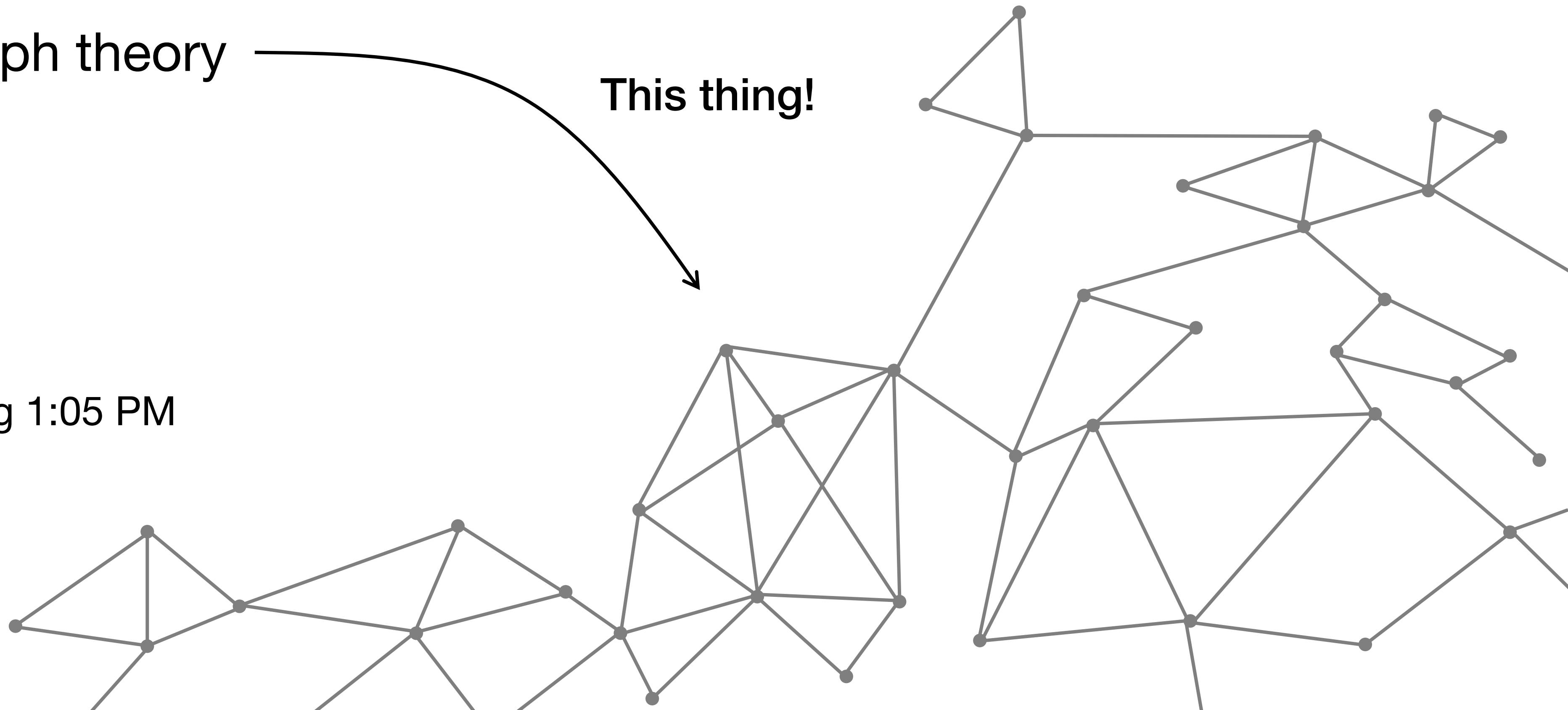
Modeling Markets, Pandemics, and Peace: The Mathematics of Multi-Agent Systems

Lecture 5

Introduction to graph theory

This thing!



MIT HSSP

August 6th, 2022. Starting 1:05 PM

Recap of game theory

Game theory is the study of multi-person decision problems (games).

A **Nash equilibrium** is an outcome where no player wants to deviate

Prisoner's dilemma

Player 2 \ Player 1	Cooperate	Defect
Cooperate	3 / 3	1 / 4
Defect	4 / 1	2 / 2

Nash equilibria:
[Defect, Defect]

Coordination game

Player 2 \ Player 1	A	B
A	1 / 1	0 / 0
B	0 / 0	1 / 1

Nash equilibria:
[A, A], [B, B],
$[\ (\frac{1}{2}A, \frac{1}{2}B), (\frac{1}{2}A, \frac{1}{2}B)\ ]$

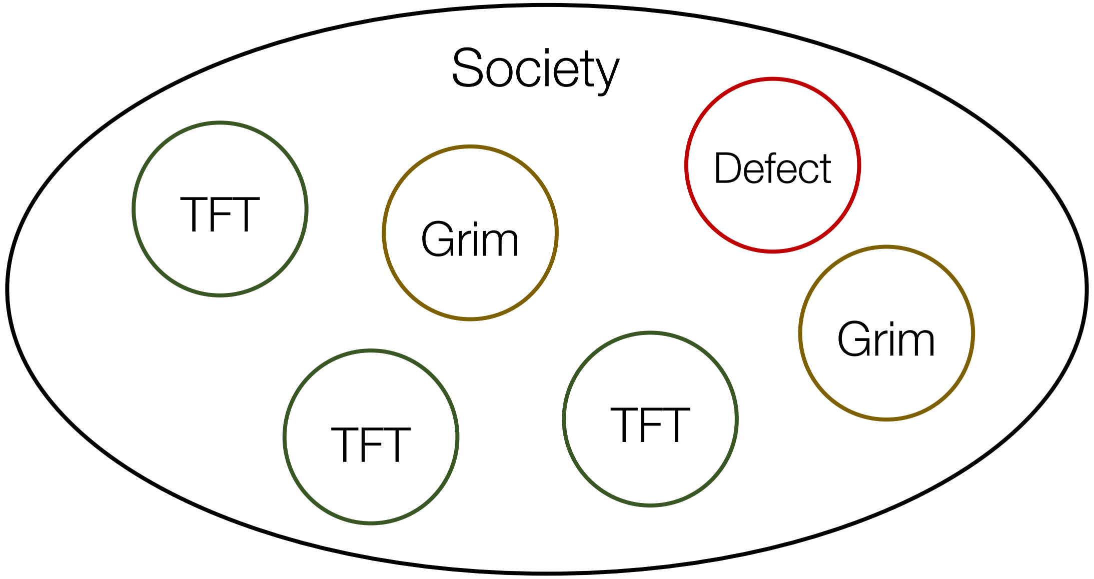
Hide-and-seek

Player 2 \ Player 1	Up	Down
Up	1 / 0	0 / 1
Down	0 / 1	1 / 0

Nash equilibria:
$[\ (\frac{1}{2}U, \frac{1}{2}D), (\frac{1}{2}U, \frac{1}{2}D)\ ]$

The evolution of cooperation

Axelrod's tournament: make strategies "reproduce" according to how much they win.



Grim Trigger (punish forever)

P1	3	3	1	2	4	2	2
P2	3	3	4	2	1	2	2

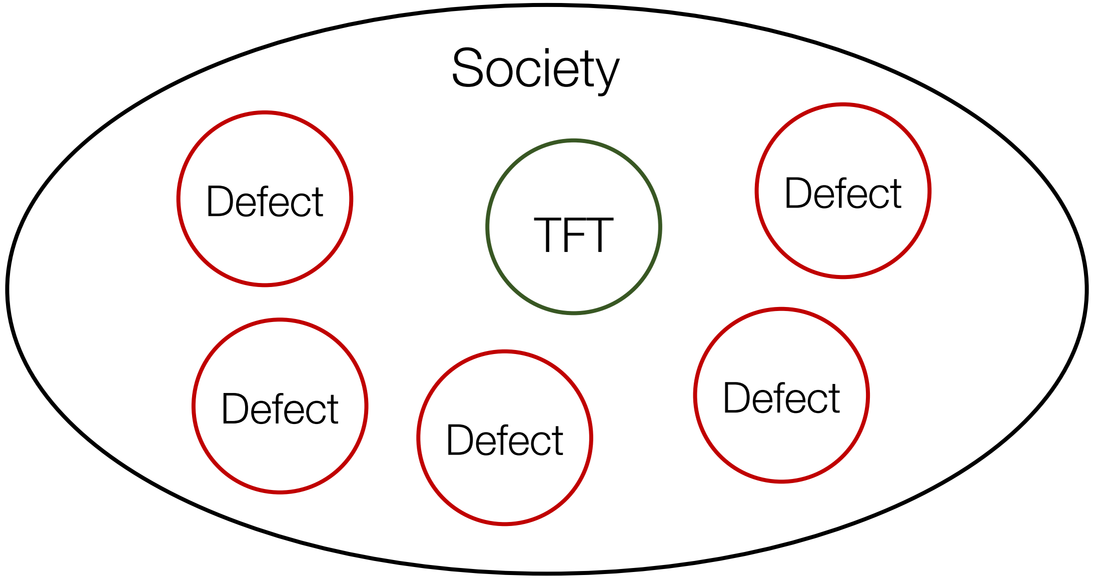
Tit for Tat (punish, but forgive)

P1	3	3	1	2	4	3	3
P2	3	3	4	2	1	3	3

Cooperation can begin with small clusters and thrive in neighborhoods that are "nice," protecting themselves from invasion. **But they can also go extinct with bad neighbors!**

The evolution of cooperation

Axelrod's tournament: make strategies "reproduce" according to how much they win.



Grim Trigger (punish forever)

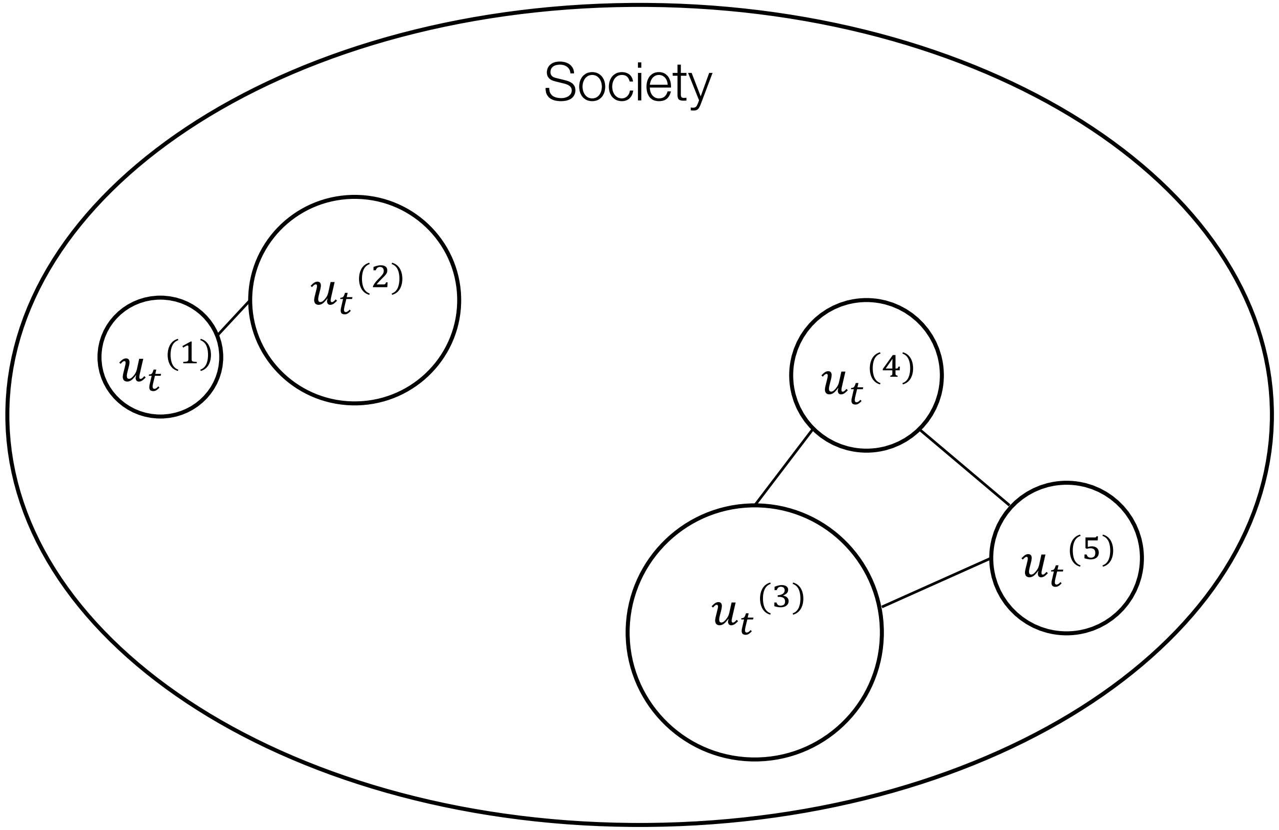
P1	3	3	1	2	4	2	2
P2	3	3	4	2	1	2	2

Tit for Tat (punish, but forgive)

P1	3	3	1	2	4	3	3
P2	3	3	4	2	1	3	3

Cooperation can begin with small clusters and thrive in neighborhoods that are "nice," protecting themselves from invasion. **But they can also go extinct with bad neighbors!**

Systems of interacting agents

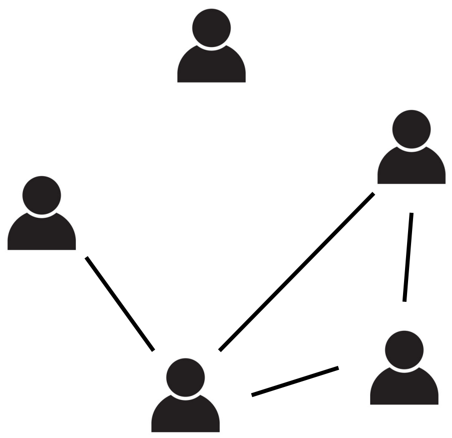


How do we formally study the nature and structure of these interactions?

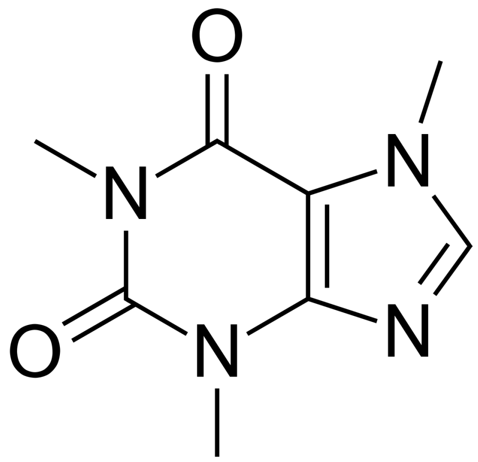
Graph theory

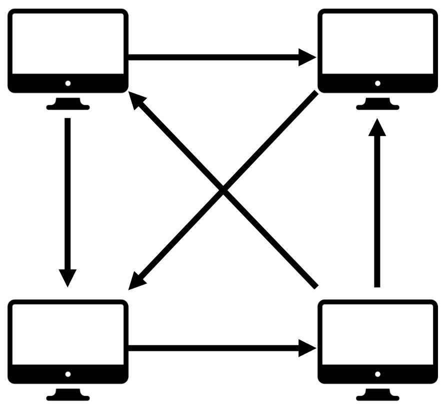
A **graph (network)** is a set of **nodes** connected by **edges**

Edges represent the **interaction or relationship** between two nodes



Disconnected graph

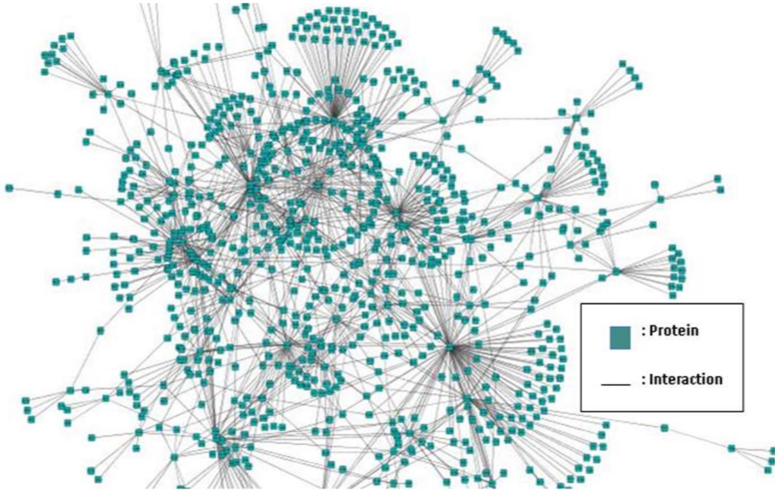


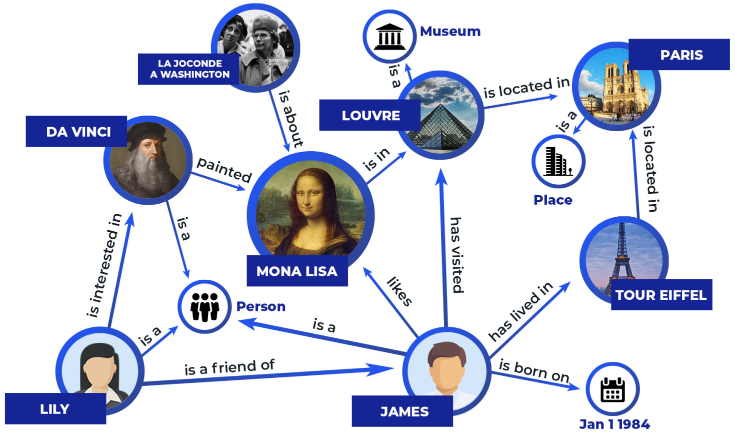


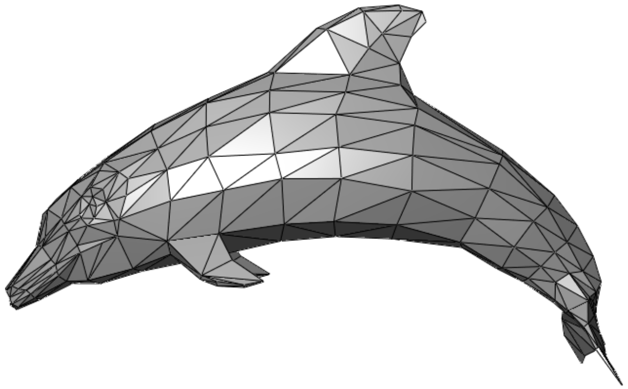
Directed graph

Degree of a node = # of edges connected to that node

Graphs are everywhere!





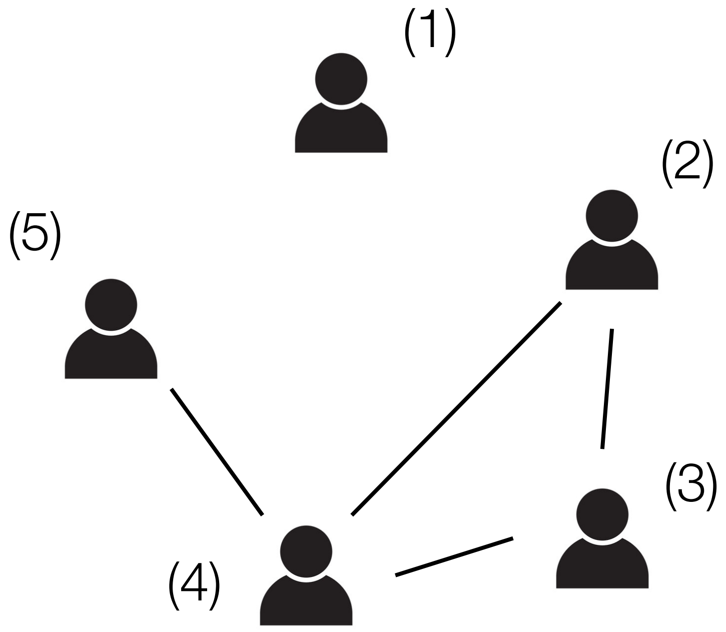


Can you think of any others?

Networks of intelligent agents

Analyzing network structure can help us:

- Identify the most influential agents
- Study properties of network formation
- Predict outcomes of multi-agent games



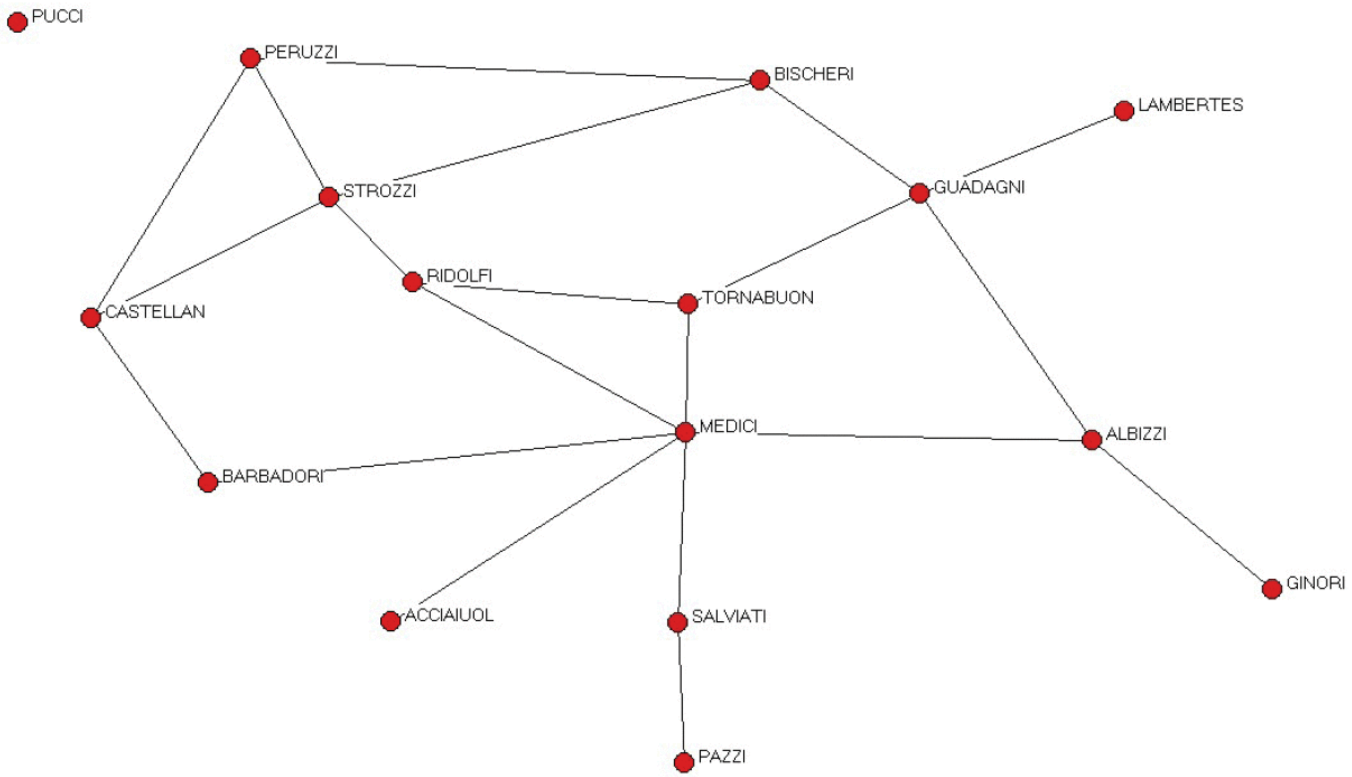
Which person is the most influential?

Power of the Medici family

The Medici family was one of the most powerful families in Renaissance Florence.

But they weren't always the wealthiest or most politically dominant. How did they gain power?

Padgett and Ansell (1993): **The key is their position in the Florence marriage network**



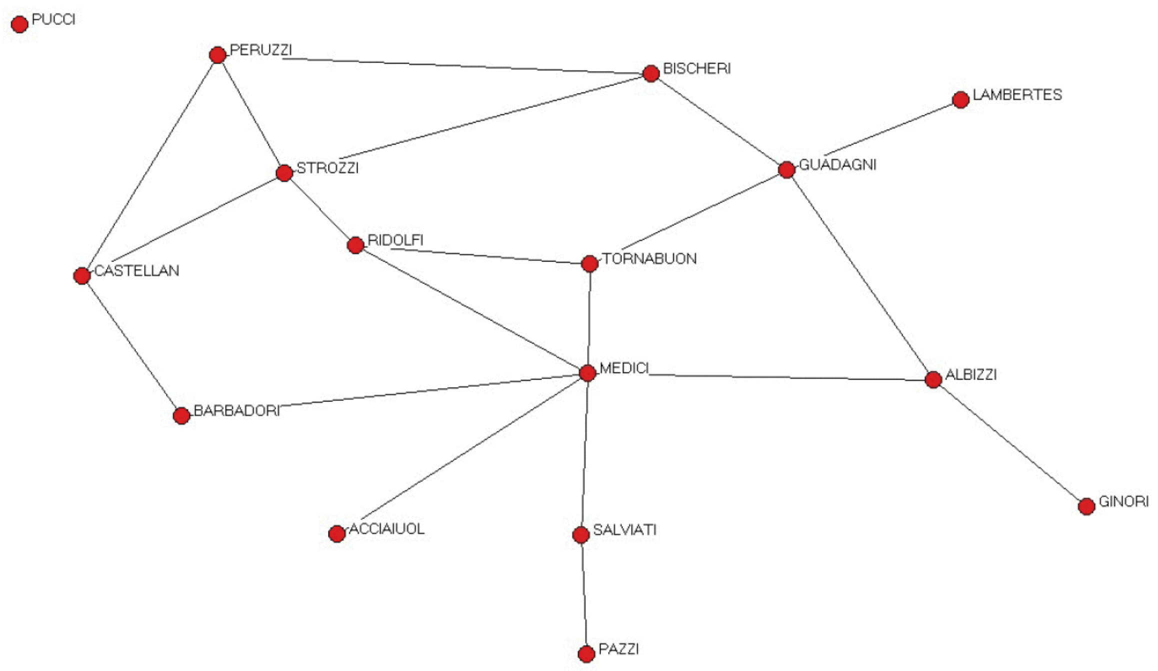
Metrics of node centrality

1) **Degree centrality**: $D_k := \frac{\deg(k)}{N-1}$
where N is the total # of nodes

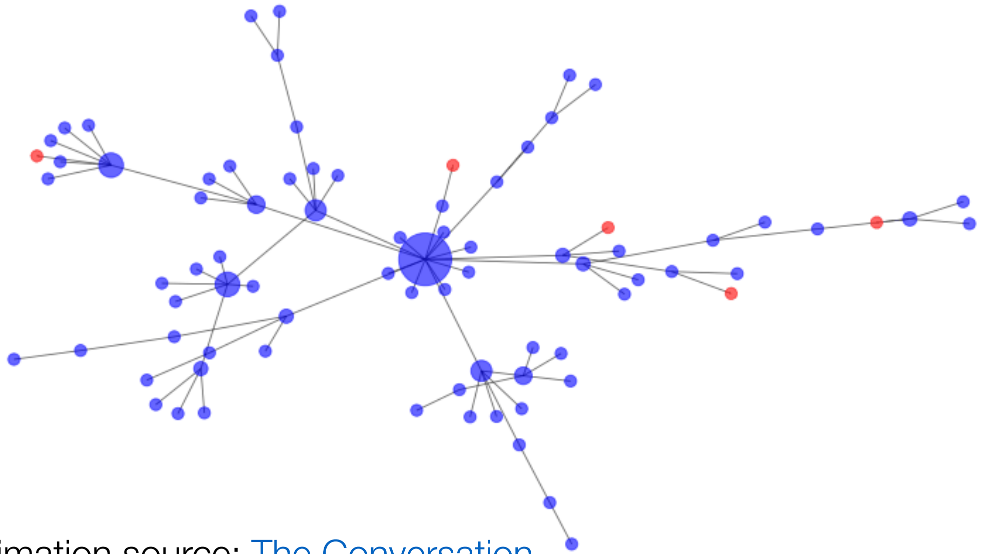
2) **Closeness centrality**: $C_k :=$ "reciprocal of avg length of shortest path between k and all others"
   - Let $\ell(i,j) :=$ length of shortest path from i to j
   - Then $C_k = \left(\frac{1}{N-1}\sum_{i\neq k}\ell(i,k)\right)^{-1} = \frac{N-1}{\sum_i \ell(i,k)}$

3) **Betweenness centrality**: $B_k :=$ "avg fraction of shortest paths that go through k"
   - Let $\mathrm{P}(i,j) :=$ # of shortest paths connecting nodes i to j
   - Let $P_k(i,j) :=$ # of shortest paths connecting nodes i to j through k
   - Then $B_k := \frac{1}{(N-1)(N-2)}\sum_{(i,j):i\neq j,\, k\neq i,j}\frac{P_k(i,j)}{P(i,j)}$

4) **Other more sophisticated metrics...**

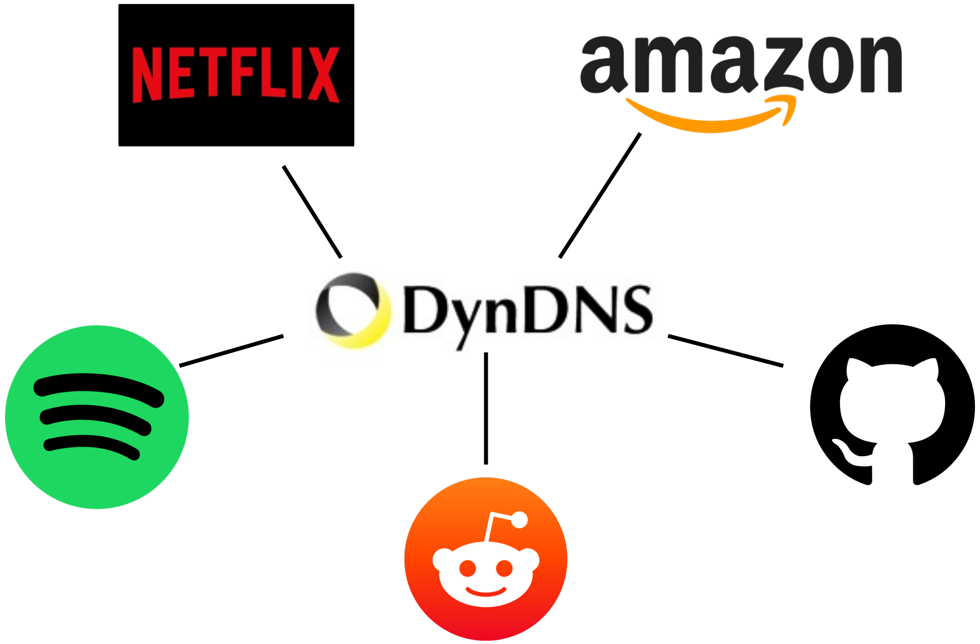


Is node centrality always a good thing?



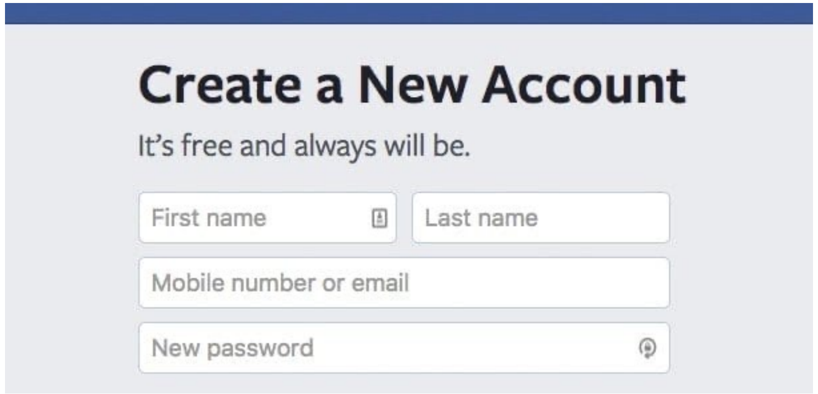
Animation source: [The Conversation]

Disease super-spreader



Systemic cyberattack risk

POV: you're making your first Facebook account



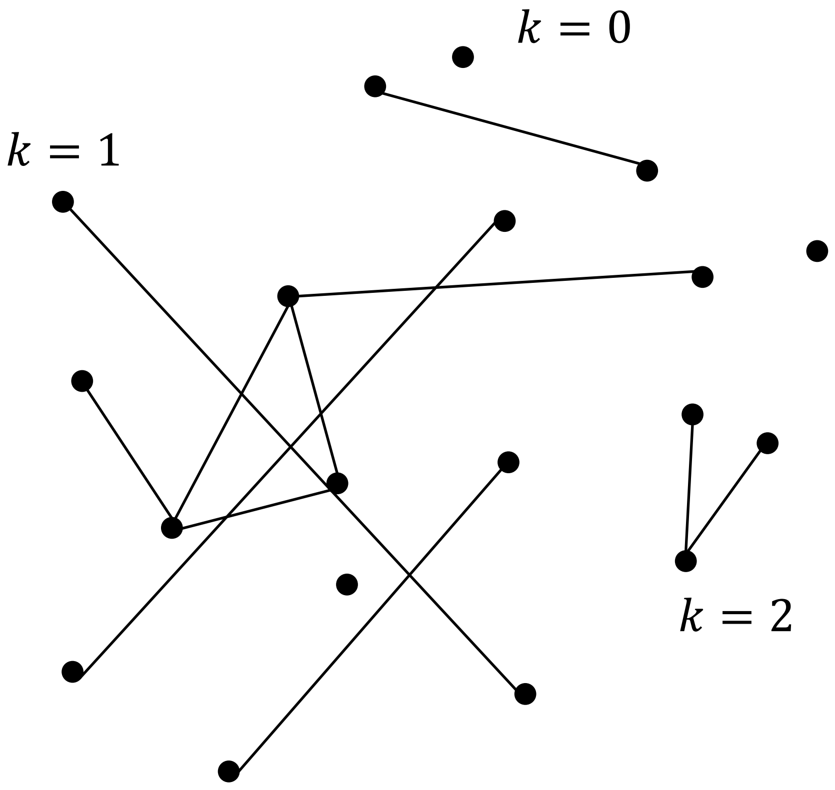
Who will you follow first? How do you branch out and find friends?

What will the structure of the social network eventually look like?

1. Erdos-Renyi random graphs

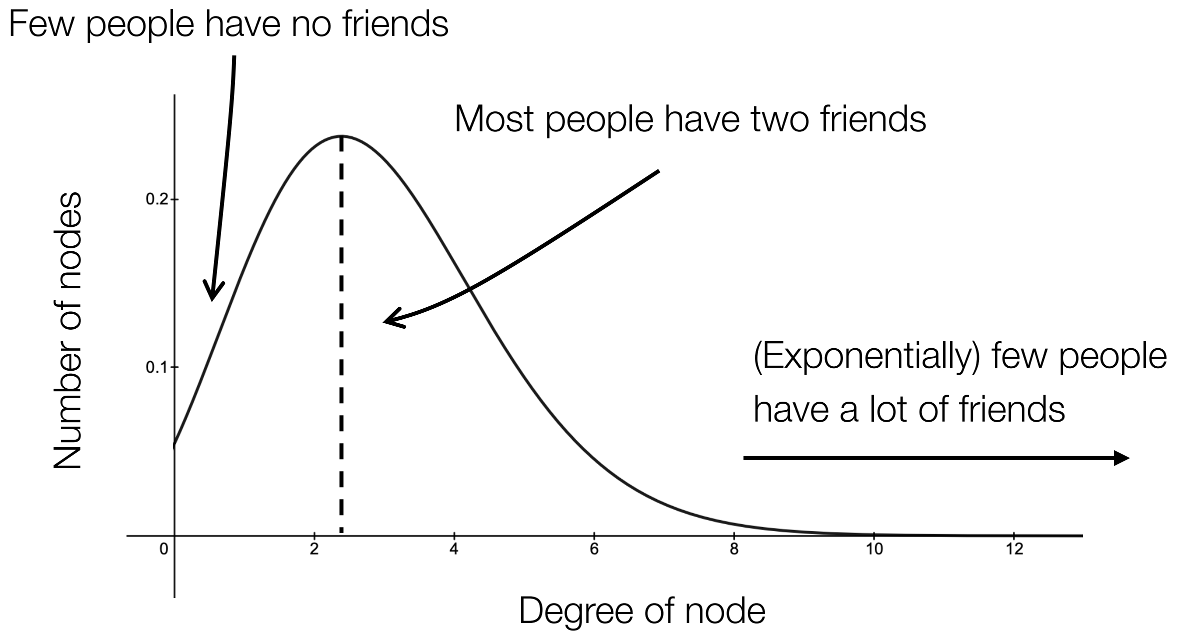
First, assume that you sent friend requests to everyone on Facebook randomly.

Given n nodes, each possible link forms with a random and independent probability p.



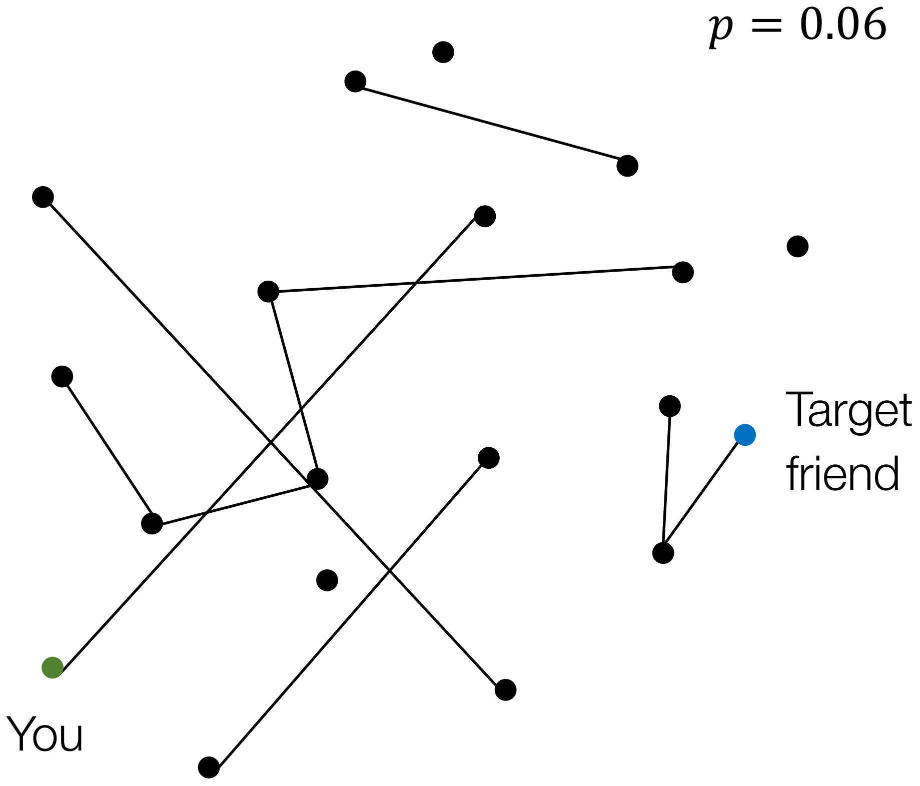
Expected number of links: $\mathbb{E}[\#\ of\ links] = \frac{n(n+1)}{2}\ \mathrm{p}$

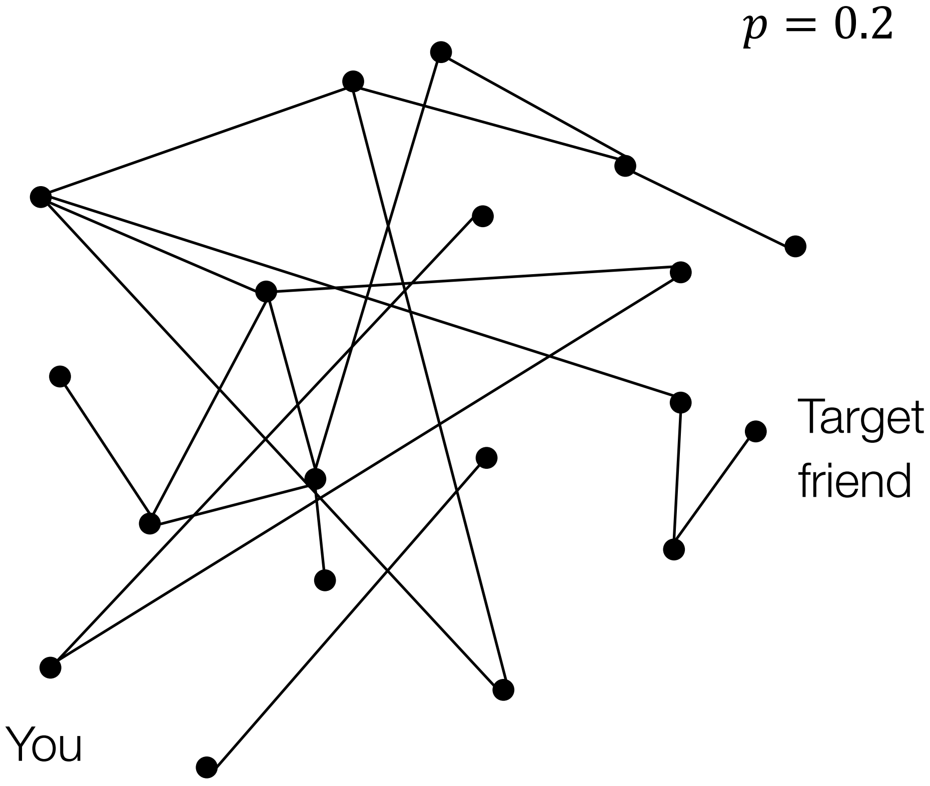
Degree distribution: $P[D = k] \approx \frac{e^{-\lambda}\lambda^k}{k!}$ $\quad \lambda = p(n-1)$



Emergence of a giant component

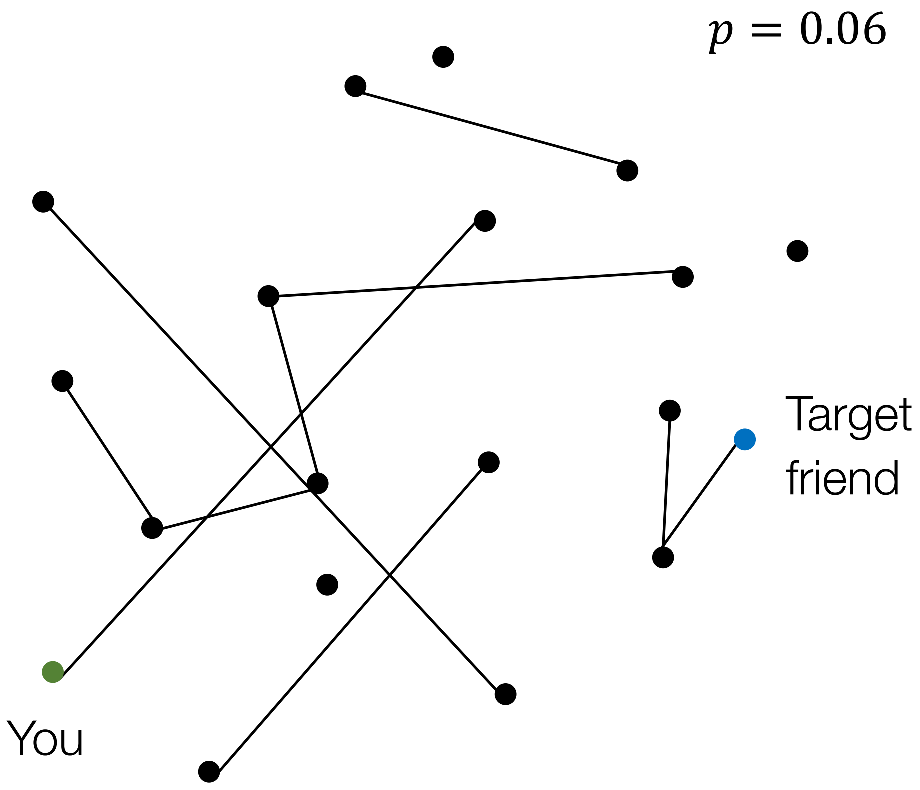
Given a node formation probability p, will you be able to reach any other person on the network just by going through mutuals?

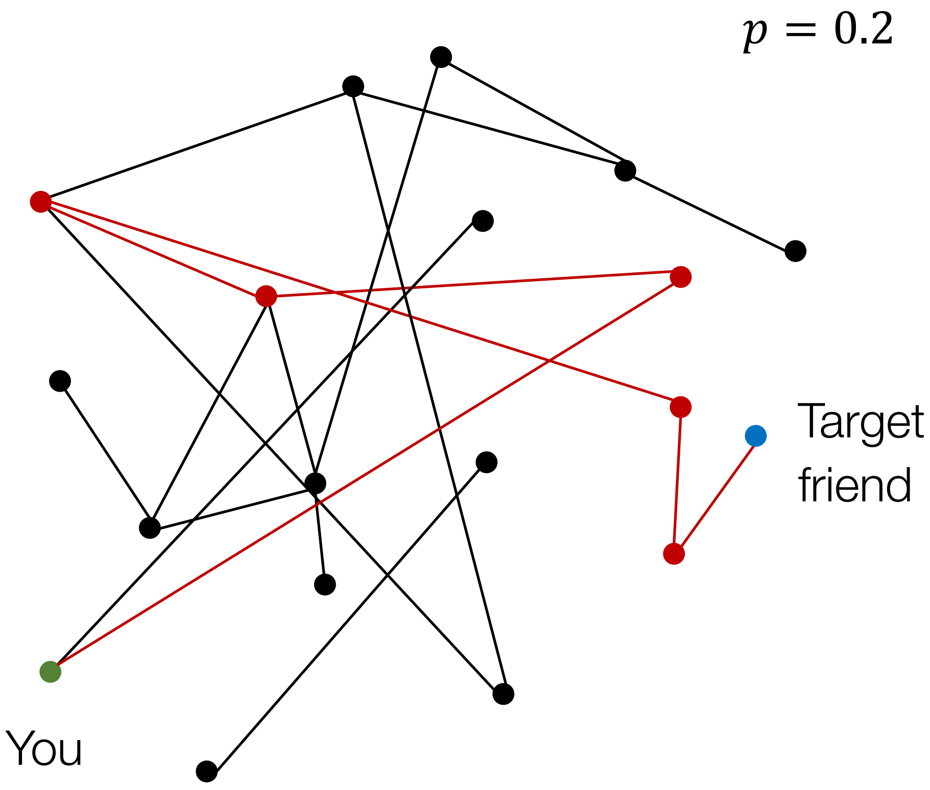




Emergence of a giant component

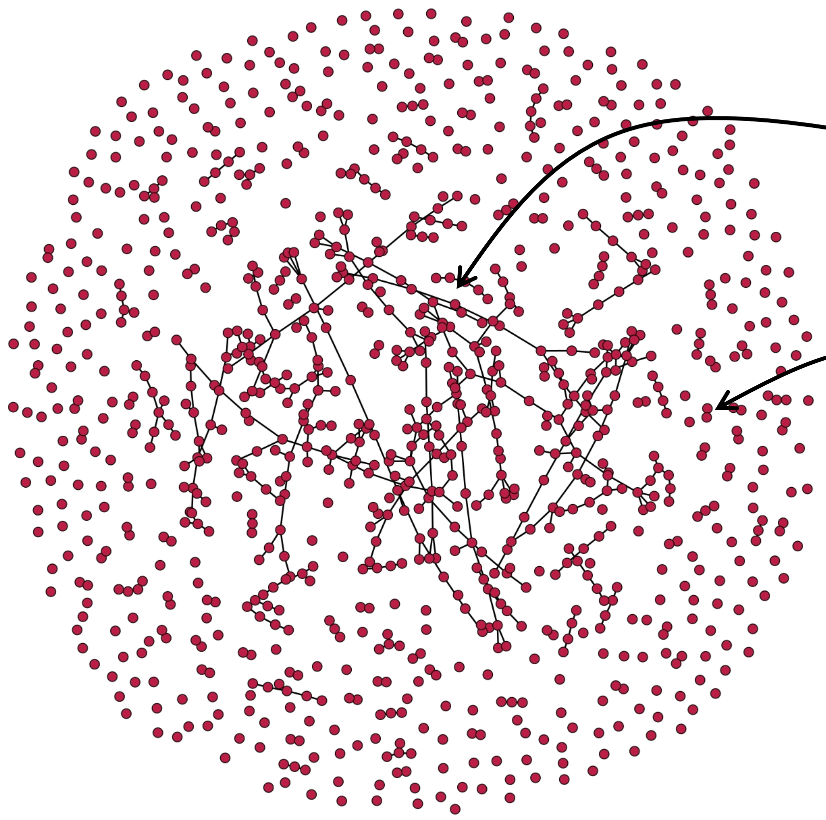
Given a node formation probability p, will you be able to reach any other person on the network just by going through mutuals?





Emergence of a giant component

Given a node formation probability p, will you be able to reach any other person on the network just by going through mutuals?

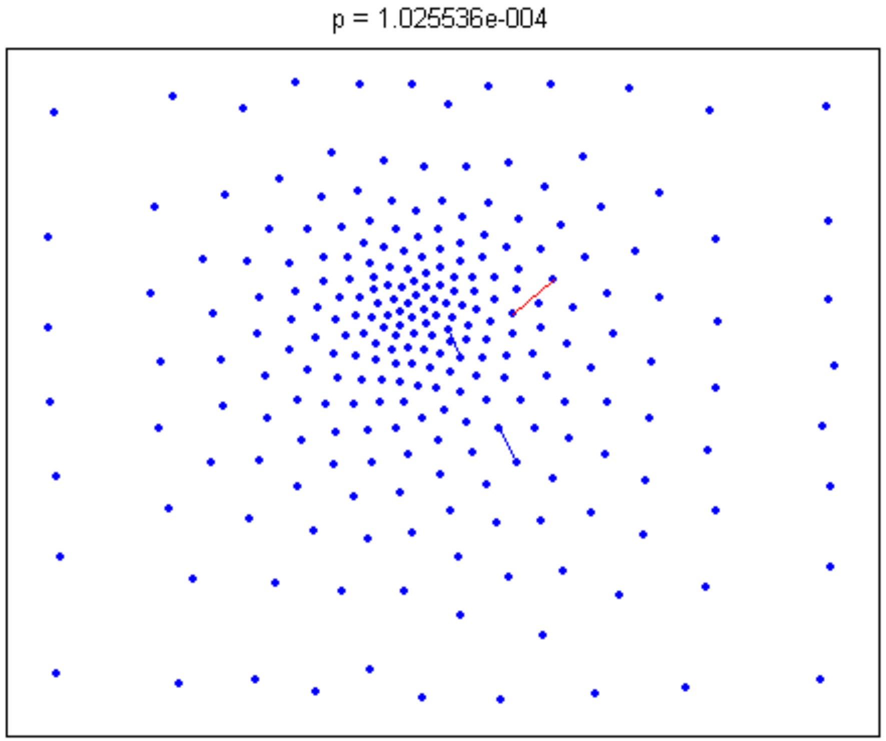


“Giant component”: everyone here can find each other via mutuals

People disconnected from the giant component

The bigger the giant component gets, the more likely it is that more links will attach to it, so it grows even bigger!

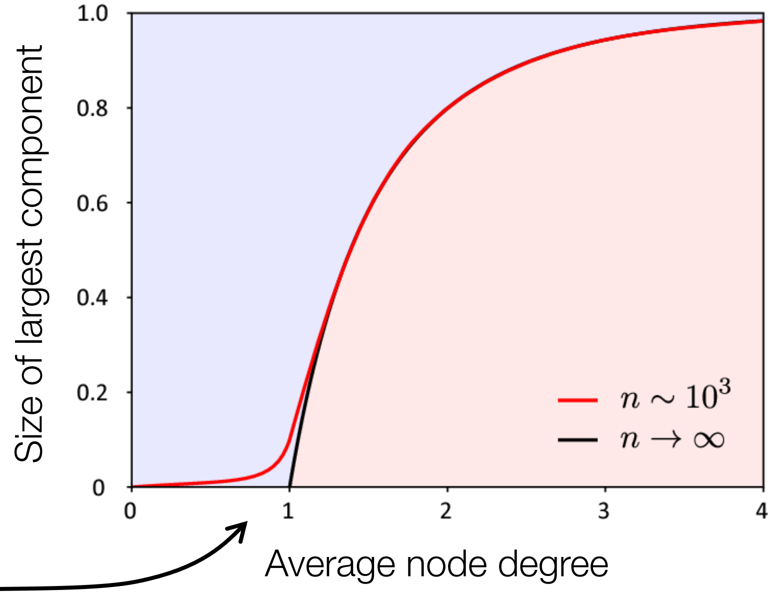
Phase transition



Animation by [David Gleich]

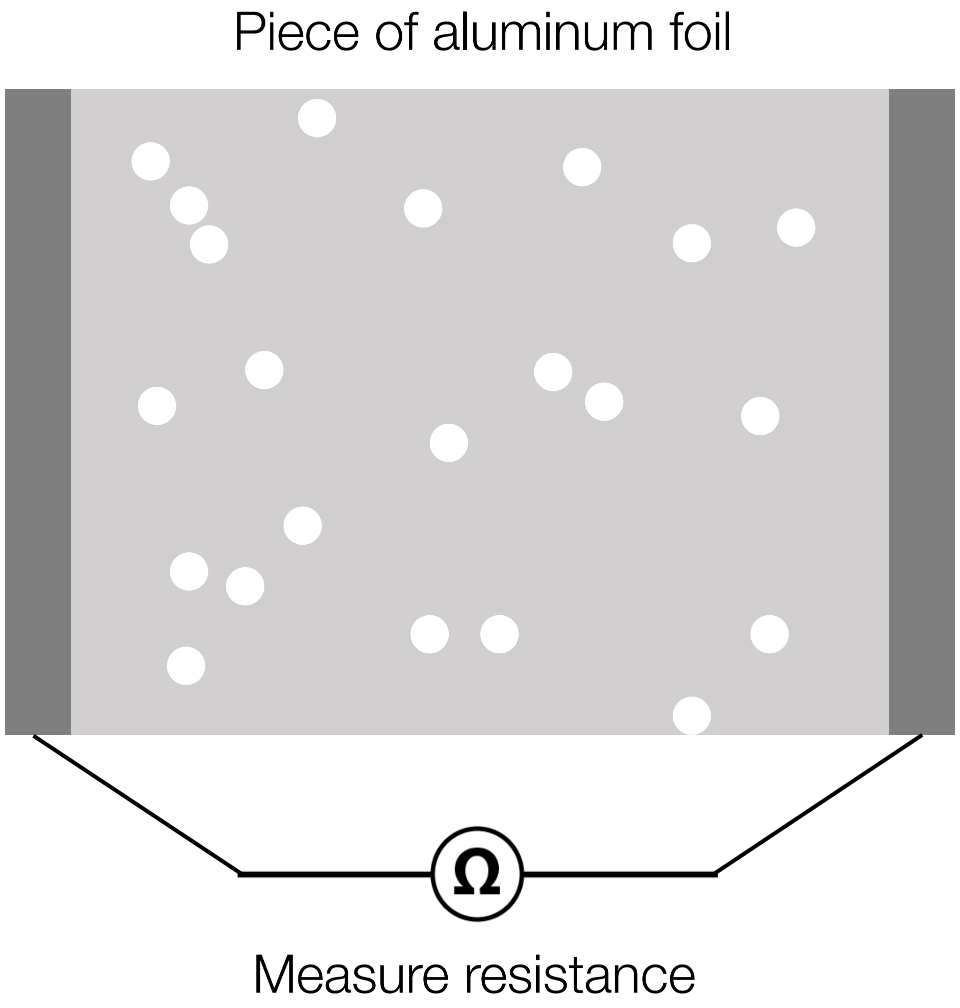
If $p < \frac{\log(n)}{n}$, P(fully connected) goes to 0 as $n \to \infty$.

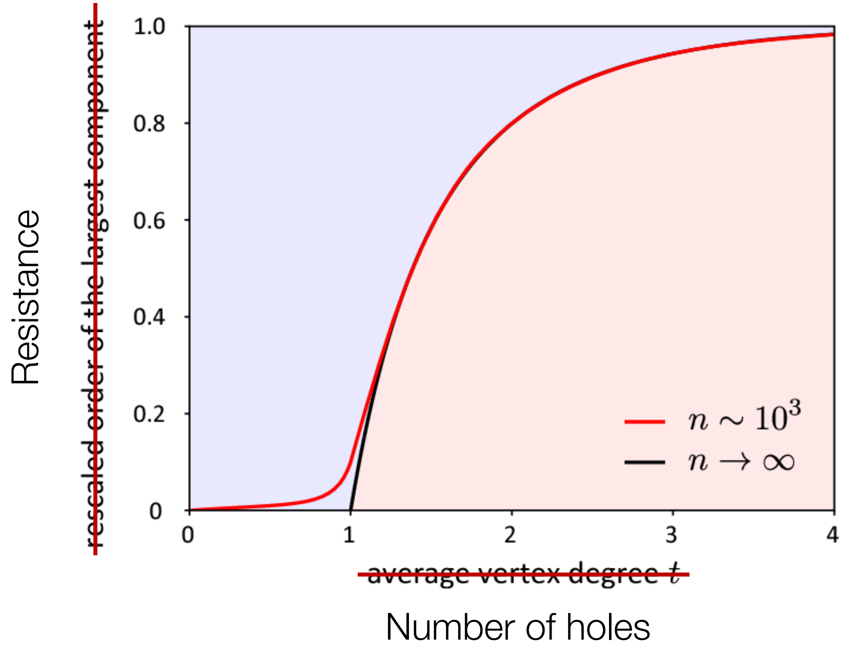
If $p > \frac{\log(n)}{n}$, P(fully connected) goes to 1 as $n \to \infty$.

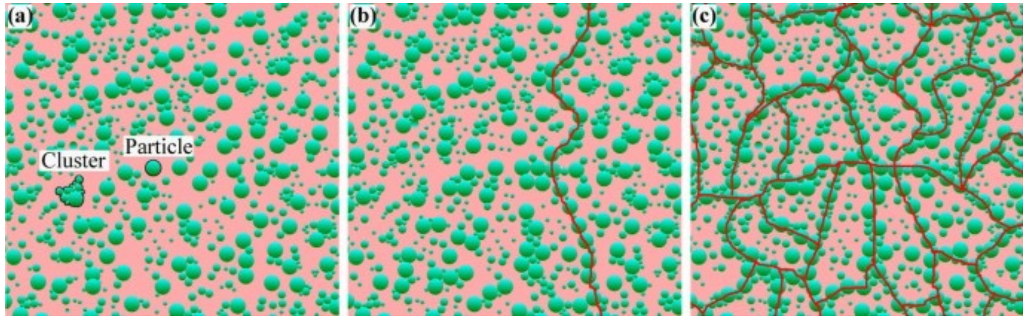


Discontinuous phase transition: the network suddenly gets fully connected past the threshold value.

(Side note: phase transitions happen everywhere)

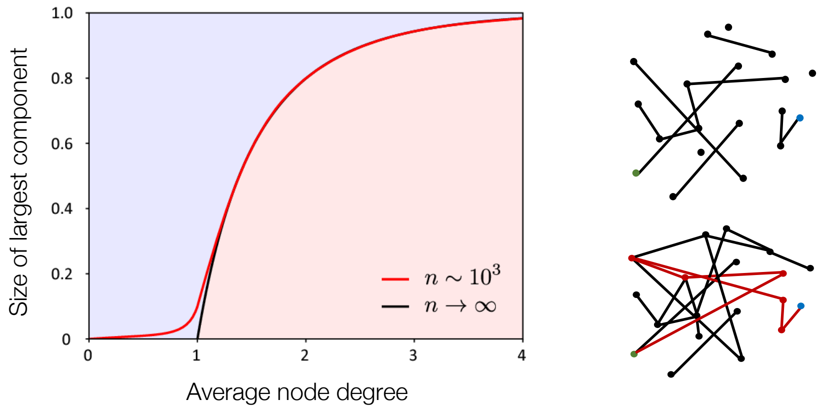




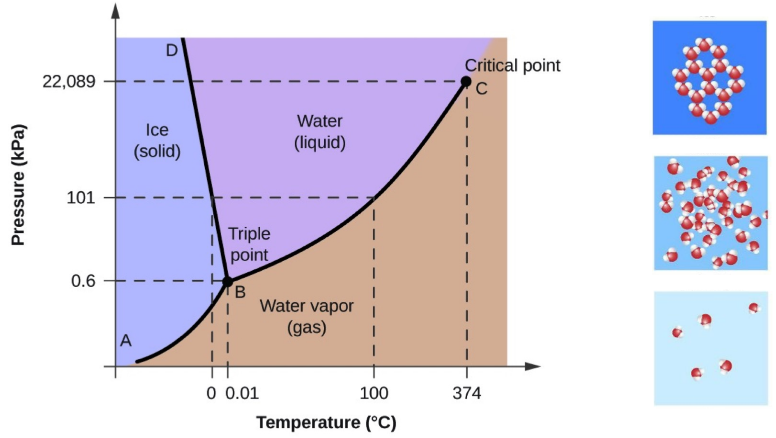


(Side note part II: you can get Nobel Prizes studying this!)

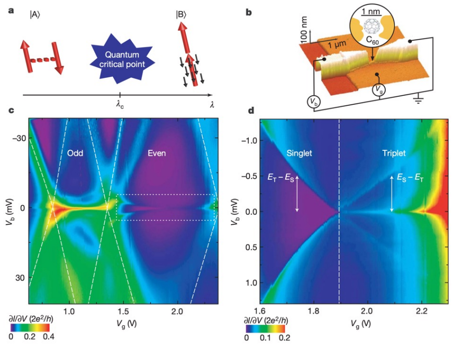
Emergence of a giant component



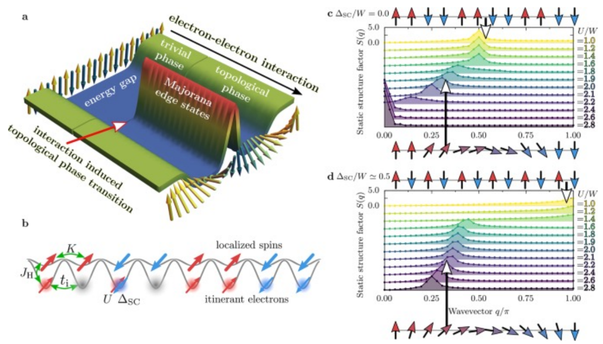
Solid-liquid-gas transition



Phase transitions in quantum matter

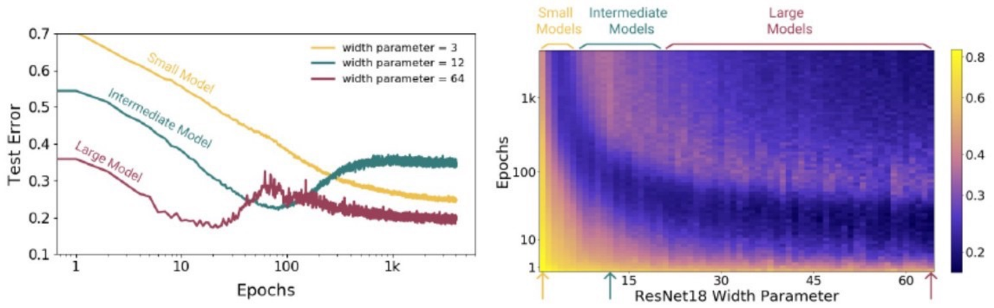


Quantum dots, Roch et al. (2008)



Topological states, Herbrych et al. (2021)

Phase transitions in machine learning

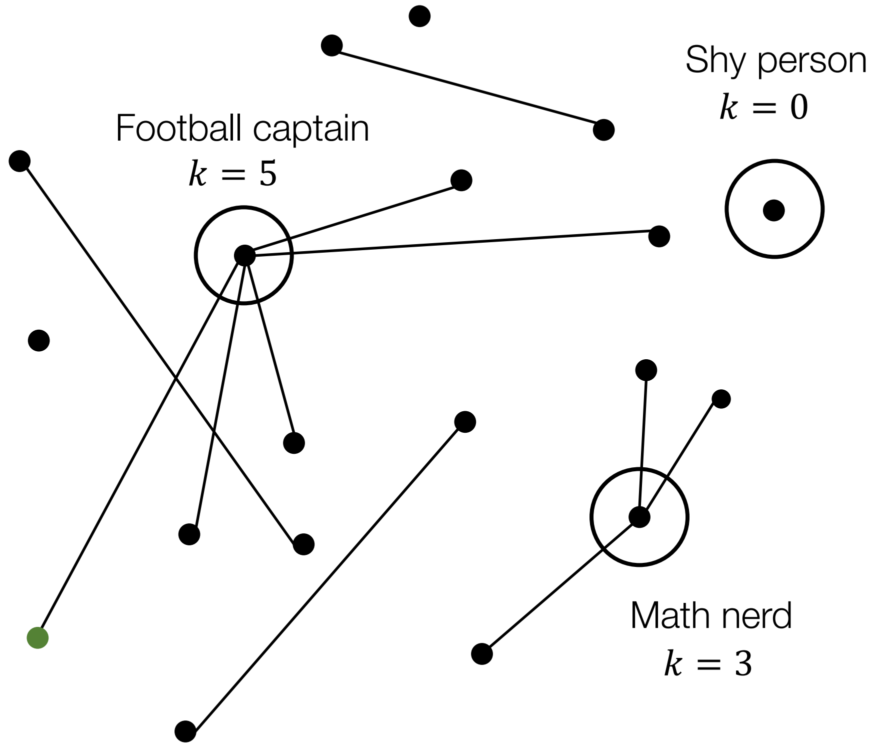


Deep Double Descent, (OpenAI, 2019)

2. Preferential attachment (Barabasi-Albert) model

Now instead of doing it randomly, you send friend requests to people who are more popular.

Starting with m_0 connected nodes. At every time step, add a new node which connects to old ones with a probability proportional to how well connected they are, $p_i = k_i / \sum_j k_j$



After t timesteps, there are $m_0 + t$ nodes and $m_0 + mt$ edges. The change in node degree is given by:

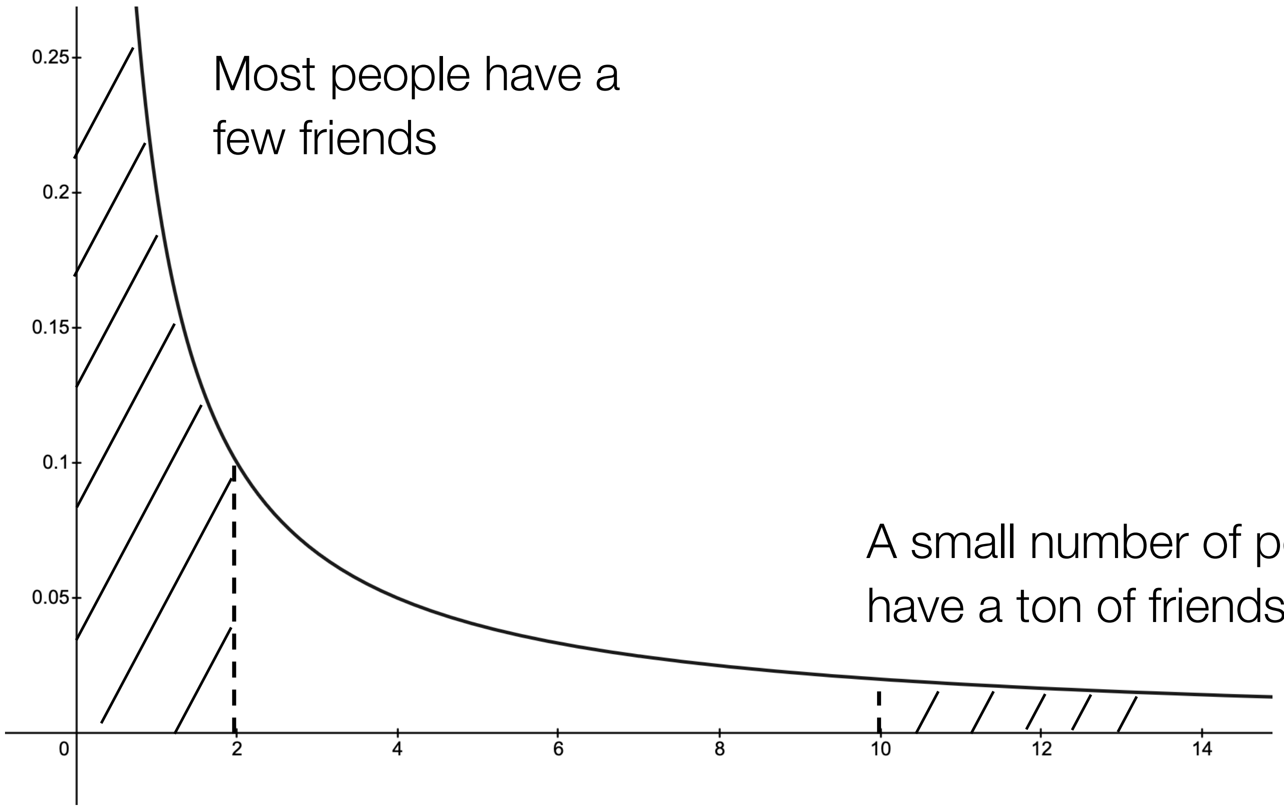
$$\frac{dk_i}{dt} = m\left(\frac{k_i}{\sum_j k_j}\right) = \frac{k_i}{2t - 1} \approx \frac{k_i}{2t} \qquad k_i(t) = m\left(\frac{t}{t_i}\right)^{1/2}$$

In other words, older nodes have a first-mover advantage. The rich get richer!

Salganik, Dodds, and Watts created 8 parallel music streaming site with 48 obscure songs which show the download count. Each ended up drastically different.

Compare and contrast

We saw that $k_i(t) = m\left(\frac{t}{t_i}\right)^{1/2}$. The number of nodes with $k < k_i$ is $t\left(\frac{m}{k}\right)^2$, so $P(k) = 1 - \left(\frac{m}{k}\right)^2$. Taking the derivative gives us the degree distribution.



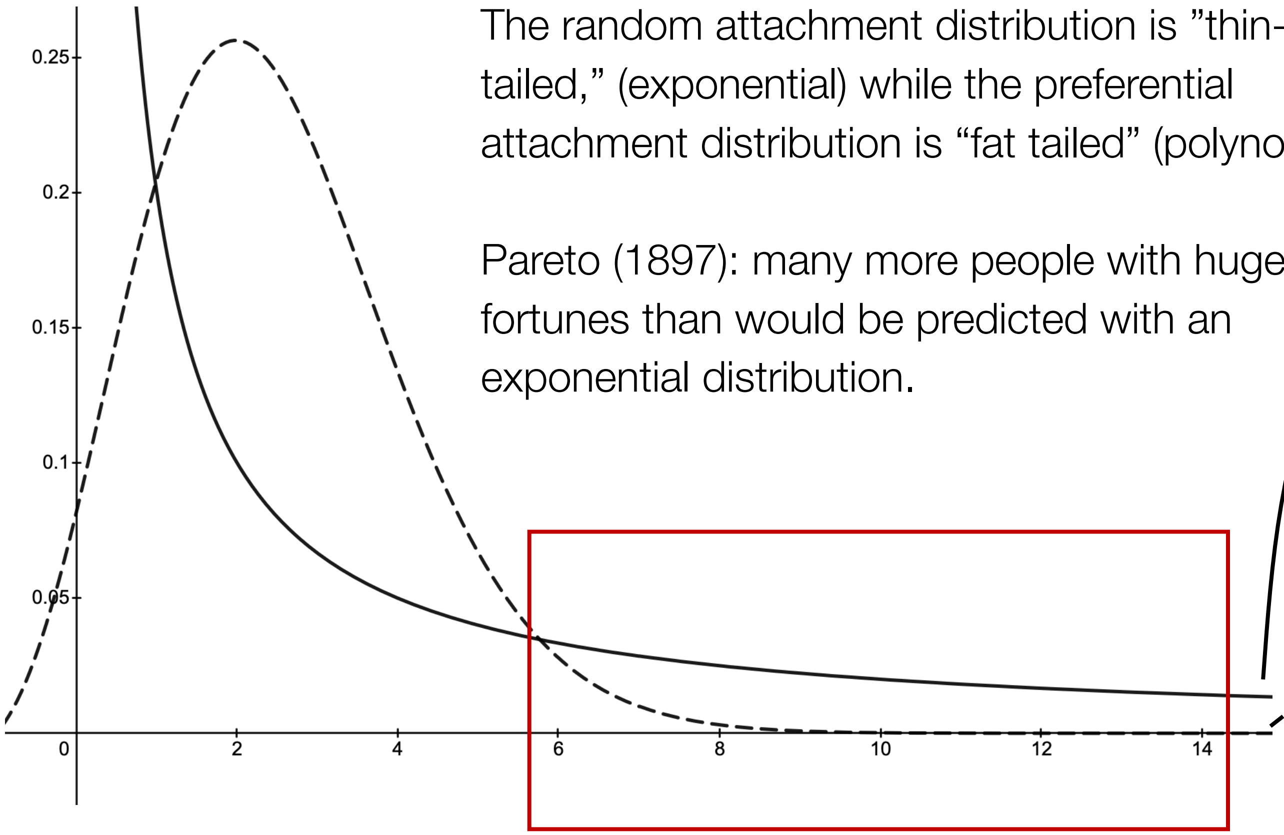
"Pareto distribution" / power law

Preferential attachment

$$p(k) = 2m^2k^{-3}$$

Compare and contrast

We saw that $k_i(t) = m\left(\frac{t}{t_i}\right)^{1/2}$. The number of nodes with $k < k_i$ is $t\left(\frac{m}{k}\right)^2$, so $P(k) = 1 - \left(\frac{m}{k}\right)^2$. Taking the derivative gives us the degree distribution.



The random attachment distribution is ”thin-tailed,” (exponential) while the preferential attachment distribution is “fat tailed” (polynomial).

Pareto (1897): many more people with huge fortunes than would be predicted with an exponential distribution.

“Pareto distribution” / power law

Preferential attachment

$$p(k) = 2m^2k^{-3}$$

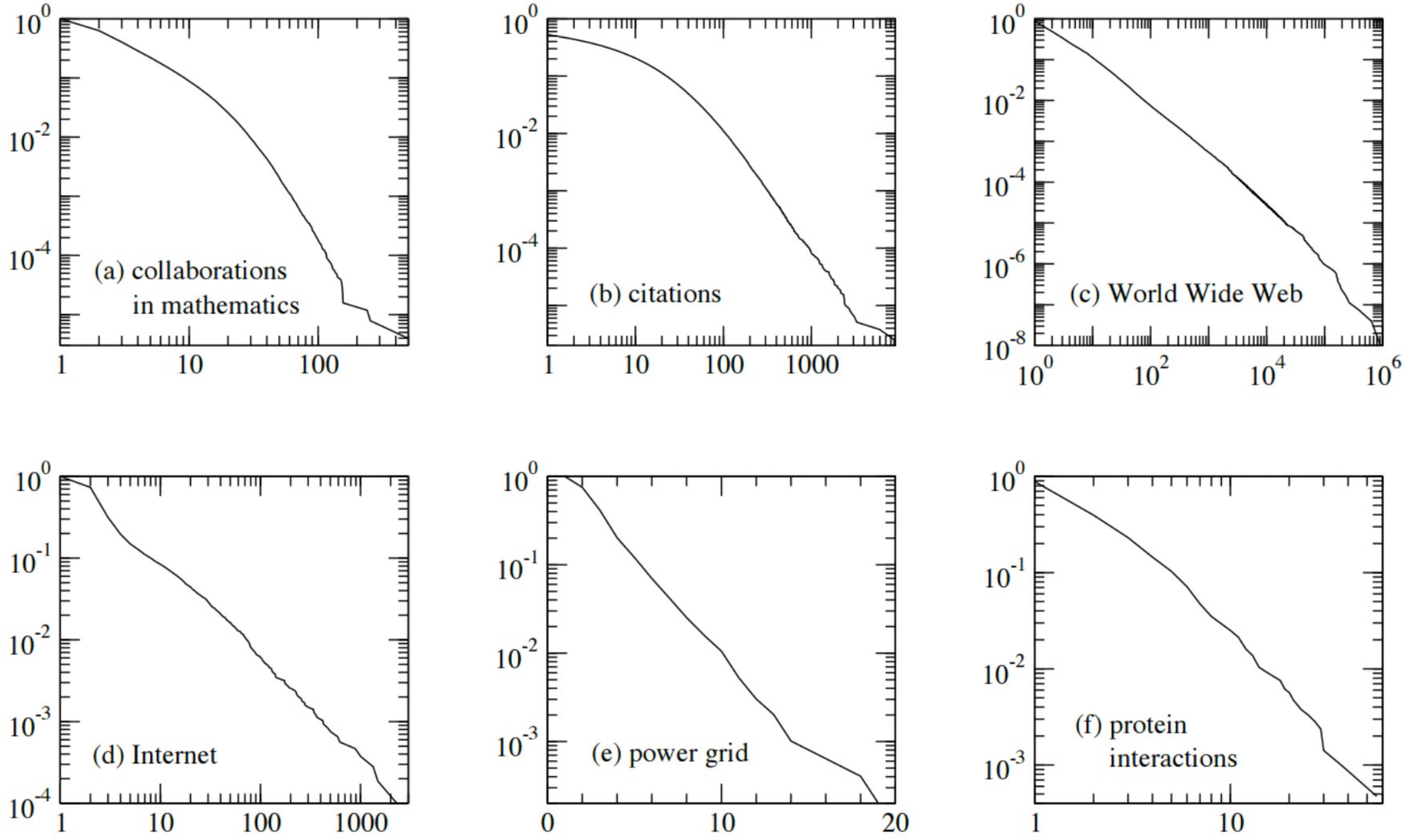
Poisson distribution

Random attachment

$$p(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

The universality* of power laws

$P(X > x) = cx^{-\alpha}, \log P(X > x) = \log(c) - \alpha \log x \longrightarrow$ A power law should be a straight line on a log-log plot



$$P(X > x) = cx^{-\alpha}$$

Power law distributions can be found* in:

- Income distributions
- Populations of cities
- Copies of genes in a genome
- Citation networks
- The structure of the internet

*With caveats: see Power law distributions in empirical data (2009), Scale-free networks are rare (2021)

3. Strategic network formation

Finally, you decide to be smart with friending people by maximizing your utility from the network.

There are n players who come together to form a network $\mathcal{G}$. Each player i receives a utility u_i as a result of some social/economic process that unfolds on the network*.

Just like in game theory, we need some notion of a "stable solution" (recall Nash equilibria). A common notion is pairwise stability: (i) no individual agent can gain by severing their link, (ii) no pair of agents can gain by linking up.

Consider a distance based utility function:

$$u_i = \sum_{j \neq i} b\left(l_{ij}(\mathcal{G})\right) - k_i(\mathcal{G})c$$

Benefit function

Distance between you and j

Cost of maintaining your links

*Yes, you can study this with reinforcement learning! See Yuan et al. (2018), Trivedi et al. (2020), Meirom et al. (2021).

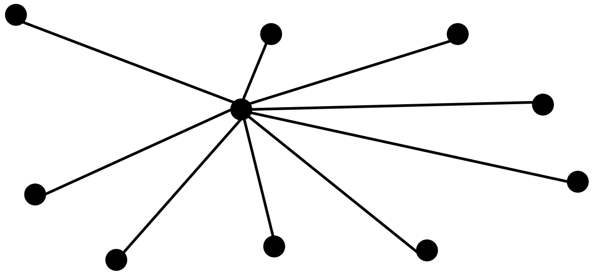
Strategic network formation

Everyone wants to have a big professional network

But no one likes "networking" or keeping up ties very much

$$u_i = \sum_{j \neq i} b\left(l_{ij}(\mathcal{G})\right) - d_i(\mathcal{G})c$$

Solution: a few "super networkers" take up the burden with networking with everyone.

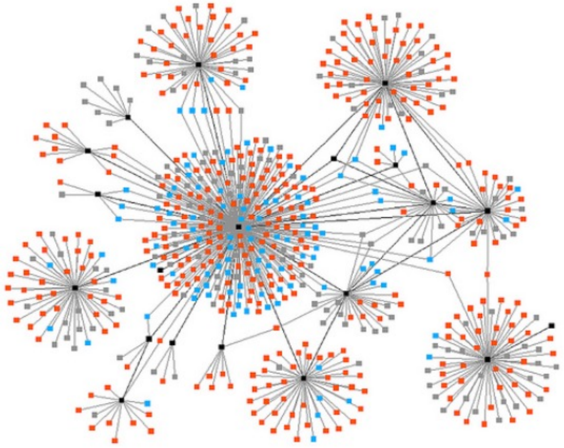


Stars are efficient if:

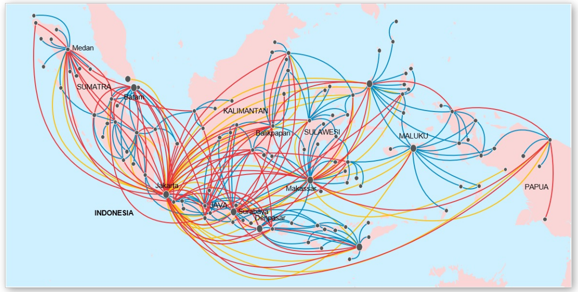
$$b(1) - b(2) < c < b(1) + \frac{n+2}{2} b(2)$$

Stars are pairwise stable if:

$$b(1) - b(2) < c < b(1)$$



Internet routing structure



Lion Air flight routes

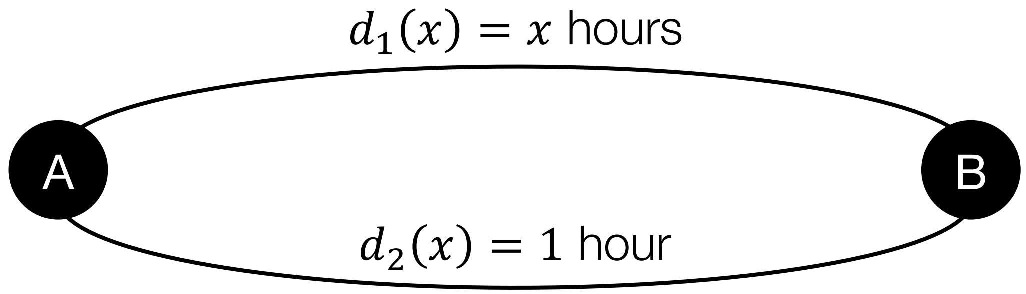
Routing games

Multiple agents need to get from point A to point B on a network
Each agent wants to minimize their own travel time given others' routes (i.e. avoid traffic)

Ex: Suppose n cars need to get from A to B in the below road network, consisting of 2 routes:

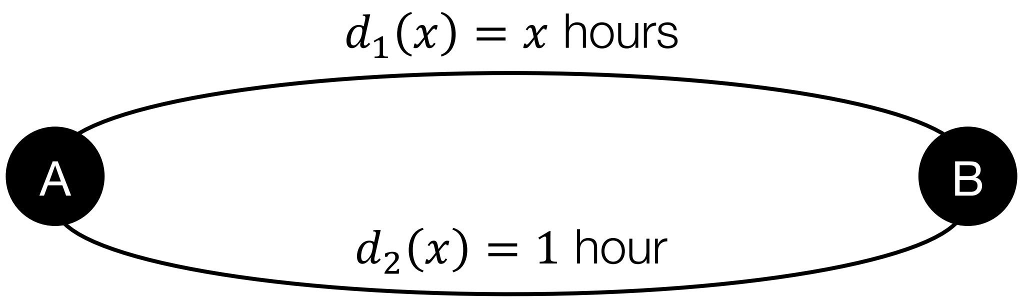
1) Delay on route 1 depends on the fraction of cars x taking that route: $d_1(x) = x$ hours
2) Delay on route 2 is constant and independent of traffic: $d_2(x) = 1$ hour

What is the **Pareto optimal** routing of cars that minimizes total delay?



Car traffic example

What is the **Pareto (socially) optimal** routing of cars that minimizes total delay?



If fraction x takes route 1, then total delay is

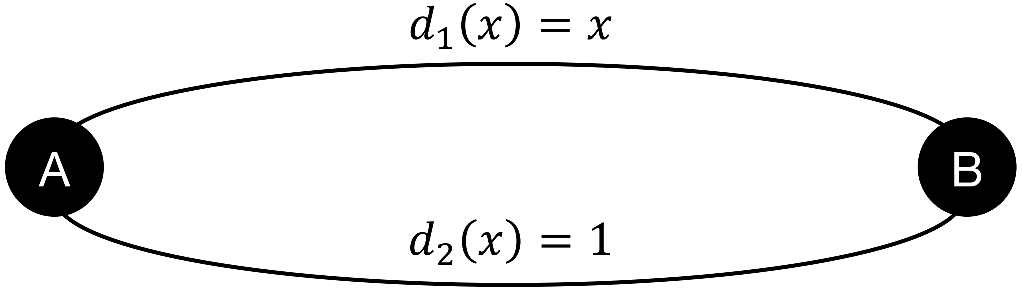
$$d_1(x) + d_2(x) = x(x) + (1 - x)(1) = x^2 - x + 1$$

This is minimized at $x = \frac{1}{2}$, which gives average delay of $\frac{1}{2}\left(\frac{1}{2}\right) + \frac{1}{2}(1) = \frac{3}{4}$ hours

The Pareto optimal routing is ½ of cars on route 1 and ½ on route 2

Car traffic example

What is the **Nash equilibrium** routing of cars (i.e. each driver minimizes their own delay)?



For any fraction of cars $x < 1$ taking route 1, $d_1(x) < d_2(x)$

Each driver will decide to take route 1, which gives average delay of $1(1) + (0)(1) = 1$ hour

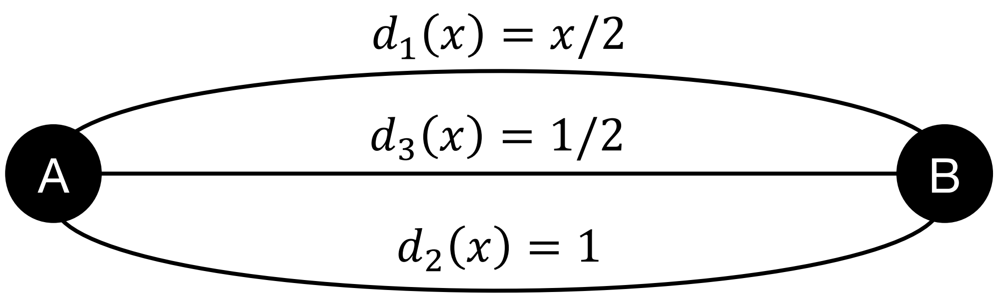
The Nash equilibrium routing is all cars on route 1

The Nash equilibrium routing results in higher delay than Pareto optimal

Making the Nash equilibrium more efficient

Option 1: Change network structure

Increase edge capacity or add more edges



Side note: adding routes can sometimes be harmful! See Braess's paradox

Option 2: Congestion pricing

Increase cost of low-delay routes

NYC will implement this starting 2023

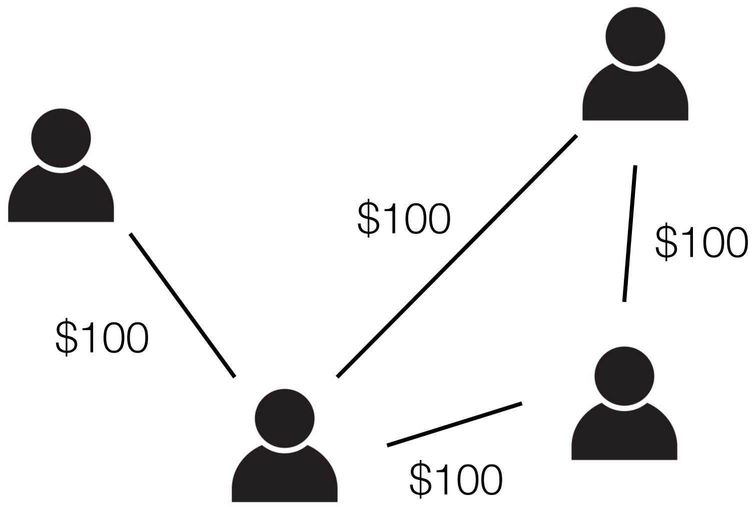
Bargaining games on networks

Social bargaining experiment (Lucas et al., 2001):

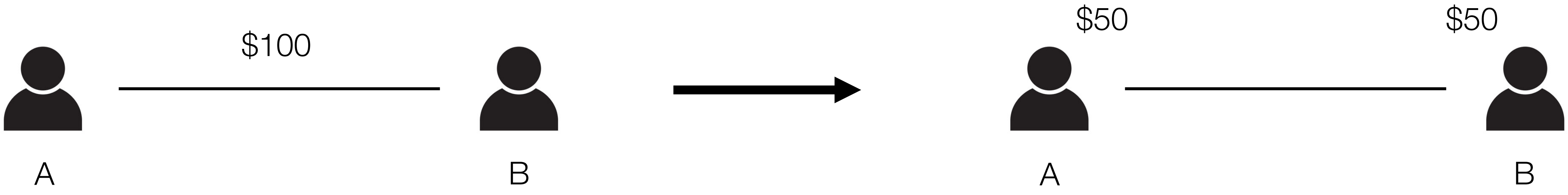
Individuals decide how to exchange fixed amount of $$ between themselves

Each node can only decide to take part in an exchange with **one neighbor**

They have to make a decision in a **fixed amount of time**

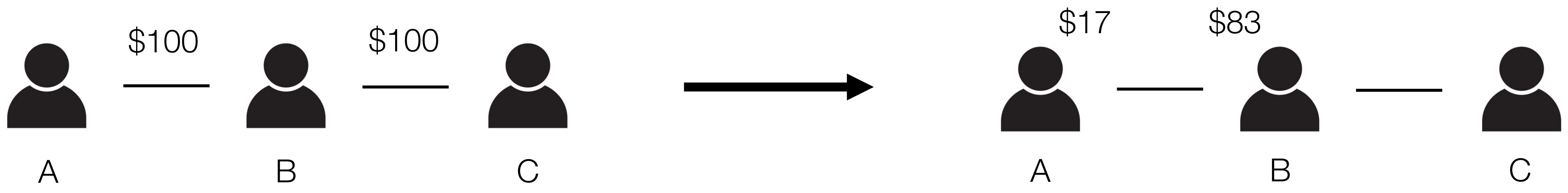


Bargaining on 2-node network



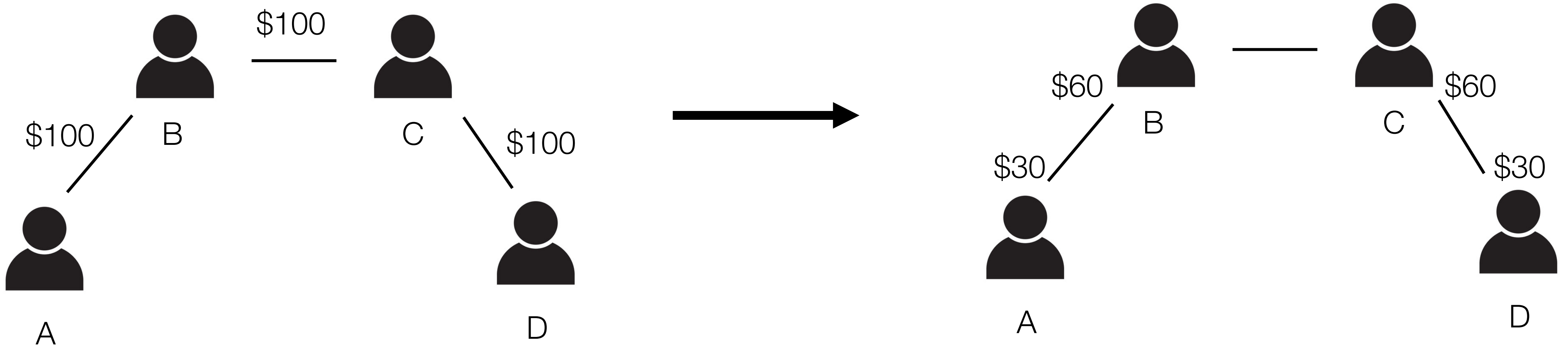
Most people agree to split 50-50

Bargaining on 3-node line network



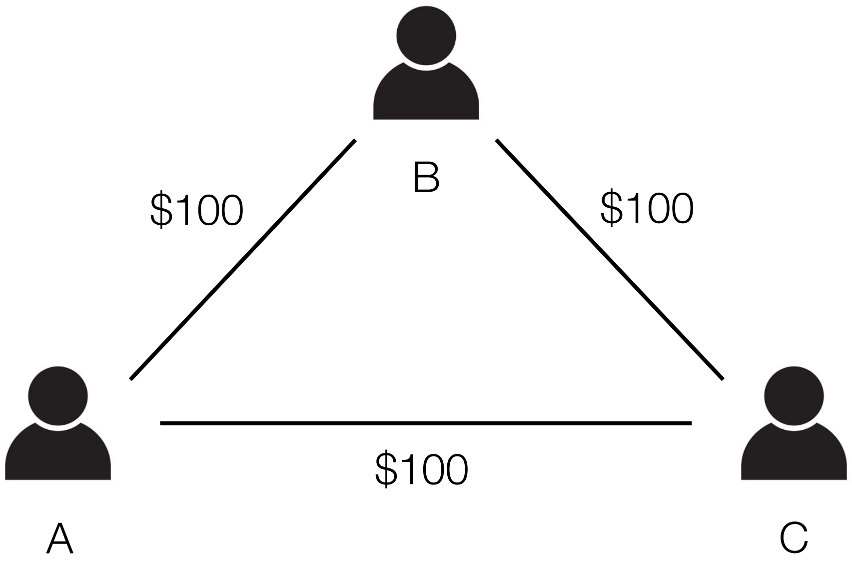
B can charge A more by threatening to trade with C instead

Bargaining on 4-node line network



B isn't as powerful since C might decide to trade with D instead

Unstable bargaining networks



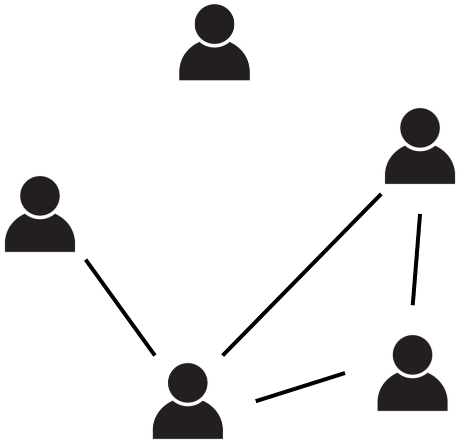
Since there can only be 1 trade, the node that is excluded will try to offer a more favorable deal

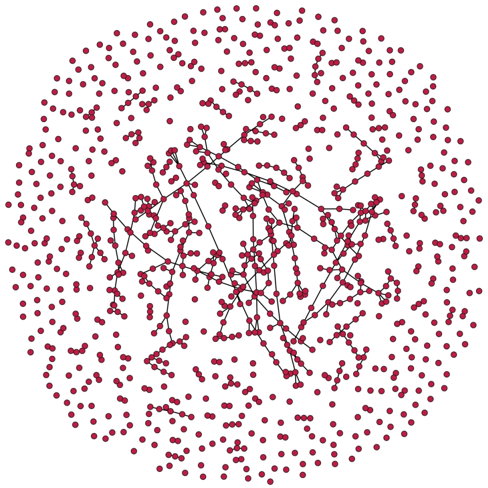
But then another node is excluded and offers an even more favorable deal

Cycle only ends due to fixed time limit: **outcome is unpredictable**

Recap

- Graph theory
- Measures of node centrality
- Graph formation and phase transitions
- Games on networks (routing)
- Networked games (bargaining)





Lecture 1	Lecture 2	Lecture 3	Lecture 4	Lecture 5	Lecture 6
Introduction and the RL problem	How computers learn	How people learn	Multi-agent systems	Interactions on graphs	Complex systems science

References and additional resources

- Network Science by Albert-Laszlo Barabasi
- MIT 6.207/14.15 Networks lecture notes
- Braess's paradox wiki
- Systemic Risk and Stability in Financial Networks by Acemoglu, Ozdaglar, and Tahbaz-Salehi
- Lines of Power in Exchange Networks by Jeffrey W. Lucas, C. Wesley Younts, Michael J. Lovaglia, and Barry Markovsky
- Math Doesn't Have to be Boring: the Pokemon Type Network video by Not David
- For those interested in ML on graph-structured data, check out A Gentle Introduction to Graph Neural Networks