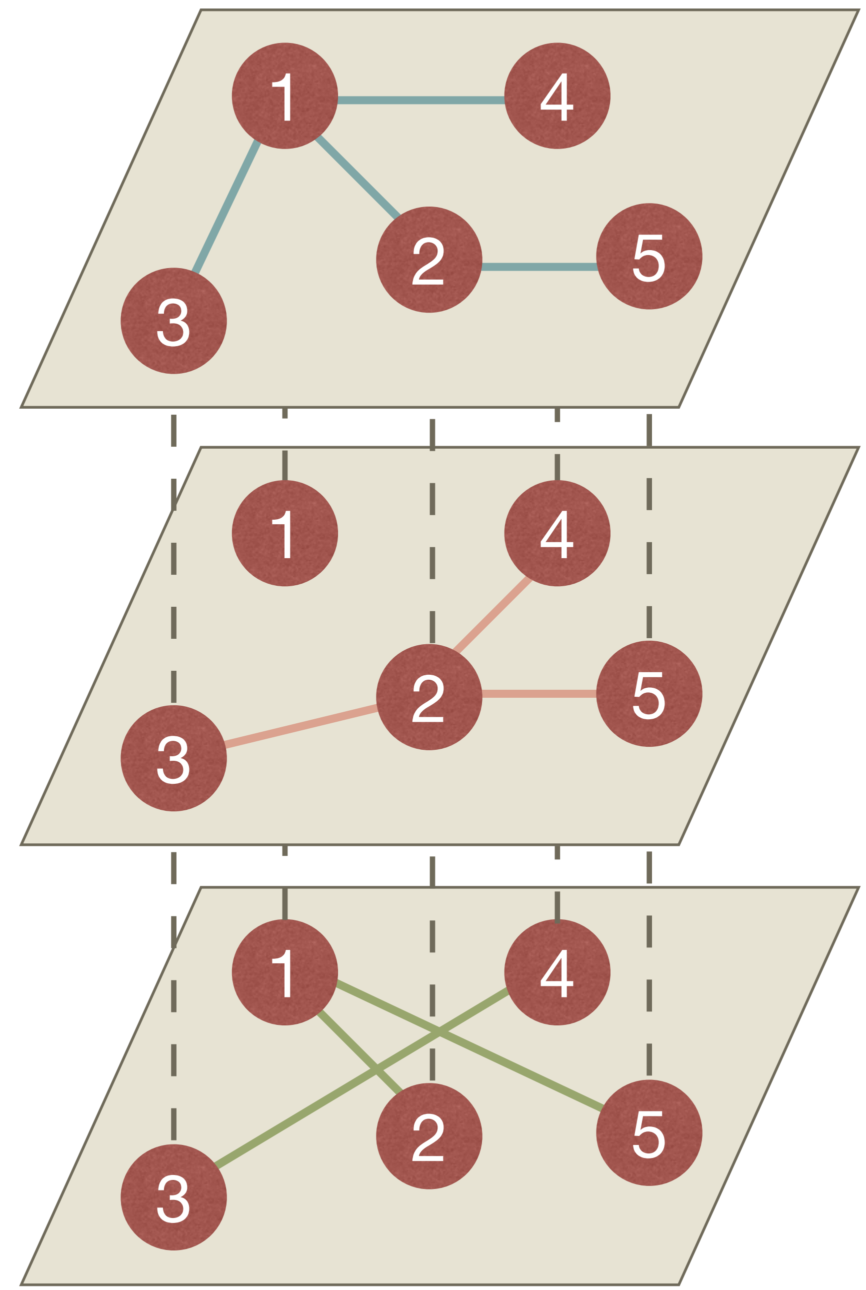
SOME OTHER TYPES OF GRAPHS

Graphs that describe relations between two different classes of objects are known as Bipartite graphs.

Graphs in which there may be different types of links between nodes are known as Multiplex graphs.

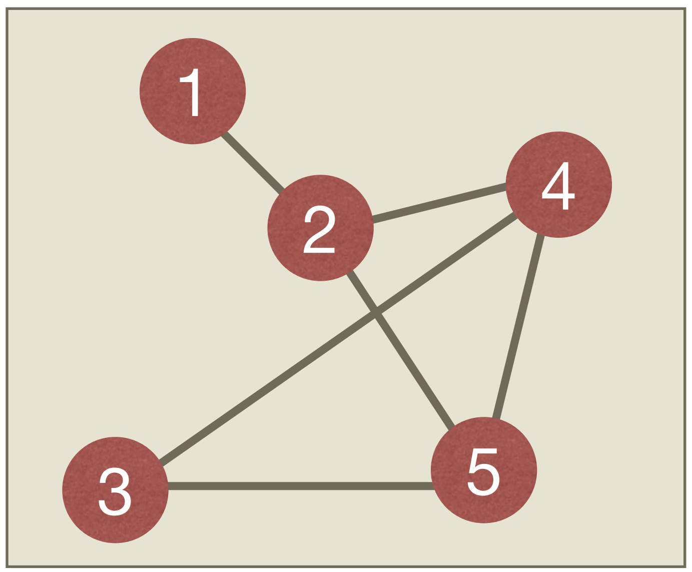


Bipartite graphs



Multiplex graph

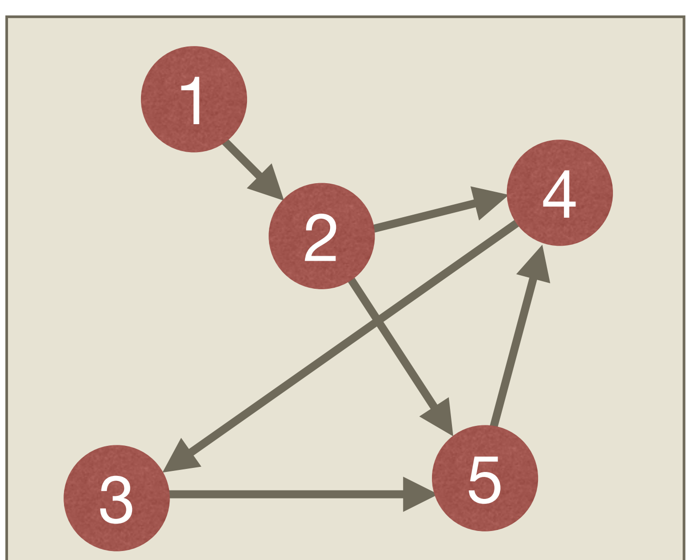
FUNDAMENTAL CONCEPTS: ADJACENCY MATRIX



Adjacency matrix

	1	2	3	4	5
1	0	1	0	0	0
2	1	0	0	1	1
3	0	0	0	1	1
4	0	1	1	0	1
5	0	1	1	1	0

The adjacency matrix $\mathbf{A}$ specifies all connections in the graph. If nodes i and j are connected then $A_{ij} = 1$ else $A_{ij} = 0$.



target →

source ↓

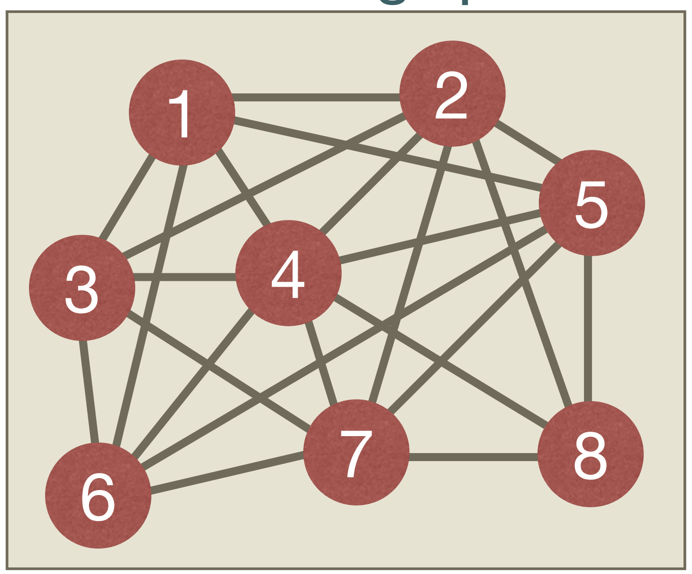
	1	2	3	4	5
1	0	1	0	0	0
2	0	0	0	1	1
3	0	0	0	0	1
4	0	0	1	0	0
5	0	0	0	1	0

In an undirected graph, the degree k_i of a node i can be obtained via:

$$k_i = \sum_i A_{ij} = \sum_j A_{ij}$$

FUNDAMENTAL CONCEPTS: DENSITY AND SPARSITY

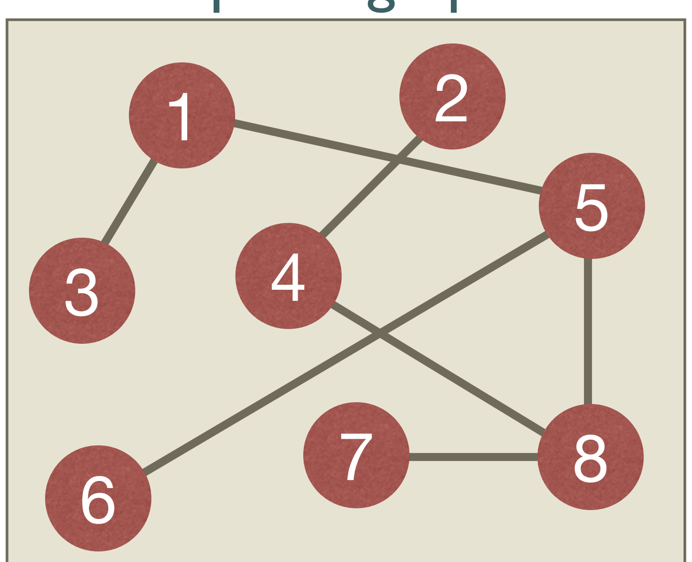
Dense graph



	1	2	3	4	5	6	7	8
1	0	1	1	1	1	1	0	0
2	1	0	1	1	1	0	1	1
3	1	1	0	1	0	1	1	0
4	1	1	1	0	1	1	1	1
5	1	1	0	1	0	1	1	1
6	1	0	1	1	1	0	1	0
7	0	1	1	1	1	1	0	1
8	0	1	0	1	1	0	1	0

The density ρ is the fraction of connected node pairs that exist in the graph.

Sparse graph

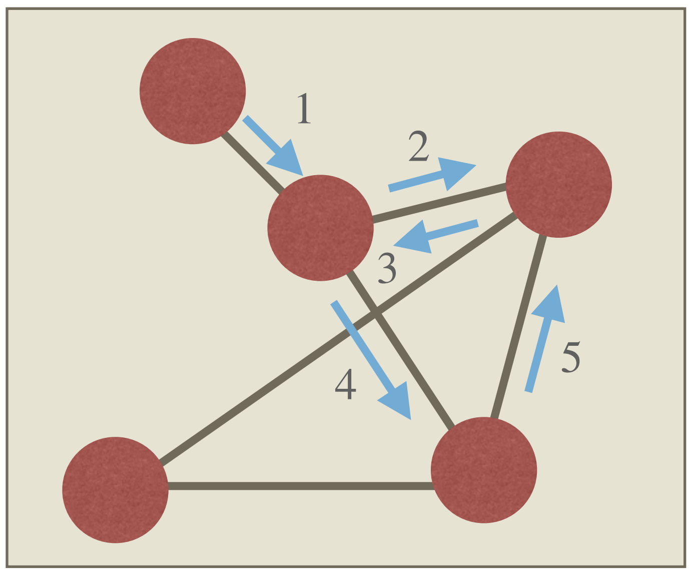


	1	2	3	4	5	6	7	8
1	0	0	1	0	1	0	0	0
2	0	0	0	1	0	0	0	0
3	1	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	1
5	1	0	0	0	0	1	0	1
6	0	0	0	0	1	0	0	0
7	0	0	0	0	0	0	0	1
8	0	0	0	1	1	0	1	0

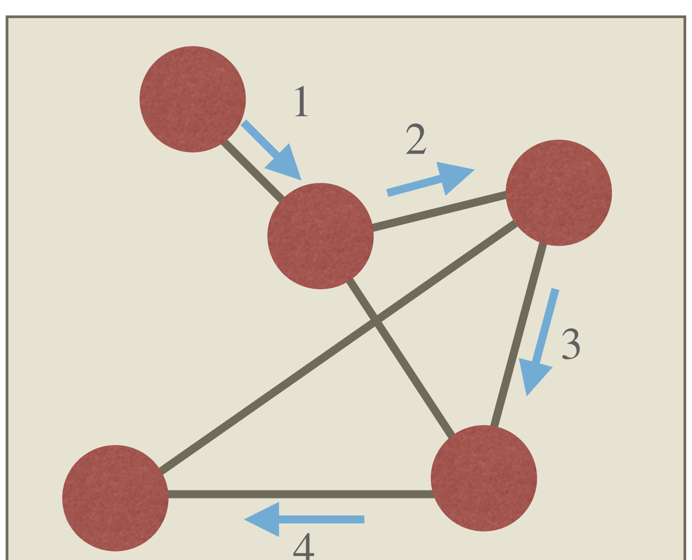
A graph is said to be dense if "most" of the possible links are present, and sparse if "most" are absent.

FUNDAMENTAL CONCEPTS: WALKS AND PATHS

Walk



Path

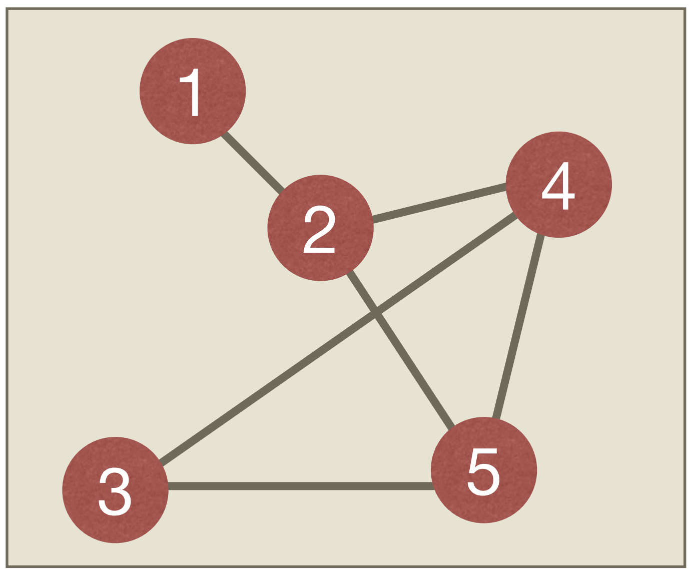


A walk is a route along the edges of a graph. In an undirected graph, an edge can be crossed in either direction.

The length of a walk is the number of hops taken along the route.

A path is a self-avoiding walk, i.e. one in which no edge is traversed twice.

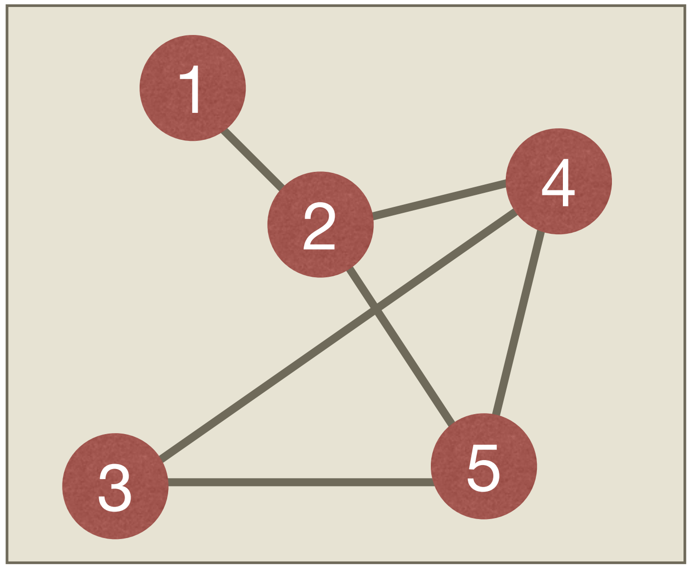
FUNDAMENTAL CONCEPTS: SHORTEST PATH LENGTH & DIAMETER



i-j	1-2	1-3	1-4	1-5	2-3	2-4	2-5	3-4	3-5	4-5
d_{ij}	1	3	2	2	2	1	1	1	1	1

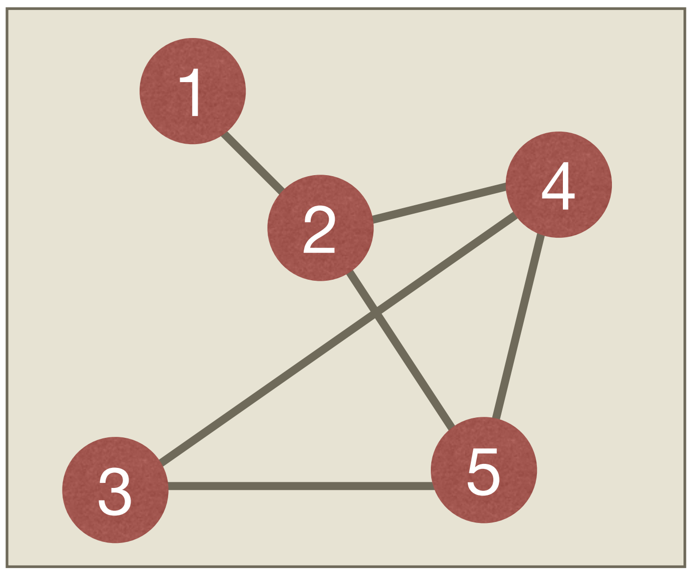
The shortest path length d_{ij} between two nodes i and j is the minimum number of links one has to cross to travel between them.

$d_{max} = 3$



The diameter d_{max} of a network is the "longest shortest path" between all pairs of nodes i and j in the graph : $\max_{(i,j)}(d_{ij})$.

FUNDAMENTAL CONCEPTS: AVERAGE PATH LENGTH



i-j	1-2	1-3	1-4	1-5	2-3	2-4	2-5	3-4	3-5	4-5
d_{ij}	1	3	2	2	2	1	1	1	1	1

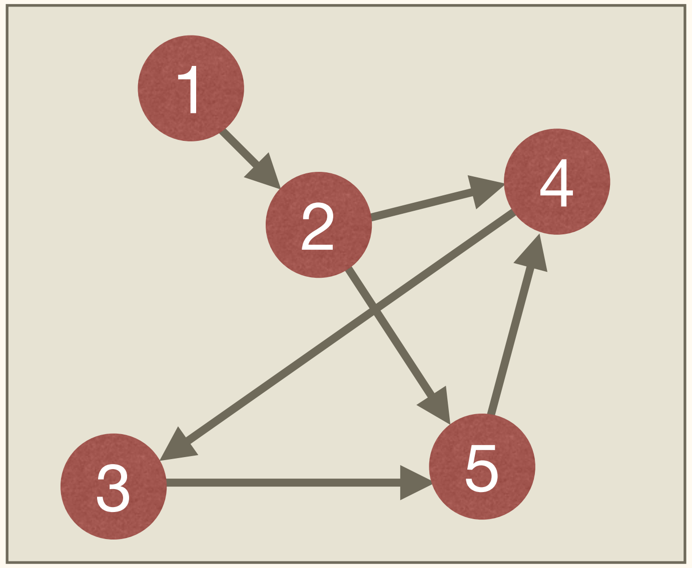
i-j	2-1	3-1	4-1	5-1	3-2	4-2	5-2	4-3	5-3	5-4
d_{ij}	1	3	2	2	2	1	1	1	1	1

$$\sum_{i \neq j} d(i,j) = 30$$

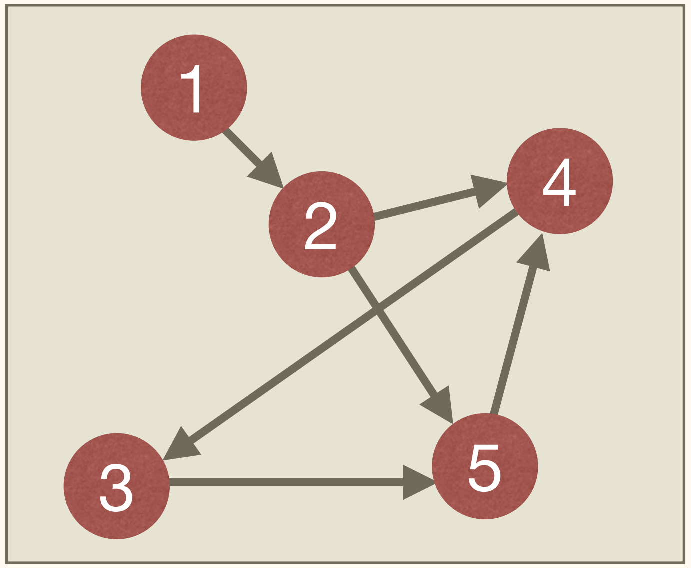
- The average path length is the average of the shortest path lengths between every pair of nodes in the graph.
- For a graph comprising N nodes, the average path length is:

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d(i,j)$$

QUESTIONS

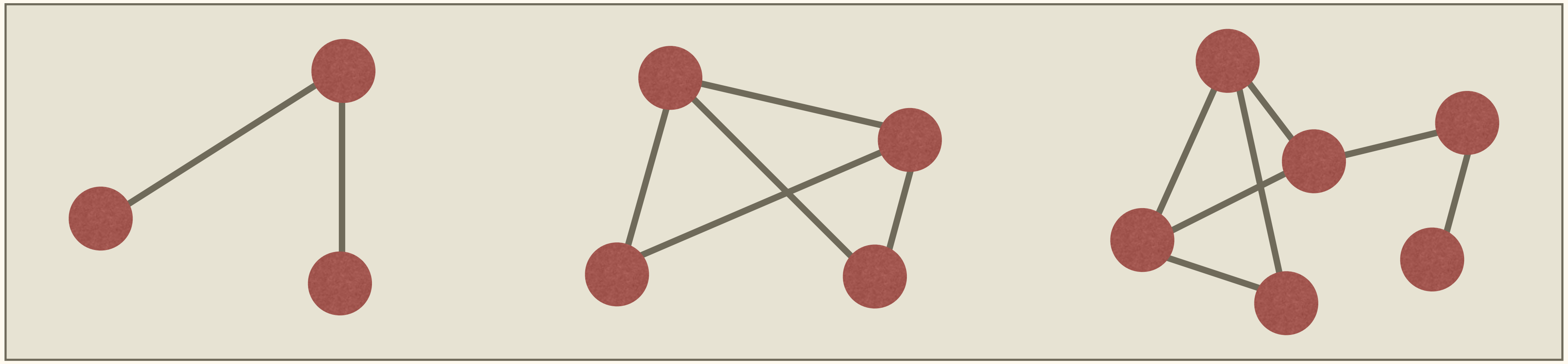


What is the shortest path length between every pair of nodes?

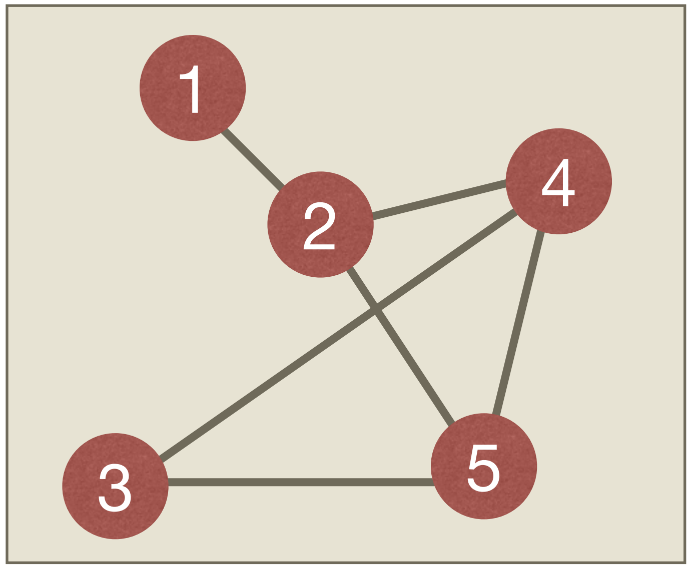


What is the diameter ($d_{\max}$)?

What is the average path length of these graphs?



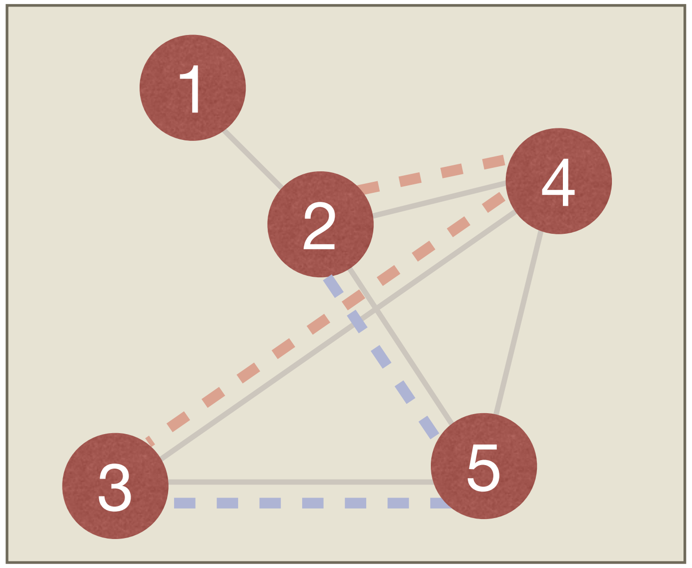
MORE ON PATH LENGTHS: TOTAL NUMBER OF WALKS OF A GIVEN LENGTH



#walks of length 1

	1	2	3	4	5
1	0	1	0	0	0
2	1	0	0	1	1
3	0	0	0	1	1
4	0	1	1	0	1
5	0	1	1	1	0

The total number of walks $N_{ij}^{(1)}$ of length 1 between a pair of nodes (i, j) is just A_{ij}

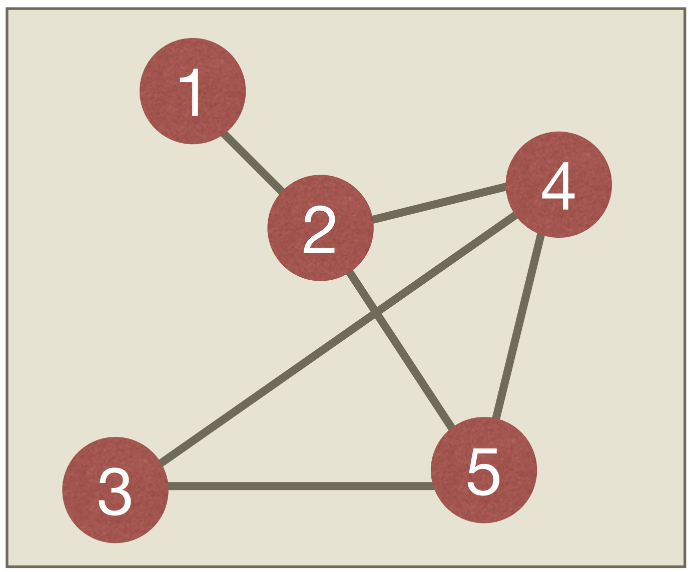


#walks of length 2

	1	2	3	4	5
1	1	0	0	1	1
2	0	3	2	1	1
3	0	2	2	1	1
4	1	1	1	3	2
5	1	1	1	2	3

One can have multiple walks $N_{ij}^{(2)}$ of length 2 between a pair of nodes (i, j).

MORE ON PATH LENGTHS:
TOTAL NUMBER OF WALKS OF A GIVEN LENGTH



	1	2	3	4	5
1	0	1	0	0	0
2	1	0	0	1	1
3	0	0	0	1	1
4	0	1	1	0	1
5	0	1	1	1	0

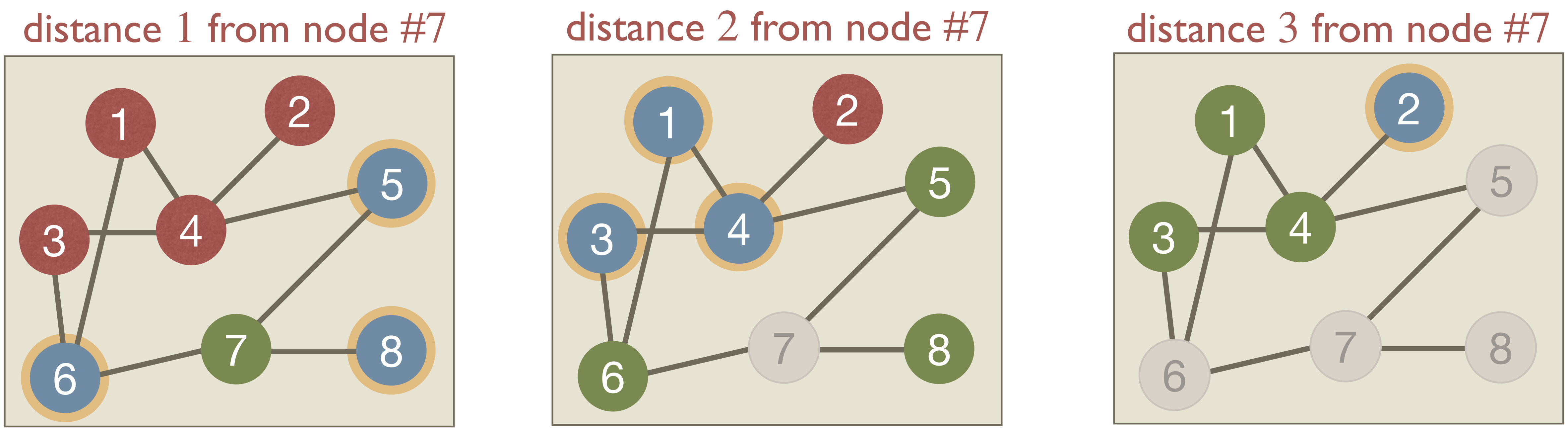
	1	2	3	4	5
1	0	1	0	0	0
2	1	0	0	1	1
3	0	0	0	1	1
4	0	1	1	0	1
5	0	1	1	1	0

In order to have a walk of length 2 between nodes $(2,3)$, we consider all the nodes of distance 1 from node #2, and count how many of them are distance 1 from node #3.

$$N_{23}^{(2)} = \sum_k A_{2k} A_{k3}$$

i.e. $N^{(d)} = \underbrace{A\, A \ldots\, A}_{d} = A^d$

MORE ON PATH LENGTHS: BREADTH-FIRST SEARCH

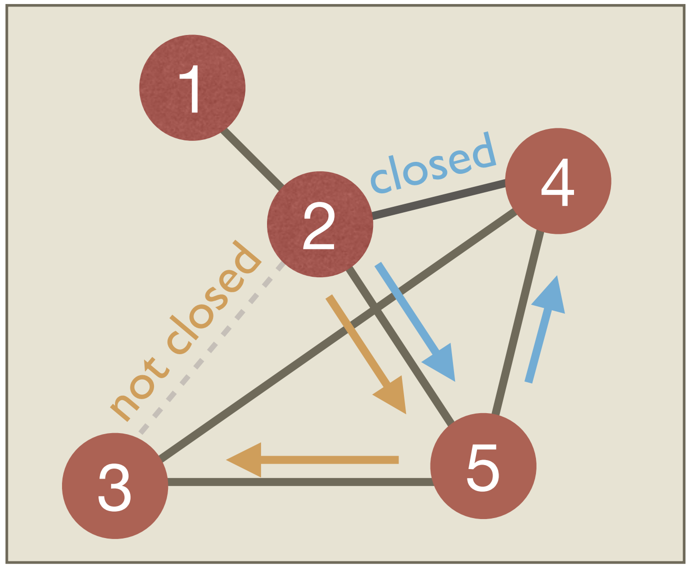


To find the shortest path between nodes (i, j), we can follow the breadth-first search algorithm:

1. Find the neighbours (blue) of node i (green) from the adjacency matrix A.
2. Remove the green node and make the blue nodes green.
3. Find the neighbours of the green nodes (excluding removed ones).
4. Repeat as long as there are neighbours.

FUNDAMENTAL CONCEPTS: CLUSTERING COEFFICIENT

- In real networks, one often finds that nodes that form links with one another also form links with those that the neighbour link to.
- This can be measured by the (global) clustering coefficient: the fraction of paths of length 2 that are "closed" (the three nodes of the path are all connected).
- A triangle of nodes connected to each other contain 3 closed paths.



Thus, the global clustering coefficient is:

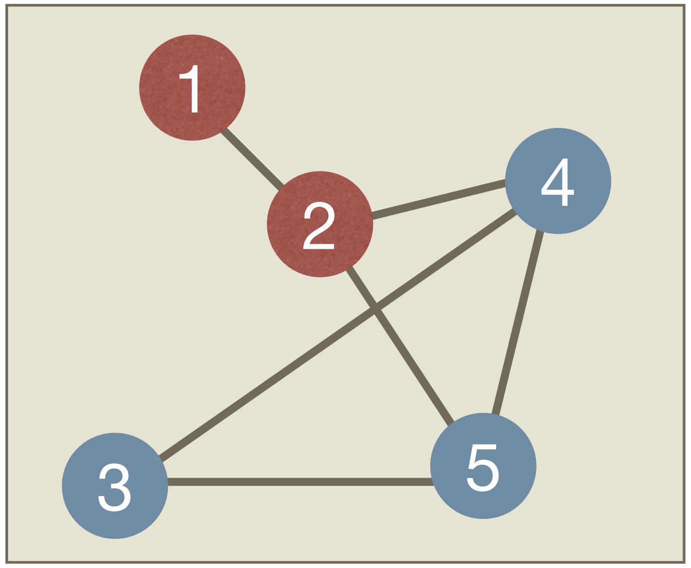
$$C = \frac{\#\text{triangles} \times 3}{\#\text{connected triples}}$$

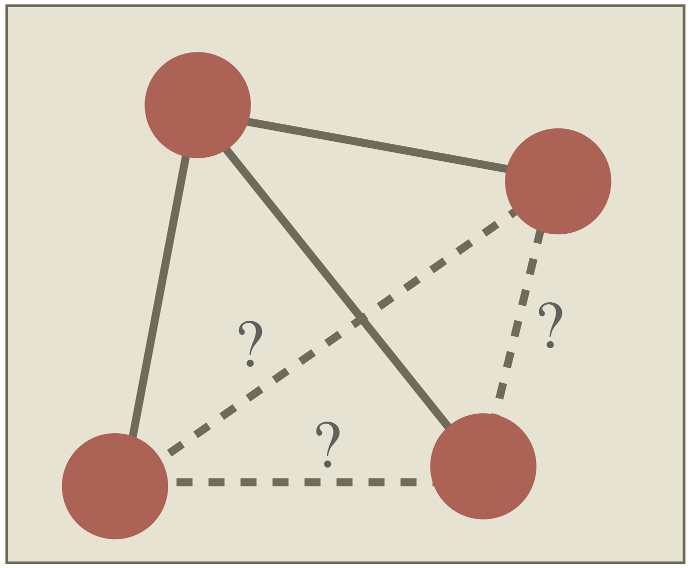
where a connected triple is an ordered set of three nodes **abc**, where both **a** and **c** have links to **b**.

FUNDAMENTAL CONCEPTS: LOCAL CLUSTERING COEFFICIENT

- The (local) clustering coefficient of a node measures the extent of connectivity of its local neighbourhood, i.e. how close they are to being a "clique" or a complete subgraph.
- If a node i in an undirected graph has k_i neighbours, there can be a maximum of $k_i(k_i - 1)/2$ links between them.
- The local clustering coefficient C_i of node i is the <u>fraction of these links that exist</u>.

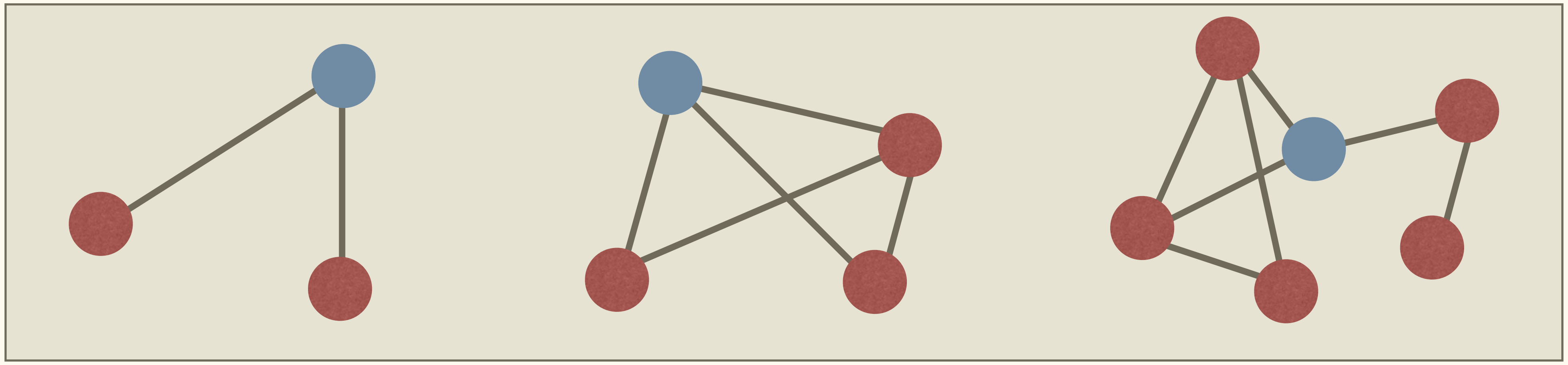
blue nodes form a clique



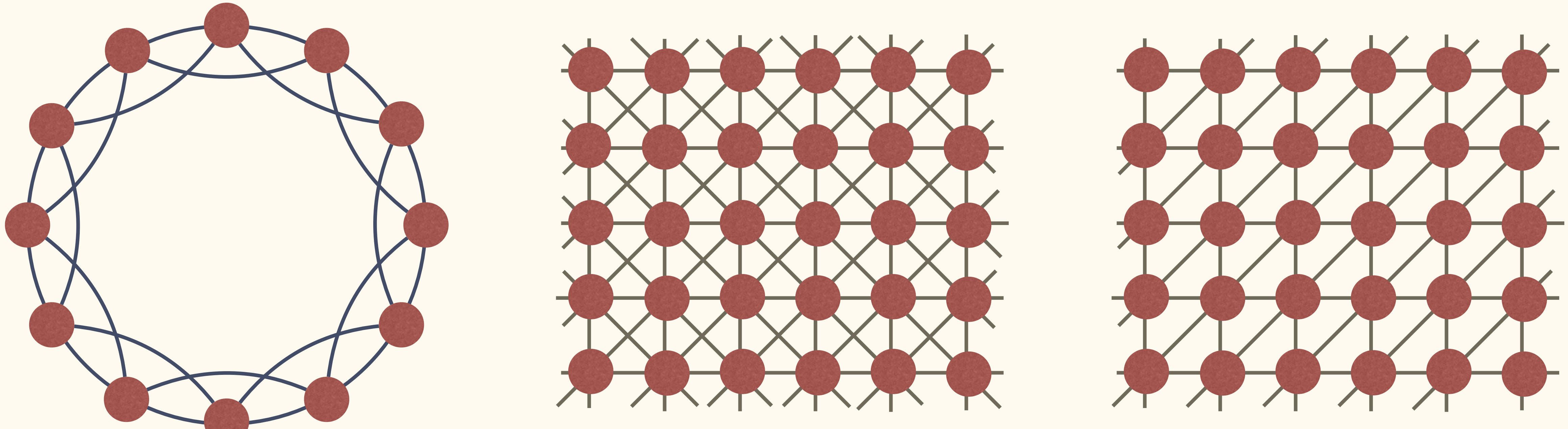


QUESTIONS

What is the clustering coefficient of the blue nodes?

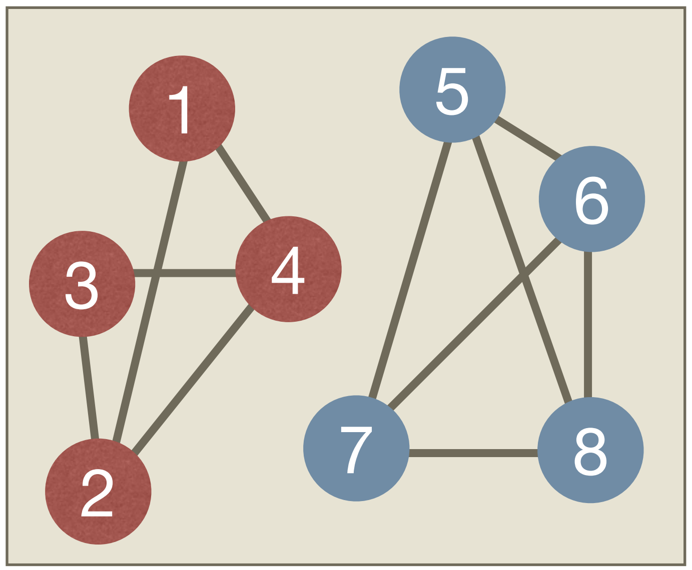


Calculate the clustering coefficients for a node in the following graphs:

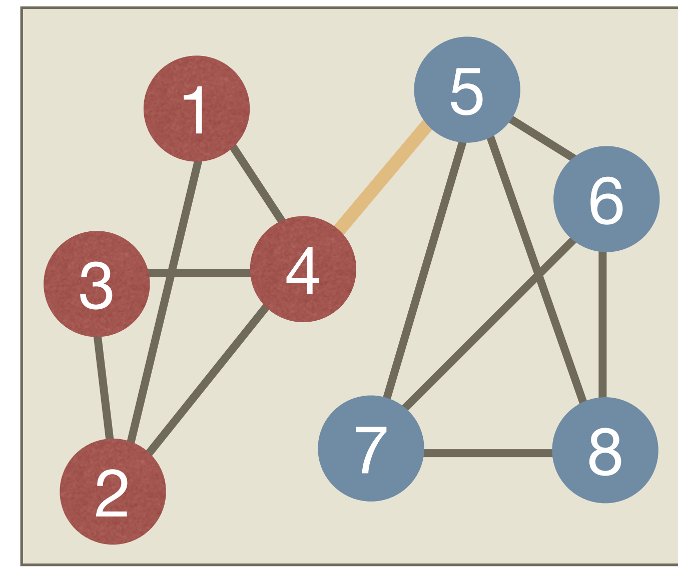


$k_i = 4$ $k_i = 8$ $k_i = 6$

FUNDAMENTAL CONCEPTS: COMPONENTS



	1	2	3	4	5	6	7	8
1	0	1	0	1	0	0	0	0
2	1	0	1	1	0	0	0	0
3	0	1	0	1	0	0	0	0
4	1	1	1	0	0	0	0	0
5	0	0	0	0	0	1	1	1
6	0	0	0	0	1	0	1	1
7	0	0	0	0	1	1	0	1
8	0	0	0	0	1	1	1	0

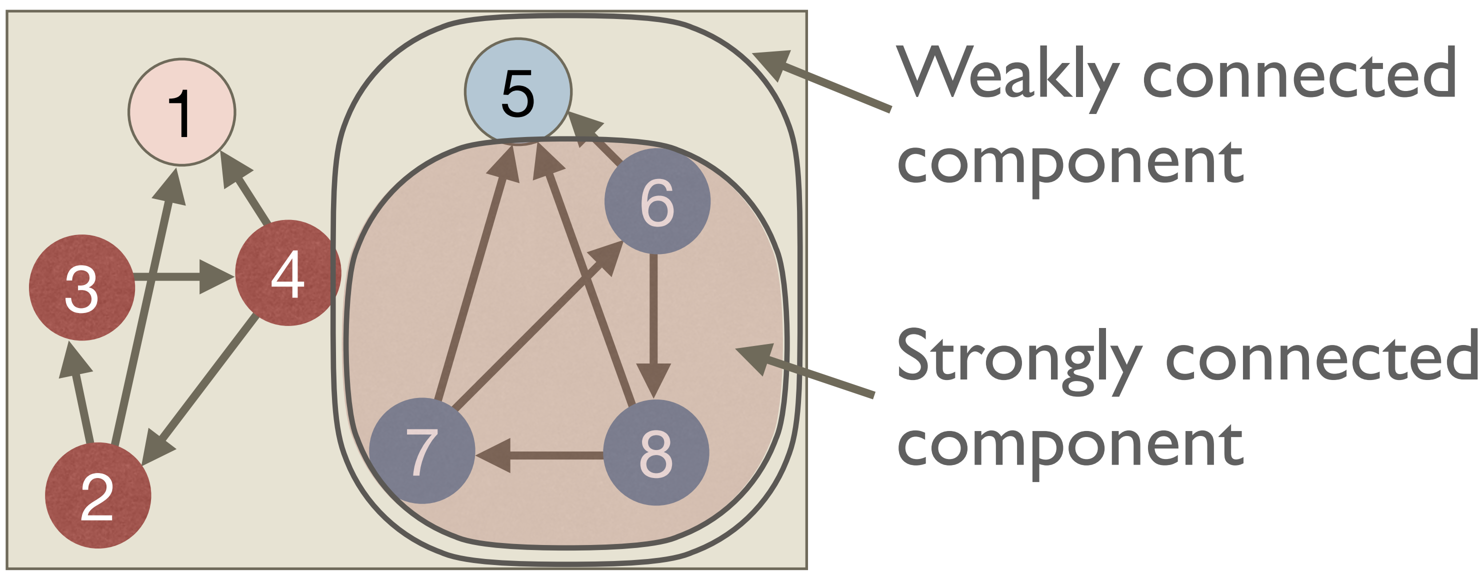


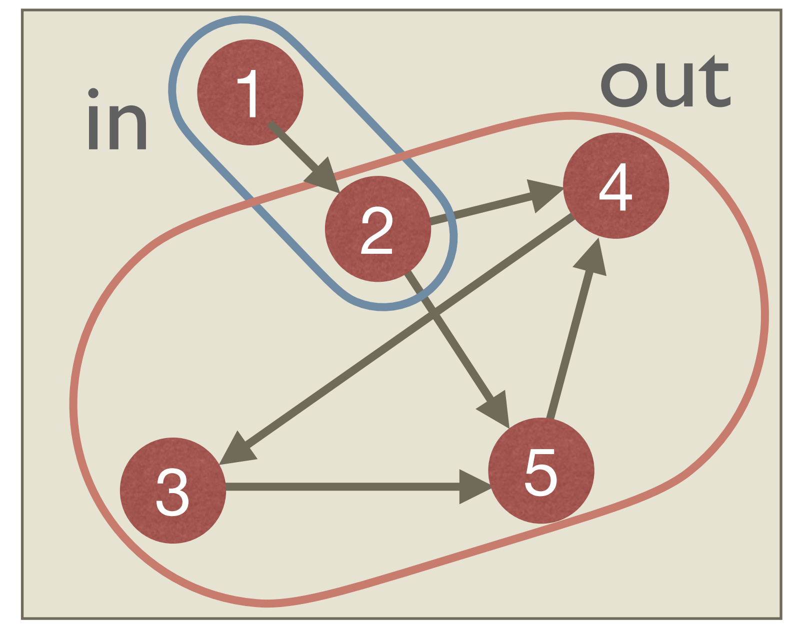
In an undirected network, a pair of nodes (i, j) are connected if there exists a path (of any length) between them.

A component is a subset of the network in which all nodes are connected.

A bridge is a link that, when cut, causes the network to be disconnected.

FUNDAMENTAL CONCEPTS: COMPONENTS



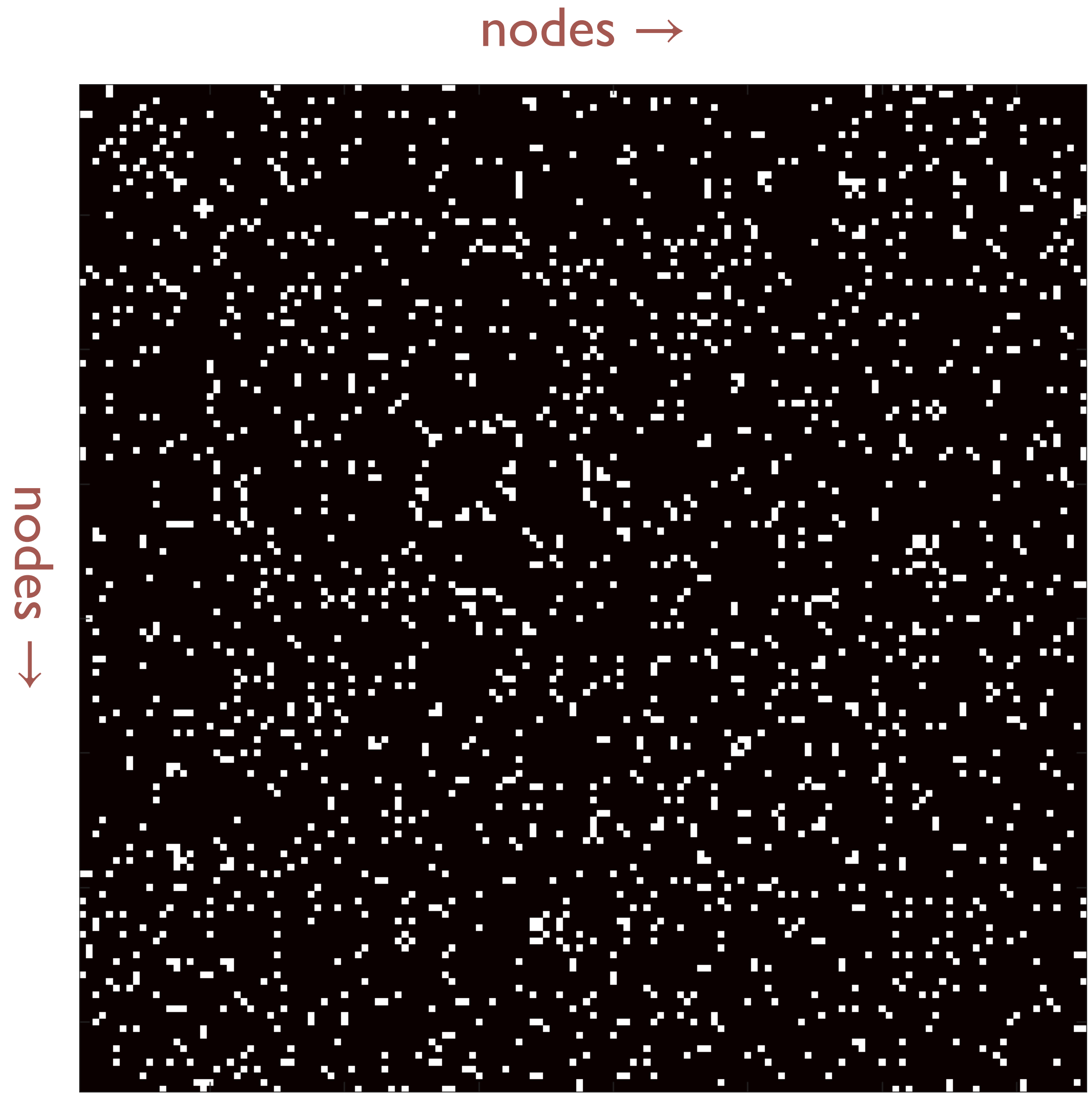


In an directed network, a strongly connected component is one where exists a path between all constituent nodes.

A weakly connected component is a connected component that exists if one were to ignore the directed nature of the edges.

The in-component of a node in a directed network is the set that can reach it, and its out-component is the set that can be reached from it.

IDENTIFYING CONNECTED COMPONENTS

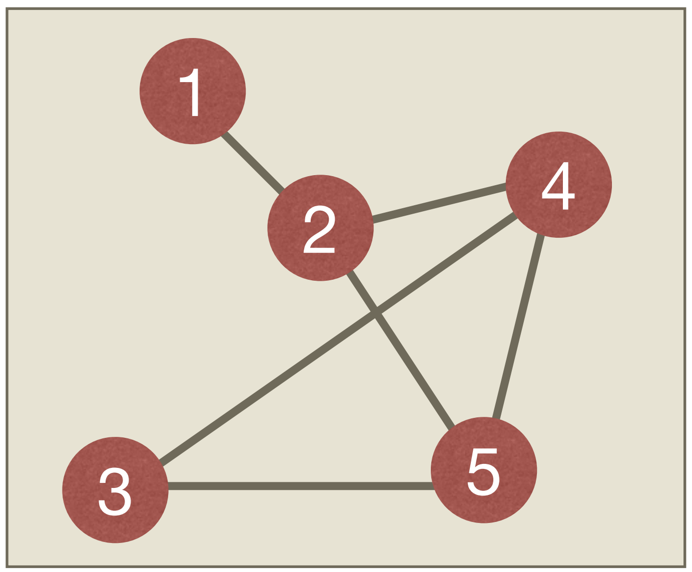


The image on the right displays the adjacency matrix of a large undirected network.

White squares represent connections between nodes, while black represents the absence of a link.

Can you guess the number of connected components of this network? 😳

FUNDAMENTAL CONCEPTS: THE GRAPH LAPLACIAN



Adjacency matrix

	1	2	3	4	5
1	0	1	0	0	0
2	1	0	0	1	1
3	0	0	0	1	1
4	0	1	1	0	1
5	0	1	1	1	0

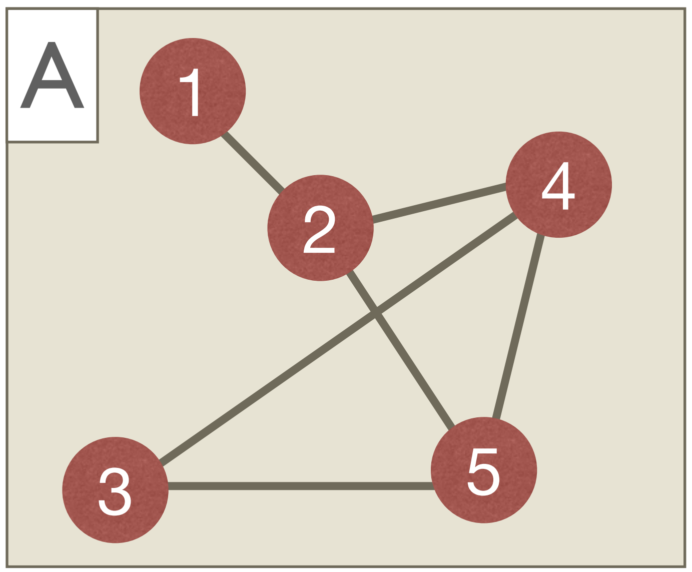
graph Laplacian

	1	2	3	4	5
1	1	-1	0	0	0
2	-1	3	0	-1	-1
3	0	0	2	-1	-1
4	0	-1	-1	3	-1
5	0	-1	-1	-1	3

For the case of undirected networks with no self-edges, one can define the graph Laplacian $\mathbf{L}$ as follows: $L_{ij} = k_i \delta_{ij} - A_{ij}$ or $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where the degree matrix $D_{ij} = k_i \delta_{ij}$ contains the degree along the diagonal and $0s$ elsewhere.

If the network is weighted, the definition is as follows: $L_{ij} = \sum_j A_{ij} \delta_{ij} - A_{ij}$

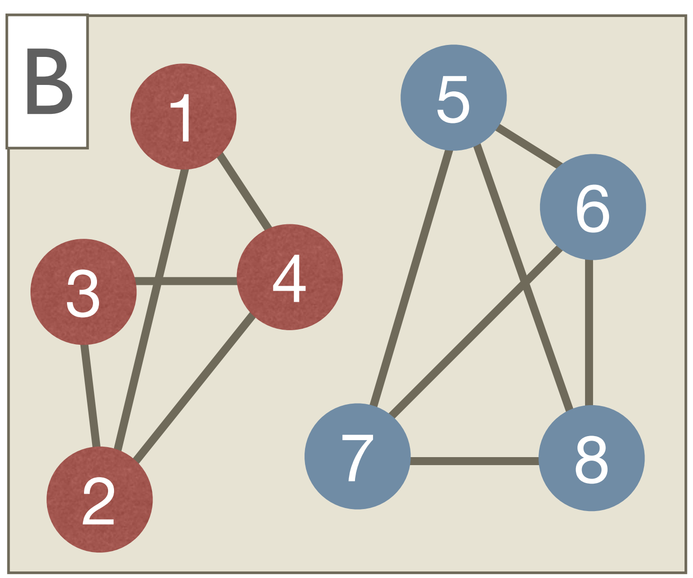
FUNDAMENTAL CONCEPTS: THE GRAPH LAPLACIAN



L_A

	1	2	3	4	5
1	1	-1	0	0	0
2	-1	3	0	-1	-1
3	0	0	2	-1	-1
4	0	-1	-1	3	-1
5	0	-1	-1	-1	3

The number of zero eigenvalues of the laplacian indicate the number of connected components of the network.



L_B

	1	2	3	4	5	6	7	8
1	2	-1	0	-1	0	0	0	0
2	-1	3	-1	-1	0	0	0	0
3	0	-1	2	-1	0	0	0	0
4	-1	-1	-1	3	0	0	0	0
5	0	0	0	0	3	-1	-1	-1
6	0	0	0	0	-1	3	-1	-1
7	0	0	0	0	-1	-1	3	-1
8	0	0	0	0	-1	-1	-1	3

$$\lambda_A = eig(L_A) =$$
$$\{0, 0.83, 2.69, 4, 4.48\}$$

$$\lambda_B = eig(L_B) =$$
$$\{0, 0, 2, 4, 4, 4, 4, 4\}$$

FUNDAMENTAL CONCEPTS: THE GRAPH LAPLACIAN

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 3 & 0 & -1 & -1 \\ 0 & 0 & 2 & -1 & -1 \\ 0 & -1 & -1 & 3 & -1 \\ 0 & -1 & -1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 3 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

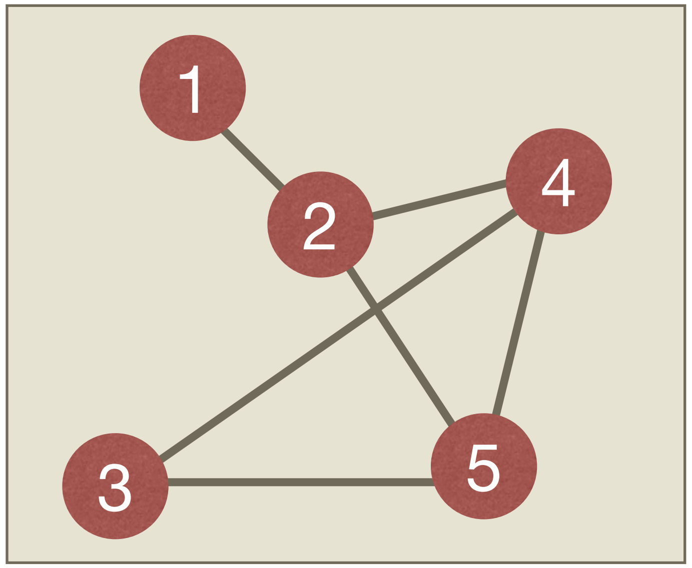
If $\mathbf{v}$ is an eigenvector of the Laplacian and λ is its associated eigenvalue, then

$$\mathbf{Lv} = \lambda\mathbf{v}$$

A Laplacian with a single component, has an eigenvector $\mathbf{v} = [1,1,\ldots]^T$ with $\lambda = 0$.

A Laplacian with a two components, has $\mathbf{v} = [1,1,\ldots,0,0,\ldots]^T$ and $\mathbf{v} = [0,0,\ldots,1,1,\ldots]^T$ both with $\lambda = 0$.

FUNDAMENTAL CONCEPTS: BETWEENNESS CENTRALITY



The table shows the list of all possible shortest paths between every pair of nodes in the above network.

	SHORTEST PATHS
1-2	{1,2}
1-3	{1,2,5,3}, {1,2,4,3}
1-4	{1,2,4}
1-5	{1,2,5}
2-3	{2,4,3}, {2,5,3}
2-4	{2,4}
2-5	{2,5}
3-4	{3,4}
3-5	{3,5}
4-5	{4,5}

The betweenness centrality of a node measures the extent to which it controls information flow, or acts as a bottleneck.

To calculate a node's betweenness centrality, we count the fraction of times it appears in the shortest paths between other nodes.

FUNDAMENTAL CONCEPTS: BETWEENNESS CENTRALITY

	SHORTEST PATHS
1-2	{1,2}
1-3	{1,2,5,3}, {1,2,4,3}
1-4	{1,2,4}
1-5	{1,2,5}
2-3	{2,4,3}, {2,5,3}
2-4	{2,4}
2-5	{2,5}
3-4	{3,4}
3-5	{3,5}
4-5	{4,5}

	OCCURRENCES	C_B
1	0	0
2	2/2 + 1 + 1	3
3	0	0
4	1/2 + 1/2	1
5	1/2 + 1/2	1

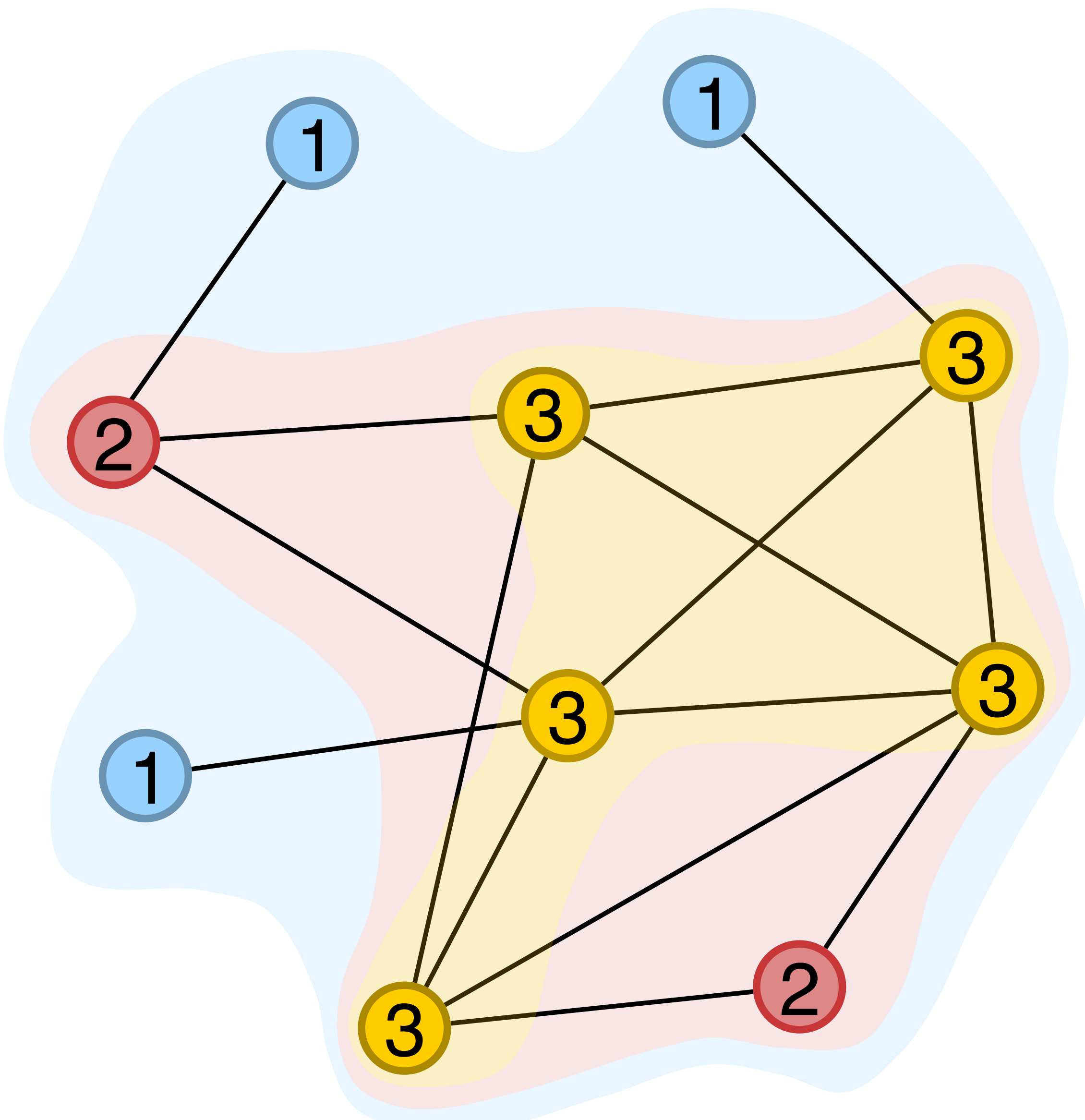
If σ_{st} is the no. of shortest paths from s to t, and $\sigma_{st}(v)$ is the number of these containing node v, then:

$$C_B(v) = \sum_{s \neq t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

is the betweenness centrality.

MESOSCALE STRUCTURAL FEATURES

CORE-PERIPHERY ORGANIZATION



The *k*-core is the defined* as the set of nodes of the network within which each node has *k* links with each other.

Stephen Seidman

The nodes of the highest *k*-core are referred to as core nodes and others are peripheral nodes.

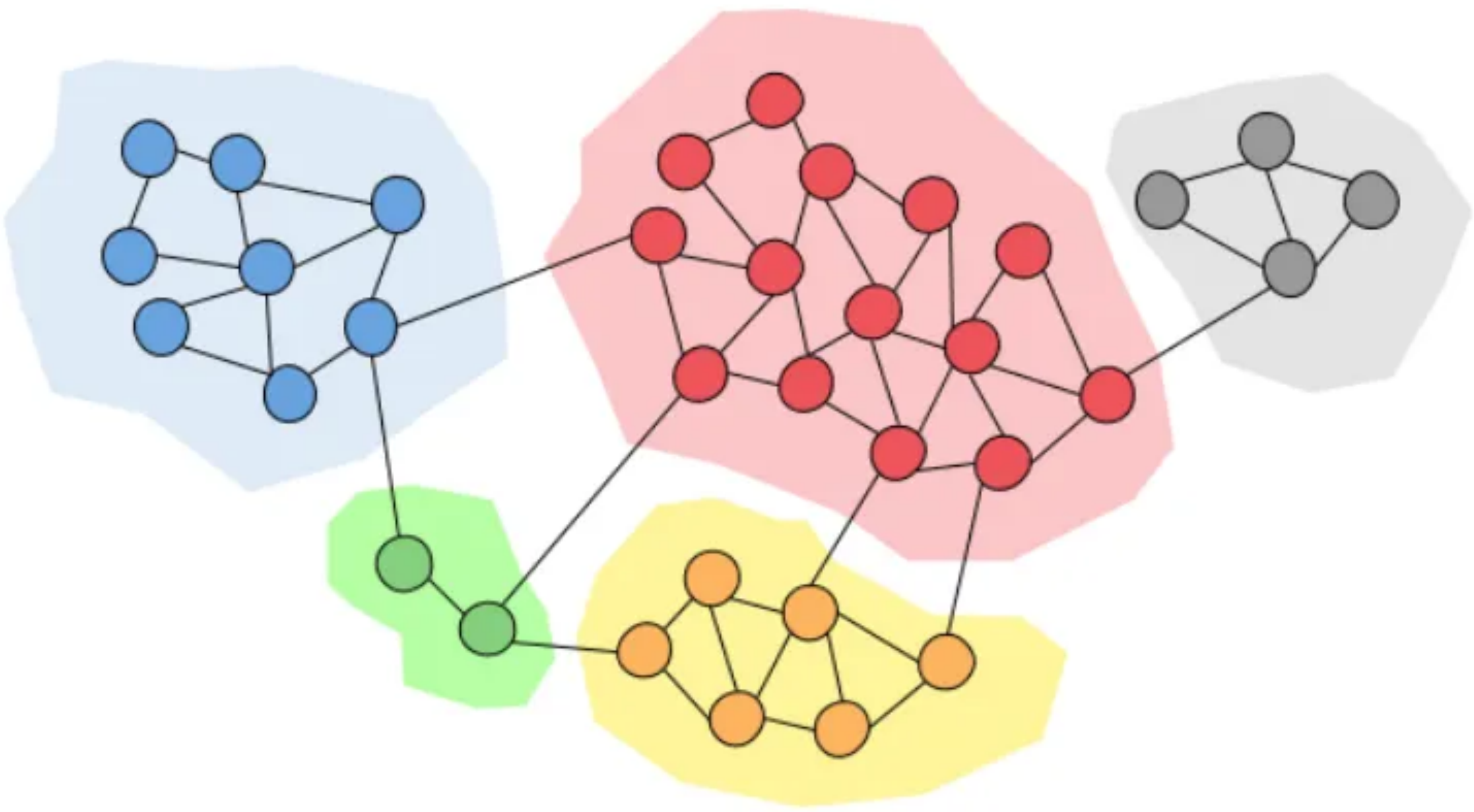
A core need not be a single connected component.

* SB Seidman, *Soc. Networks* **5**(3), 269 (1983)

MESOSCALE STRUCTURAL FEATURES

MODULARITY

A network is said to have a modular structure if there exist groups (or "communities") of nodes that have a higher density of connections than that between groups.



There is typically no single modular structure for a given network. Upon positing a potential "decomposition", one then measures the extent to which such an organisation is modular. Obtaining the globally optimal decomposition requires techniques such as simulated annealing.

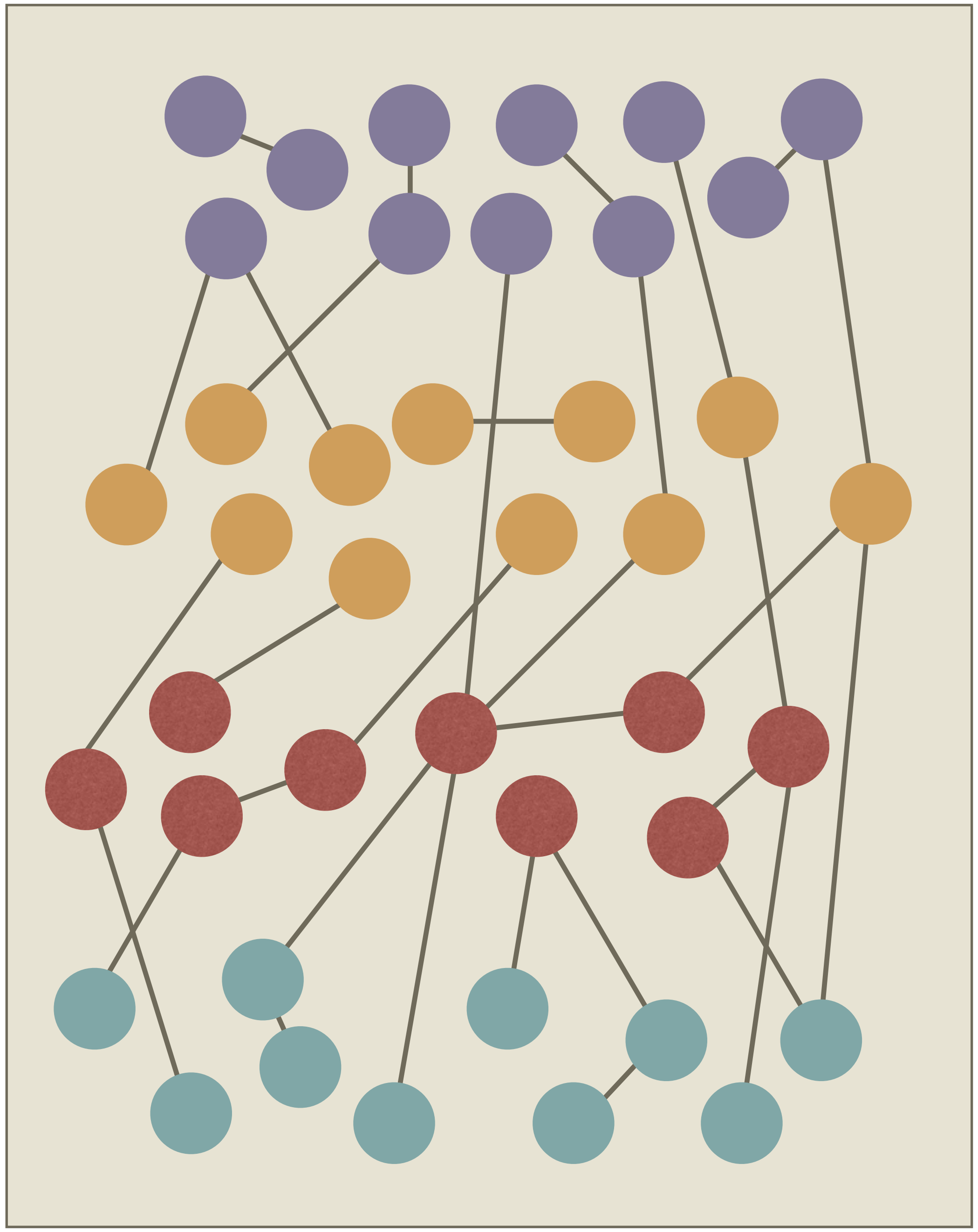
image: https://towardsdatascience.com/community-detection-algorithms-9bd8951e7dae

MESOSCALE STRUCTURAL FEATURES

HIERARCHY

A network is said to have a hierarchical structure if there exists "layers" of nodes, such that the density of connections *between* consecutive layers is higher than that *within* layers, or between non-consecutive layers.

Similar to the process of modularity detection one can use heuristic algorithms* to determine the hierarchical levels of a given network.

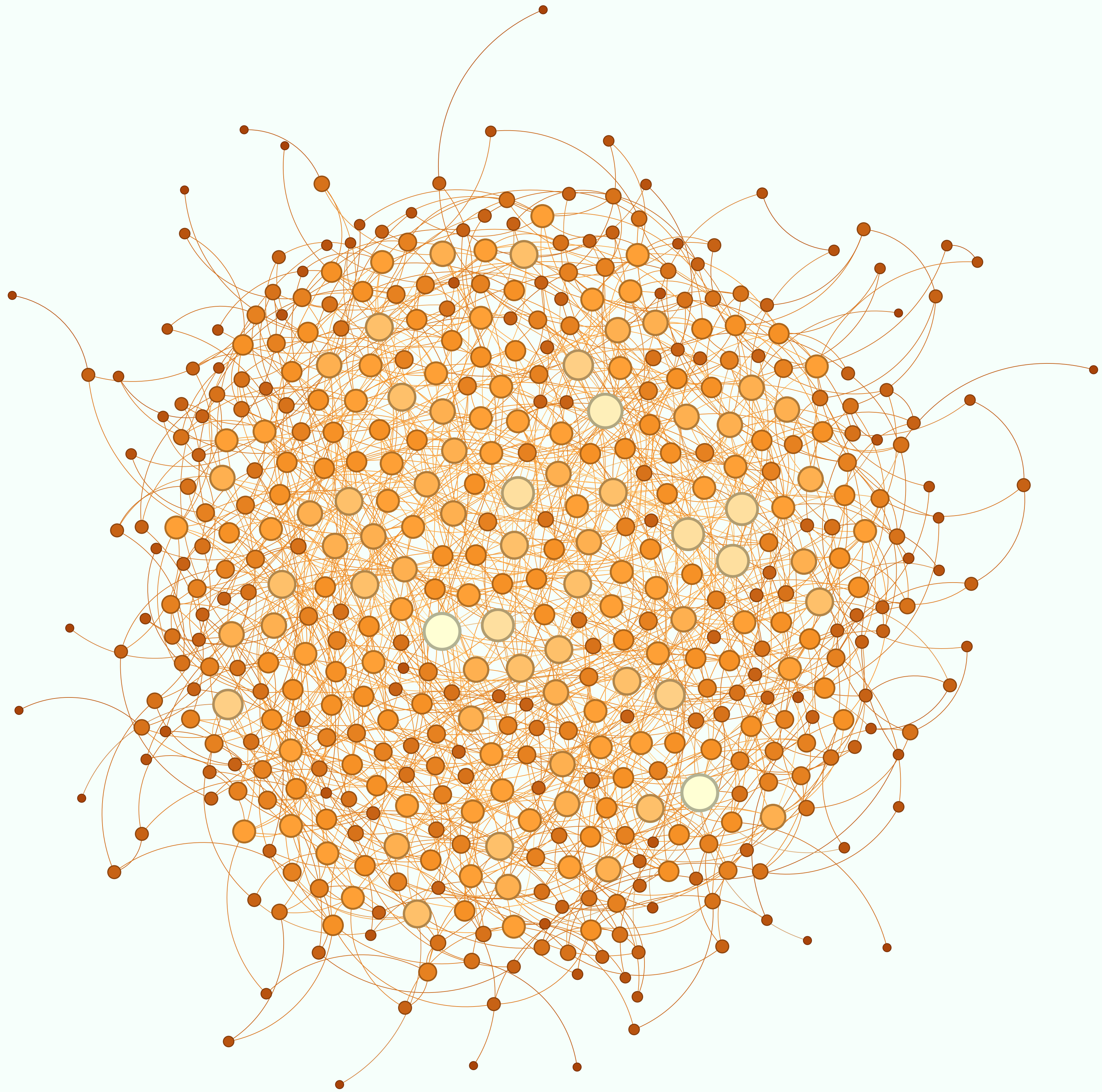


* Pathak, A., Menon, S. N. and Sinha, S., *PNAS* **121**, e2314291121 (2024).

WHY STUDY RANDOM GRAPHS?

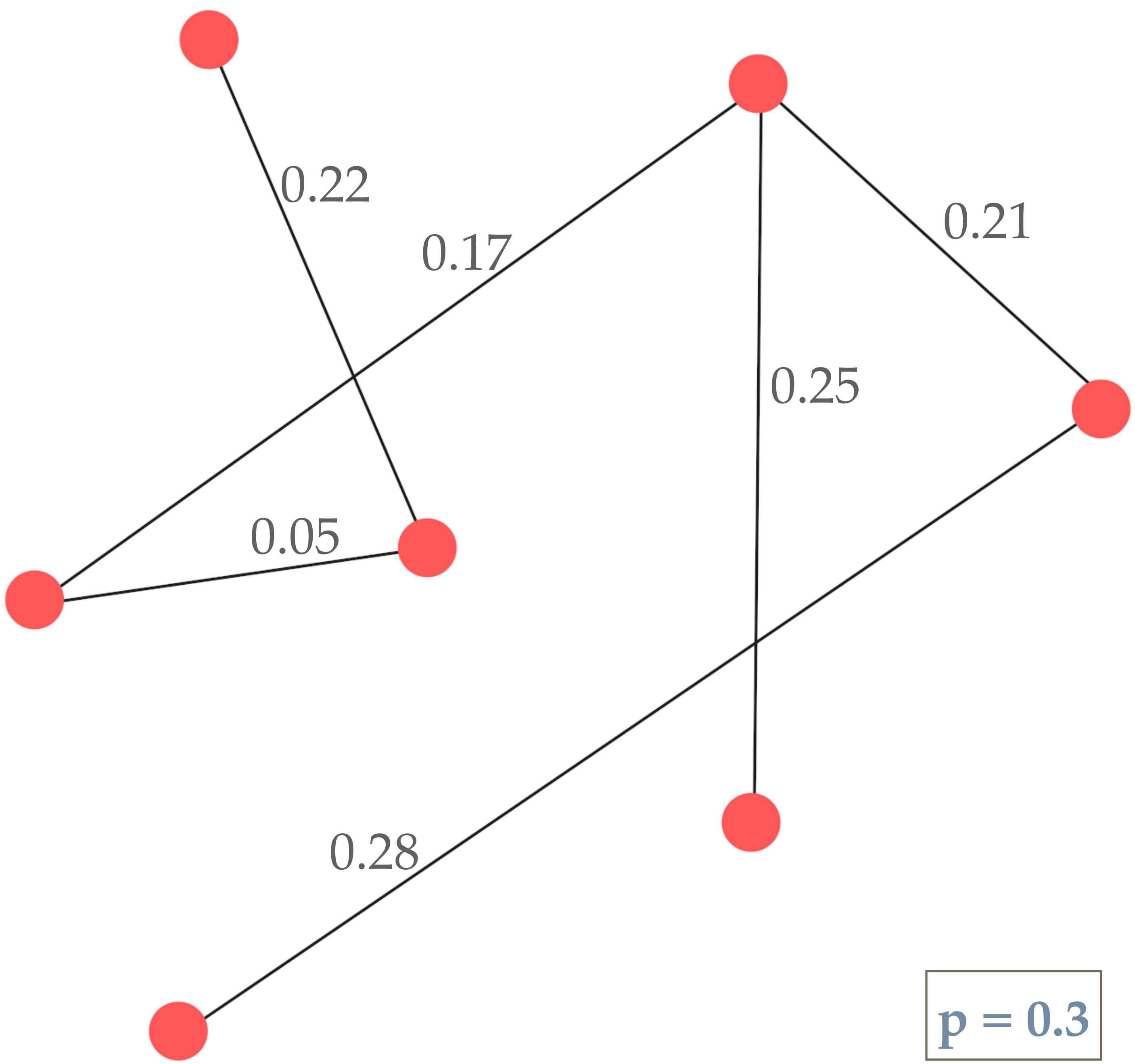
Random graphs provide *null models* against which we can test certain hypotheses.

Once the key attributes of the network responsible for certain properties have been identified it is then possible to generate numerous *surrogate* networks that can be used in place of the empirical network for further study.



GENERATIVE NULL MODELS

CREATING A RANDOM GRAPH



In 1959 a model was proposed for generating a random graph comprising n nodes.

Assign a random number between [0,1] to every potential link.

Keep only those links whose values are less than a specified threshold $p \in [0,1]$.

GENERATIVE NULL MODELS

ERDŐS-RÉNYI RANDOM GRAPHS

- This is known as the $G(n, p)$ model, and the resulting graphs are commonly referred to as Erdős-Rényi (ER) random graphs

- For certain choices of (n, p), the resulting graph may have multiple connected components.

- The degree distribution (the probability $p(k)$ that a randomly selected node in the network has degree k) is a binomial distribution.

Pál Erdős Alfréd Rényi Edgar Gilbert

GENERATIVE NULL MODELS

ERDŐS-RÉNYI RANDOM GRAPHS

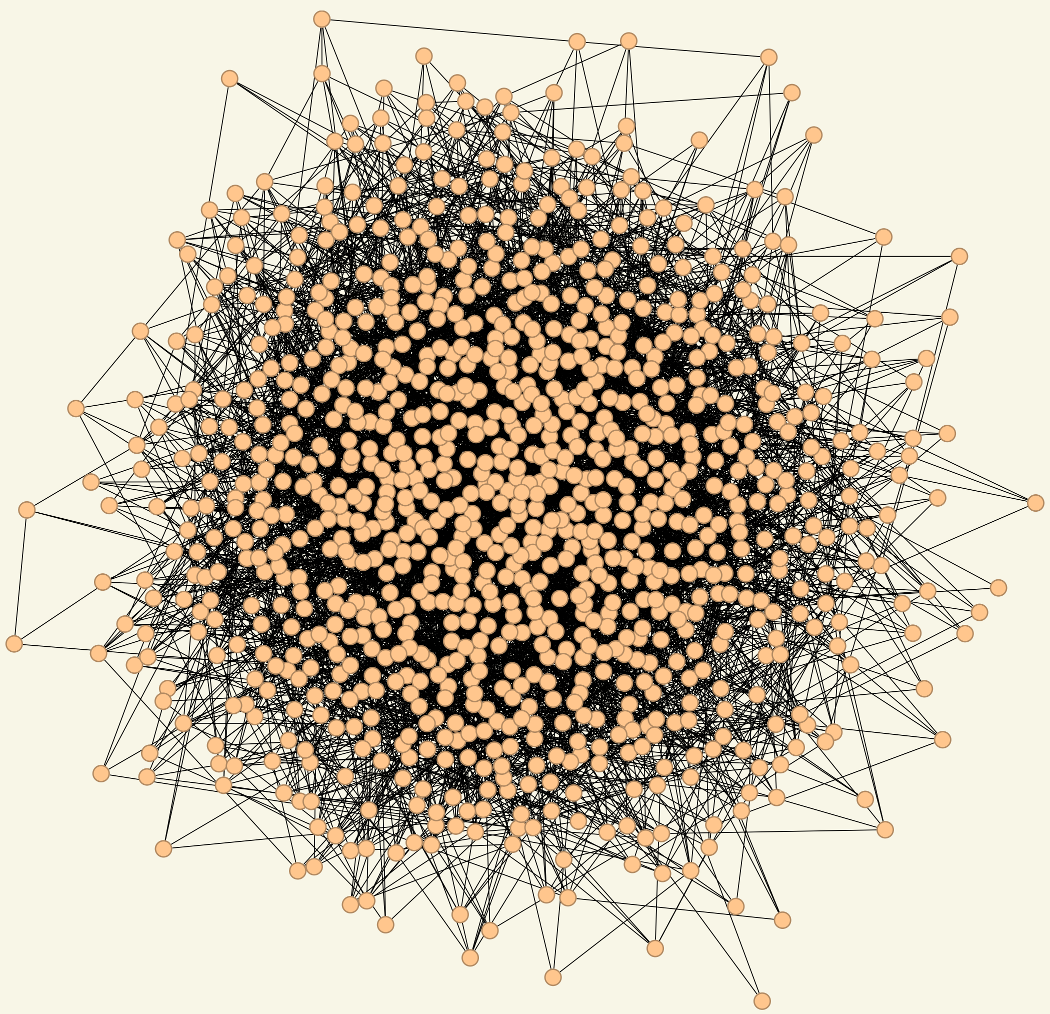
- In a network of n nodes, a node has independent probabilities of connecting to each of the other $n-1$ nodes.

- There are $\binom{n-1}{k} = \frac{(n-1)!}{k!(n-1-k)!}$ ways of choosing k out of $n-1$ nodes.

- The probability of connecting to k nodes and not to the other $n-1-k$ nodes is hence: $p^k(1-p)^{n-1-k}$.

- Thus, given that, the probability that a node connects to k nodes follows a binomial distribution:

$$p(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

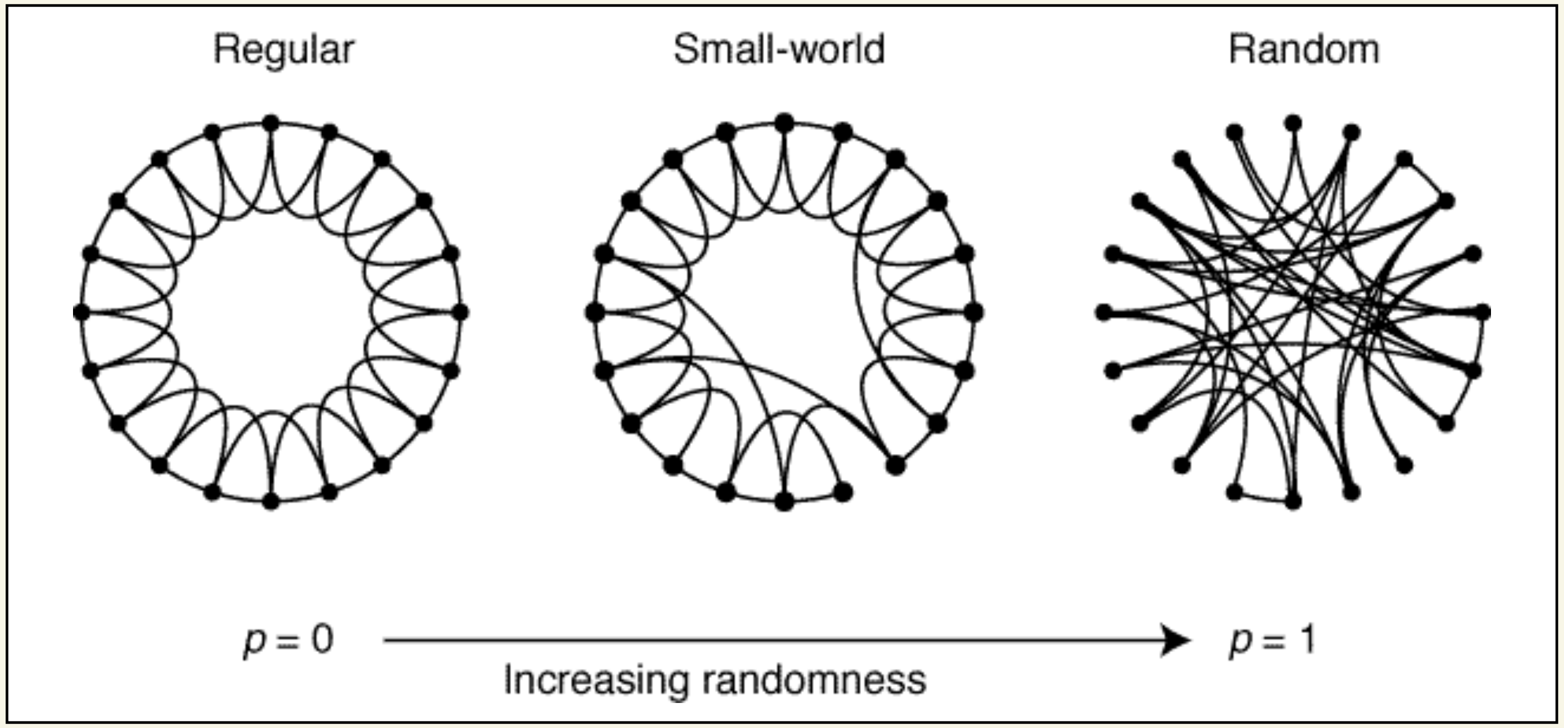


- In the limit $n \to \infty$ and for $np = \text{constant}$, we see that this expression reduces to a Poisson distribution:

$$p(k) = \frac{(np)^k e^{-np}}{k!} = \frac{\lambda^k e^{-k}}{k!}, \text{ where } \lambda = np$$

GENERATIVE NULL MODELS

WATTS-STROGATZ NETWORKS



Duncan J. Watts   Steve Strogatz

Watts and Strogatz (1998) suggested a procedure for obtaining graphs with properties of both regular and random graphs:

- Start with a regular graph where each node has K neighbours.
- Cycle through each node, and consider the $K/2$ rightward links.
- Randomly rewire each of these links with probability p, avoiding self-loops and duplicate links.

image: D Watts & S Strogatz, *Nature* **393**, 440-442 (1998).

GENERATIVE NULL MODELS

WATTS-STROGATZ NETWORKS

For an intermediate p the resulting graphs have

- low average path length L
- high clustering coefficient C

These are referred to as "**small-world**" networks.

For the *C. elegans* connectome, when comparing the empirical data to a random surrogate network of same $n(=282)$ and $\langle k \rangle (=14)$, they found that $L_{emp} > L_{rand}$ and $C_{emp} > C_{rand}$.

LEMP	LRAND	CEMP	CRAND
2.65	2.25	0.28	0.05

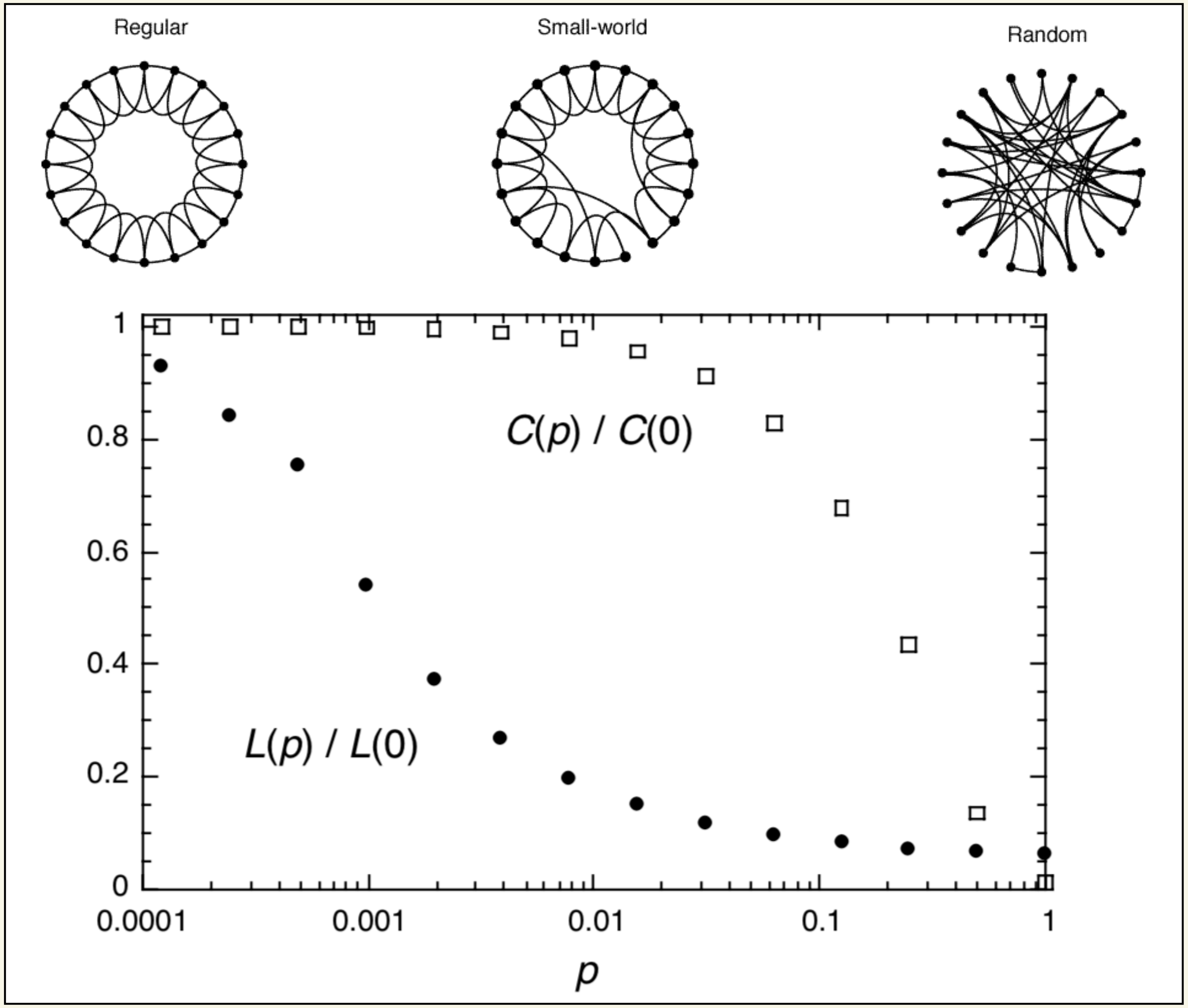
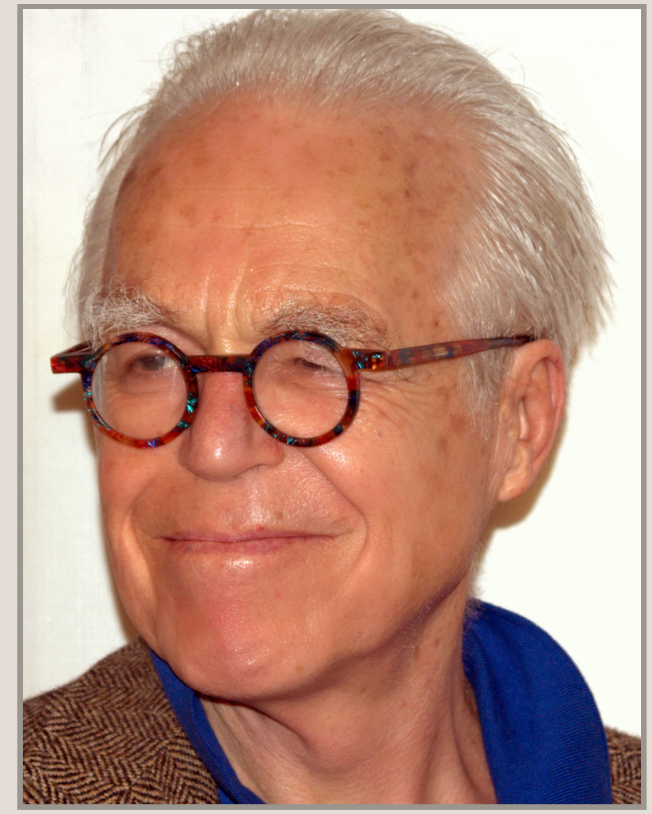


image: D Watts & S Strogatz, *Nature* **393**, 440-442 (1998).



"The worker knows the manager in the shop, who knows Ford; Ford is on friendly terms with the general director of Hearst Publications, who last year became good friends with Árpád Pásztor, someone I not only know, but is to the best of my knowledge a good friend of mine - so I could easily ask him to send a telegram via the general director telling Ford that he should talk to the manager and have the worker in the shop quickly hammer together a car for me, as I happen to need one."

Frigyes Karinthy, "*Láncszemek (Chains)*" (1929).

"Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice. Fill in the names."

John Guare, "*Six Degrees of Separation*" (1990).

Six degrees of Kevin Bacon

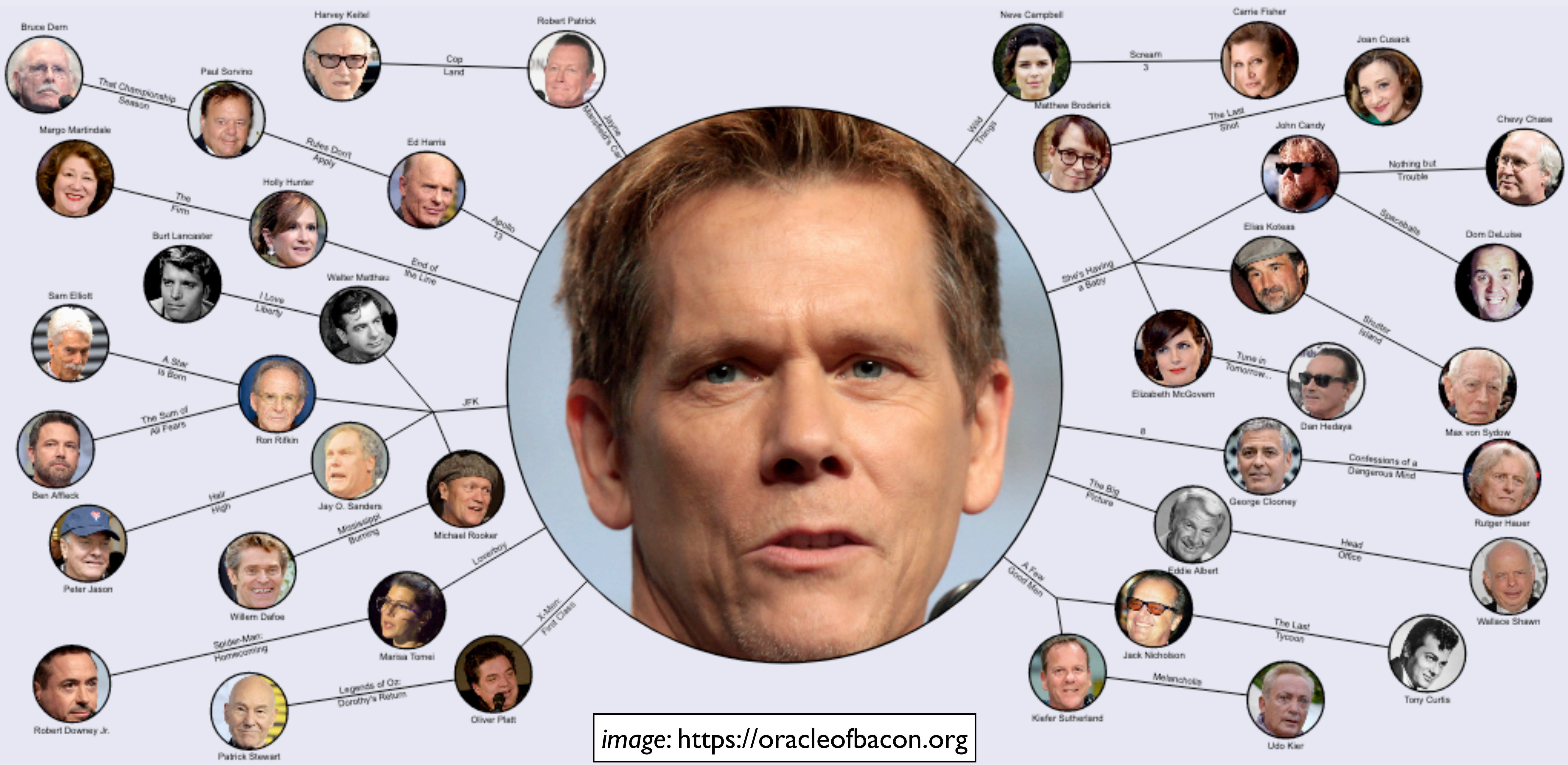
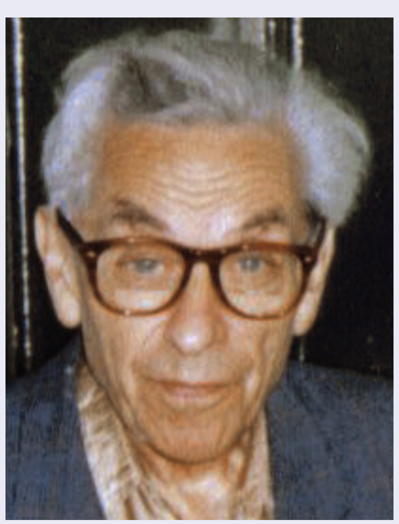
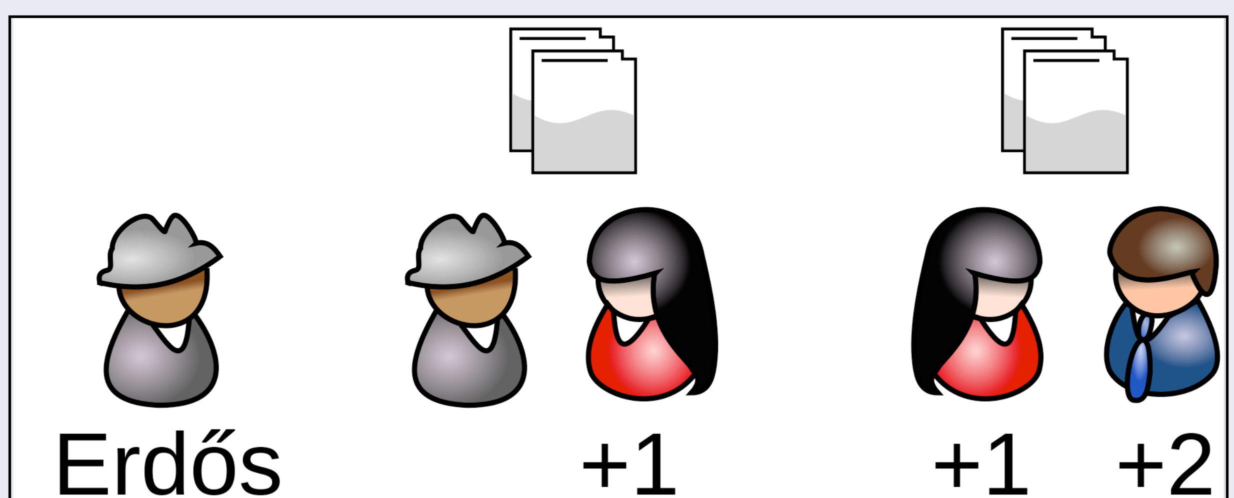


image: https://oracleofbacon.org





Erdős number

Collaborative "distance" between an author and Pál Erdős.

GENERATIVE NULL MODELS

BARABÁSI-ALBERT NETWORKS

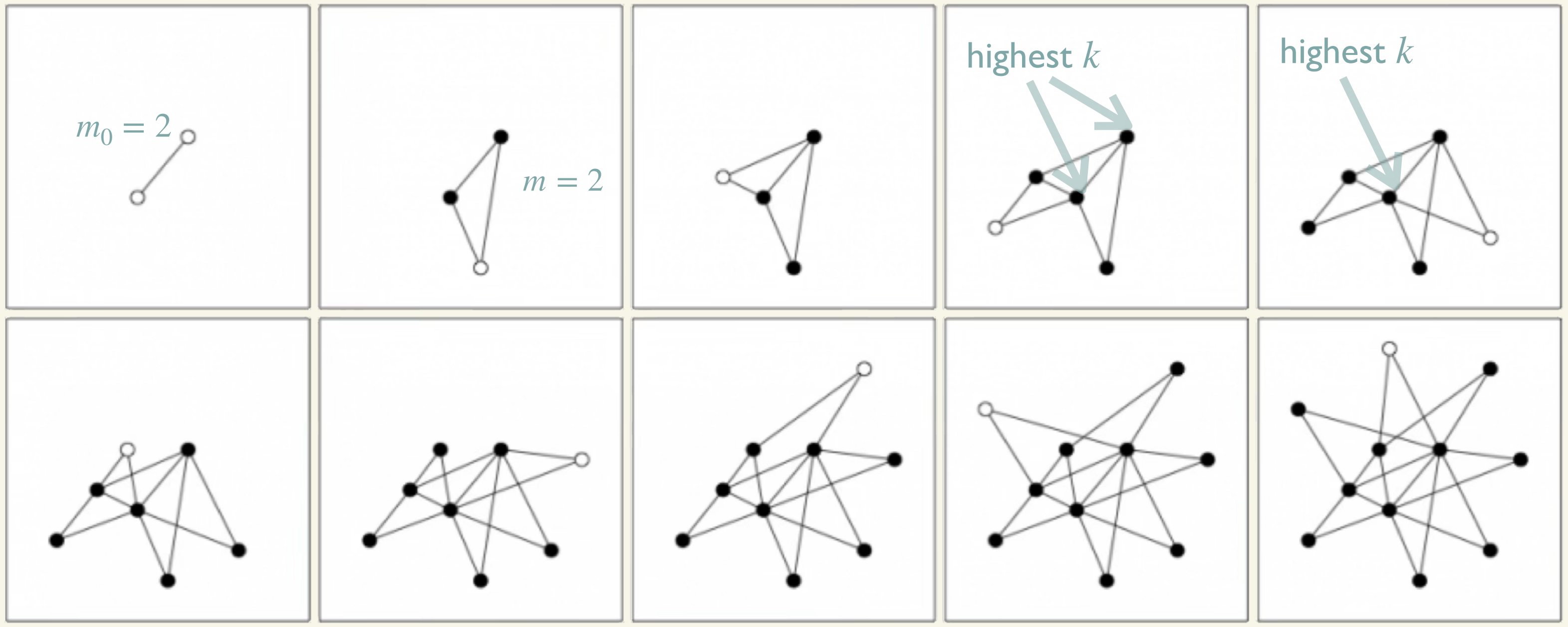
The Barabási-Albert (BA) model employs a mechanism of network growth and preferential attachment, and leads to **scale-free** random graphs with power law degree distributions.

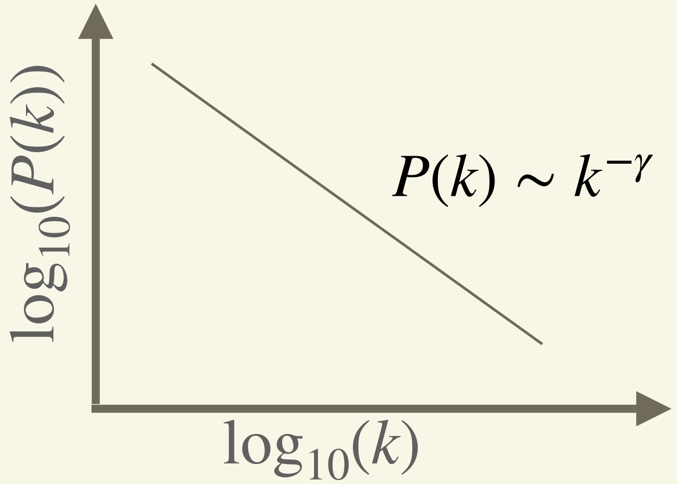
- Starting with m_0 nodes, at each step add a new node and connect it to $m(\leq m_0)$ existing nodes.
- The probability of connecting to an existing node i is $p(k_i) = k_i / \Sigma_j k_j$

Albert-László Barabási

Réka Albert

Derek J. de Solla Price





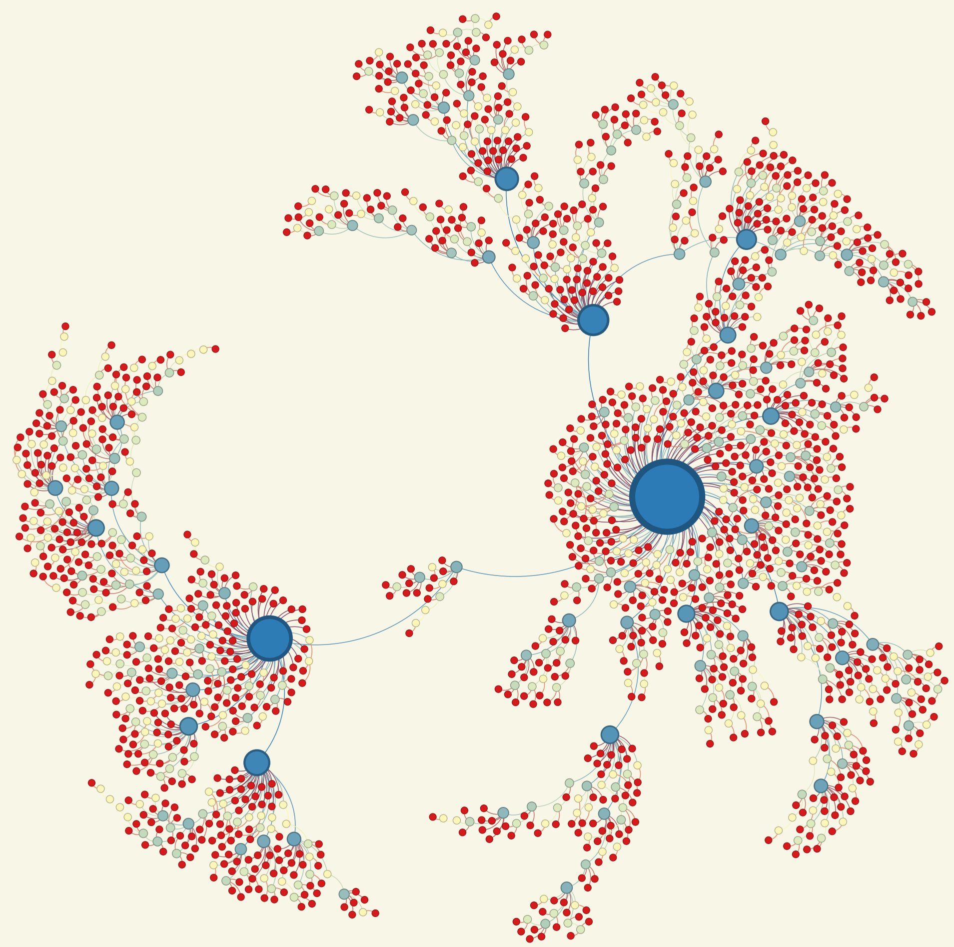
power law distribution

image: AL Barabási, "Network Science" (http://networksciencebook.com/)

GENERATIVE NULL MODELS

BARABÁSI-ALBERT NETWORKS

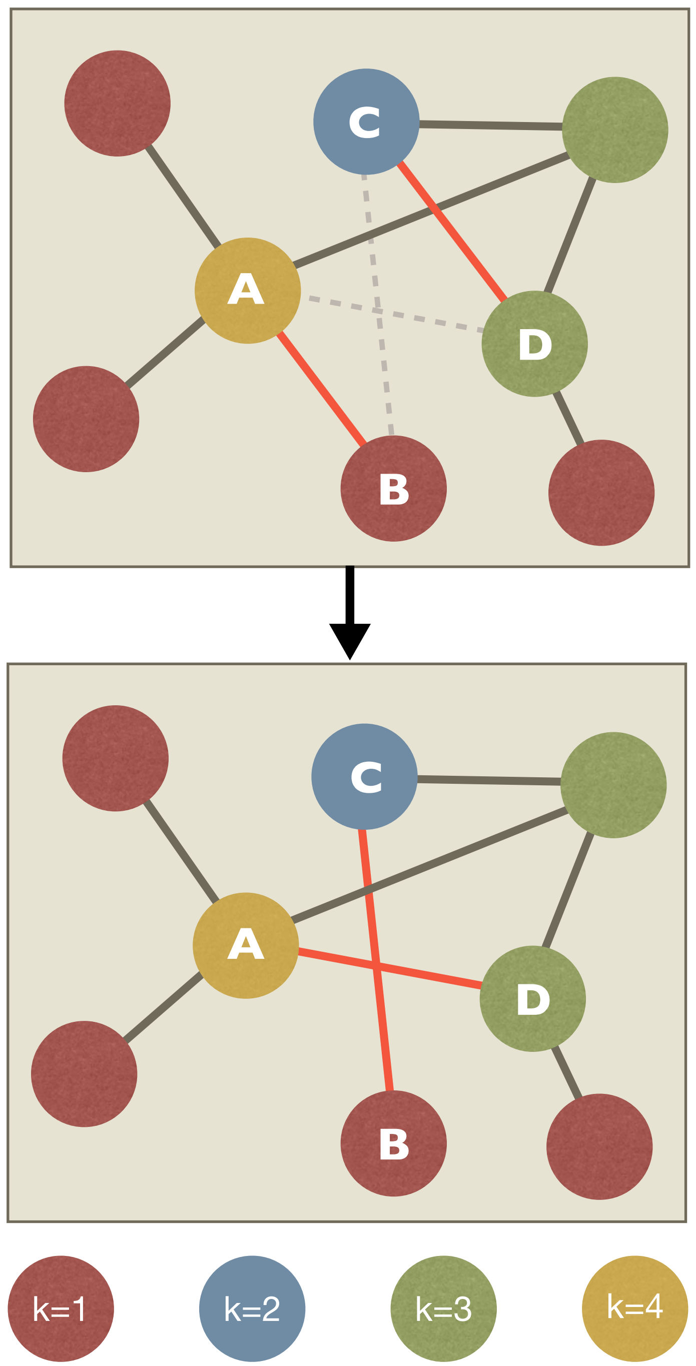
- After $t(\gg 1)$ time steps, the total number of links in the network is $\approx 2mt$.
- The rate at which the degree of node i changes is:
$$\frac{dk_i}{dt} = mp(k_i) = \frac{mk_i}{\Sigma_j k_j} \approx \frac{k_i}{2t}$$
- Solving this equation, we get
$$\ln(k_i) = \frac{1}{2}\ln(t) + C \implies k_i(t) = At^{1/2}$$
- When node i first joins the network at t_i is has degree $k_i(t_i) = m$, and so $A = m/t_i^{1/2}$. Hence,
$$k_i(t) = m(t/t_i)^{1/2}$$



- For a node to reach a degree k at time t, it would need to be added at time $t_i = t(m/k)^2$.
- The number of nodes with $k_i(t) \geq k$ is hence $N_k = t_i = t(m/k)^2$ as one node is added each unit time.
- Hence, the cumulative distribution function is $P(k_i \geq k) = N_k/t$, and by definition the probability distribution is just: $p(k) = \frac{d}{dk}\left(P(k_i \leq k)\right) = \frac{d}{dk}\left(1 - P(k_i \geq k)\right) = 2m^2k^{-3}$, i.e. a power law with degree exponent 3.

NULL MODELS FROM REWIRING

MASLOV–SNEPPEN ALGORITHM



One of the most widely used network null models, especially in the context of brain network analysis, is the degree-preserved randomized network, which is typically obtained using the Maslov–Sneppen rewiring algorithm*:

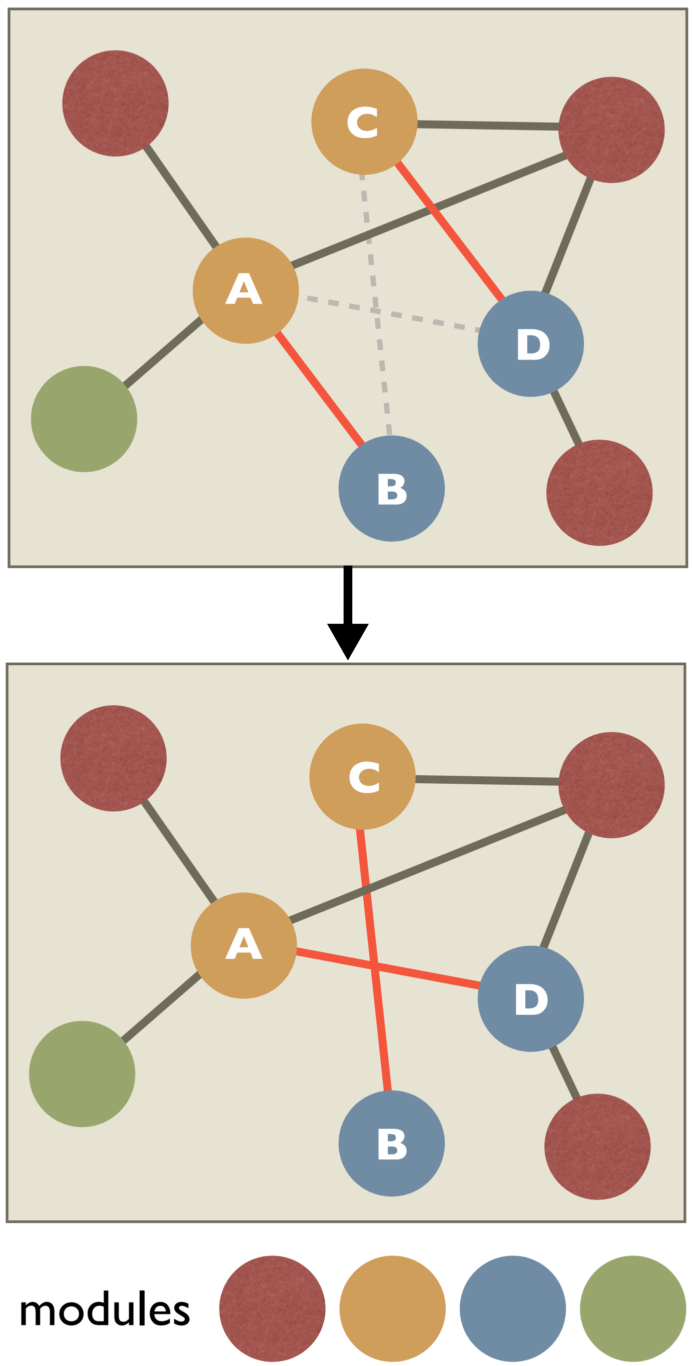
- At each step we select at random two edges AB and CD.
- If A, B, C and D are all *distinct* nodes, and neither of the links AD or BC exist, we create them and delete the links AB and CD. Otherwise we select two new random edges.
- We perform this procedure a large number of times (many more than the total no. of links).

Every node in the resulting network has the same degree sequence as before the rewiring procedure.

* S Maslov & K Sneppen, *Science* **296** 296 (2002).

NULL MODELS FROM REWIRING

MODULE-PRESERVED NETWORKS



While the Maslov-Sneppen algorithm yields random degree-preserved (RD) networks, the procedure can be modified to ensure that the number of links within each module are also preserved. To obtain these random degree-preserved module-preserved (RDM) networks:

- At each step we select at random two edges AB and CD.
- If A, B, C and D are all *distinct* nodes, where A & C are in the same module, B & D are in the same module, and neither of the links AD or BC exist, we create them and delete the links AB and CD. Otherwise we select two new random edges.

Note that does not necessarily guarantee that the module membership will remain intact, simply that the module densities are unchanged.

NULL MODELS FROM REWIRING

RANDOMIZED ENSEMBLES

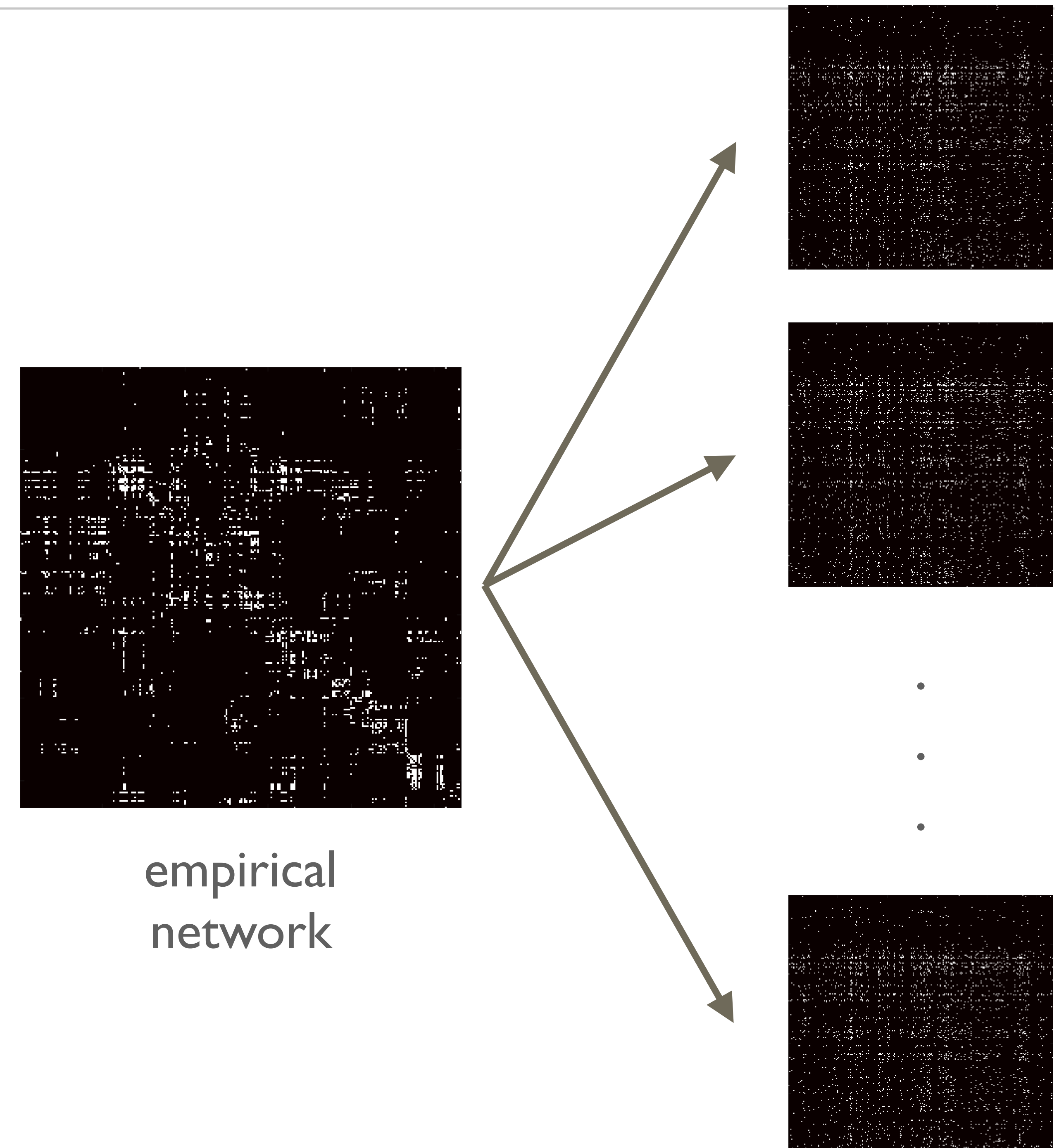
Using this procedure one can create ensembles of randomised networks, all of which have exactly the same preserved properties as the original empirical network (degree, module density, etc).

This allows us to perform our analyses on a large number of networks that are surrogates for the original one.

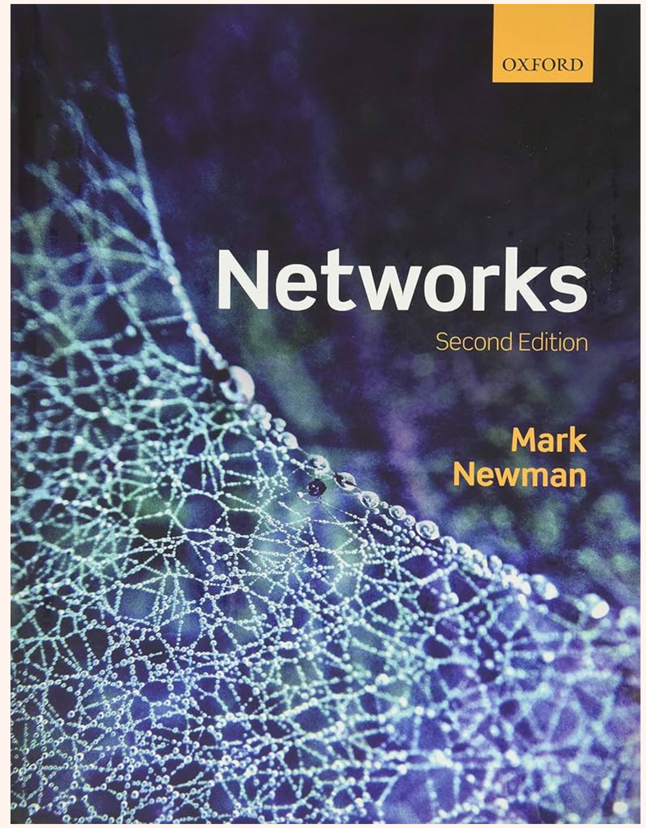
If we have some measure ϕ_{emp} on the empirical network and ϕ^i_{rnd} on the i^{th} random network, then one can measure its z-score:

$$z = \frac{\phi_{emp} - \langle \phi^i_{rnd} \rangle}{\sigma^2(\phi^i_{rnd})}$$

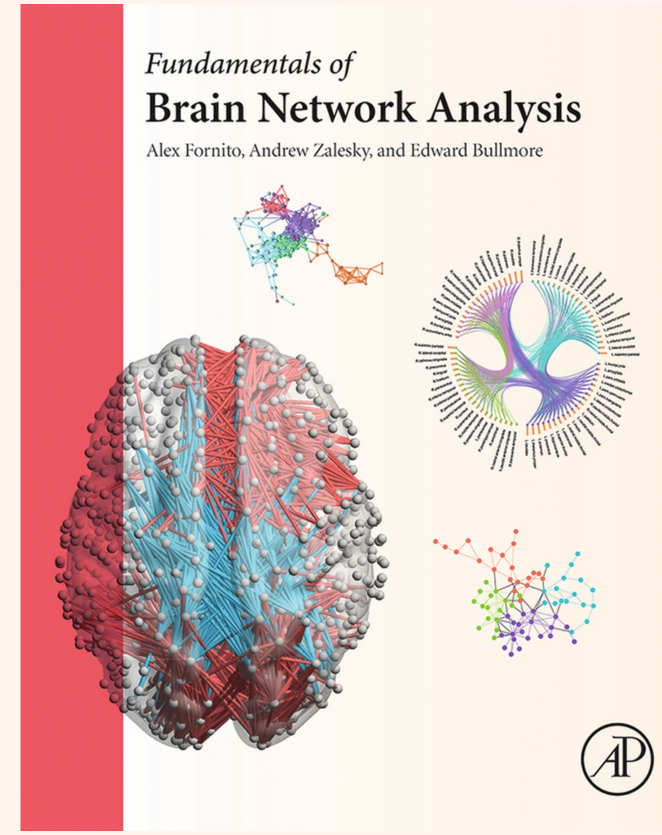
This tells us the extent to which the measured value is more or less than what would be expected from random (with certain properties preserved).



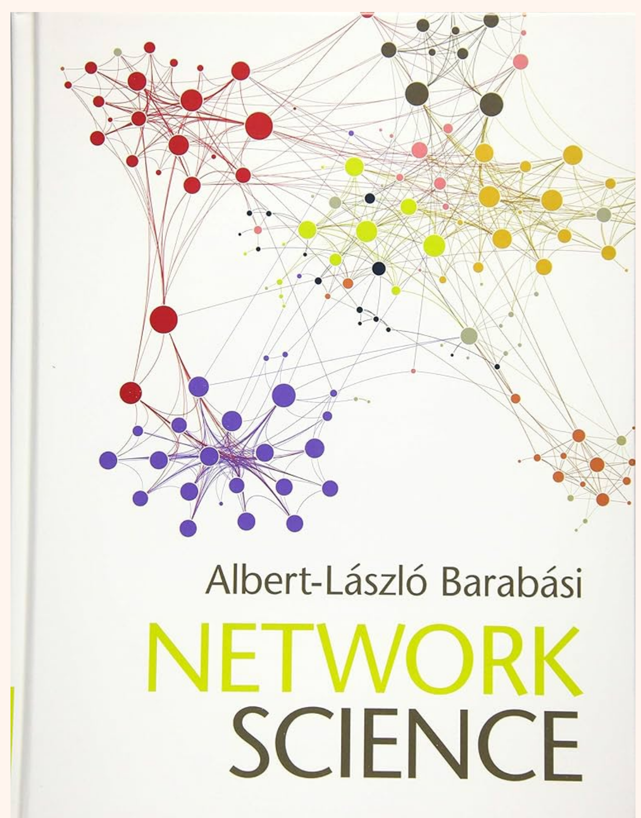
FURTHER READING



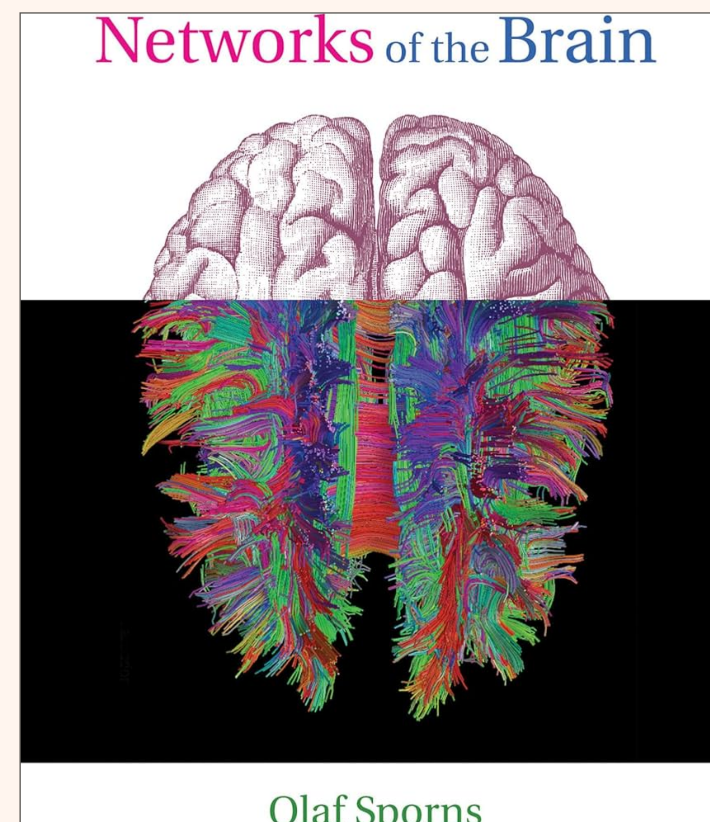
M Newman, *Networks*
(Oxford University Press, 2018).



A Fornito, A Zalesky, & E Bullmore, *Fundamentals of Brain Network Analysis* (Academic Press, 2016).



A-L Barabási, *Network Science*
(Cambridge University Press, 2018).



O Sporna, *Networks of the Brain*
(MIT Press, 2010).