Bangalore School on Statistical Physics XIV, Sept 2023

Statistical Mechanics of Complex Networks

Lecture I: Introduction (or "What ?" & "Why ?")

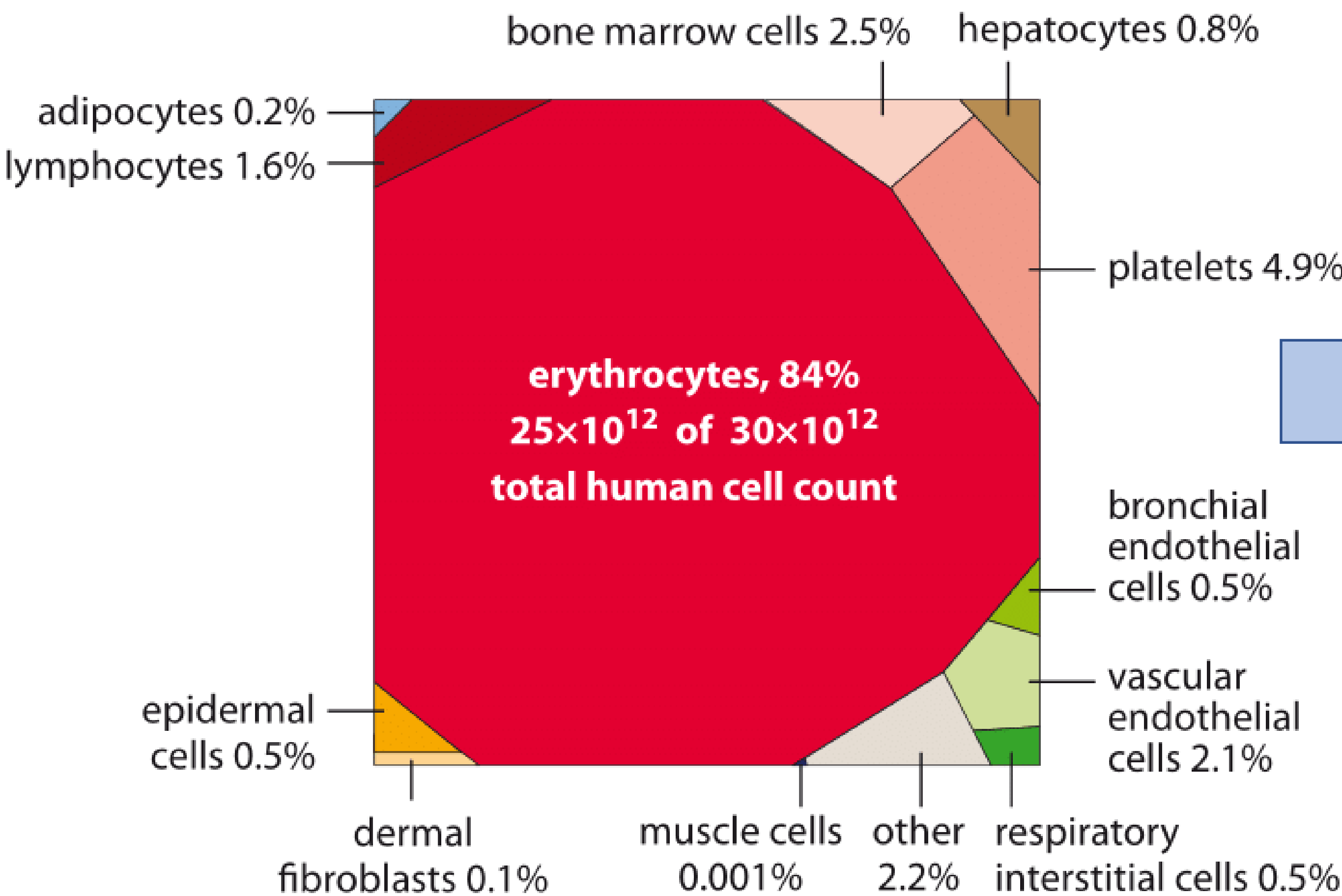
Sitabhra Sinha

The Institute of Mathematical Sciences, Chennai

How do you build a complex system?

Example: a human being

Estimate of the number of different cell types in an adult human



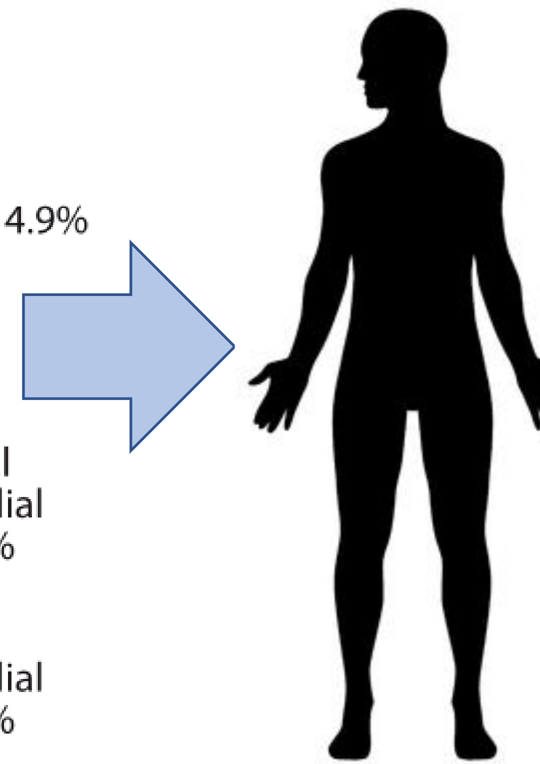


Image: Hong Li, 123RF

Image: book.bionumbers.org

Recipe for a human being

Ingredients

- 35 kg Oxygen
- 6.4 kg Hydrogen
- 17.5 kg Carbon
- 1.5 kg Nitrogen
- 1.0 kg Calcium
- 0.54 kg Phosphorus
- 110 g Sulfur
- 72 g Sodium
- 120 g Potassium
- 76 g Chlorine
- 17 g Magnesium
- 18 g Silicon
- 2.5 g Iron (Enough to make a nail!)
- 2.4 g Zinc
- 83 mg Copper
- 31 mg Iodine
- 12 mg Manganese
- 4.2 mg Fluorine
- 6.2 mg Chromium
- 5.4 mg Selenium
- 4.9 mg Molybdenum
- 1 mg Cobalt

Instructions

Collect the component parts

Mix and shake well ?

Image: Wikipedia

https://blogs.scientificamerican.com/but-not-simpler/excerpts-from-the-mad-scientiste28099s-handbook-the-human-recipe/

How do you build a complex system?

It's not enough to know the components

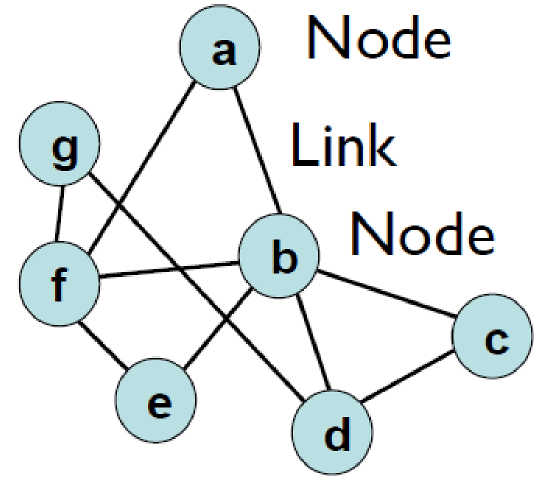
We also need to know how each component relate or interact with the others

How do we describe all these interactions ?

We use the language of Networks

What is a network ?

Network…



Components = Nodes or Vertices

Interactions = Links or Edges

System = Network or Graph

…and its adjacency matrix

Network structure is specified by

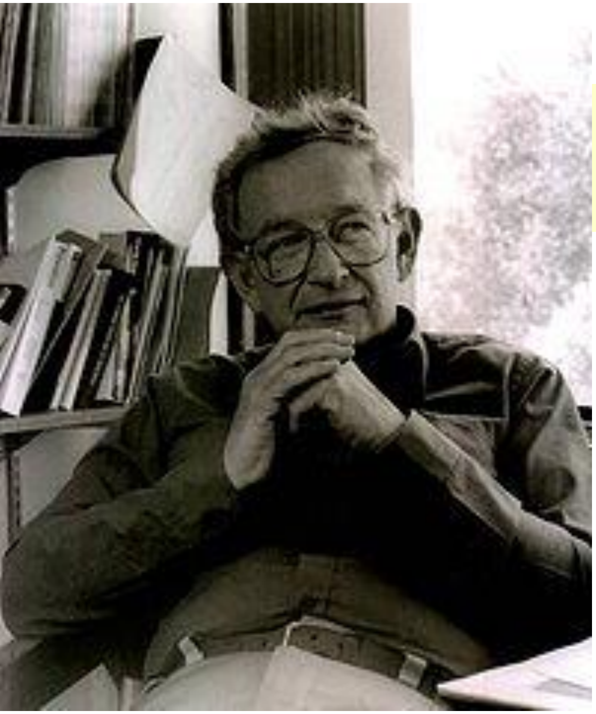
adjacency matrix A

A_{ij} = 1, if a link exists between i and j ($\neq$ i)

= 0, otherwise

a	b	c	d	e	f	g	
0	1	0	0	0	1	0	a
1	0	1	1	1	1	0	b
0	1	0	1	0	0	0	c
0	1	1	0	0	0	1	d
0	1	0	0	0	1	0	e
1	1	0	0	1	0	1	f
0	0	0	1	0	1	0	g

Why Networks ? More is Different



Philip W Anderson, *Science* (1972)

The elementary entities of science X obey the laws of science Y.

X	Y
solid state or many-body physics	elementary particle physics
chemistry	many-body physics
molecular biology	chemistry
cell biology	molecular biology
•	•
•	•
•	•
psychology	physiology
social sciences	psychology

But this hierarchy does not imply that science X is "just applied Y."

At each stage entirely new laws, concepts, and generalizations are necessary, requiring inspiration and creativity to just as great a degree as the previous one. Psychology is not applied biology, nor is biology applied chemistry.

Why Networks ?

- Emergence
- in a network of interacting components
- of qualitatively different behavior from that of the individual components.

E.g., component = neuron, system = brain

Interactions add a new layer of complexity!

The Importance of Networks

how interactions → complexity at systems-level

But wait! Isn't that …

Why statistical mechanics ?

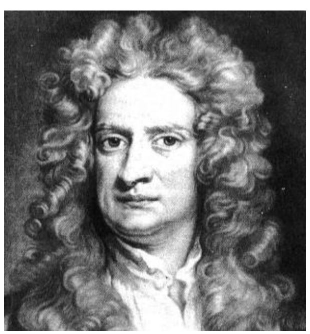
Classical Mechanics

N = 1 particle

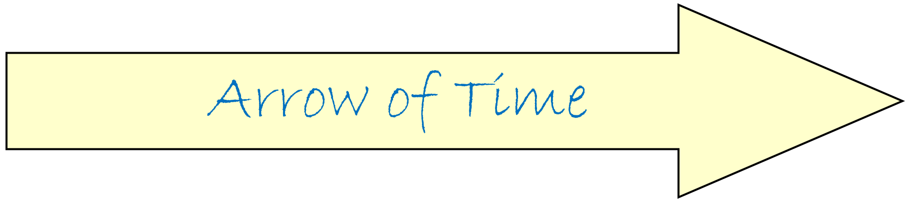
Relevant variables:

x, p

Phenomenology: Kepler's laws



Foundation: Newton's laws



Statistical Mechanics

Ludwig Botzmann

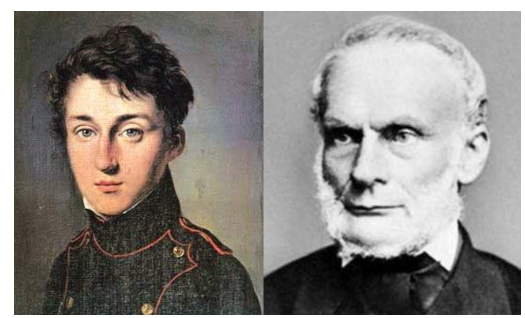
Thermodynamics

N = 10^{23} particles

Relevant variables:

P, V, T

Phenomenology: Boyle's law etc



Foundation: The laws of thermodynamics (Carnot, Clausisus)

Which brings us to

Why networks in a Statistical Physics school?

Statistical Physics aims to describe the emergent collective properties of large number of interacting elements

These properties often depend crucially on the structural arrangement of these interactions

Traditionally, the interactions occur in the neighborhood of elements arranged on a regular lattice in *d* (=1,2,3, etc.) dimensions

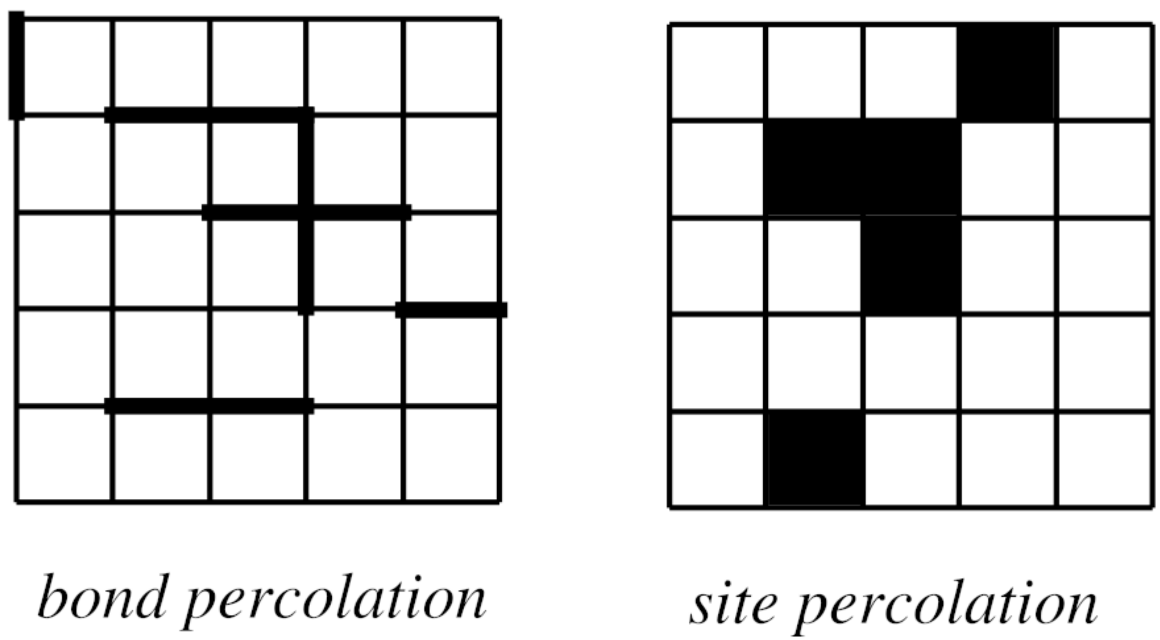
But can we have a more generalized description of the structural arrangements ? Or, as we say, have arbitrary connection topologies

Yes, by using the language of Networks

Almost every phenomenon in lattice systems studied by statistical physics can be ported to a generalized networked structure – and that conforms better to the observed arrangement of interactions in real systems

Example: Percolation

Image: Wolfram Mathworld



Example: Random Walks

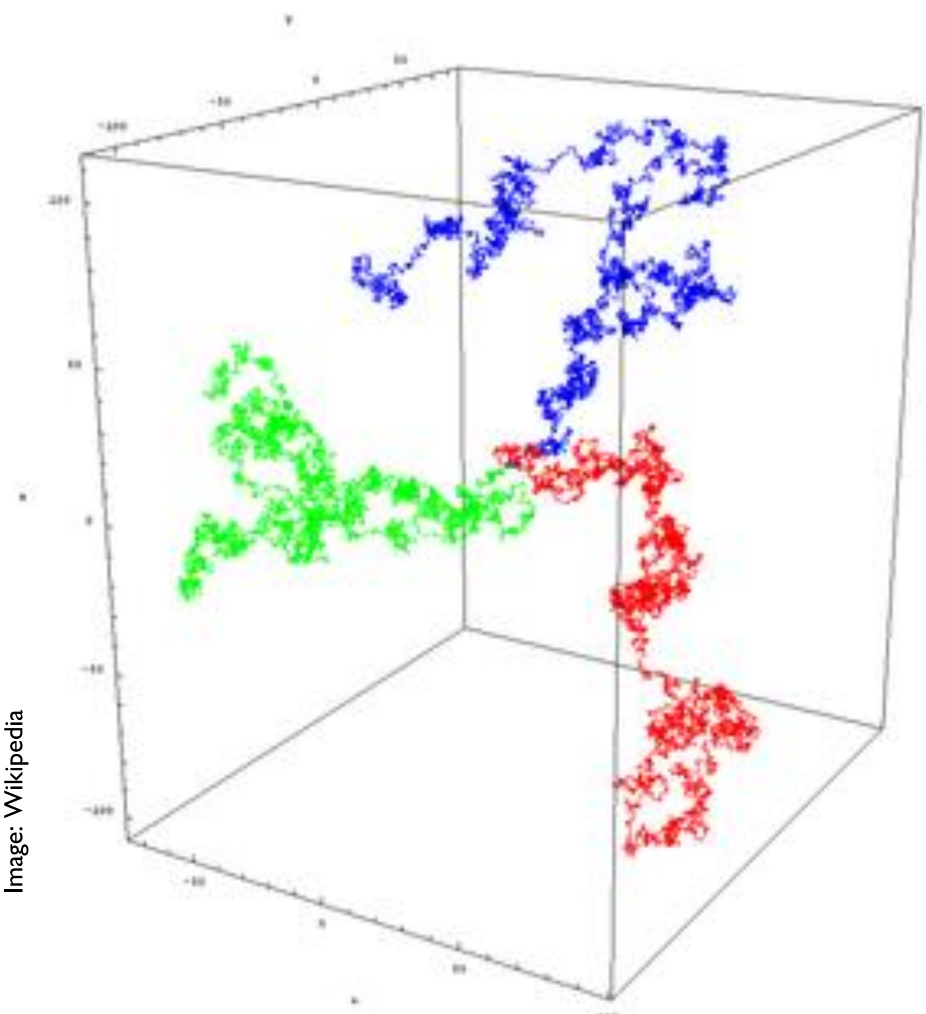


Image: Wikipedia

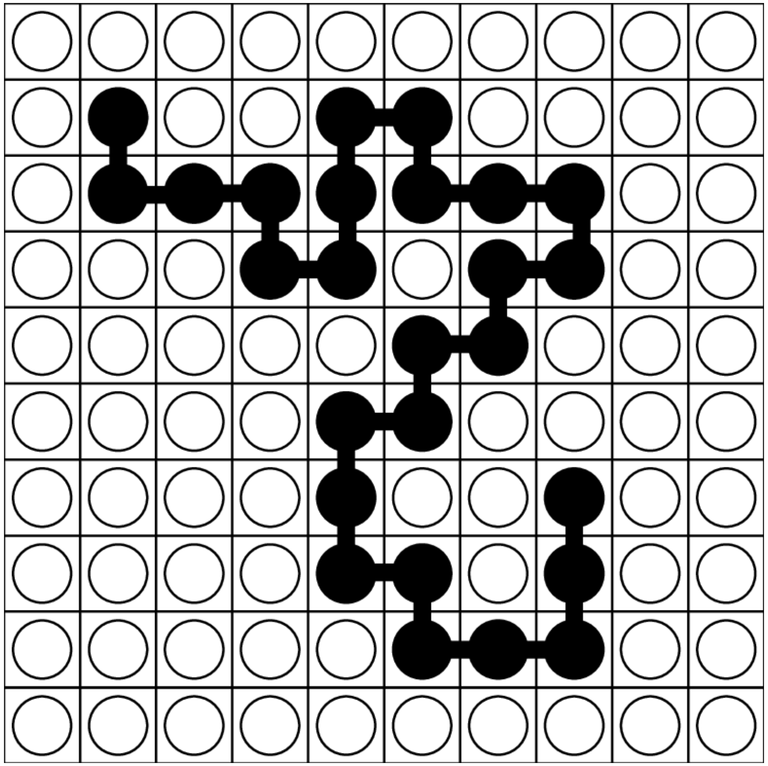


Image: Sabine Enders, Polymers (2012)

Polymers as random walks

Example: Diffusion

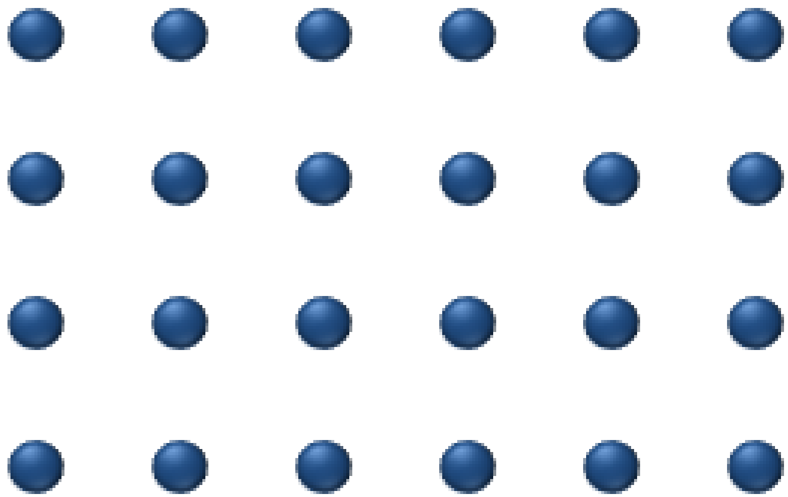


Image: Wikipedia

Deviation from a regular lattice

Diffusion in interfaces on surfaces and along dislocations

Grain Boundary Diffusion :

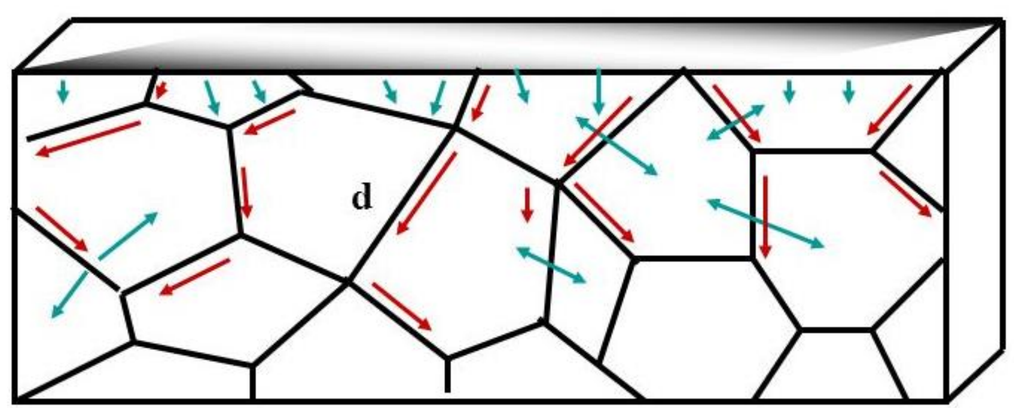


Image: Susan Nash, SlideShare

Example: Spin ordering

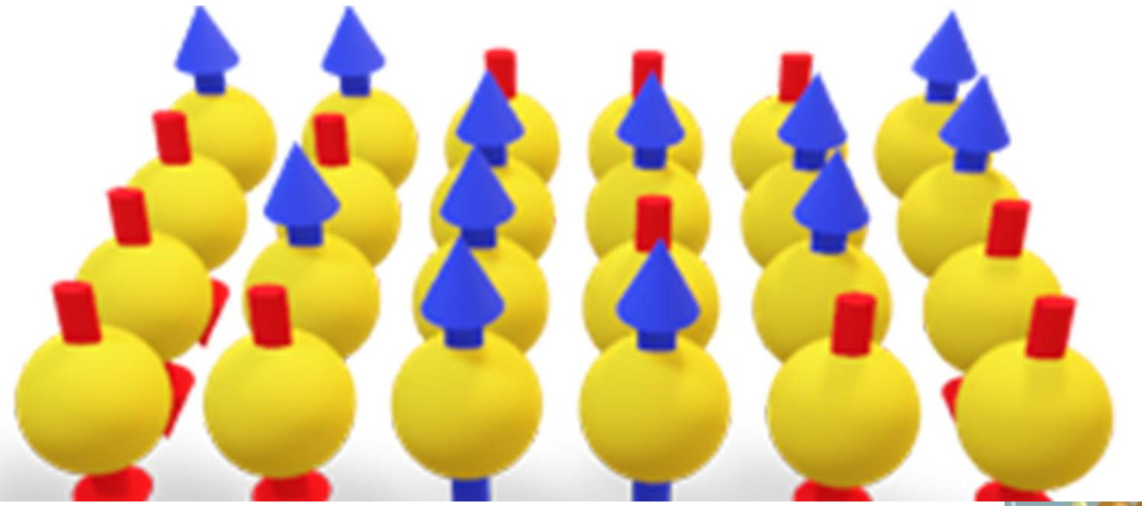


Image: Dept of Physics, University of Oxford

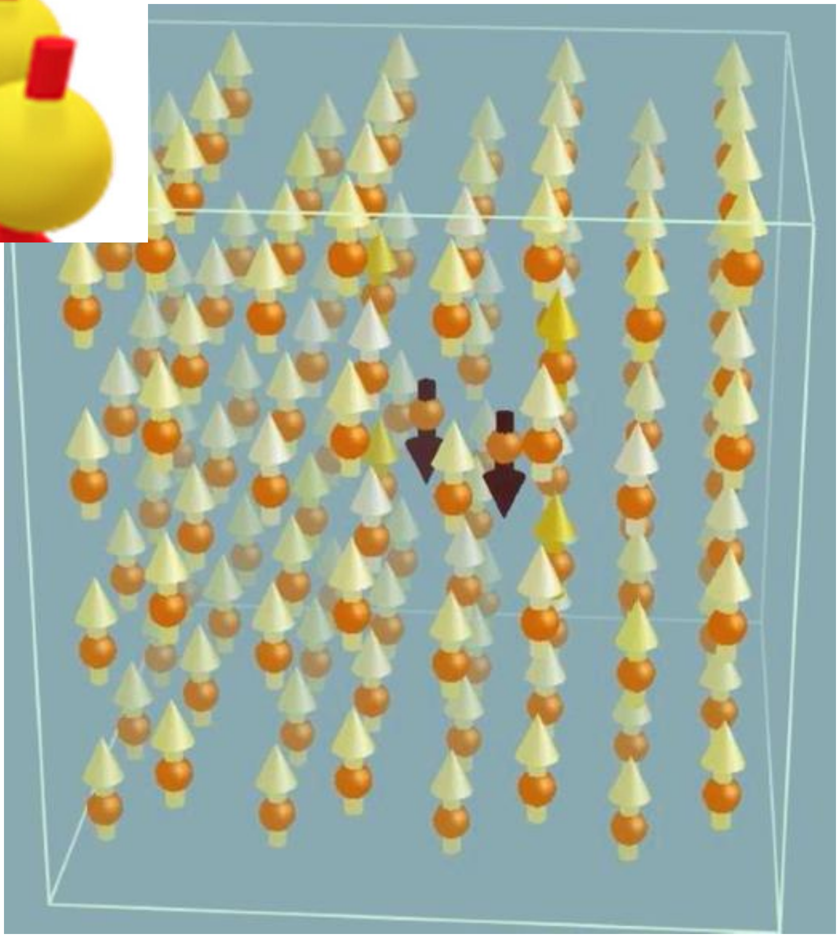


Image: P-W Ma and S L Dudarev, *Handbook of Materials Modeling* (Springer, 2018)

Example: Synchronization of oscillators

https://www.youtube.com/watch?v=T58lGKREubo

Video: uclaphysicsvideo

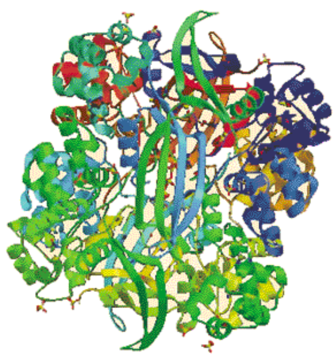
Why networks? *instead of lattice*

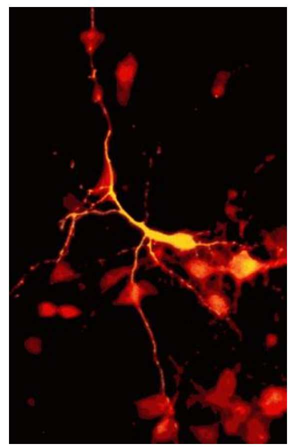
Provides a language to more accurately describe the structural arrangement of interactions in real systems across the physical, biological and social arenas

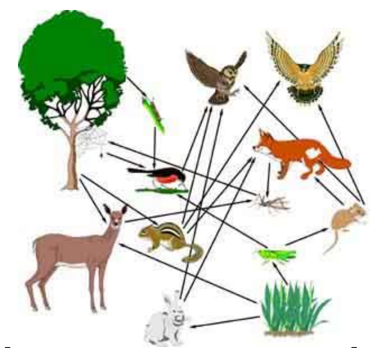
Ubiquity of Networks

Networks appear at all scales in the living world

Proteins | Intra-cellular signalling | Neuronal communication | Epidemics | Food webs



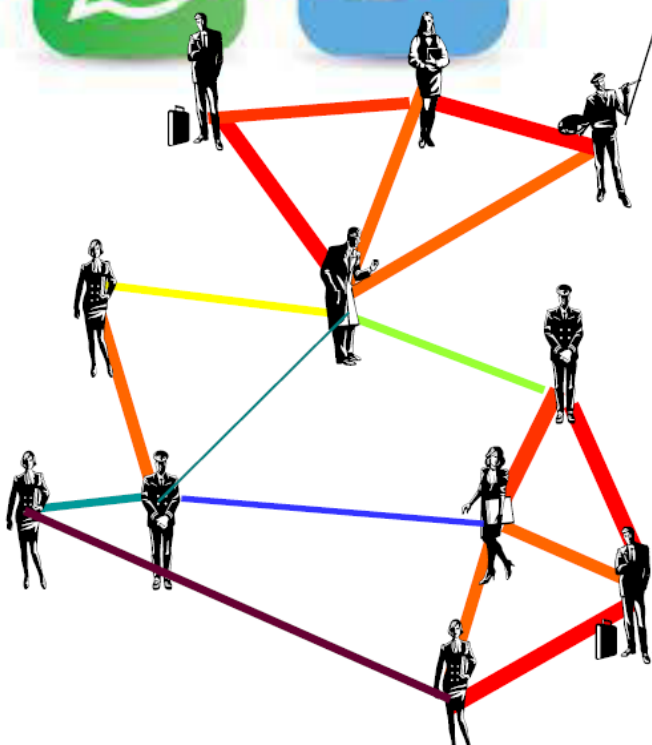




10^{-9} m | 10^{-6} m | 10^{-3} m | 1 m | 10^{3} m | 10^{6} m

Molecules | Cells | Organisms | Populations | Ecologies

Social networks



Nodes : individuals ($N = 3 - 10^9$)

Links : social interactions (< 150/node)

- ❑ Represent some underlying social relation (e.g. a "knows" b)
- ❑ Approximated by measurable interactions between individuals (e.g. via phone calls, emails, likes, physical proximity, trades, …)
- ❑ Not all relations equally strong! Weights ⇒ interaction strength

Dunbar's number

Dunbar suggested a limit to our cognitive capacity required for building stable social relations ⇒ one requires knowledge of not only all the nodes one is connected to but also their relation to each other!

Image: www.rolandberger.com

Robin Dunbar

~ 150 by extrapolating from primates, the correlation between brain size & social group size

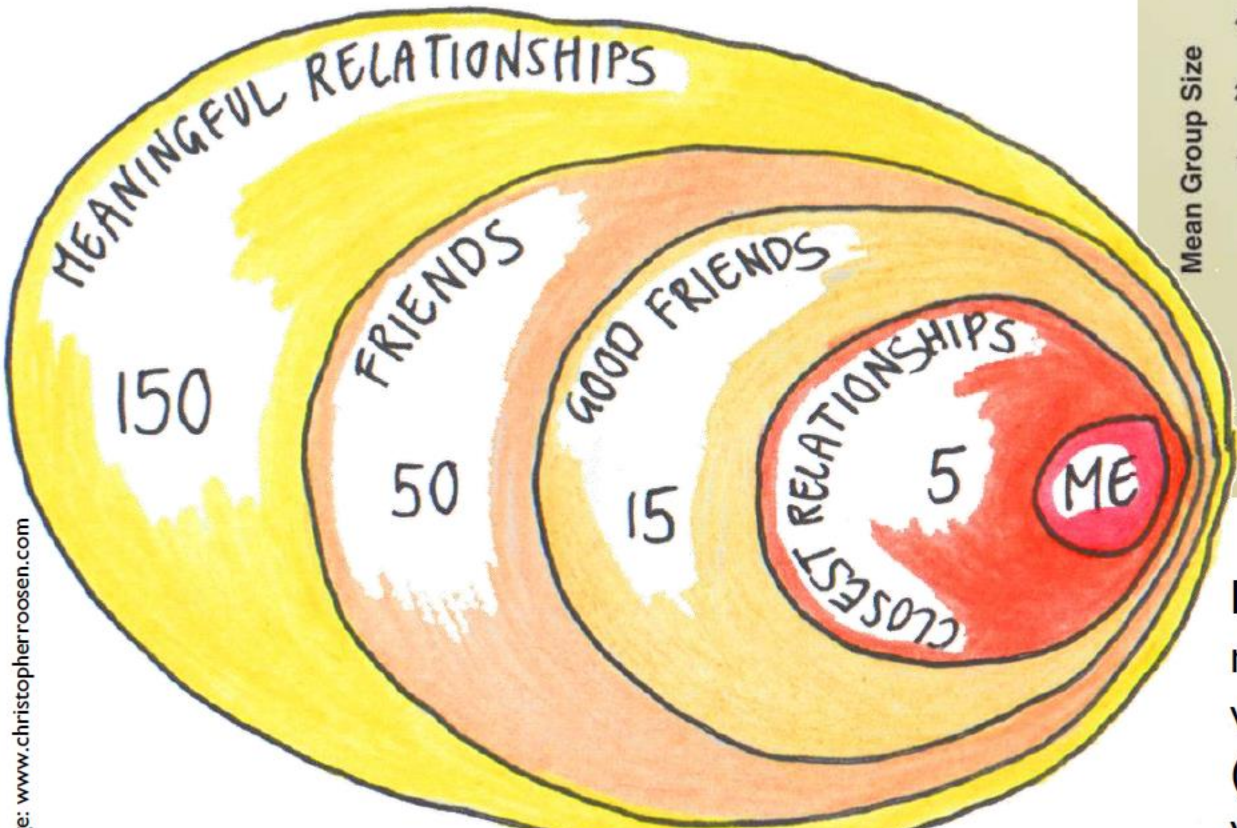


Image: www.christopherroosen.com

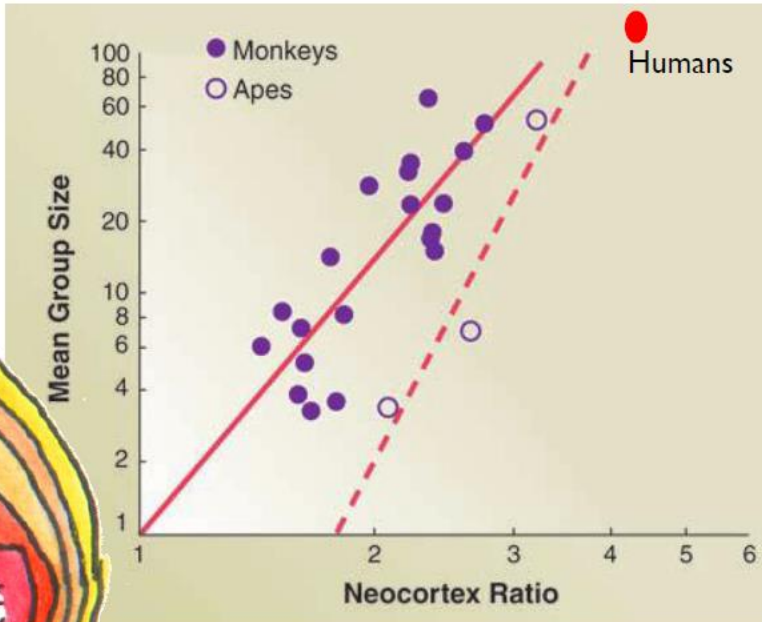


Image: R Dunbar & S Shultz, Science 317 (2007)

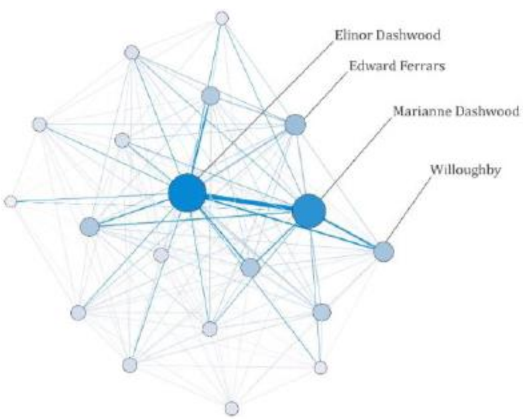
In anthropoid primates, mean social group size increases with relative neocortex volume (ratio of neocortex vol w.r.t. volume of the rest of the brain)

Social Networks in Fictional World

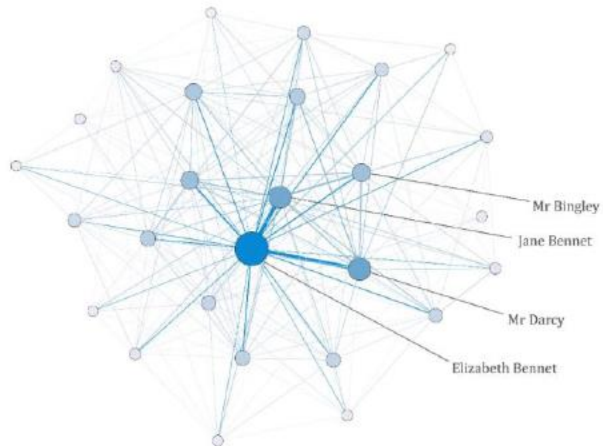
Generally tend to be much simpler than those in reality, …

Example: Jane Austen's novels

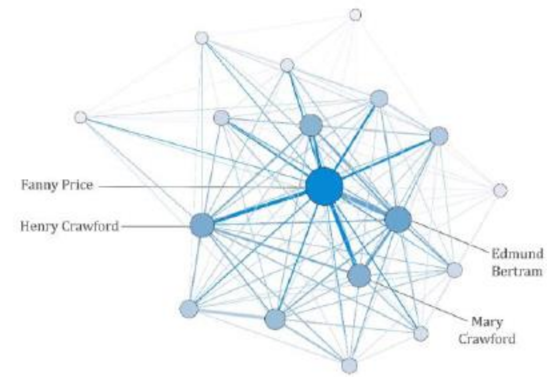
SENSE AND SENSIBILITY (1811)



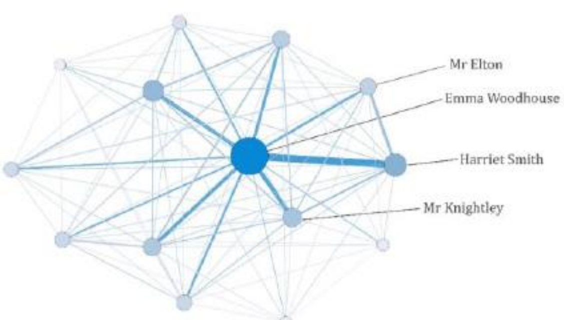
PRIDE AND PREJUDICE (1813)



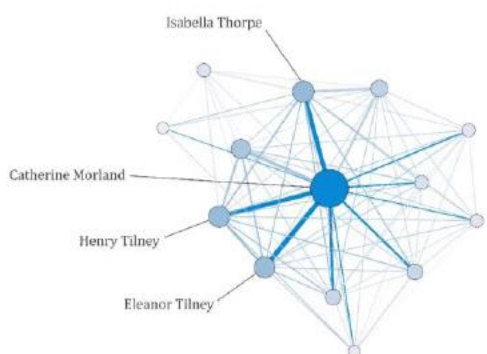
MANSFIELD PARK (1814)



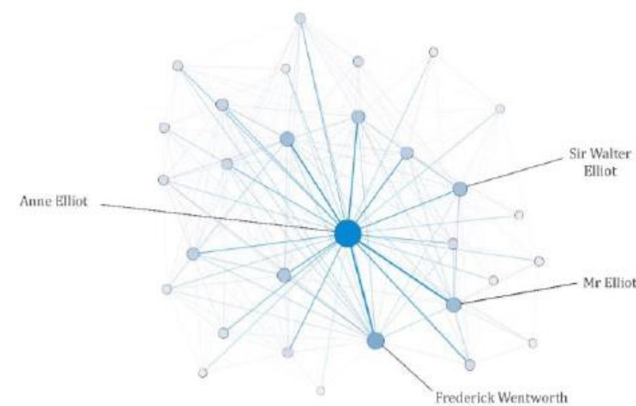
EMMA (1815)



NORTHANGER ABBEY (1817)



PERSUASION (1818)

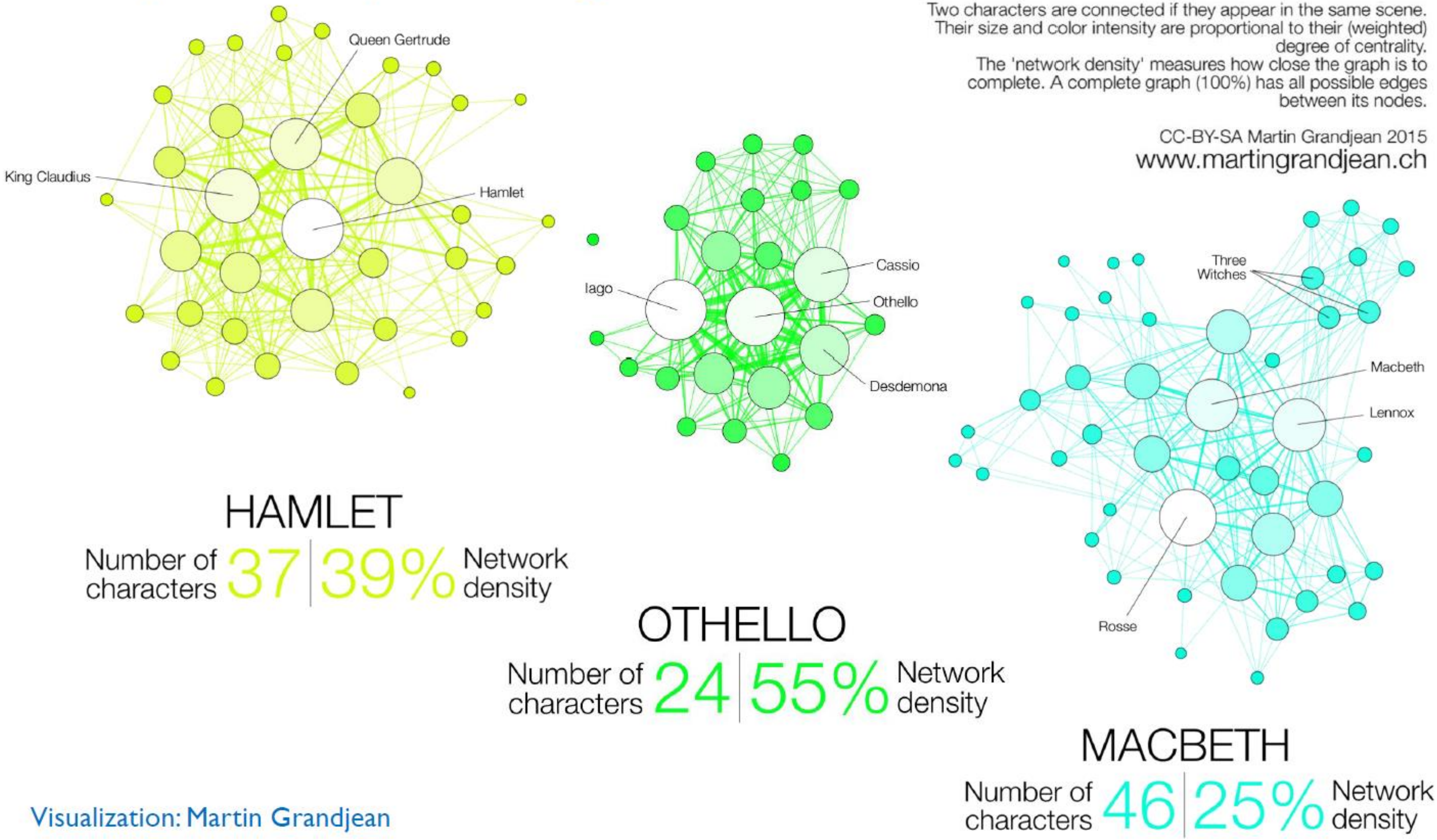


Source: https://theseaofbooks.com/ (Centre for Cultural Analytics, Univ College Dublin)

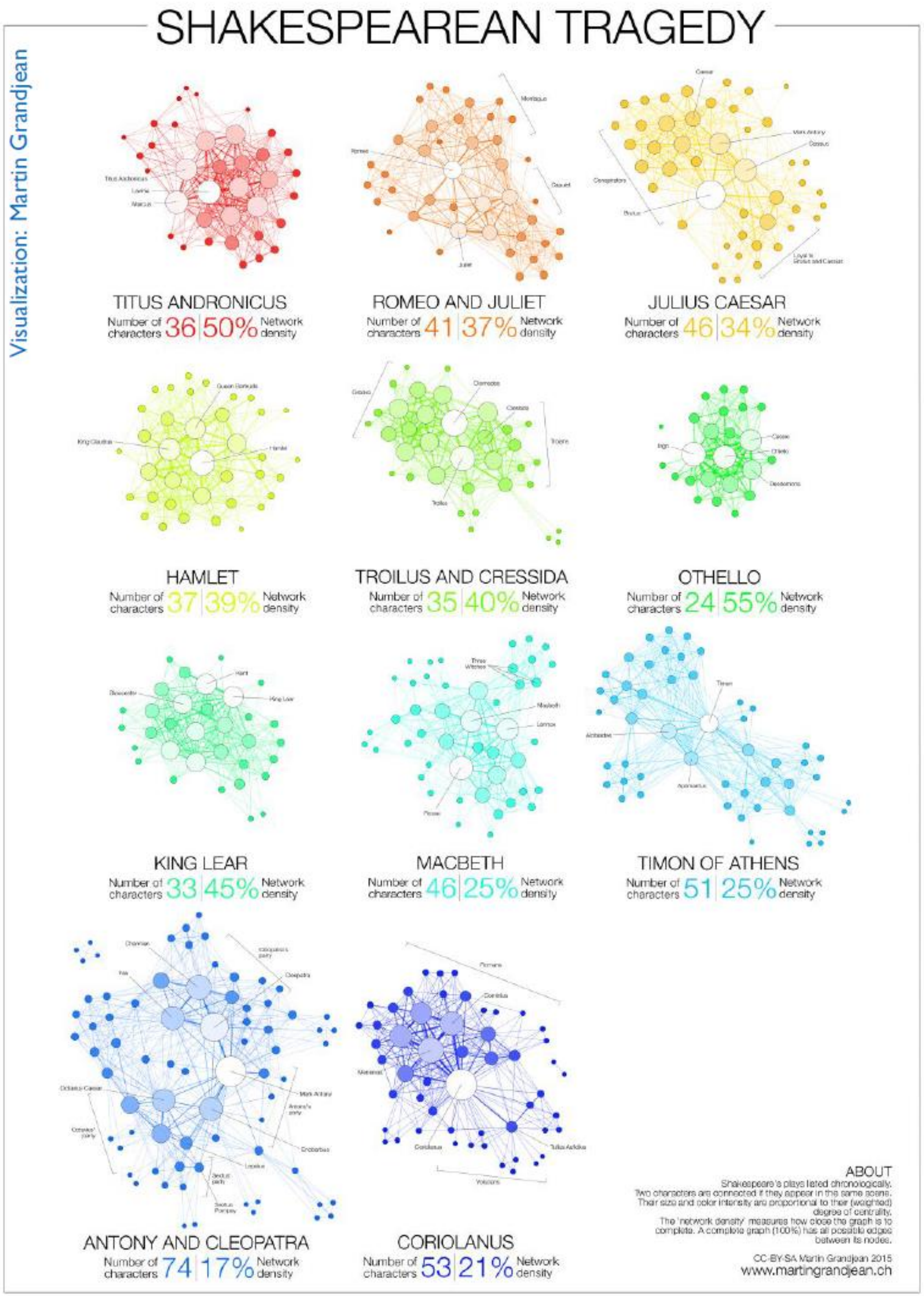
Social Networks in Fictional World

Generally tend to be much simpler than those in reality, …

Example: Shakespeare's tragedies



Visualization: Martin Grandjean



Do fictional social networks necessarily need to be simpler than those in real life?

If larger brains are primarily are primarily for negotiating our social networks over our lifetime, fictional networks have to be simpler to give the reader/listener/viewer the chance to comprehend the network in the space of hours, days or weeks!

The social networks of Shakespeare's plays show a general tendency to become larger but also sparser as we move on from earlier to later plays (?)

Similar exercise can be done with other works, e.g., Mahabharata

See https://blog.gramener.com/the-mahabharatha-in-pictures/

Social Networks in Fictional World

Generally tend to be much simpler than those in reality, but there are exceptions!

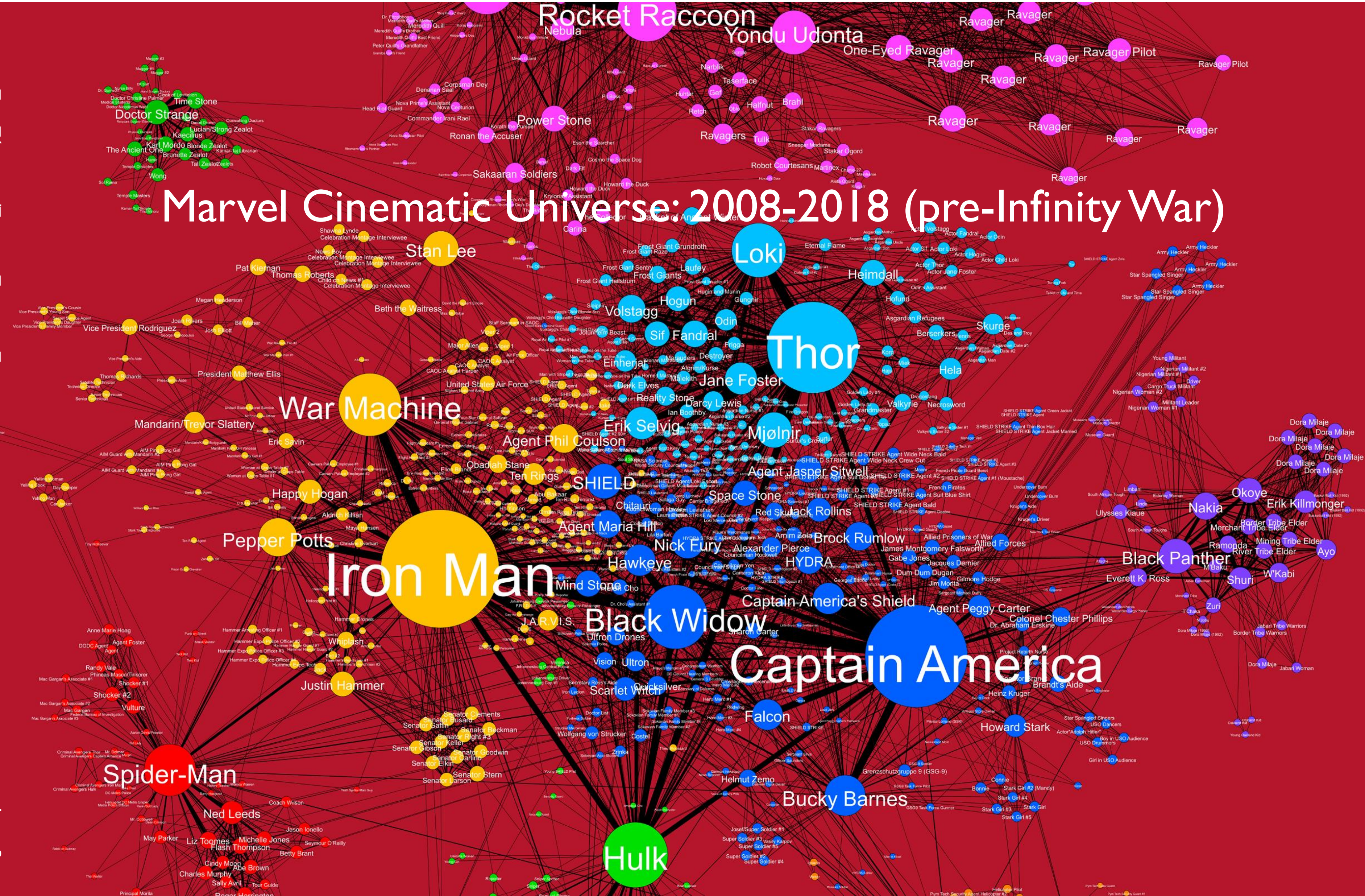


Image: https://www.reddit.com/r/marvelstudios/comments/8e36n7/marvel_cinematic_universe_preinfinity_war_20082018/

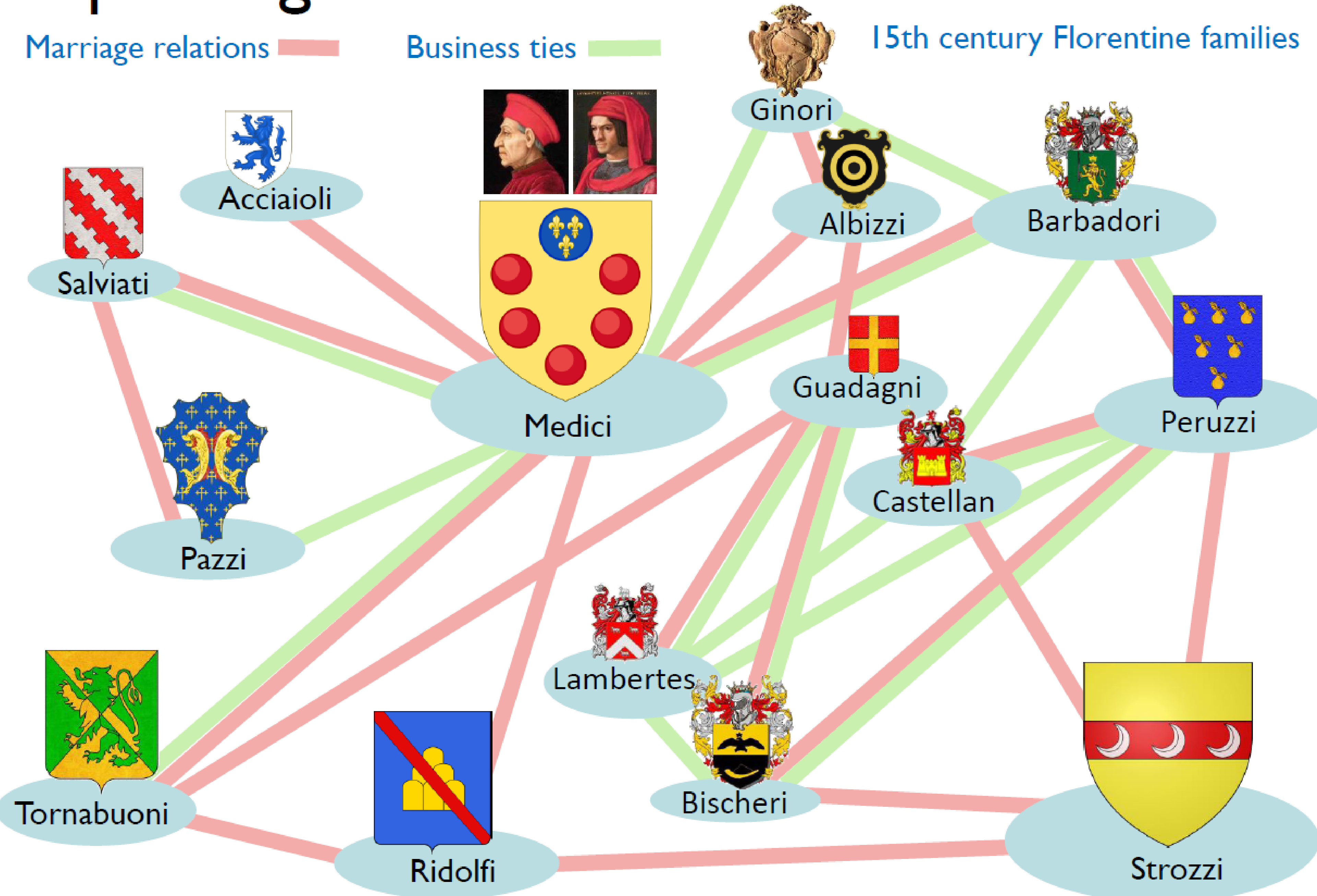
Explaining the rise of the Medici

https://www.youtube.com/watch?v=SFUWvceed5s

Medici: Masters of Florence trailer

Data: John F Padgett (1993)

Explaining the rise of the Medici



Technological networks

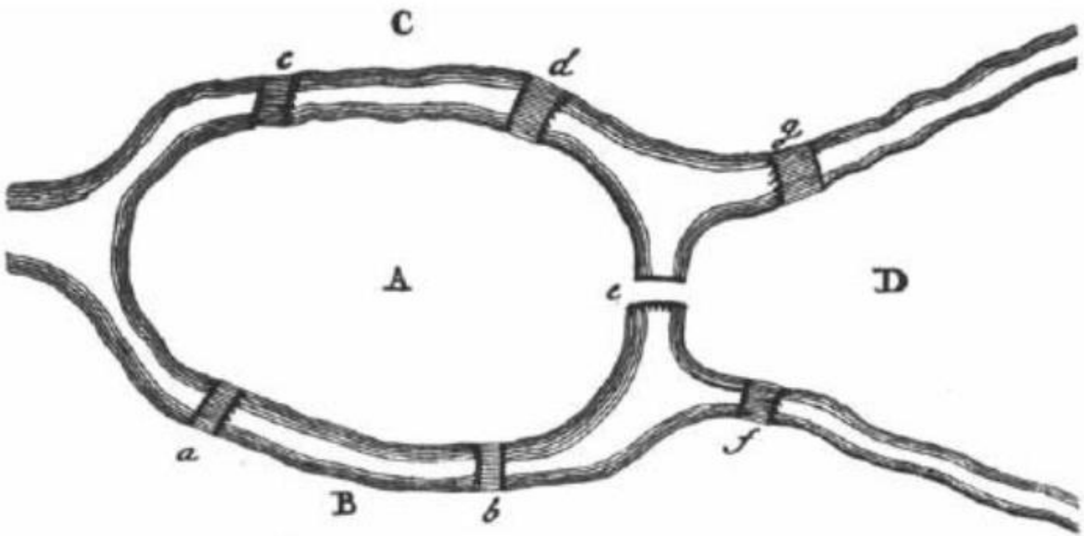
Transportation

Power Transmission

Image: www.wirecable.in

Information

Using the concept of networks:
From the 7 bridges of Konigsberg…



Seven bridges of Koenigsberg crossed the River Pregel

Source: Scientific American

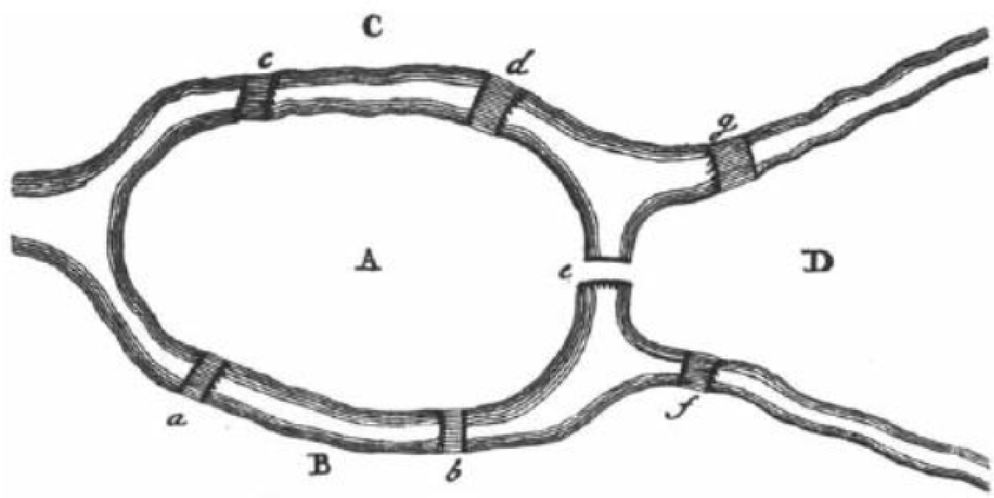
Problem: to find a walk through the city that would cross each bridge once and only once.

Source: Wikipedia

Leonard Euler (1707-1783)

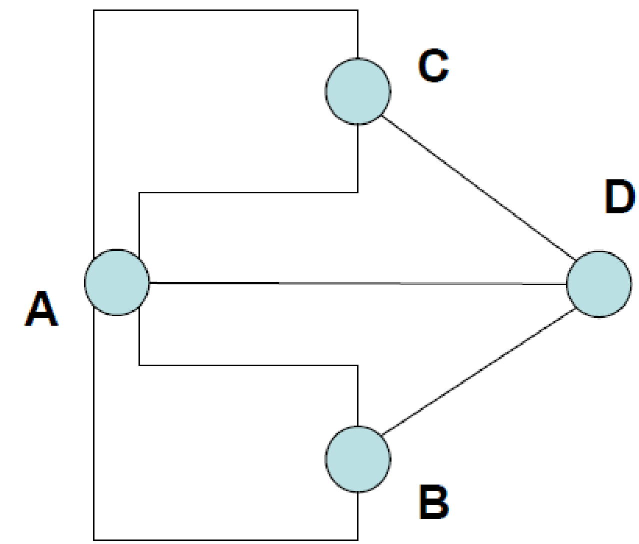
Konigsberg Bridge problem

Source: Scientific American



Seven bridges of Koenigsberg crossed the River Pregel

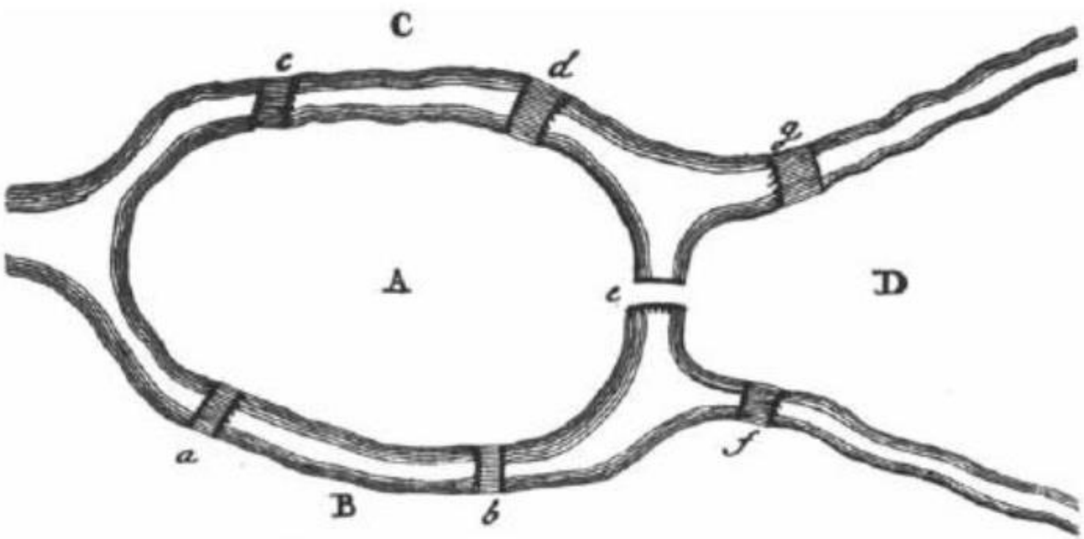
≡



"Does there exist any walking route that crosses all seven bridges exactly once each?" ≡ finding an Eulerian path (i.e., a path that traverses each link in a network exactly once) on the equivalent network

Any Eulerian path must enter as well as exit each node it passes through except the first and last noes ⇒ at most two nodes in the network with odd degree if such a path is to exist.
As all four nodes in the Konigsberg network have odd degree, the bridge problem necessarily has **no solution**.

Using the concept of networks:
From the 7 bridges of Konigsberg…



Seven bridges of Koenigsberg crossed the River Pregel

Source: Scientific American

Problem: to find a walk through the city that would cross each bridge once and only once.

Source: Wikipedia

Leonard Euler (1707-1783)

Led to foundation of Graph theory: the study of mathematical structures called graphs that model pairwise relations between objects

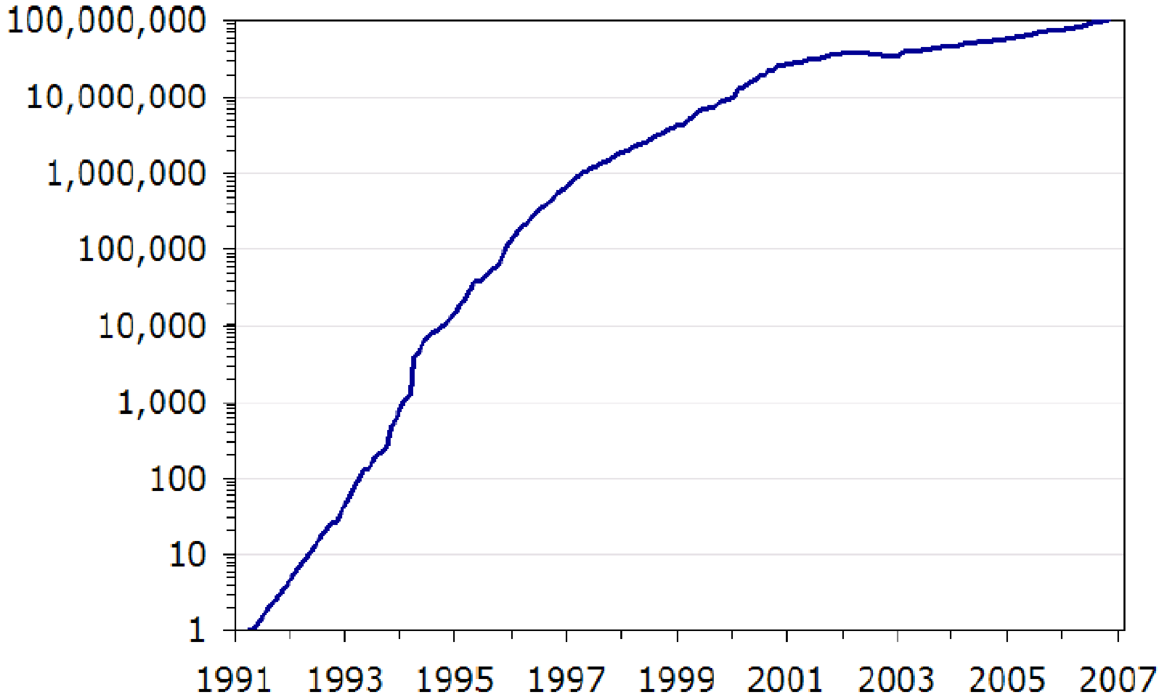
Using the concept of networks:

…to searching the World Wide Web

When the World Wide Web (WWW) was first proposed, it was widely assumed that its utility was limited because it would be impossible to efficiently access the vast quantity of information distributed throughout the entire network

Like searching for a needle in a haystack most of the time what you will find will be completely irrelevant to what you want !

Growth of number of web pages between 1991-2007



Source:www.useit.com/alertbox/web-growth.html

Using the theory of networks

Google's answer to data deluge

Manually indexed search engines were clearly incapable of handling the exponentially increasing amount of information in the web…

….until Google came up with an efficient automated search algorithm that weights the importance of each page according to how many other pages are pointing to it

Image: Time.com

How to measure the importance of a node ?

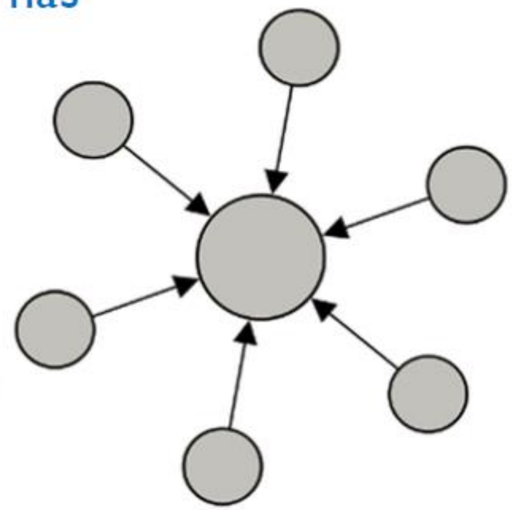
One possibility is to ask how many does a node "know"?

Degree of a node: The number of connections a node has with other nodes

In-Degree

'Effect'

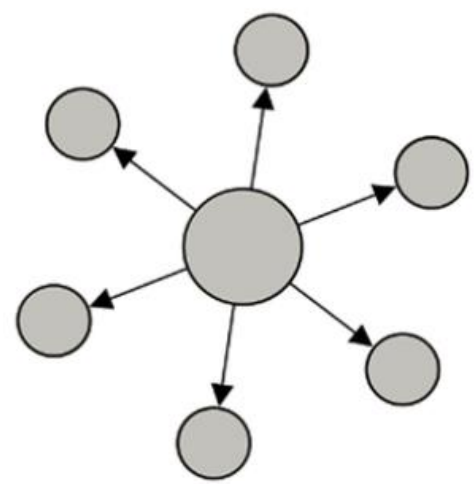
Variables with high in-degree are impacted by multiple other variables. An in-degree of 0 means a variable is not influenced by others in the system.



Out-Degree

'Cause'

High out-degree variables have an ability to change many others in the system. Variables with an out-degree of 0 do not directly influence others.



But if your webpage links to only that of your friends and they link back to you, you may have many links (as will your friends) but your page may not be of much interest to most users

Source: J Mcglashan et al, PLoS One 11 (2016) 10.1371/journal.pone.0165459

The basic idea of PageRank

“It’s not just how many you know, it’s <u>who</u> you know!”

The <u>importance</u> of a page is determined by how many other <u>important</u> pages link to it

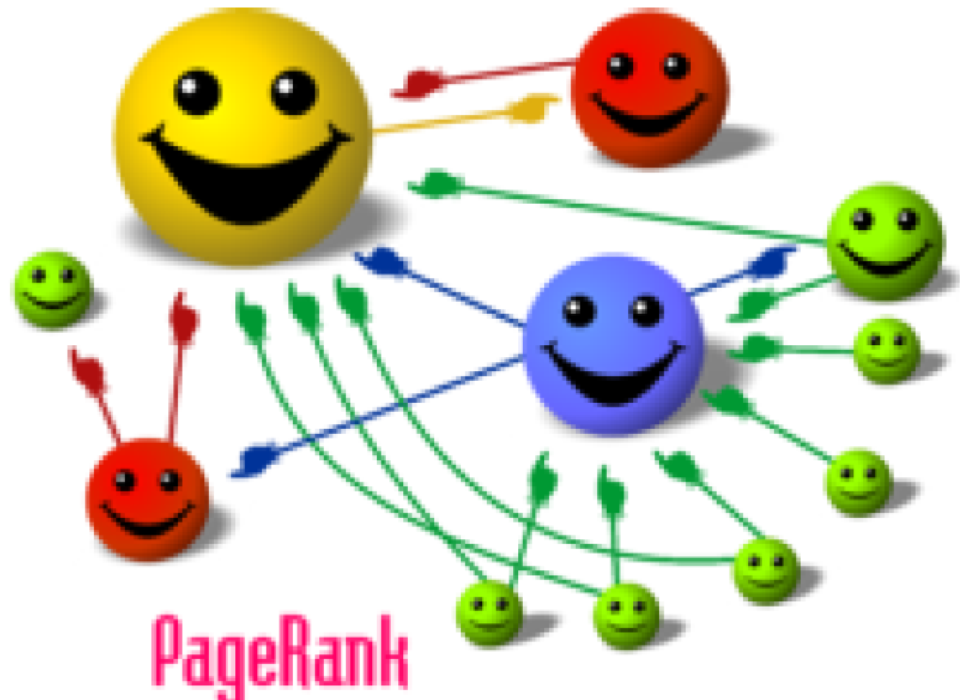


Image: Wikipedia

If that seems a circular definition, turns out that we can define it rigorously using linear algebra

Enter

Eigenvector centrality

Eigenvector Centrality

Each node given a score proportional to sum of scores of its neighbors

- Let each node i be given a initial score $x_i(0)$ e.g., = 1 for all i
- Starting from an initial guess, a better value of the centrality is calculated
  $x_i(1) = \Sigma_j A_{ij} x_j(0)$ from definition of centrality as sum of neighbors centralities
  $\mathbf{x}(1) = \mathbf{A}\,\mathbf{x}(0)$ in matrix notation

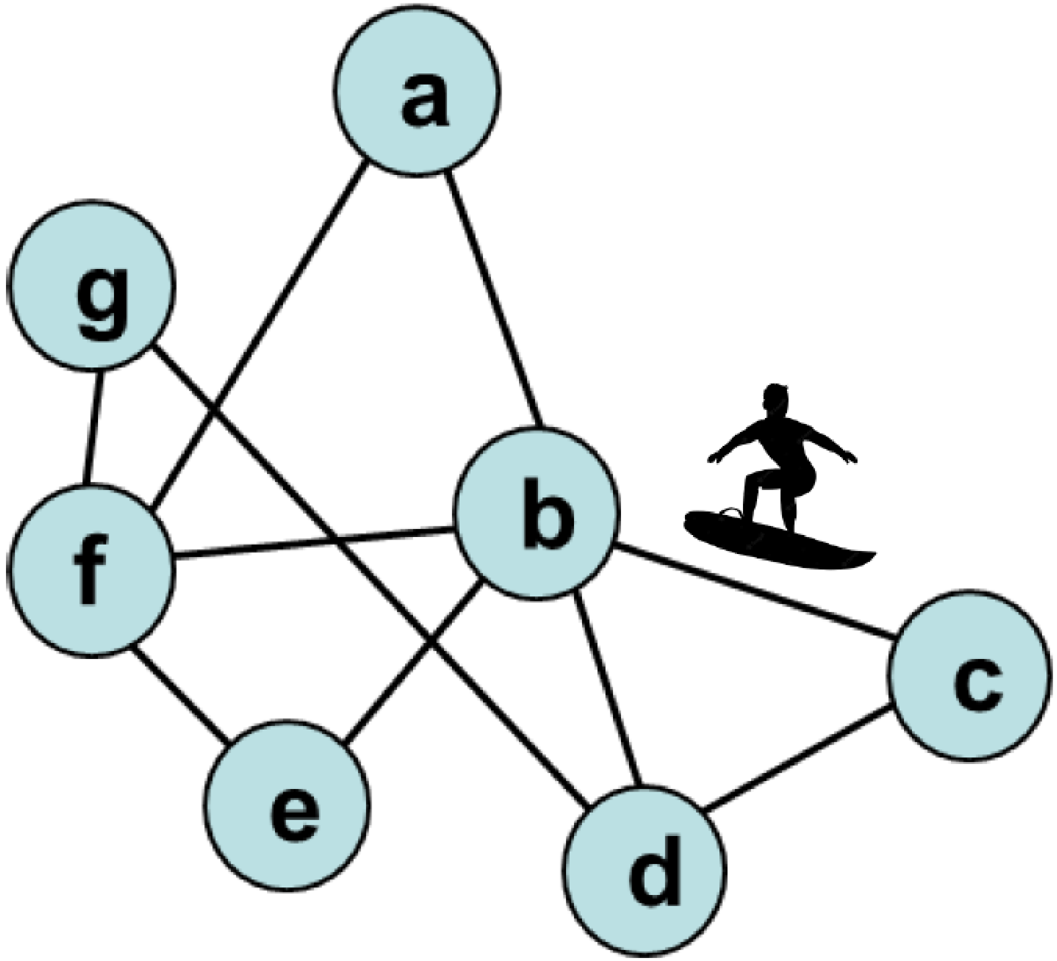
- Repeating this process iteratively for t steps,
  $\mathbf{x}(t) = \mathbf{A}^t\,\mathbf{x}(0)$

- Expressing $\mathbf{x}(0) = \Sigma_i c_i \mathbf{v}_i$ a linear combination of the eigenvectors $\mathbf{v}_i$ of $\mathbf{A}$
  $\mathbf{x}(t) = \mathbf{A}^t \Sigma_i c_i \mathbf{v}_i$
  $\mathbf{x}(t) = \Sigma_i c_i \lambda_i^t \mathbf{v}_i = \lambda_1^t \Sigma_i c_i [\lambda_i/\lambda_1]^t \mathbf{v}_i$
  where $\lambda_1 > \ldots > \lambda_i > \ldots > \lambda_N$ are eigenvalues of $\mathbf{A}$

- As $\lambda_i/\lambda_1 < 1$, all terms other than the first decay as $t \to \infty \Rightarrow \mathbf{x}(t) \to c_1 \lambda_1^t \mathbf{v}_1$
  $\Rightarrow$ centrality x satisfies $\mathbf{A}\,\mathbf{x} = \lambda_1 \mathbf{x}$
  $\Rightarrow$ is proportional to the leading eigenvector of the adjacency matrix $\mathbf{A}$

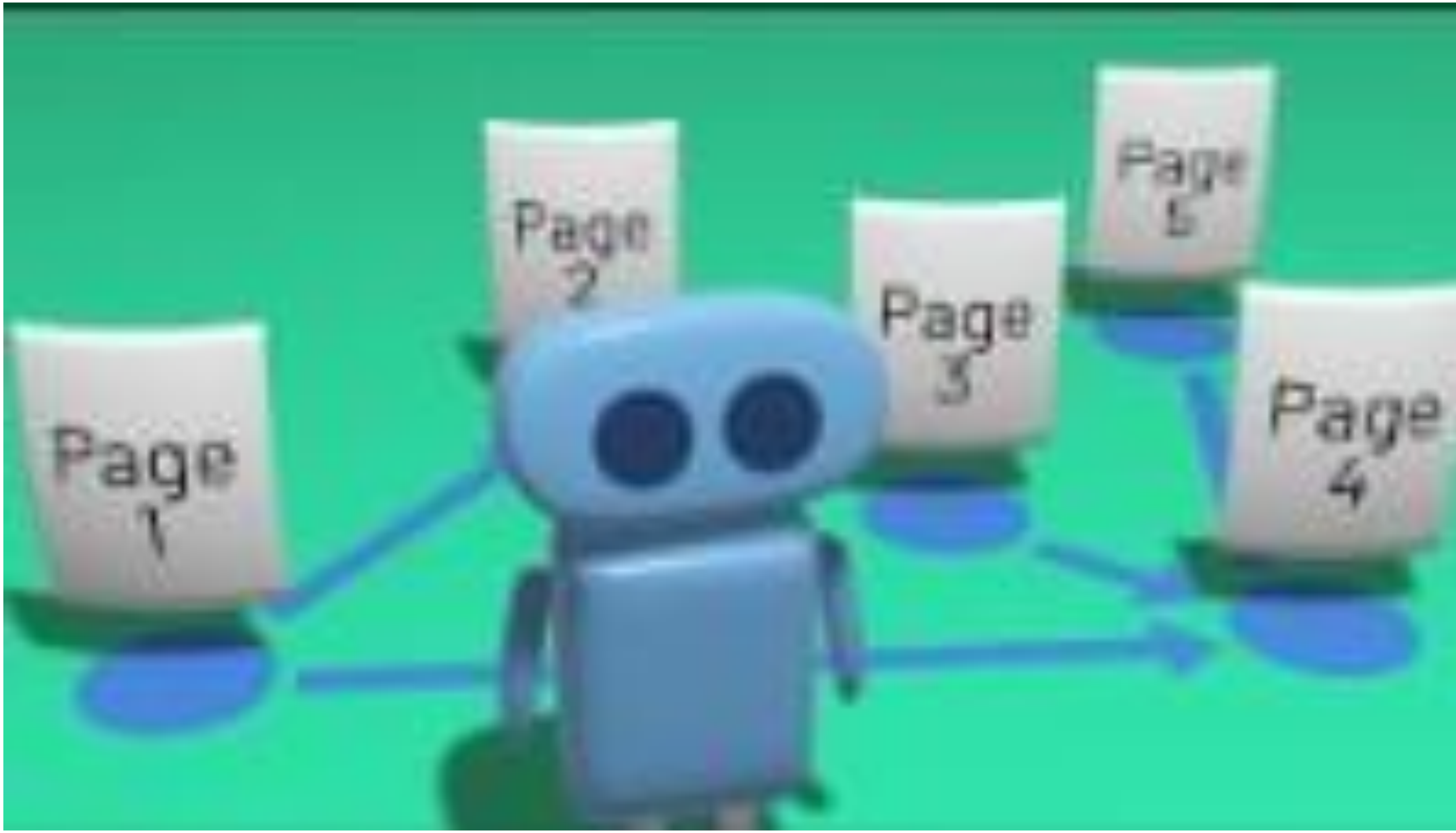
If that sounds too complicated, you can just randomly surf the net

Random Surfer Model



Score or rank of a node (e.g., a webpage) will be related to the frequency with which a random surfer will visit the page

How Google's PageRank Algorithm Works



https://www.youtube.com/watch?v=meonLcN7LD4

Beyond networks: Hypergraphs

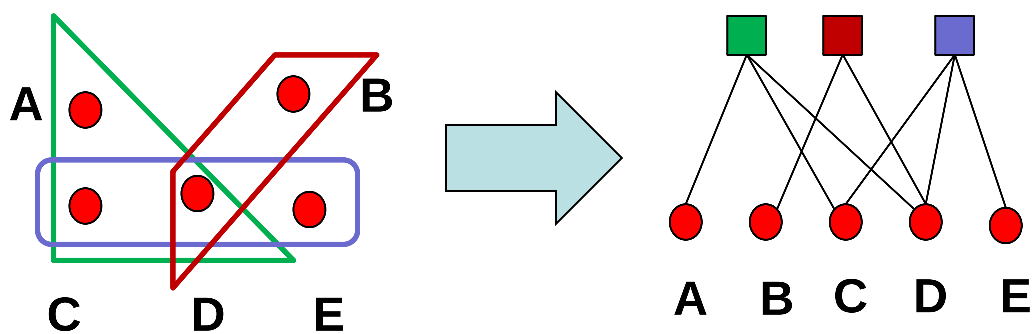
Networks → pairwise interactions between nodes.
But
Relations may be defined in terms of multilateral rather than just bilateral relations
Many processes involve several components participating together in an interaction, e.g.,

(i) substrate A is converted to product B on coming in contact with enzyme C
(ii) a protein complex that comprises more than 2 proteins

A generalized link connecting more than two nodes is a *hyperedge*

Hypergraph: A network with hyperedges

Can be represented by a bipartite network – a network consisting of two different types of nodes, with links occurring only between nodes of unlike type



Multilayer Networks

Multiplex

Inter-dependent

Example: The cell

- ❑ Protein-protein interaction networks
- ❑ Metabolic networks
- ❑ Gene transcriptional regulatory networks
- ❑ Cell signalling networks

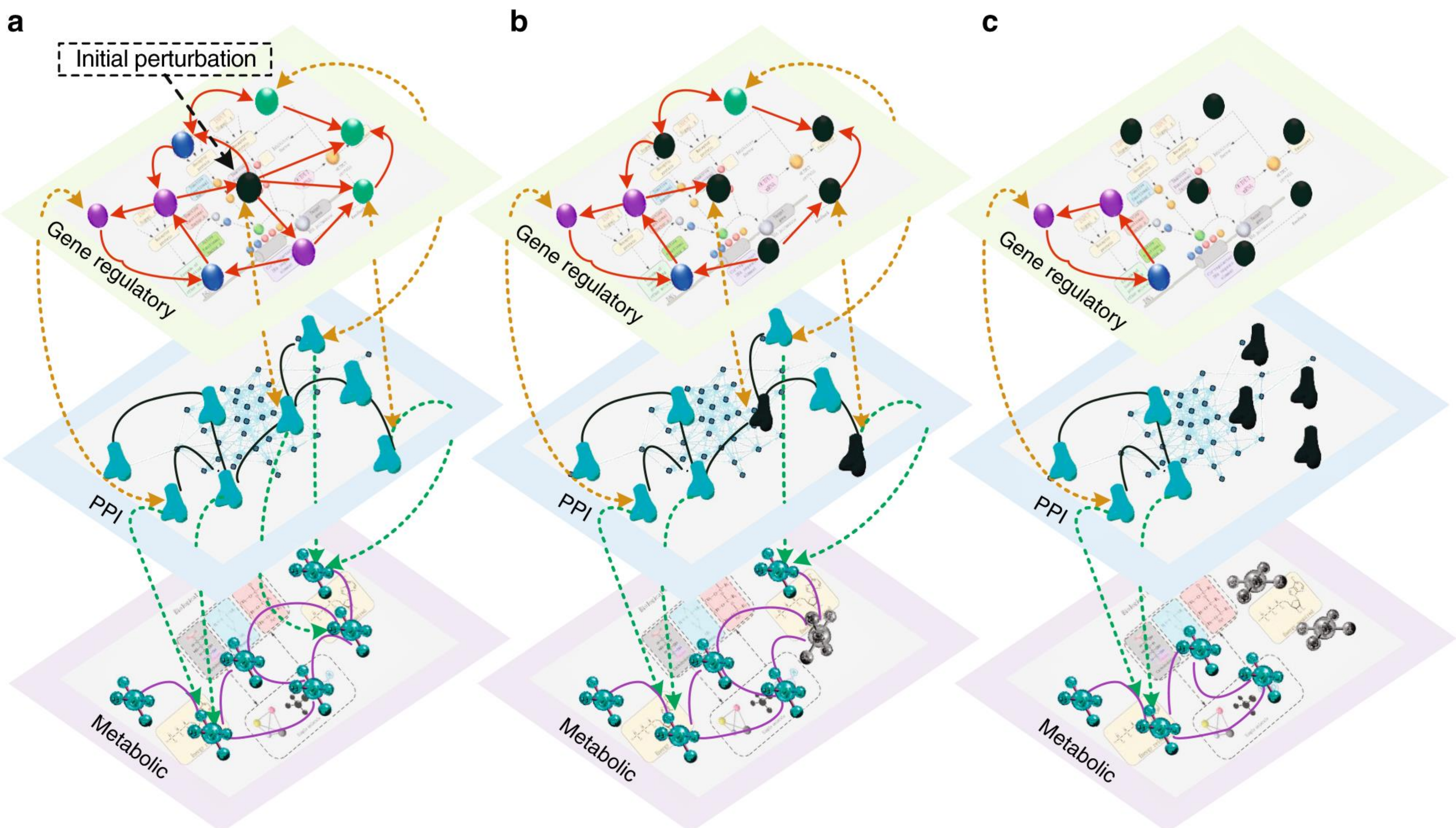
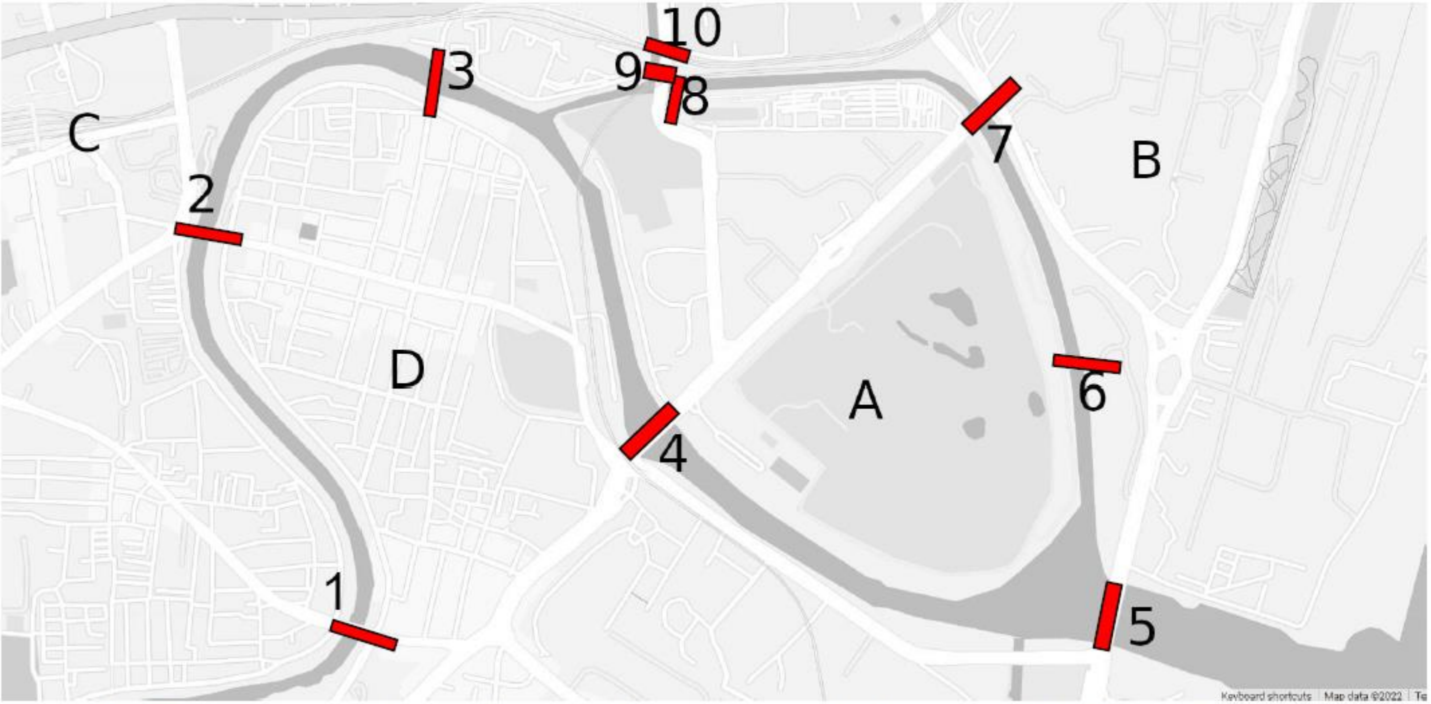


Image: X Liu et al. *Nat Commun* 11, 6043 (2020)

Introductory Papers & Books

- M. E. J. Newman, *Networks: An Introduction* (2nd edition, 2018) Oxford U Press
- A. Barrat, M. Barthelemy and A. Vespignani, *Dynamical Processes on Complex Networks* (2008) Cambridge U Press
- U. Alon, *An Introduction to Systems Biology: Design principles of biological circuits* (2007) CRC Press
- S. H. Strogatz: "Exploring complex networks", *Nature* 410:268 (2001)
- M. E. J. Newman: "The structure and function of complex networks", *SIAM Review* 45: 167 (2003)
- S. N. Dorogovtsev & J.F.F. Mendes, "Evolution of networks", *Advances in Physics* **51**, 1079 (2002)
- S. Boccaletti et al, "Complex networks: Structure and dynamics", *Physics Reports* (2006)
- M. E. J. Newman , A.-L. Barabási and D. J. Watts (eds) *The Structure and Dynamics of Networks* (2006) Princeton Univ Press
- D. J. Watts, *Six degrees* (2004) Penguin
- A.-L. Barabási, *Linked: The New Science of Networks* (2002) Perseus

Assignment The Ten Bridges of Chennai



Consider the map given above showing the area around the Central Railway Station of Chennai (part of Old Madras). There are 4 land masses (A,B,C,D) separated by waterways and linked to each other by10 bridges (indicated by the numbers 1,…,10).

(a) Draw a graph representation of the above by showing land masses as vertices (nodes) & bridges as edges (links).
(b) Write down the corresponding adjacency matrix, with A as the first node, B as the second node, C as the third node and D as the fourth node.
(c) What is the average path length in the network?
(d) Is it possible to find an Eulerian walk through the 10 bridges, i.e., one which goes through all the bridges, visiting each of them once and once only (analogous to the problem of finding a walk through the seven bridges of Konigsberg solved by Euler)?
If yes, mention the sequence in which each bridge can be traversed in such a walk.
If no, provide reasons why there may not be such a path.