Social and Information Networks

CSCC46H, Fall 2022
Lecture 12

Prof. Ashton Anderson
ashton@cs.toronto.edu

Today

Voting
Summary
Emphasis on final help

Today

Final info:
Tuesday, Dec 13 7–10pm in IC130

Today

Missed a blog post? Submit by this Friday and email Richard and Conroy with a link to the post, your utorid, and which blog post you missed (1 or 2)

Late penalty will apply

Voting

Why have voting?

Synthesize the preferences of a group

Aggregate information, preferences, beliefs, decisions

Voting on:

- Candidates
- Laws
- Verdicts for trials
- Awards

Simple example

Say you want to pick the fairest outcome for the group

Everyone has their preferred number (e.g. price)

What should you do?

Easy…take the average

Why fair?

Minimizes the squared loss

Why voting is hard

But in many situations there is no natural **"average"**!

Voting on:

- Candidates
- Laws
- Verdicts for trials
- Awards

Averaging fails here…

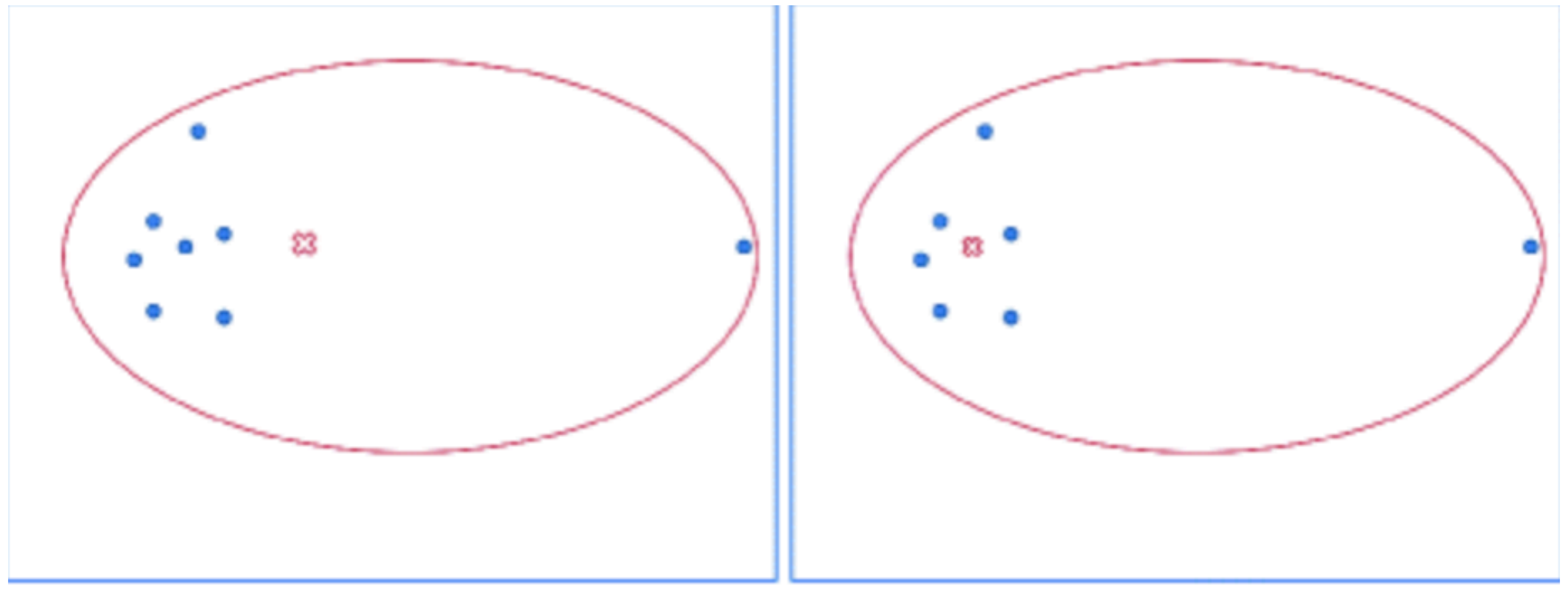
Why voting is hard

Often need to pick a **single winner** that becomes **binding for the group**

- President
- Award-winner
- Policy decision

Voting as **group decision making**

Parallels to clustering: finding the centre vs finding the "medioid"—the best representative element



Mean Medioid

Individual preferences

We want to **aggregate many individuals' preferences**

What are individual preferences?

Setup: a group of k people are evaluating a finite set of possible *alternatives*

Individual preferences

The people want to produce a single **group ranking** that orders the alternatives from best to worst

The ranking should **reflect the collective opinion of the group**

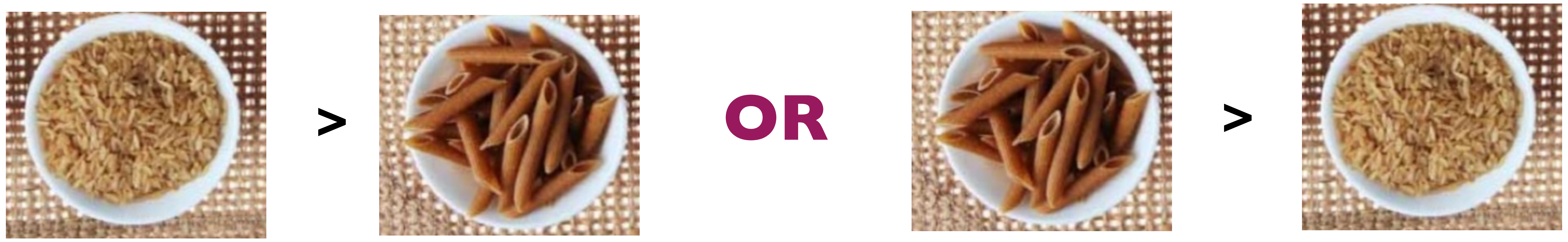
The challenge: how do we define what it means to reflect multiple, potentially contradictory opinions?

Individual preferences

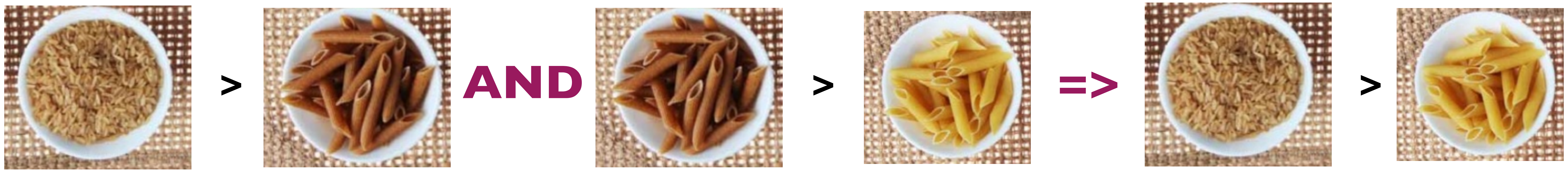
Every person has a preference relation over the alternatives, denoted $>_i$ for player i

Must satisfy two properties:

Complete: all pairs of distinct alternatives X and Y, either $X >_i Y$ or $Y >_i X$



Transitive: if $X >_i Y$ and $Y >_i Z$ then $X >_i Z$



Individual preferences

A way to think about preference relations: as a **graph**

Nodes: alternatives

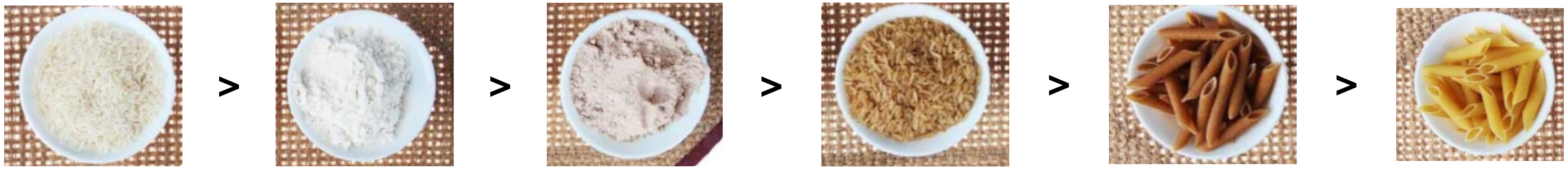
Directed edges: Y⟶X if X $>_i$ Y

(complete and transitive example)

Individual preferences

Another way of expressing preferences: ranked list

For example:



Ranked list → preference relation

Obviously complete and transitive

Preference relation → ranked list

Less obvious but still true

Individual preferences

Claim: Ranked list → Preference relation

Proof:

A ranked list is complete, since for any pair of alternatives X and Y, either X>Y or Y>X

A ranked list is transitive, since if X is higher than Y and Y is higher than Z, then X is also higher than Z.

Individual preferences

Claim: Preference relation → ranked list

Proof:

Identify the alternative X that wins the most pairwise comparisons

Claim: X actually beats **every** other alternative

Why? Suppose Y $>_i$ X. Then Y would beat everything X beats (by transitivity), and also X. Therefore beats more than X. **Contradiction!**

Put X at the top of the list, remove it from the set of alternatives, and recurse

- Relation is **still complete and transitive** over remaining alternatives
- Construct a list by **repeatedly finding the alternative** that beats everyone else

Individual preferences

Summary:

Preference relation → Ranked list

Ranked list → Preference relation

Therefore preference relations and ranked lists are equivalent!

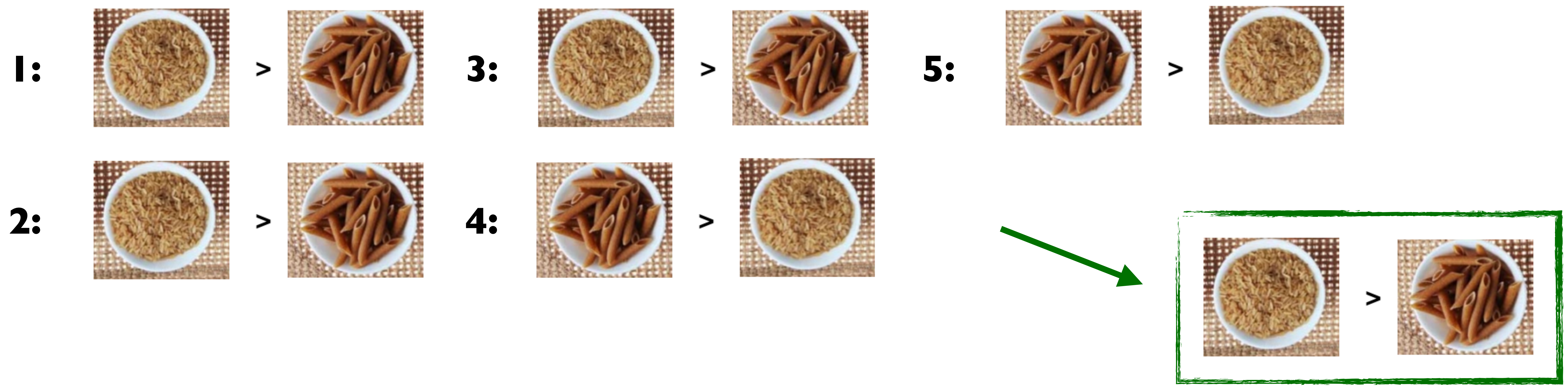
Voting Systems

<u>Voting system</u>: a method that takes a set of complete and transitive individual preference relations (or ranked lists) and **outputs a group ranking**

When there's only two alternatives, what should we do?

Majority Rule: whoever is preferred by a majority of the voters wins, other one is second

(let k be odd to avoid ties)



Majority Rule

Easy enough, what about majority rule with more than two alternatives?

What's a natural way to extend it?

Majority rule on every pair of alternatives: X > Y if a majority of voters have $X >_i Y$

Is this complete?

Everyone has a preference for every pair, and there's always a majority (assume k is odd). So this is **complete**

Is this transitive?

Majority Rule

Is majority rule on at least 3 alternatives transitive?

1:  $>_1$  $>_1$ 

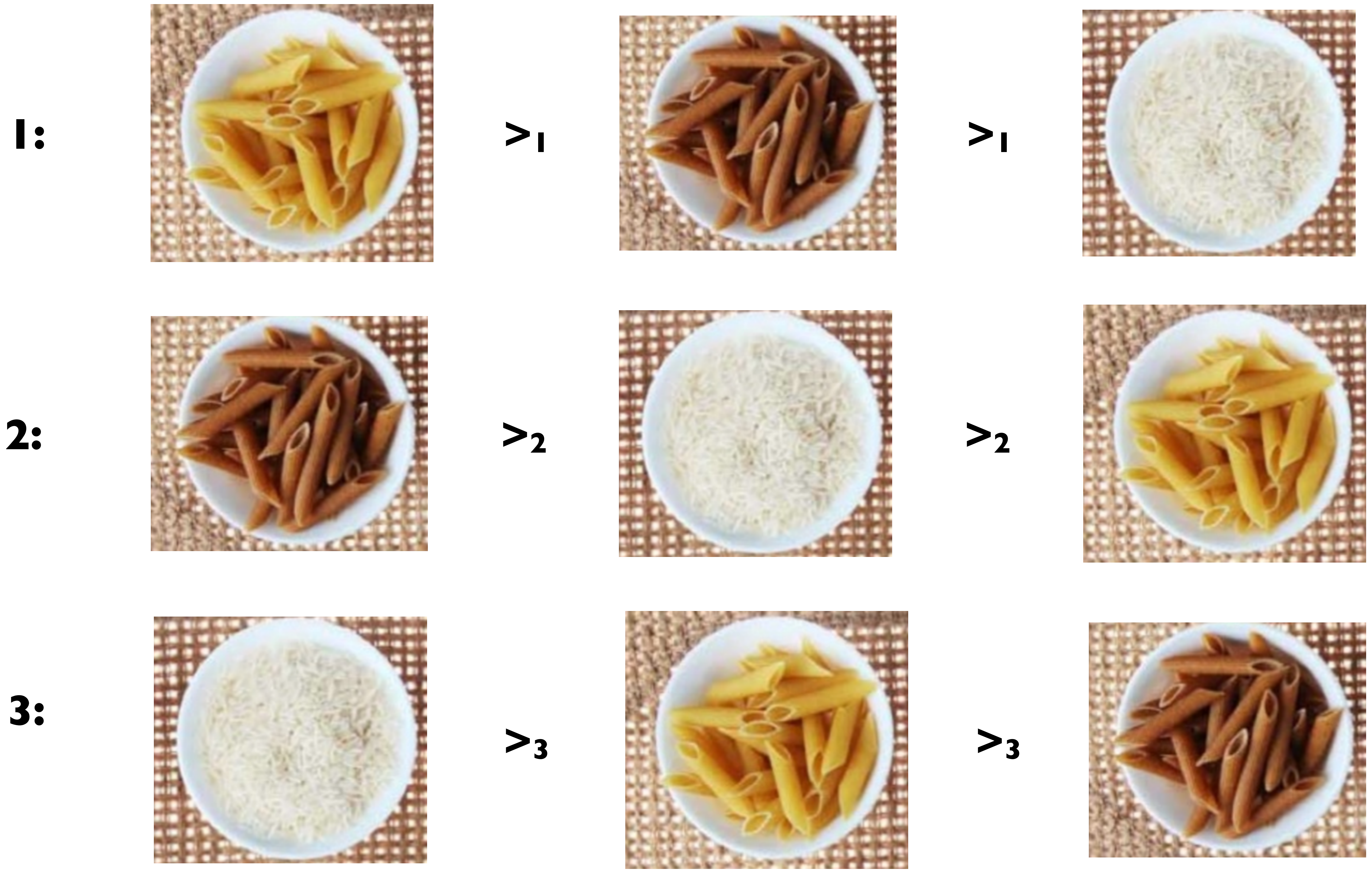
2:  $>_2$  $>_2$ 

3:  $>_3$  $>_3$ 

What does majority rule do here?

Majority Rule

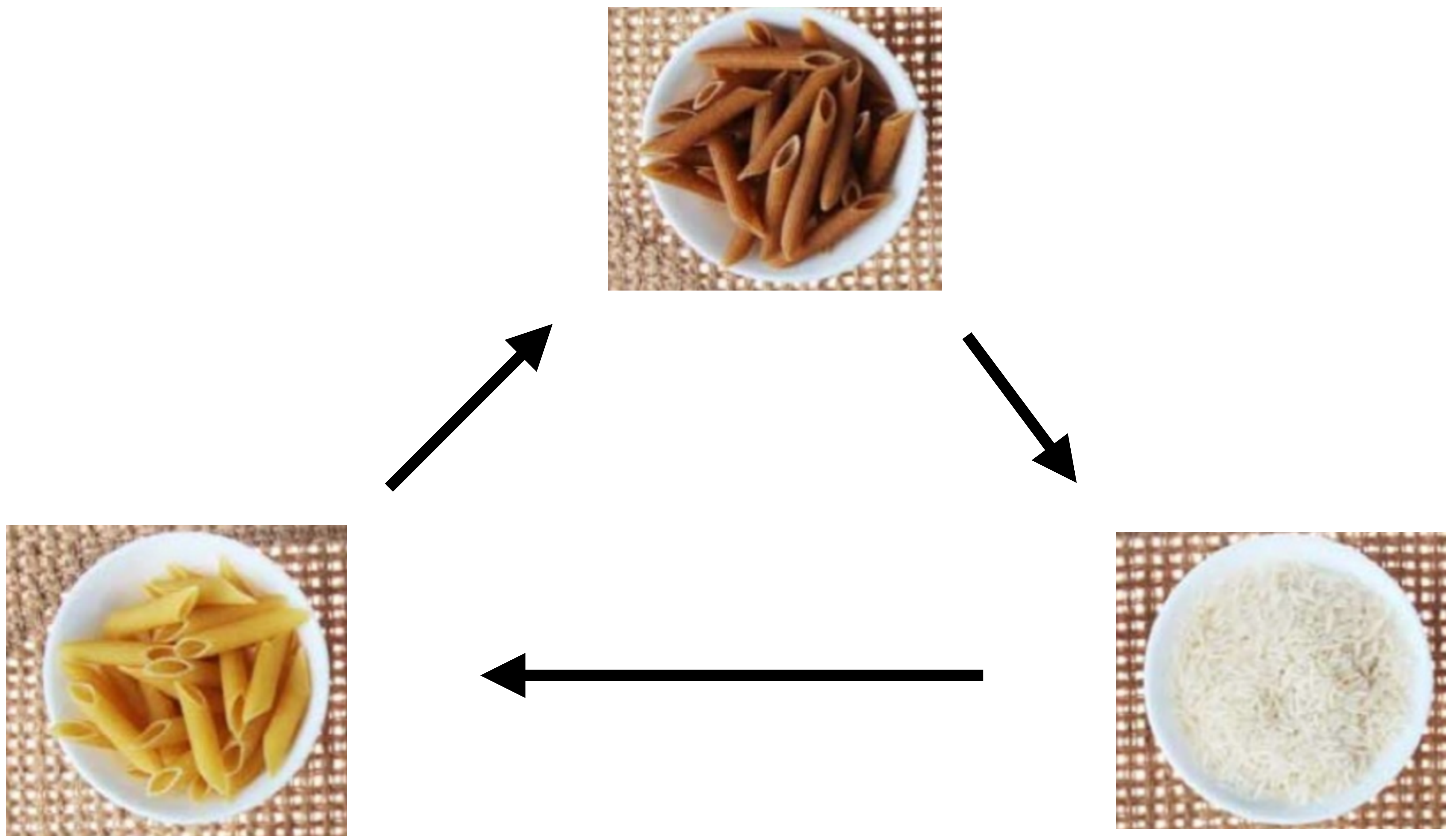
Is majority rule on at least 3 alternatives transitive?



Y pasta > B pasta, B pasta > rice, rice > Y pasta!

Majority Rule

Majority rule with at least three alternatives can produce a *non-transitive* group ranking



Cycle on preferences => non transitive => bad!

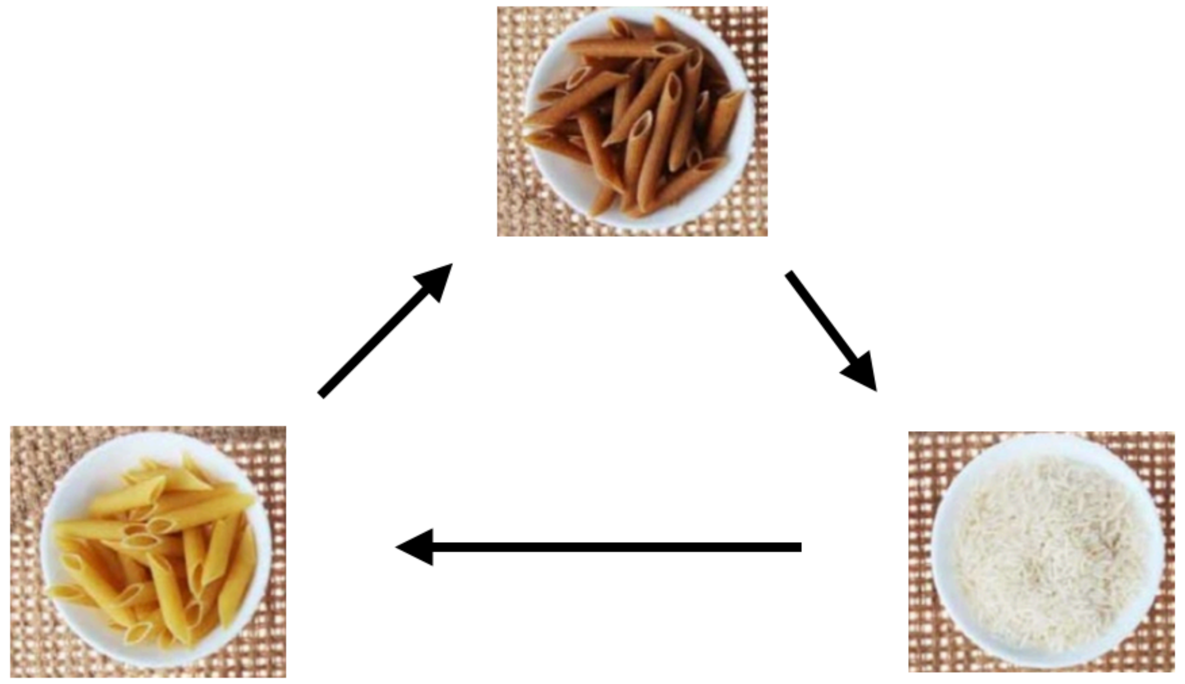
Condorcet Paradox

Majority rule with at least three alternatives can produce a *non-transitive* group ranking

Called the "Condorcet Paradox"

Really bad news!

- Everyone had **perfectly plausible preferences**
- But they behave incoherently as a group, can't even decide on a favourite



Condorcet Paradox

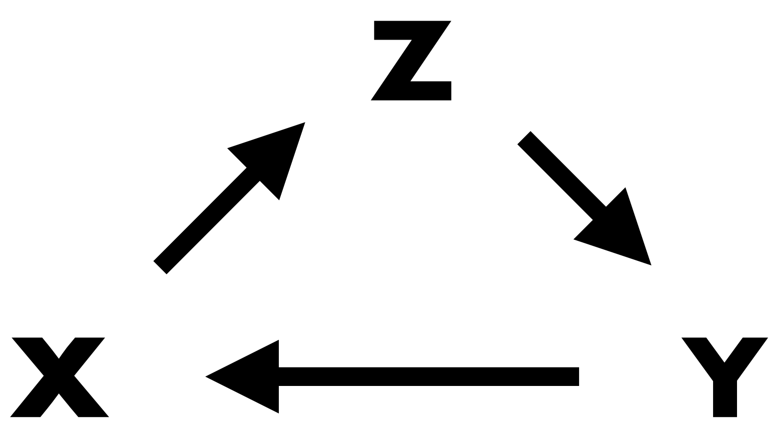
Condorcet Paradox can even happen within a single individual person

Consider a student deciding which college to attend

Wants a highly-ranked college, a small average class size, and maximum scholarship money

Plans to decide between pairs by favouring the one does better on the most criteria

College	National Ranking	Average Class Size	Scholarship Money Offered
X	4	40	$3000
Y	8	18	$1000
Z	12	24	$8000



Majority Rule: Other Ideas

What about using majority rule another way?

Iterative approach: find a winner, remove from the list, and recurse

One idea: arrange alternatives in some order, then compare by majority vote, compare the winner to the third alternative, and so on.

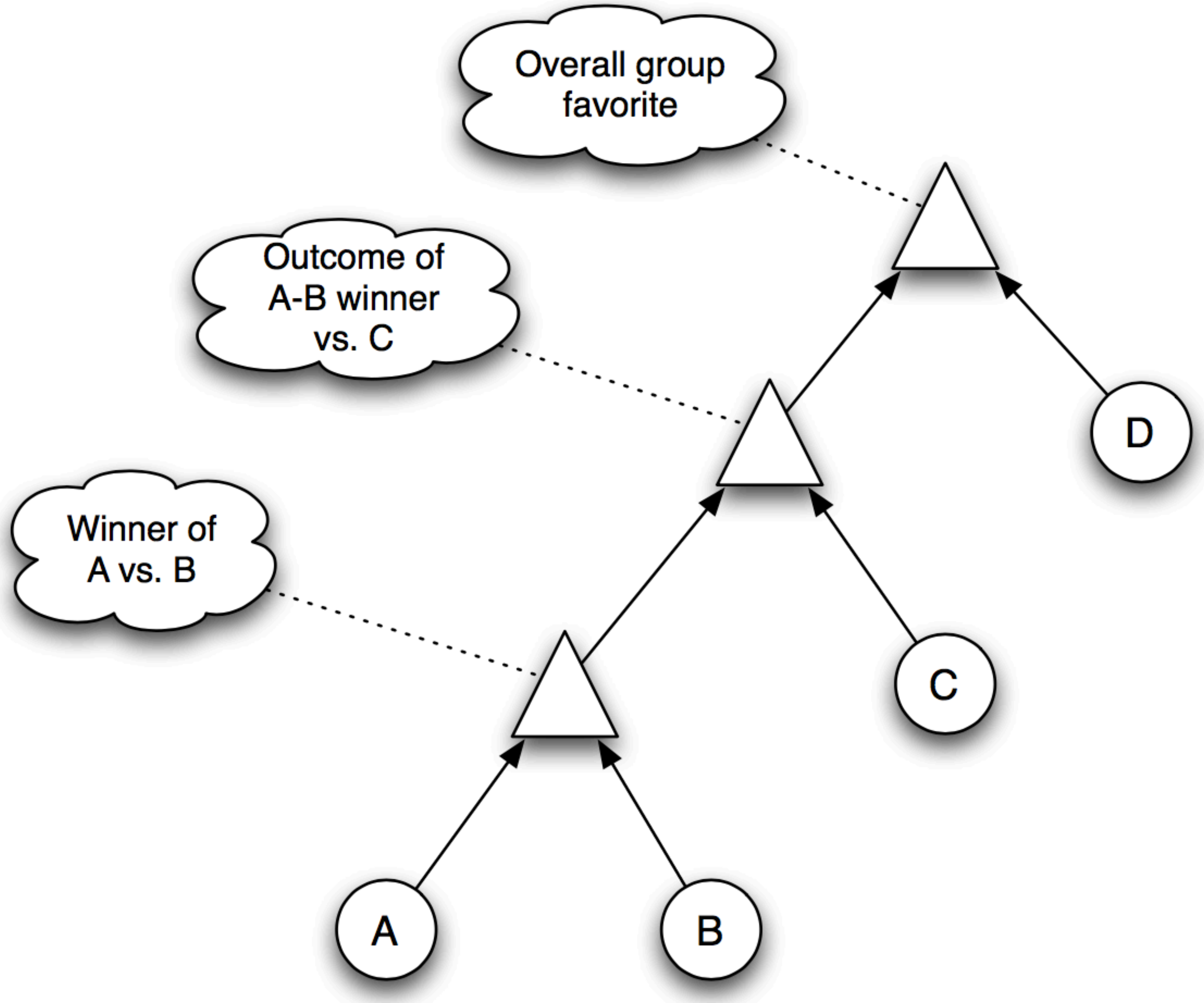
Winner of the final comparison is the group favourite

More generally, we can schedule any kind of elimination tournament to determine the favourite

→ Then recurse!

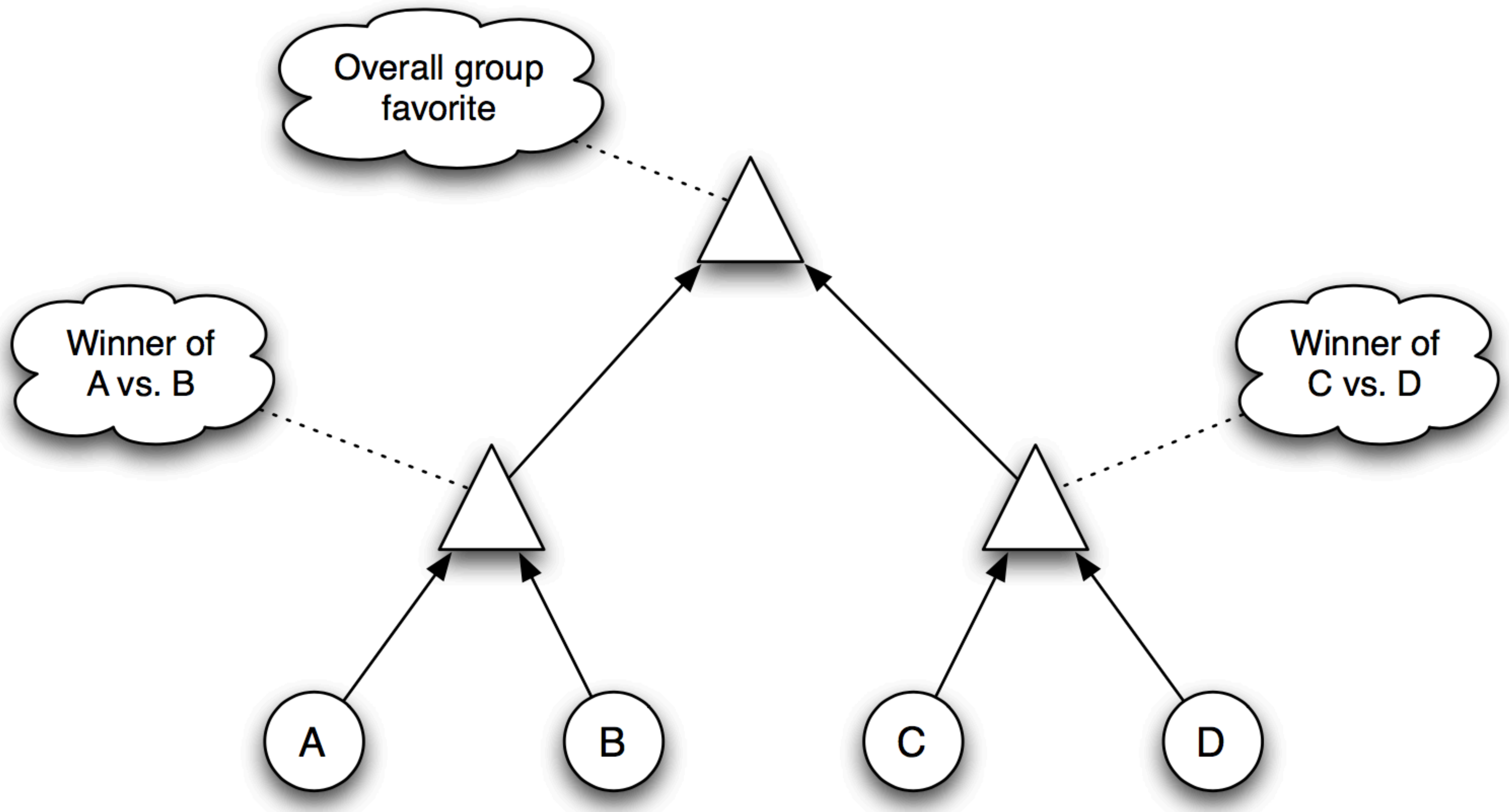
Majority Rule: Other Ideas

Graphically:



Majority Rule: Other Ideas

Other kind of elimination tournament:

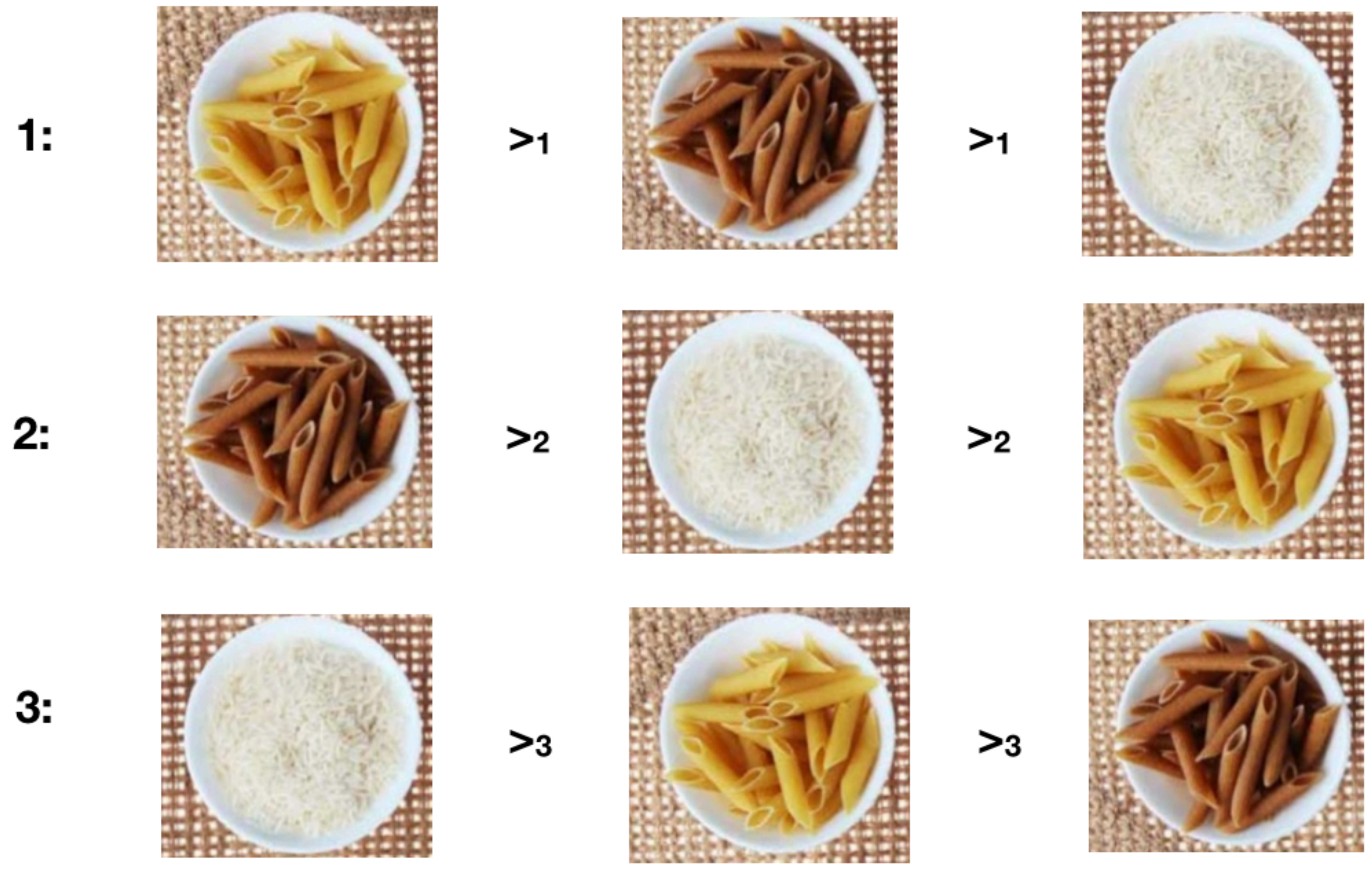


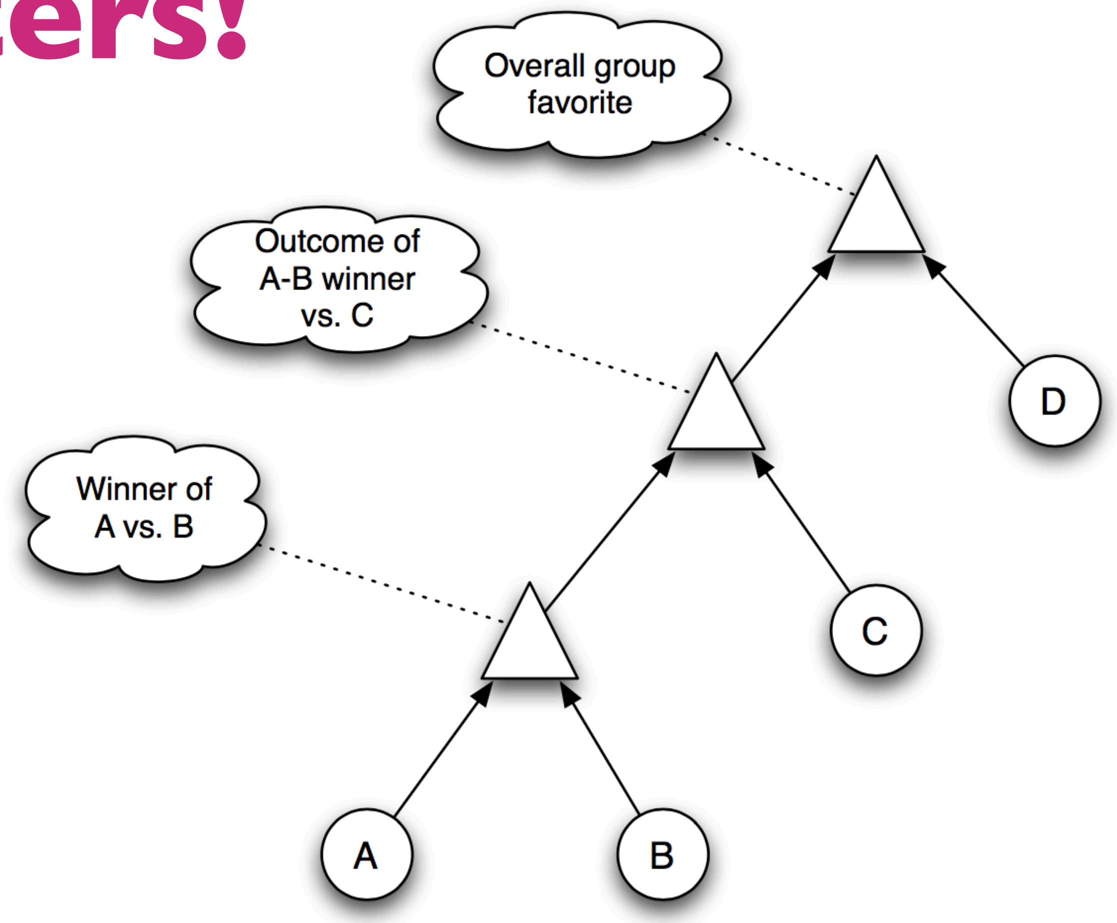
Majority Rule: Other Ideas

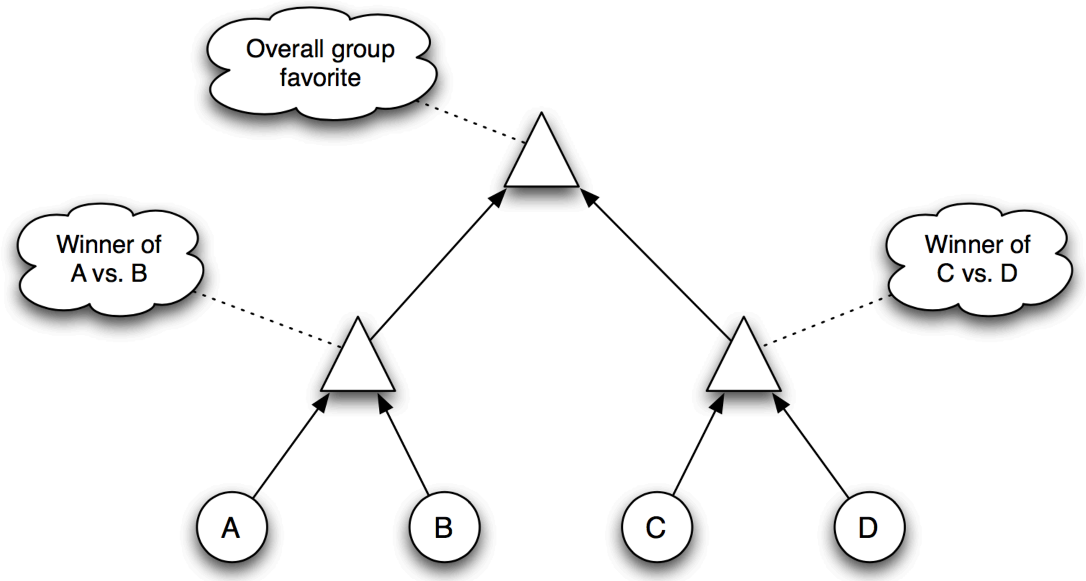
What's wrong with this?

Strategic agenda setting: order matters!

Consider example from before:



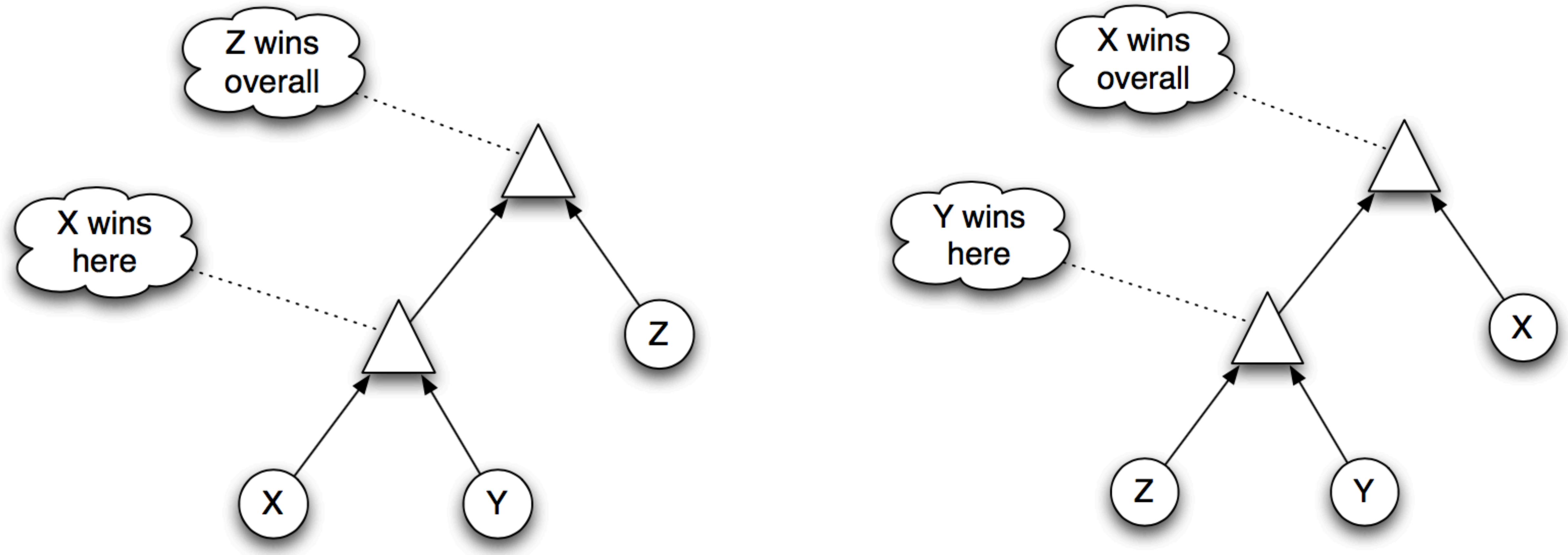




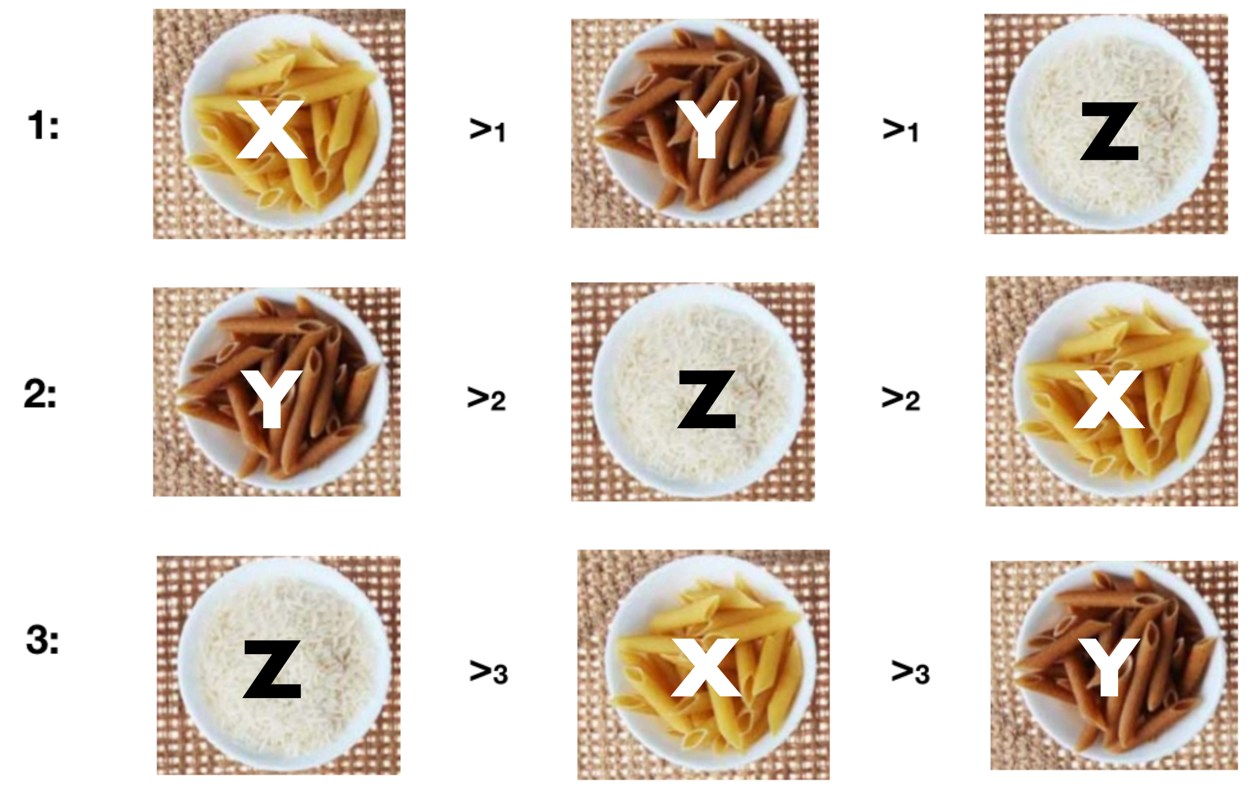
In what order do we evaluate the alternatives?

Majority Rule: Other Ideas

In what order do we evaluate the alternatives?



Entire ranking is entirely determined by the order in which we evaluate!



Other systems?

Majority rule led to some **bad outcomes**

What about other strategies?

Positional voting: produce a group ranking directly from the individual rankings

- Forget pairwise comparisons
- Each alternative receives a certain **weight** based on its positions in all the individual rankings

Borda count

Heisman trophy in college football (and NBA MVP, etc.) all use the following method: get weight 0 for being picked last, 1 for being second last, …, k-1 for being picked first

Repeat for each voter, tally up the scores, and rank

Example: two voters, four alternatives

Voter 1: $A >_1 B >_1 C >_1 D$

Voter 2: $B >_2 C >_2 A >_2 D$

A: 3 + 1 = 4

B: 2 + 3 = 5

C: 1 + 2 = 3

D: 0 + 0 = 0

Group ranking: B > A > C > D

Called the "Borda Count"

Borda count

You can create your own variants (and many have) by changing the number of points per position

Example: if only top 3 matter, you could assign 3 for first place, 2 for second place, 1 for third place, and 0 otherwise

Any such system is a **"positional voting system"**

Ignoring ties, Borda Count always produces a complete, transitive ranking!

Borda count

But the Borda Count **has its own problems**

Magazine tries to rank greatest movie of all time, asks five film critics to rank Citizen Kane and The Godfather

Three prefer CK, two prefer TG => CK>TG => **all good!**

At the last second, they want to inject some modernity into the discussion, so they include Frozen

First three only like old movies, so they vote:

$CK >_i TG >_i F$

Critics 4 and 5 only like past 40 years, so:

$TG >_i F >_i CK$

What is the Borda Count now?

Borda count

First three only like old movies, so they vote:

CK $>_i$ TG $>_i$ F

Critics 4 and 5 only like past 40 years, so:

TG $>_i$ F $>_i$ CK

Borda:

CK: 6, TG: 7, F: 2 => TG > CK > F

But before Frozen was introduced it was CK > TG!

TG and CK flip because of Frozen??

Both TG and CK beat Frozen head-to-head

Yet still Frozen influenced CK > TG

Borda count

Borda Count is susceptible to "irrelevant alternatives"

What voters think of Frozen **should be irrelevant** to how they feel about relative ranking of TG and CK

But it isn't

This gives rise to another problem: voters can **strategically misreport their preferences**

For example, say voters 4 and 5 actually had the true ranking TG > CK > F

1,2,3: CK $>_i$ TG $>_i$ F

4,5: TG $>_i$ CK $>_i$ F

Borda: CK $>_i$ TG $>_i$ F

By lying and reporting TG $>_i$ F $>_i$ CK, they get TG to win

Irrelevant Alternatives in Politics

These problems with "irrelevant alternatives" and strategic misreporting have happened in elections around the world

Most vote with **plurality voting:** the candidate ranked at the top by most voters wins

Q: **is this a positional voting system?**

A: **Yes: 1 for winner, 0 otherwise**

"Third-party effects"/"spoiler effects": if very few people favour some candidate, this can swing outcome of two leading contenders

In response, some people strategically misreport their preferences

What's The Deal?

Voting is one society's **most important institutions**

On its face, seems like a relatively simple problem

But we can't find a system that doesn't have horrible pathologies!

Is there any system that is free of pathologies?

What's The Deal?

Is there any system that is free of pathologies?

Let's define "Free of pathologies"

- Criterion 1 **"Unanimity"**: if there is a pair X and Y for which $X >_i Y$ for every i, then X > Y
- Criterion 2 **"Independence of Irrelevant Alternatives" (IIA)**: the ordering of X and Y should only depend on the relative positions X and Y in individual rankings

  If we have a bunch of rankings that produces a group ranking with X > Y

  Then we move some Z around in the individual rankings

  It should still be the case that X > Y
- Criterion 3 **"Non-Dictatorship"**: the group ranking should not just always be what one particular voter thinks

Independence of Irrelevant Alternatives

Good Voting Systems

What satisfies Unanimity and IIA and non-dictatorship?

With two alternatives, majority rule clearly satisfies all

Arrow's Theorem [Arrow 1953]: With at least three alternatives, no voting system satisfies Unanimity, IIA, and Non-dictatorship

In general, there is no good voting system!

In practice, this means that there will always be inherent tradeoffs we have to choose from

What Do We Do Now?

How do we vote, how do we decide on things in the presence of Condorcet's Paradox and Arrow's Theorem?

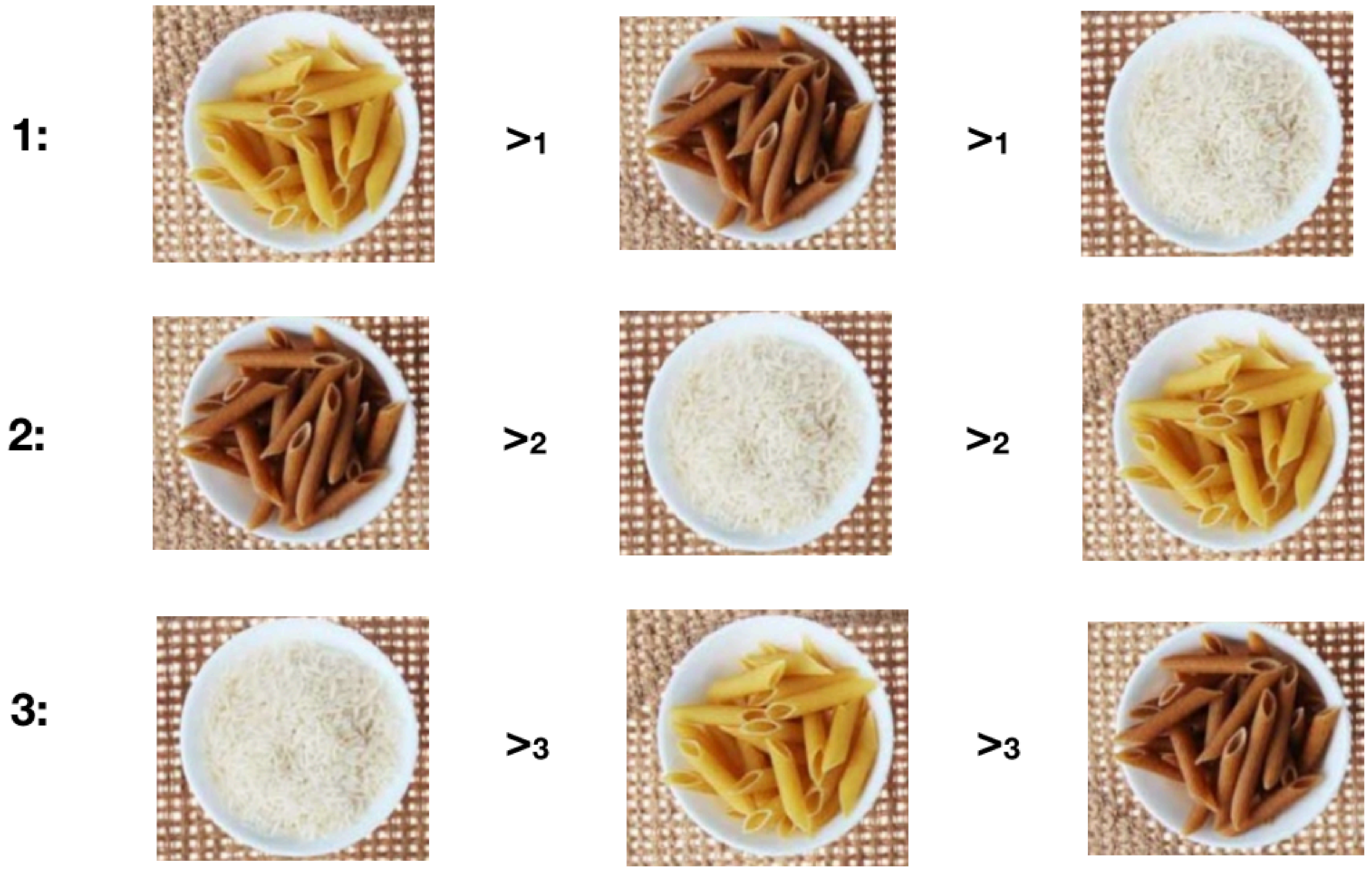
If you're faced with an impossibility result, you don't just give up

One common technique is to **look for important special cases**

Arrow's Theorem is a **general result**, so it doesn't necessarily apply if we make some additional assumptions

What Do We Do Now?

Go back to original Condorcet problem



Replace food with choices about how much money to spend on education

What Do We Do Now?

Go back to original Condorcet problem with money now:

1: X $>_1$ Y $>_1$ Z	**X: small**	**Amount to spend on education**
2: Y $>_2$ Z $>_2$ X	**Y: medium**	
3: Z $>_3$ X $>_3$ Y	**Z: a lot**	

Voter 1's preferences "make sense"

Voter 2's preferences do too: prefer between Y and Z, so say Y then Z then X

Voter 3's preferences are **harder to justify**

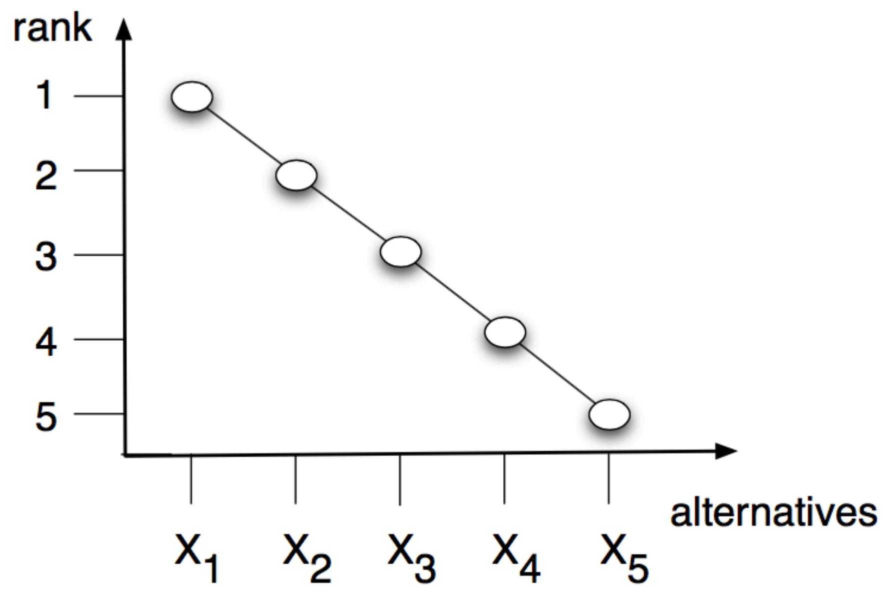
Not impossible, but they're more unusual

Ideal Points

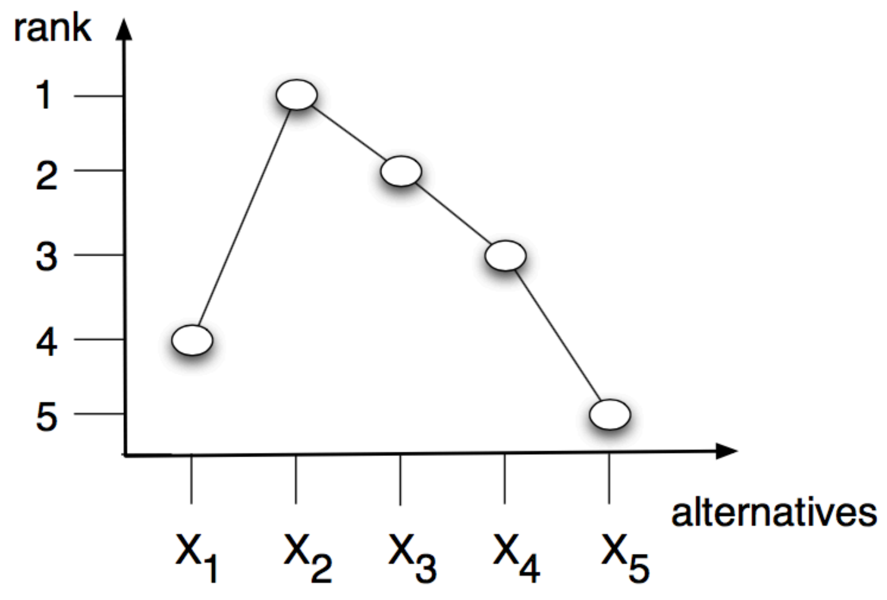
Assume the preferences lie on a one-dimensional spectrum, and each voter has an "ideal point" on the spectrum

They evaluate alternatives by proximity to this ideal point

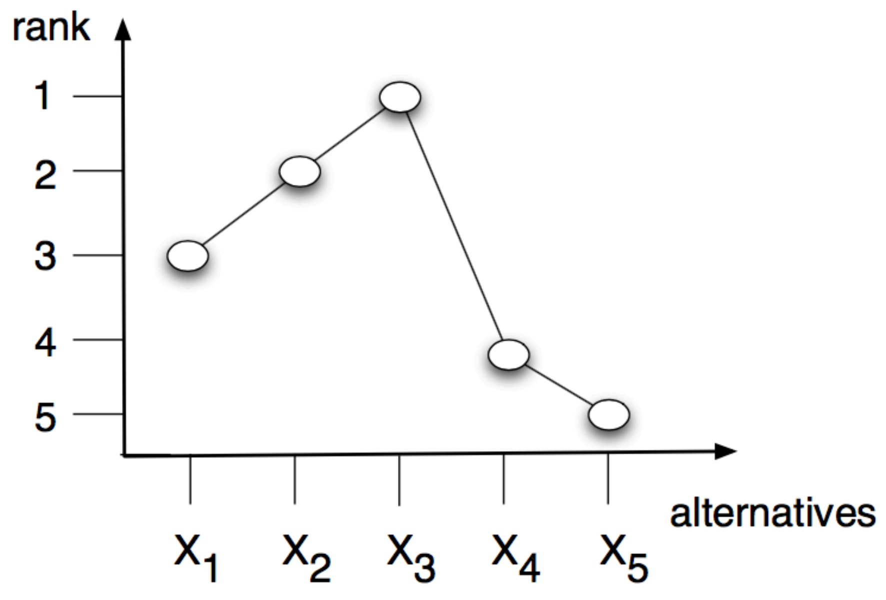
Actually we can assume something weaker: each voter's preferences "fall away" consistently on both sides of their favourite alternative



(a) *Voter 1's ranking.*



(b) *Voter 2's ranking.*



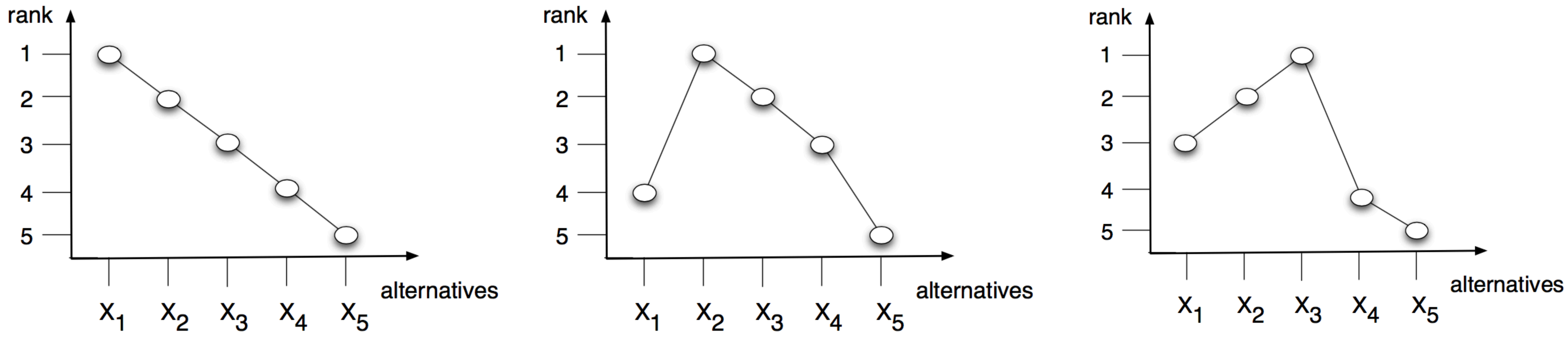
(c) *Voter 3's ranking.*

Single-Peaked Preferences

Definition: a voter has "single-peaked preferences" if there is no alternative X_s for which both neighbouring alternatives X_{s-1} and X_{s+1} are ranked above X_s

Equivalent to: every voter i has a top-ranked option X_t, and her preferences fall off on both sides of t:

$$X_t \succ_i X_{t+1} \succ_i X_{t+2} \succ_i \cdots \quad \textbf{and} \quad X_t \succ_i X_{t-1} \succ_i X_{t-2} \succ_i \cdots$$



Single-Peaked Preferences

Majority rule with single-peaked preferences

Recall majority rule: compare every pair of alternatives X and Y, and decide X > Y or Y > X by the majority of voters

Claim: If all individual rankings are single-peaked, then majority rule applied to all pairs of alternatives produces a group preference relation that is **complete** and **transitive**.

In other words, majority rule works!

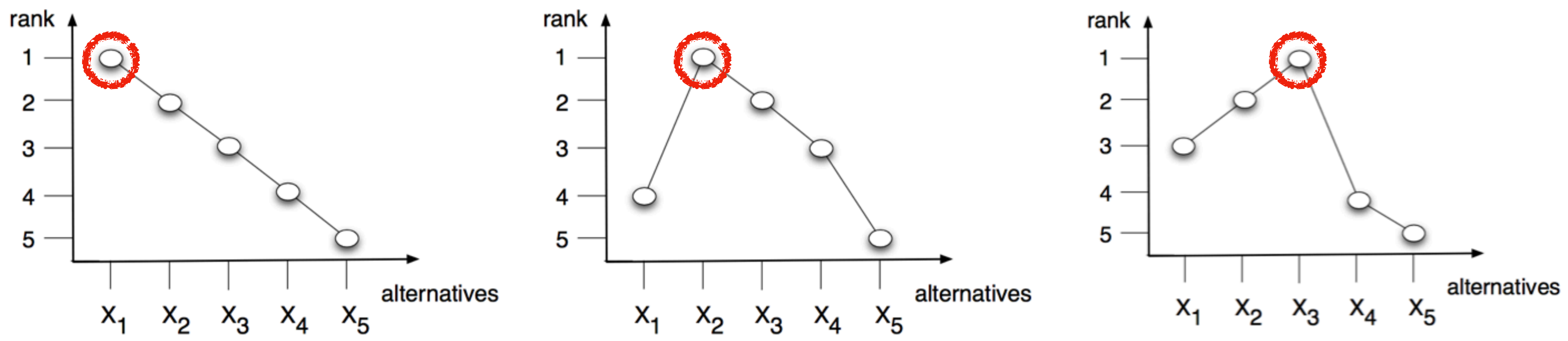
Median Voter

Start off by trying to find a group favourite, then proceed by recursion on the rest of the alternatives

Consider every voter's top-ranked alternative — their peak — and sort this set of favourites from left to right along the spectrum

- A popular alternative can show up many times

Now consider the **median** of these favourites



Favourites: X_1, X_2, X_3 Median: X_2

Median Voter

The median individual favourite is a natural candidate for potential group favourite

Strikes a compromise between more extreme favourites on either side

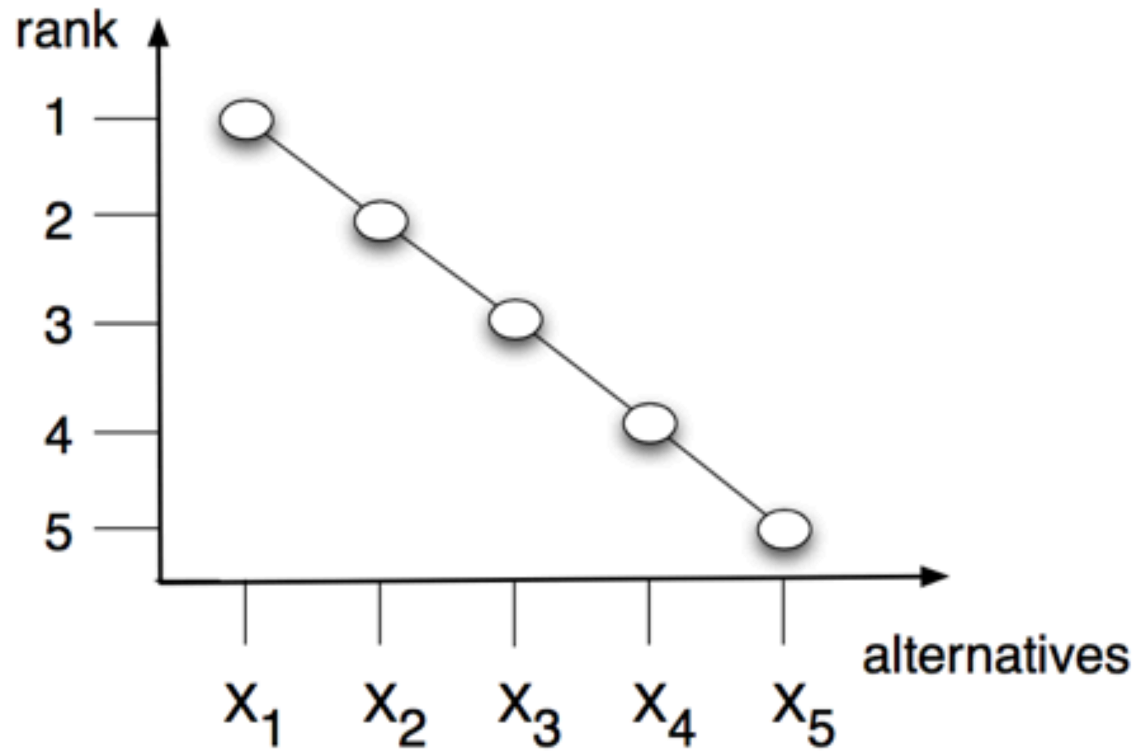
Median Voter Theorem: With single-peaked preferences, the median individual favourite defeats every other alternative in a pairwise majority vote.

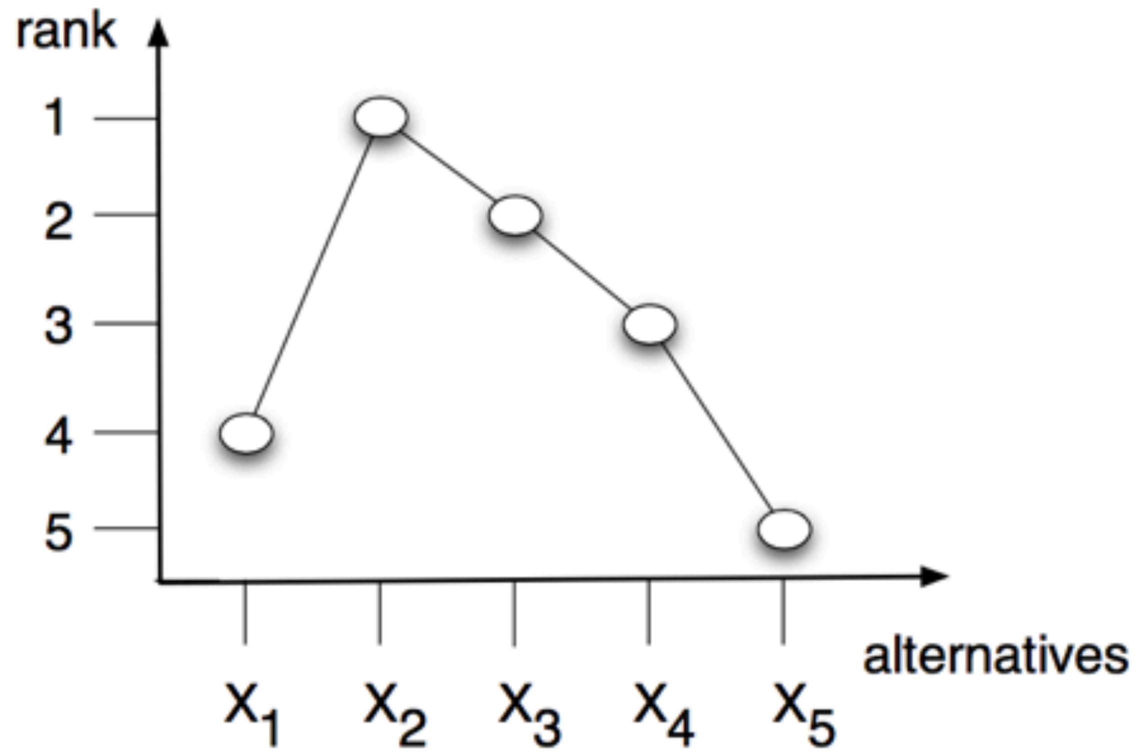
Example

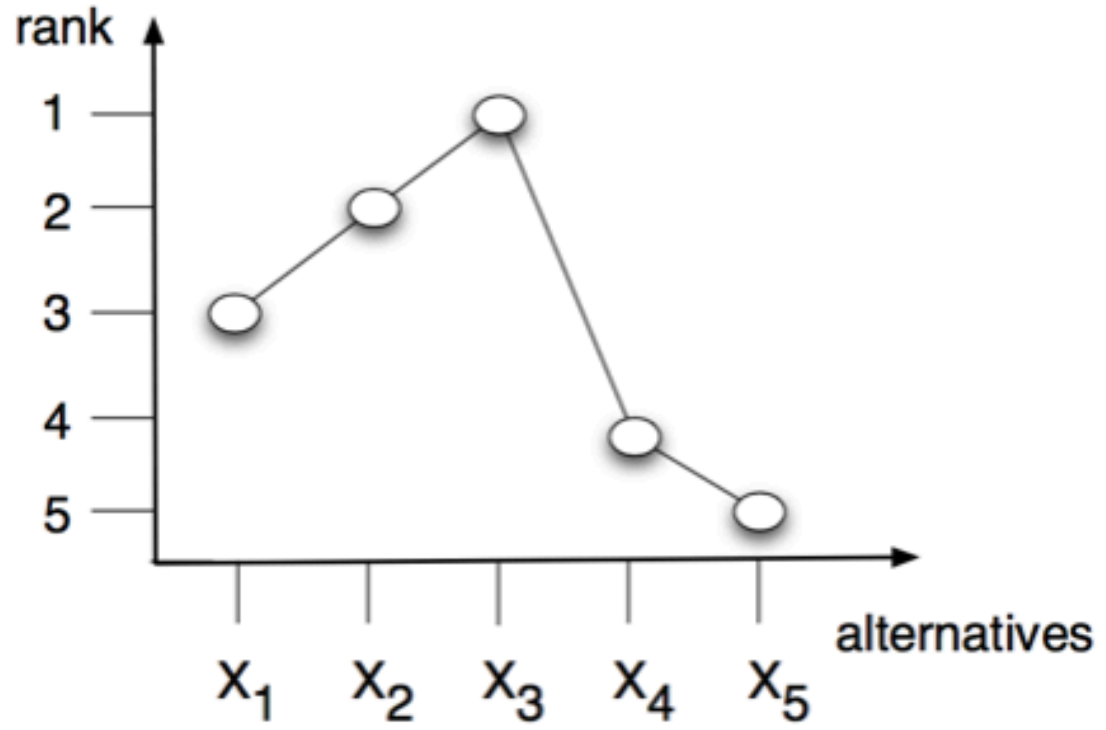
X_2 is global median favourite

Then favourites are X_1, X_3, X_3 => X_3 median favourite

Eventually we get $X_2 > X_3 > X_1 > X_4 > X_5$







Voting as Information Aggregation

So far, trying to come up with **methods for people who have different preferences**

Sometimes there is a "true" underlying ranking and people with different information are trying to uncover it

Examples:

Jury deliberation

Board of advisors to a company

Simple Case: Simultaneous, Sincere Voting

Simple setting, two alternatives X and Y

One is genuinely the best choice, each voter casts vote on what she thinks the right choice is

Assume everyone votes sincerely

Model: similar to information cascades

Prior probability that X is best is 1/2

Each voter gets a private independent signal on which is best, prob of getting right signal is q (> 1/2)

With probability q, voter should vote for what her signal says

Condorcet Jury Theorem: as the number of voters increases, probability of the majority choosing correct decision goes to 1

Oldest "wisdom of crowds" argument

Simple Case: Simultaneous, Sincere Voting

Formal Bayes argument

Recall Bayes Rule: P[A|B] = P[B|A]P[B]/P[A]

We want to compute P[X is best | X-signal]

Given: P[X is best] = 1/2 and P[X-signal | X is best] = q

Voter's strategy: evaluate P[X is best | X-signal] then vote X if this probability > 1/2

P[X is best | X-signal] = P[X-signal | X is best]P[X is best]/P[X-signal]

X-signal can be observed if X is best or if Y is best:

P[X-signal] = P[X is best] * P[x-signal observed | X is best] + P[Y is best] * P[X-signal observed | Y is best] = 1/2q + 1/2(1-q) = 1/2

So overall: P[X is best | X-signal] = (1/2)q / (1/2) = q

Voter favours the alternative that is reinforced by her signal

Insincere Voting

We just assumed sincere voting

But there are very natural situations where a voter should actually lie, even though her goal is to maximize the probability that the group chooses the right alternative!

Example, information cascades-style:

- Experimenter has two urns, 10 marbles each
- One urn has 10 white marbles ("pure") and the other has 9 green and one white ("mixed")
- Three people privately draw one marble and guess what urn it is, and all win money if the majority of them are right

Insincere Voting

Suppose you draw a white marble

→ Way more likely that urn is **pure** than **mixed**

If you draw a green marble

→ Know for sure it's **mixed**

But what should you guess?

First, when will your guess actually matter?

If the two others agree, then your guess doesn't change anything!

Only case where it matters is if they're split

If they're split, someone said mixed, so they know it's mixed!

Then you should guess mixed to break the tie the right way!

Assuming others vote sincerely, you have an incentive to vote insincerely => everyone voting sincerely is **not** a Nash equilibrium

Insincere Voting

This is very naturally thought of as a game

Voters are players, guesses are strategies, and they result in certain payoffs

This is highly stylized setting so we can see what's going on

But it happens in the real world too

Jury Deliberations

Consider a jury deliberating on a verdict: **guilty** or **innocent**

There is a "best" answer — whether the defendant is actually guilty or innocent

Compare with Condorcet Jury Theorem setup:

1. Juries require a **unanimous** vote. **Guilty** only if everyone says guilty
2. In Condorcet, evaluate alternatives just by picking most likely one (if > 1/2 sure, pick it). Here, only pick guilty if sure beyond a reasonable doubt:

$$\Pr\left[\textit{defendant is guilty} \mid \textit{all available information}\right] > z$$ **for some large z**

Jury Deliberations

Each juror gets an independent private signal: **guilty signal** (G-signal) or **innocent signal** (I-signal)

They usually get the right signal: P[G-signal | defendant guilty] = P[I-signal | defendant innocent] = q, q > 1/2

Assume prior probability of guilt of 1/2, but doesn't matter

What should a juror do?

Jury Deliberations

- **What should a juror do?**
- Say you receive an I-signal
  - At first it seems obvious that you should vote to acquit
  - But: conviction criterion is $\Pr[\textit{defendant is guilty} \mid \textit{available information}] > z$ so **if all the other jurors received G-signals you might still be above that threshold**
  - Second, ask yourself key question from before: **when does my vote actually matter?**
  - **Like before, your vote only changes the outcome if everyone except you is voting guilty!**
    - **If you vote guilty, defendant is found guilty**
    - **If you vote to acquit, defendant is found innocent**

Jury Deliberations

- **If everyone but you is voting guilty, what is the probability of defendant being guilty?**

$$\Pr[\textit{defendant is guilty} \mid \textit{you have the only I-signal}] = \frac{\Pr[\textit{defendant is guilty}] \cdot \Pr[\textit{you have the only I-signal} \mid \textit{defendant is guilty}]}{\Pr[\textit{you have the only I-signal}]}.$$

$$\begin{aligned}
\Pr[\textit{you have the only I-signal}] &= \Pr[\textit{defendant is guilty}] \cdot \Pr[\textit{you have the only I-signal} \mid \textit{defendant is guilty}] + \\
&\quad \Pr[\textit{defendant is innocent}] \cdot \Pr[\textit{you have the only I-signal} \mid \textit{defendant is innocent}] \\
&= \frac{1}{2} \cdot q^{k-1}(1-q) + \frac{1}{2}(1-q)^{k-1} q.
\end{aligned}$$

Jury Deliberations

- If everyone but you is voting guilty, what is the probability of defendant being guilty?

$$\Pr\left[\textit{defendant is guilty} \mid \textit{you have the only I-signal}\right] = \frac{\Pr\left[\textit{defendant is guilty}\right] \cdot \Pr\left[\textit{you have the only I-signal} \mid \textit{defendant is guilty}\right]}{\Pr\left[\textit{you have the only I-signal}\right]},$$

$$\Pr\left[\textit{defendant is guilty} \mid \textit{you have the only I-signal}\right] = \frac{\frac{1}{2}q^{k-1}(1-q)}{\frac{1}{2}q^{k-1}(1-q) + \frac{1}{2}(1-q)^{k-1}q} = \frac{q^{k-2}}{q^{k-2} + (1-q)^{k-2}},$$

- Since $q>1/2$, $(1-q)^{k-2}$ is super small, so the probability goes to 1
- **In only case where your vote to acquit matters, you should vote guilty despite your I-signal!**

Jury Deliberations

- Intuitively: because of the unanimity rule, you only affect the outcome when everyone else holds the opposite opinion
- Assuming everyone else is as informed as you, and **assuming independence** (remember information cascades!), then the conclusion is that they're probably collectively right
- The result is: assuming everyone else votes sincerely, you have an incentive to vote **insincerely**
  - **All-sincere voting is not an equilibrium**
- **What is the equilibrium?**
  - There are several
  - Most interesting is a mixed equilibrium (randomly disregard I-signal some fraction of the time to correct for possibility that it's wrong)
  - **In this equilibrium, probability of convicting an innocent defendant does not go to zero as #jurors goes to infinity!**

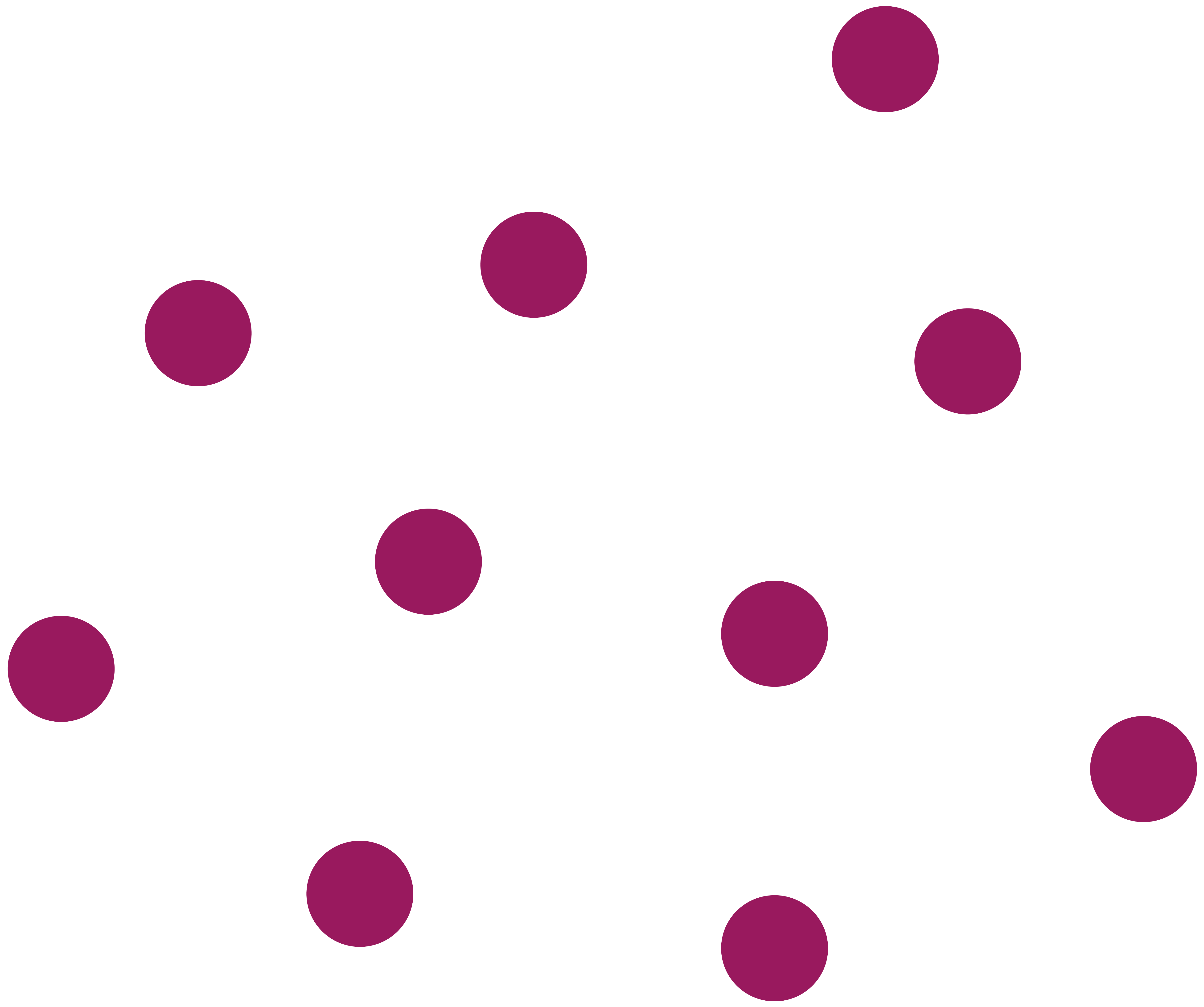
Jury Decisions

- Why do we get such a bad outcome?
- **Unanimity is a very harsh constraint.**
  - If we relax to only requiring a certain fraction f saying guilty, then the probability that we convict an innocent defendant goes to 0

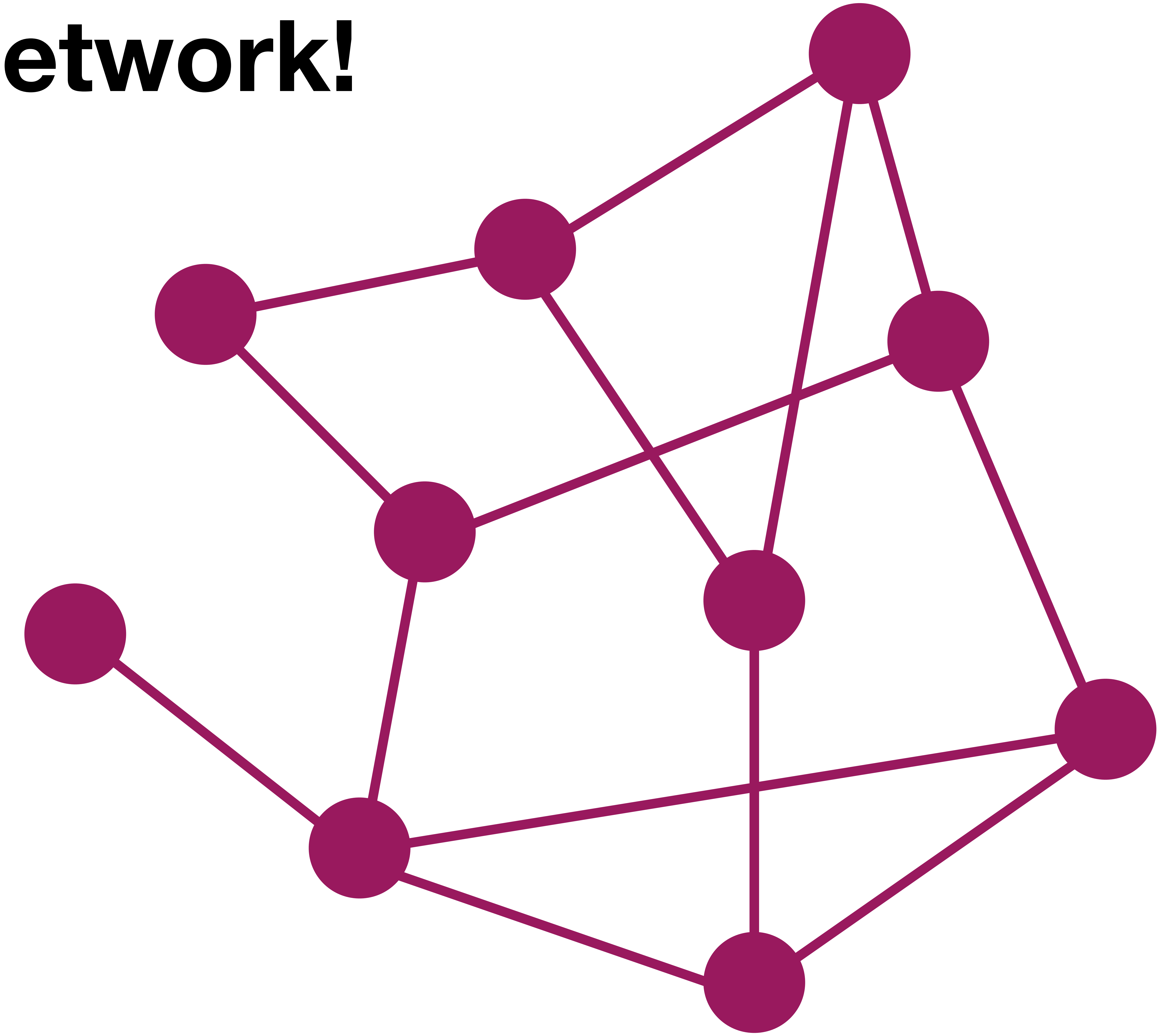
Summary

- **Voting**: synthesizing the preferences of many people into a single group preference
- Many fundamental issues:
  - **Condorcet paradox**: most natural method (majority rule) can turn a set of reasonable preference relations into an unreasonable one
  - **Arrow's Theorem**: **no general voting system** simultaneously satisfies unanimity, IIA, and non-dictatorship.
- Special case: single-peaked preferences
  - Median Voter Theorem says we can get good outcomes
- Jury deliberations: insincere voting can be incentivized

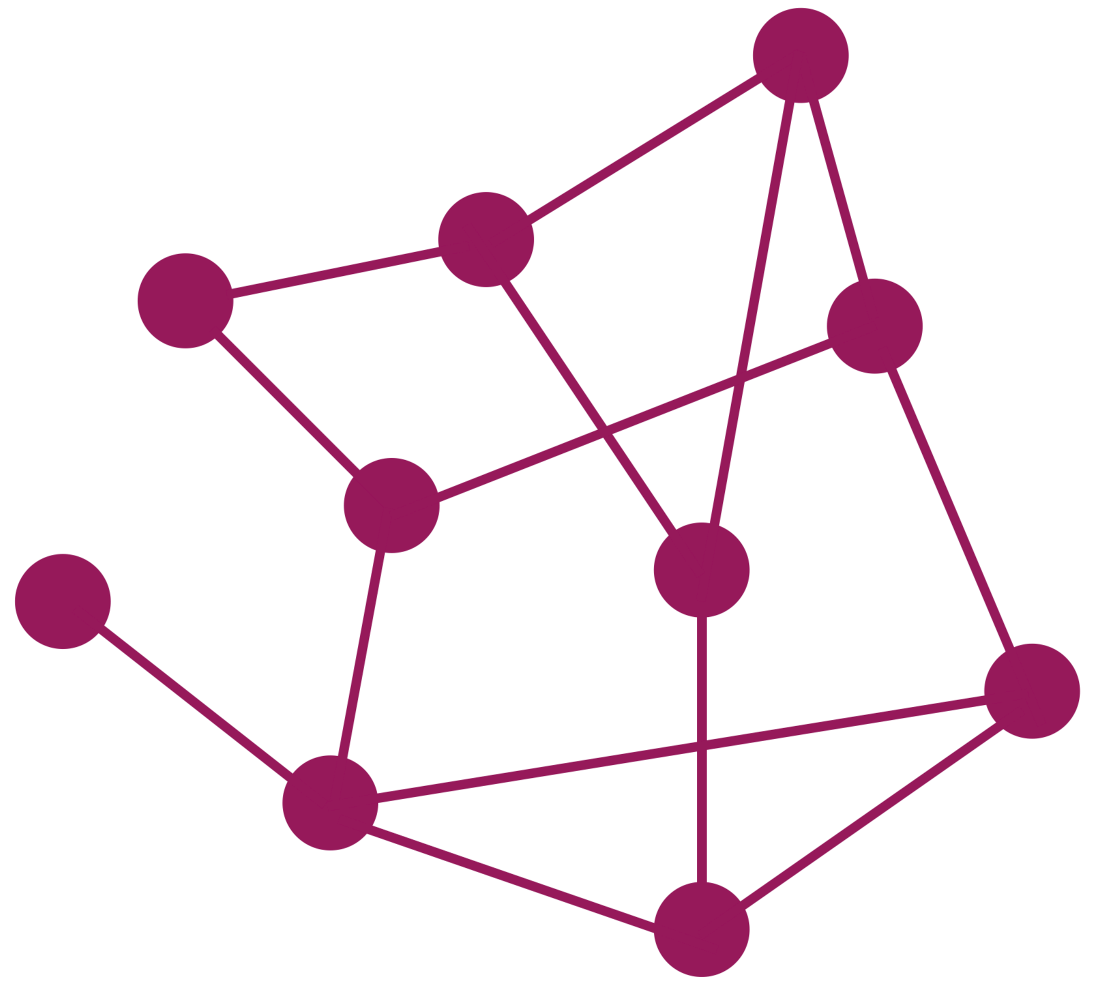
Lecture 1



A Network!



Components of a Network



Objects: nodes, vertices ***N***
Interactions: links, edges ***E***
System: network, graph ***G(N,E)***

Why study networks?

Networks are a **universal language for describing complex data**

Networks from science, nature, and technology are more similar than you might expect

Shared vocabulary between fields

CS, finance, tech, social sciences, physics, economics, statistics, biology

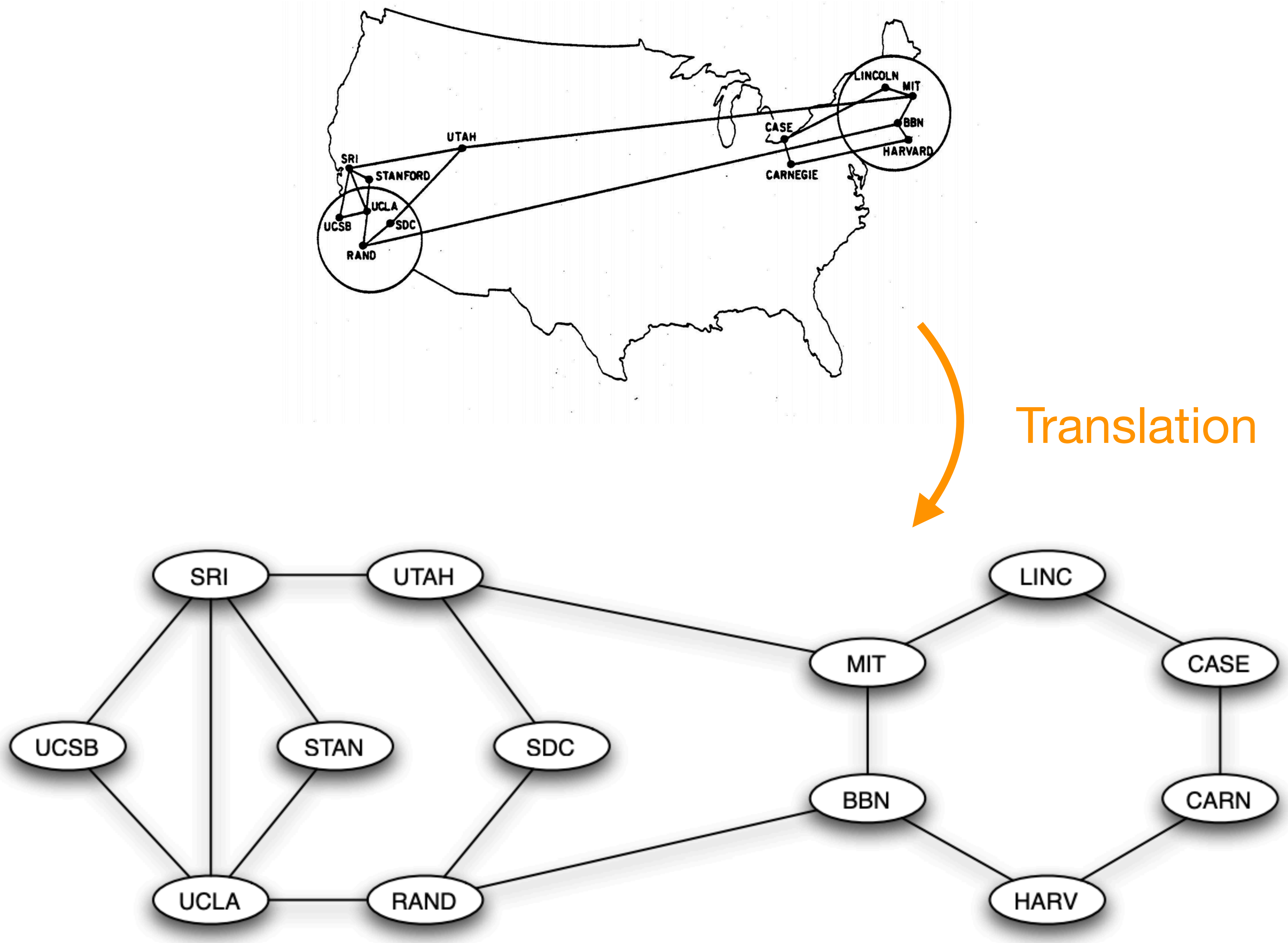
Data availability (and computational challenges)

Web/mobile, bio, health, medical

Impact!

Social networking, social media, drug design

A first example

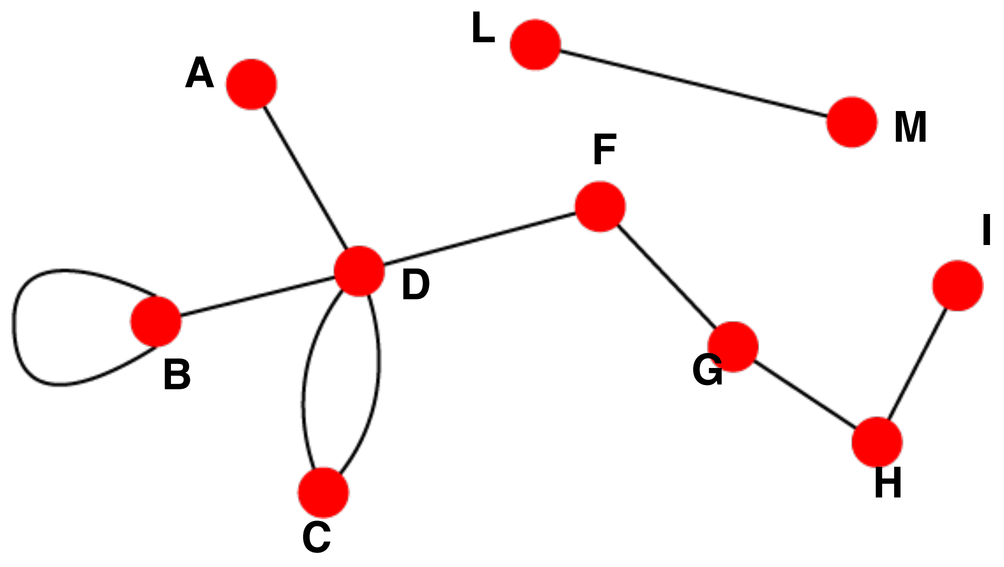


The Internet in 1970

Undirected and Directed Networks

Undirected

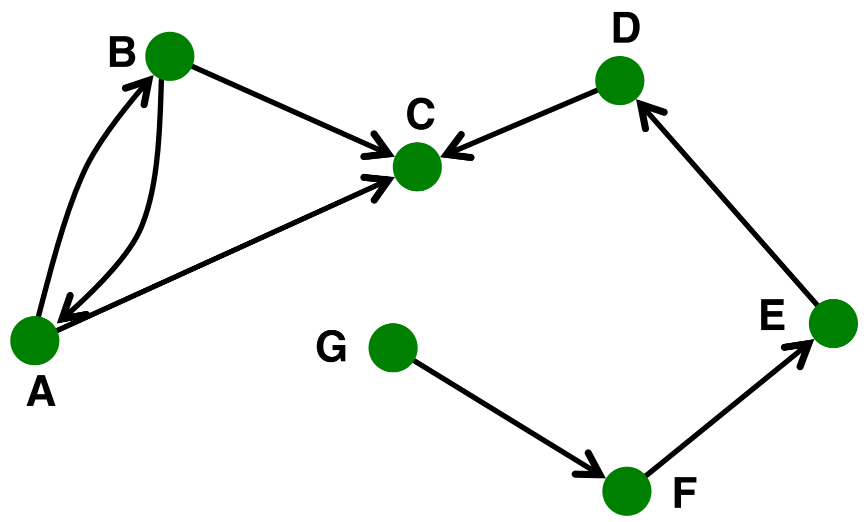
- **Links:** undirected (symmetrical, reciprocal)



- **Examples:**
  - Collaborations
  - Friendship on Facebook

Directed

- **Links:** directed (arcs)



- **Examples:**
  - Phone calls
  - Following on Twitter

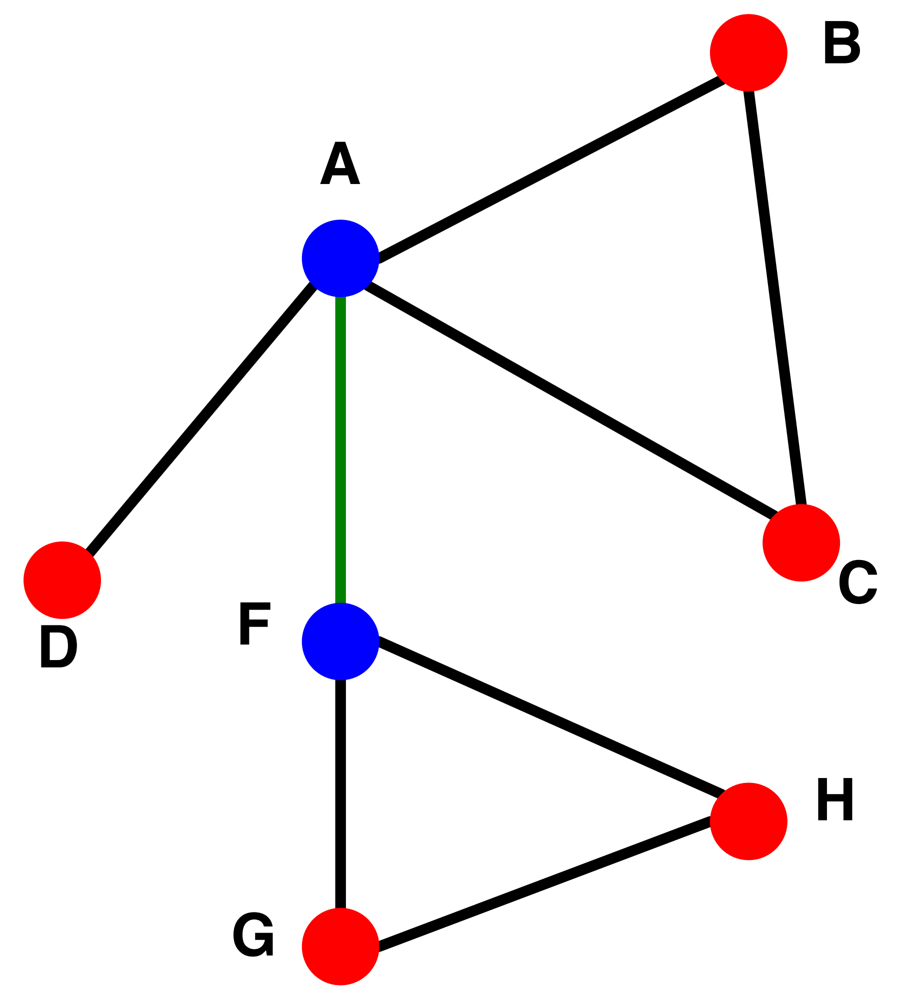
Connectivity of Graphs

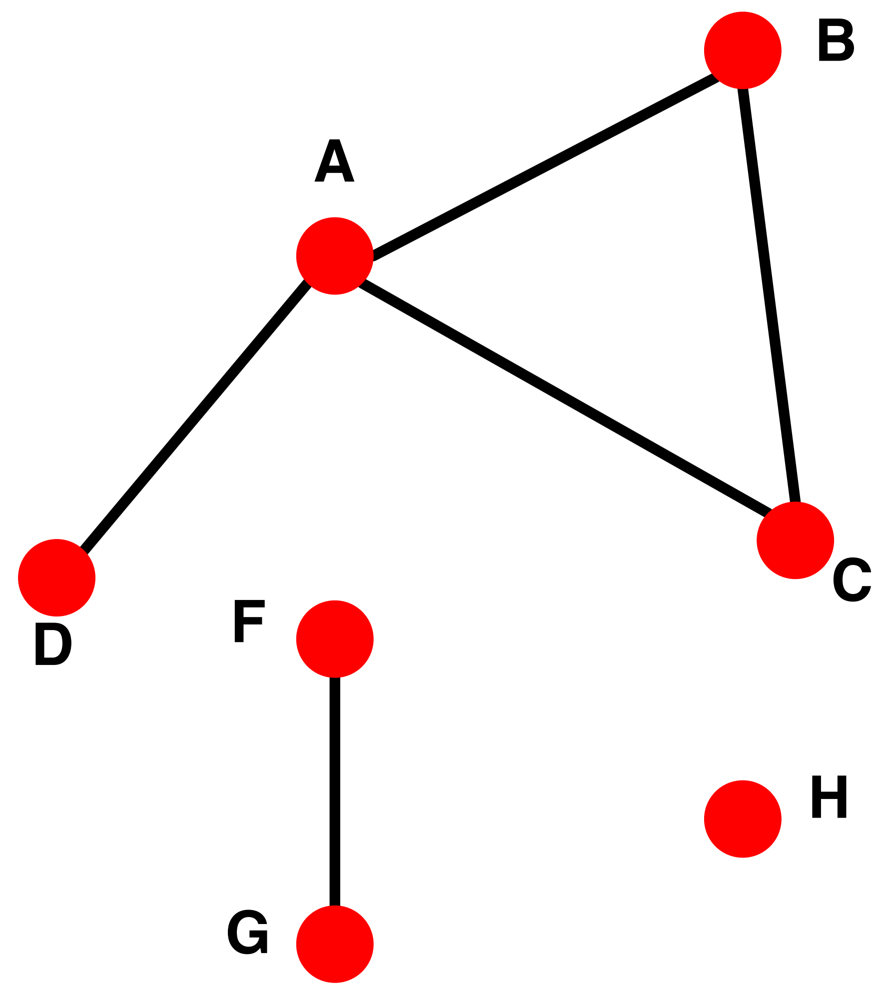
Connected component (undirected):

Any two vertices can be joined by a path

No superset with the same property

A disconnected graph is made up of two or more connected components





Largest Component:
Giant Component

Isolated node (node H)

Bridge edge: If we erase it, the graph becomes disconnected.

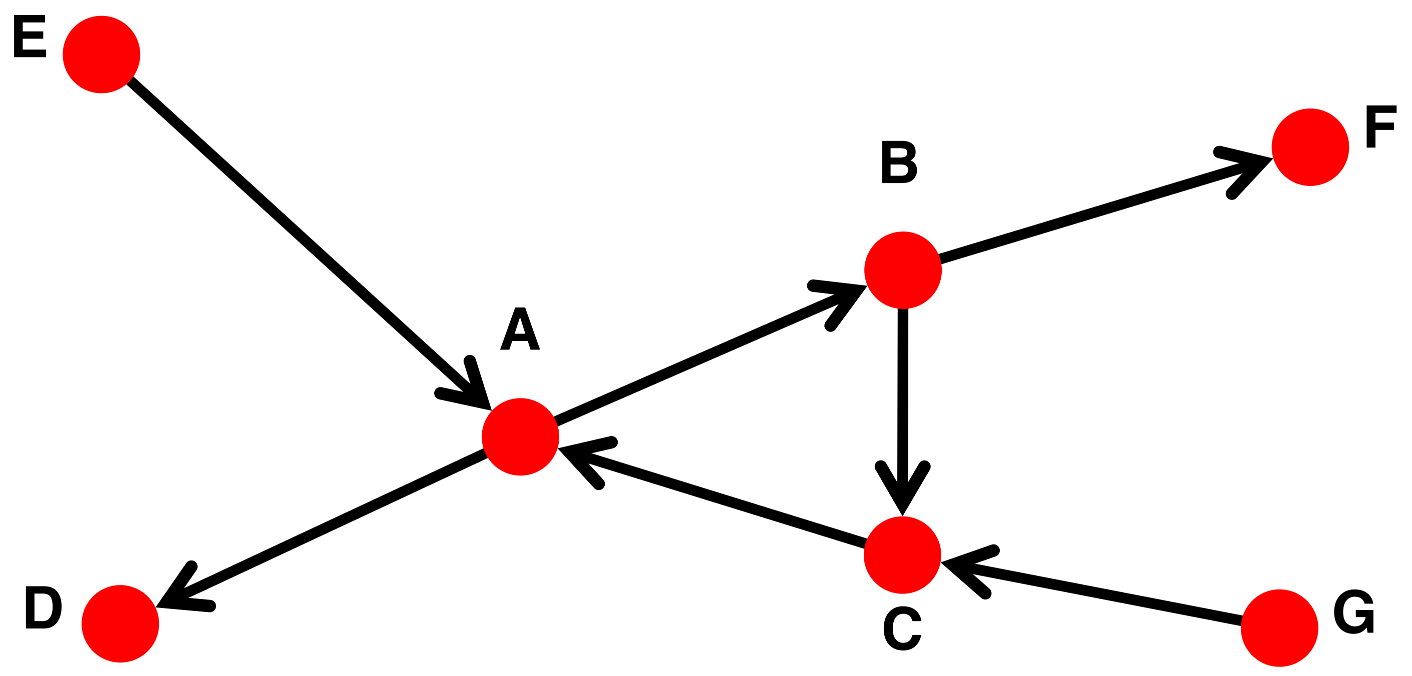
Connectivity of Directed Graphs

Strongly connected directed graph

has a path from each node to every other node and vice versa (e.g., A-B path and B-A path)

Weakly connected directed graph

is connected if we disregard the edge directions



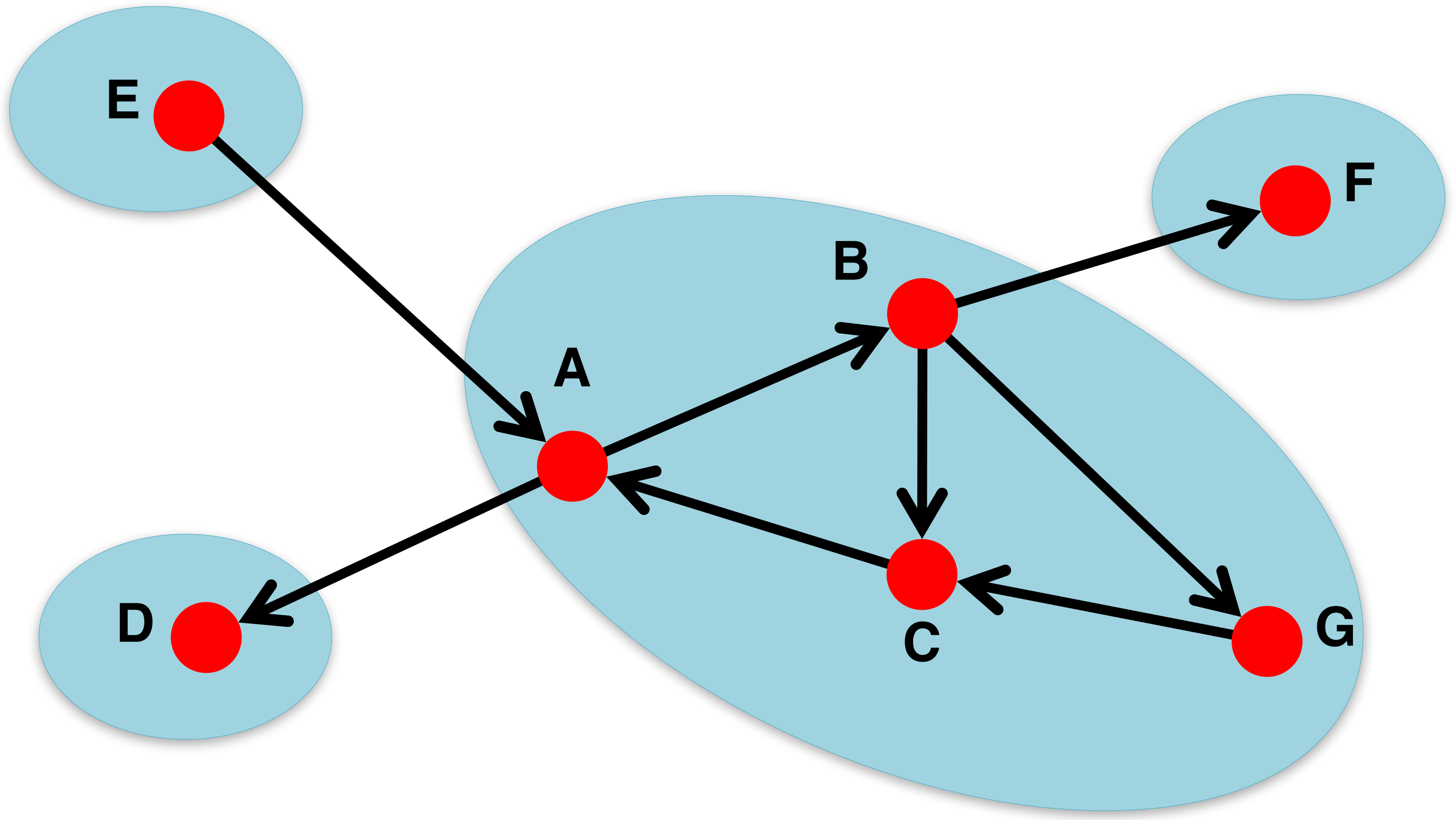
It is connected but not strongly connected (e.g., there is no way to get from F to G by following the edge directions)

Strongly Connected Component

Strongly connected component (SCC)
is a set of nodes ***S*** so that:

Every pair of nodes in ***S*** can reach each other

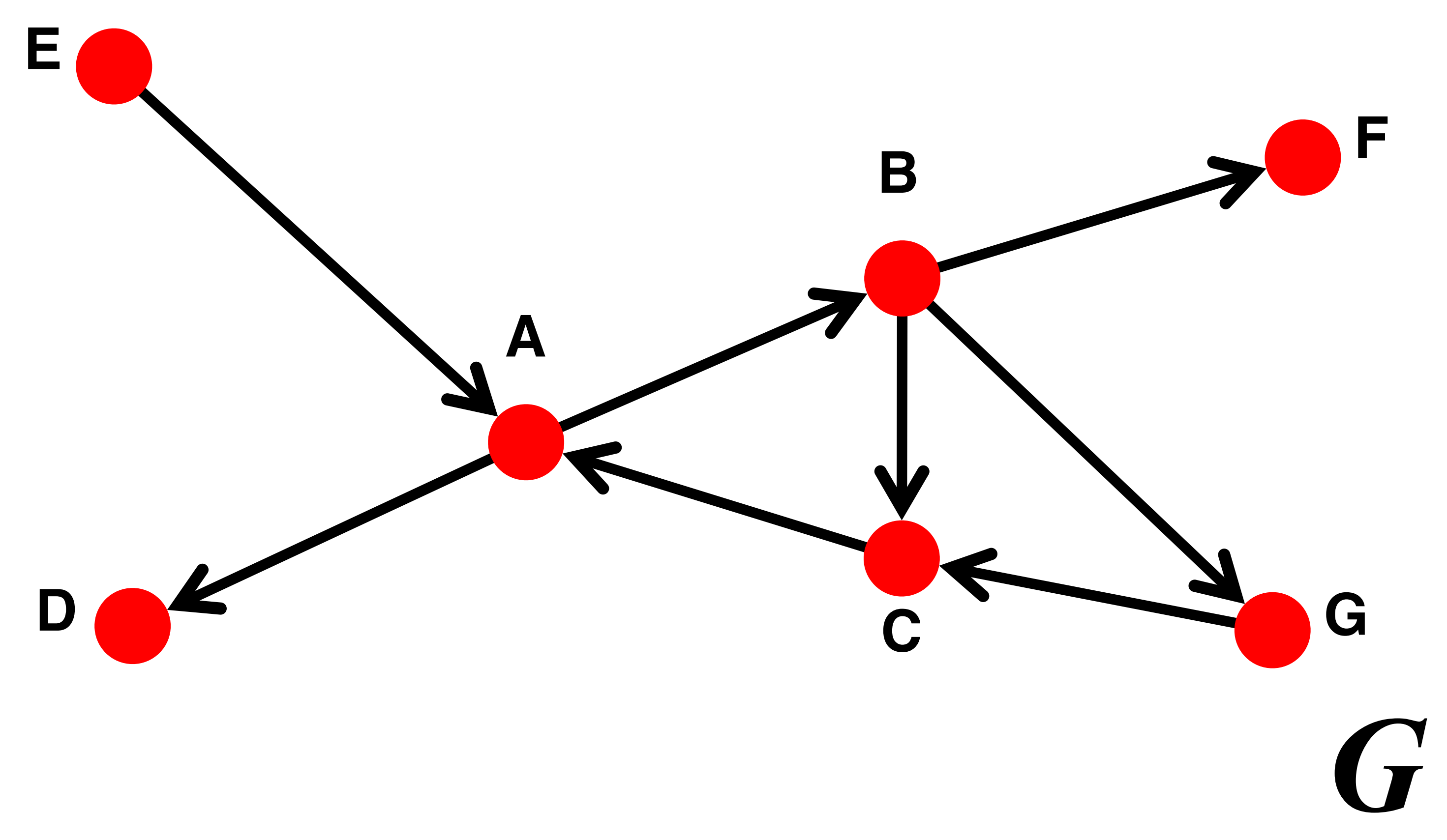
There is no larger set containing ***S*** with this property



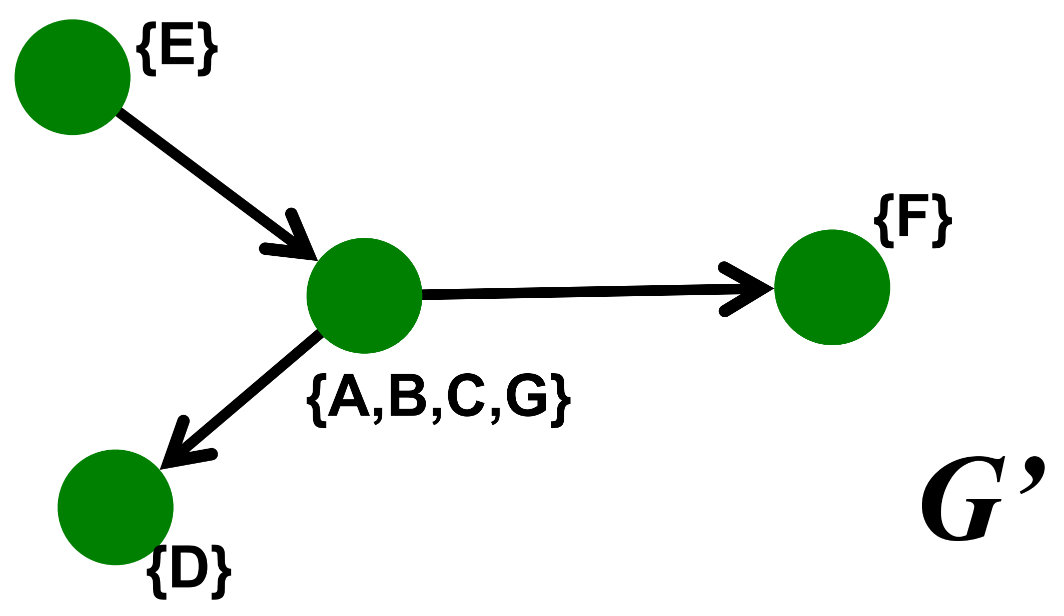
Strongly connected components of the graph: {A,B,C,G}, {D}, {E}, {F}

Strongly Connected Component

- **<u>Fact:</u>** **Every directed graph is a DAG on its SCCs**
  - **(1)** SCCs partitions the nodes of ***G***
    - That is, each node is in exactly one SCC
  - **(2)** If we build a graph ***G'*** whose nodes are SCCs, and with an edge between nodes of ***G'*** if there is an edge between corresponding SCCs in ***G***, then ***G'*** is a DAG

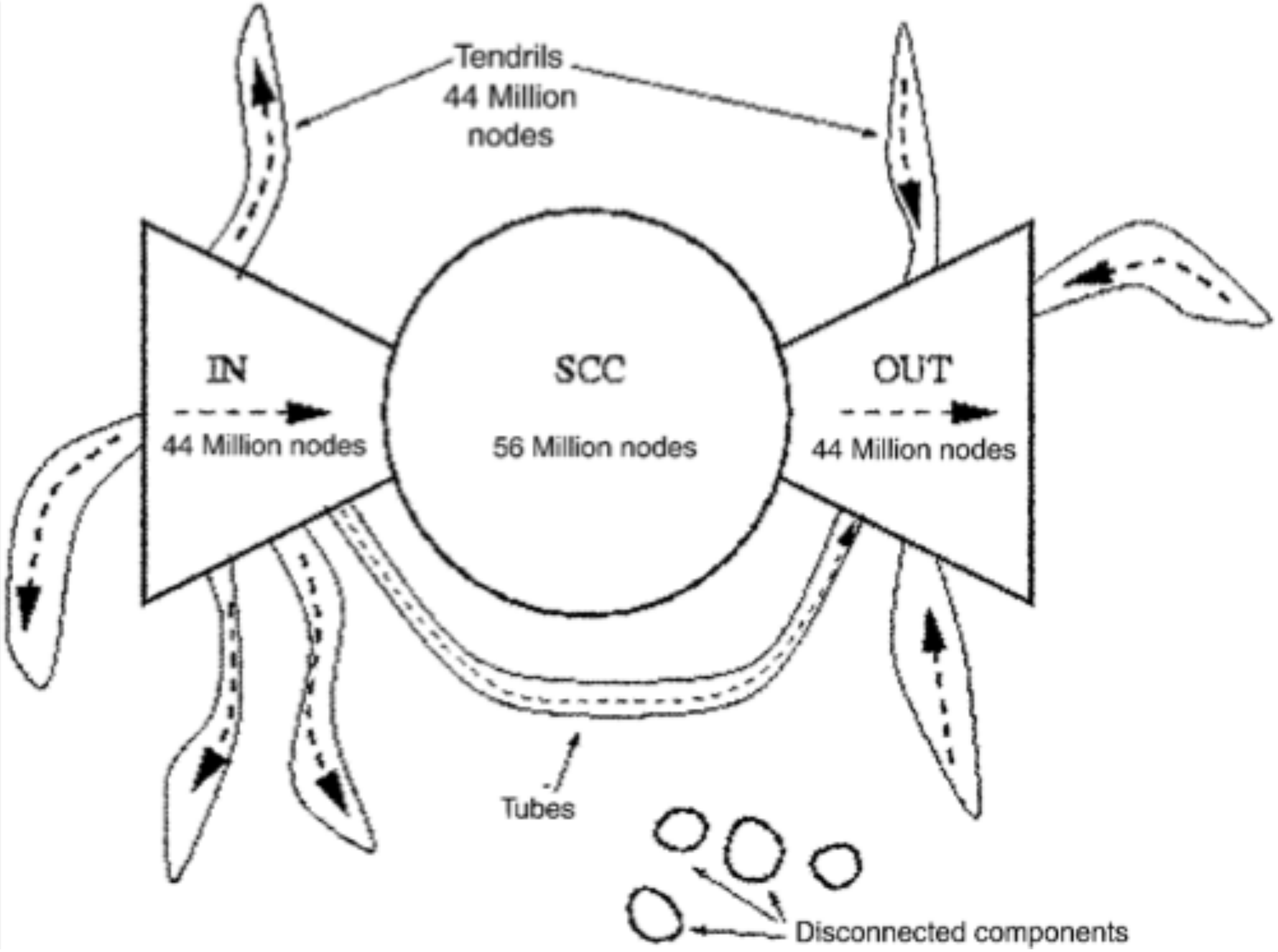


(1) Strongly connected components of graph G: {A,B,C,G}, {D}, {E}, {F}
(2) G' is a DAG:



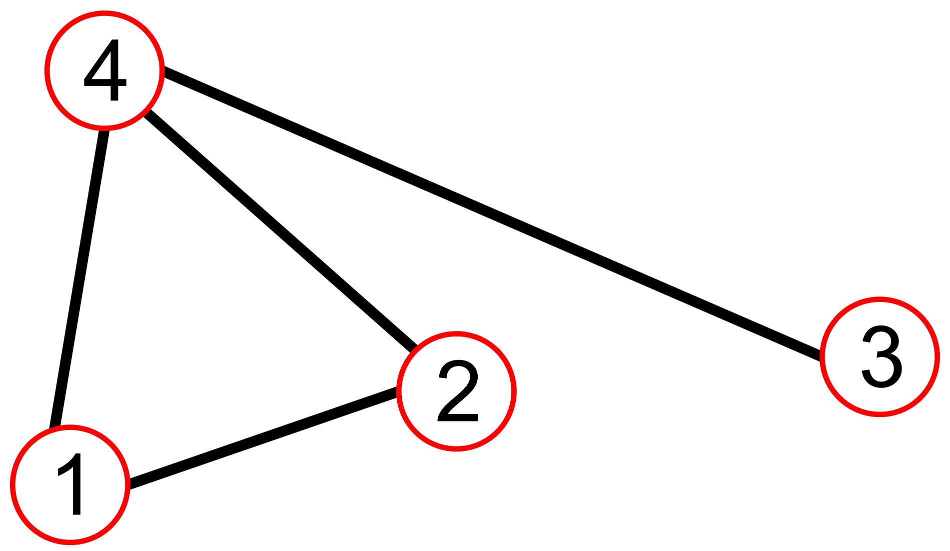
Bow-tie Structure of the Web

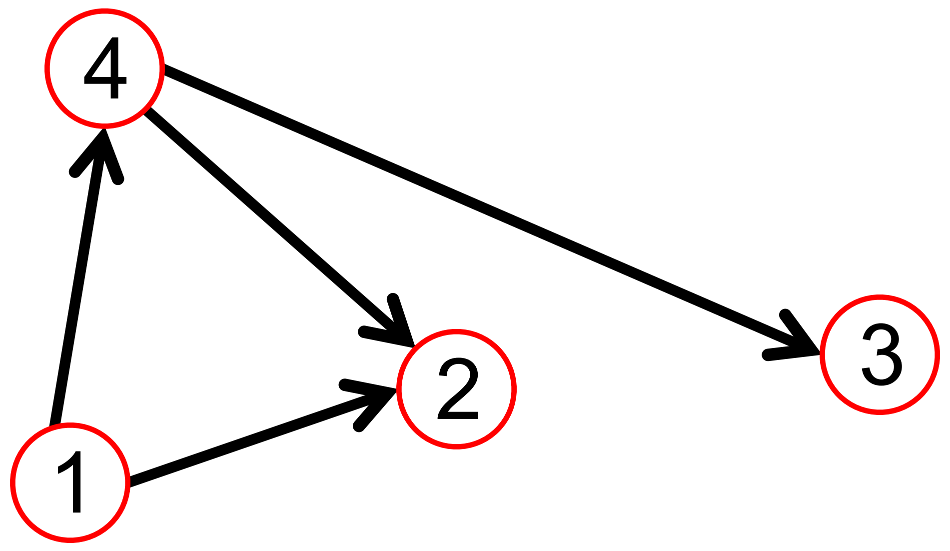
203 million pages, 1.5 billion links [Broder et al. 2000]



Lecture 2

Adjacency Matrix





$A_{ij} = 1$ if there is a link from node *i* to node *j*

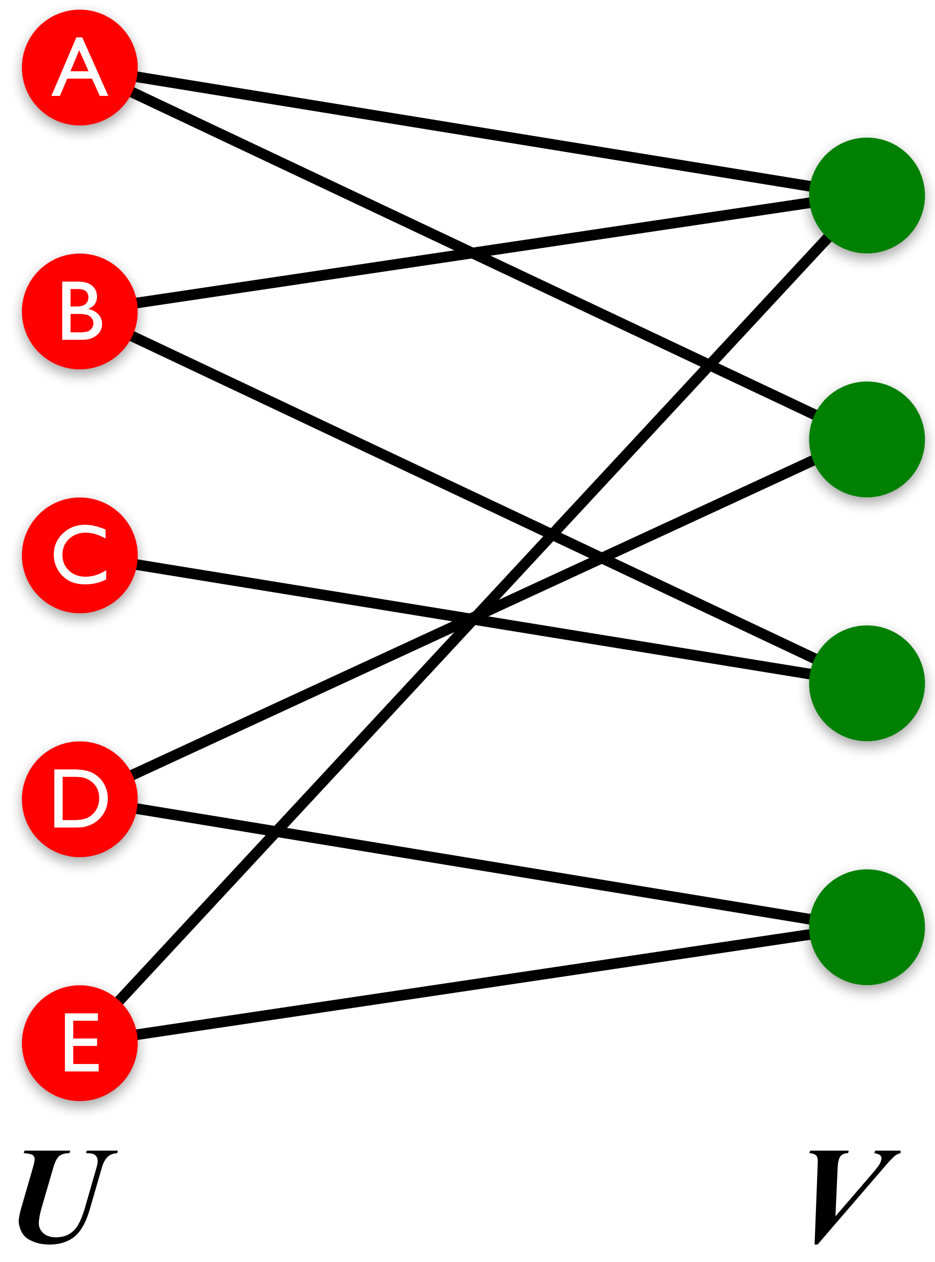
$A_{ij} = 0$ otherwise

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Note that for a directed graph (right) the matrix is not symmetric.

Bipartite Graph

Bipartite graph is a graph whose nodes can be divided into two disjoint sets U and V such that every link connects a node in U to one in V; that is, U and V are **independent sets**



Examples:

–Authors-to-papers (they authored)

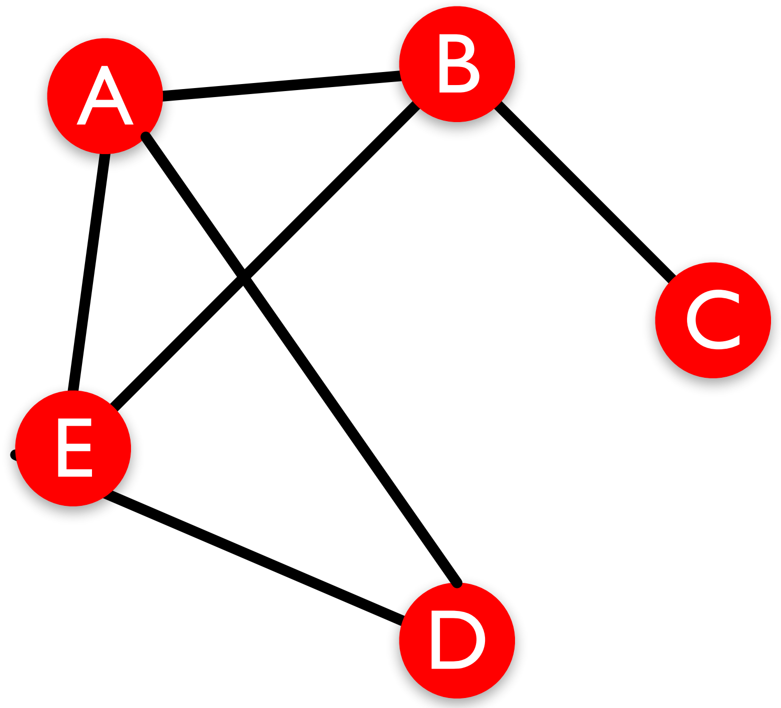
–Actors-to-Movies (they appeared in)

–Users-to-Movies (they rated)

"Folded" networks:

–Author collaboration networks

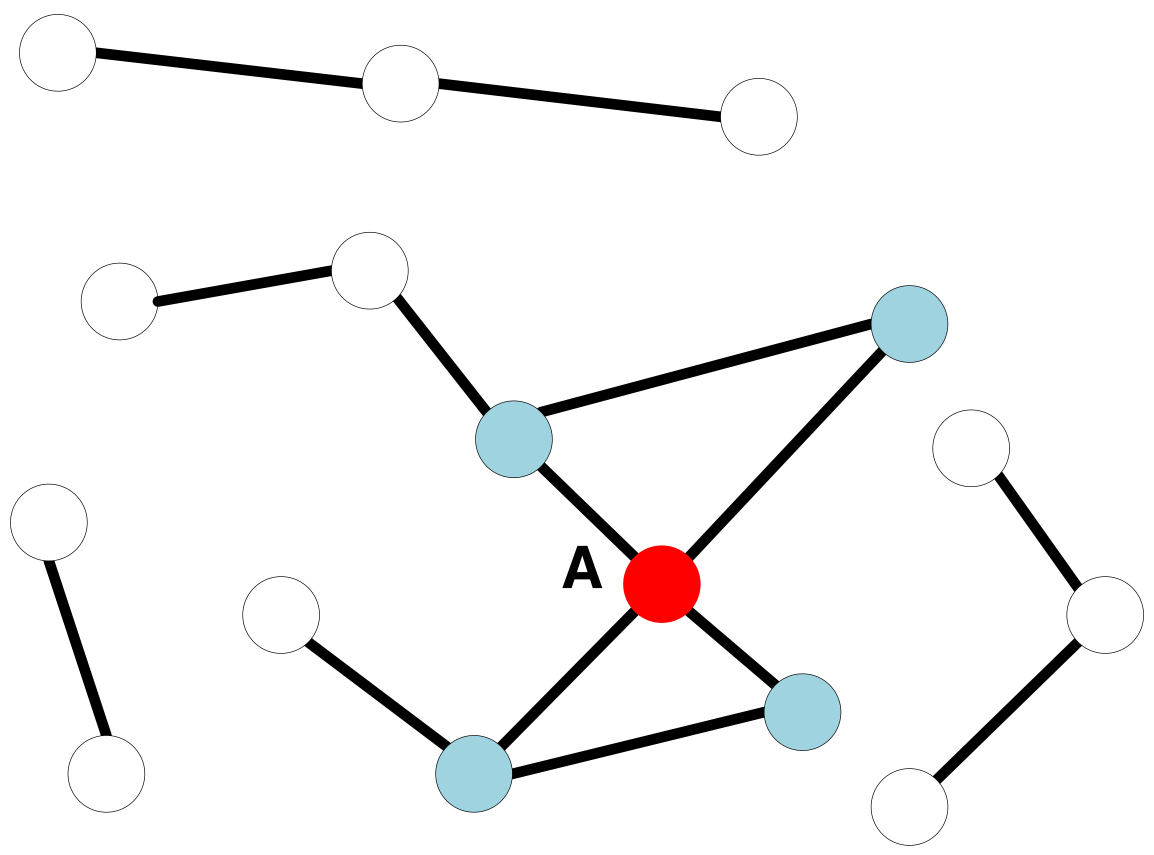
–Movie co-rating networks



Folded version of the graph above

Connectivity: Node Degrees

Undirected

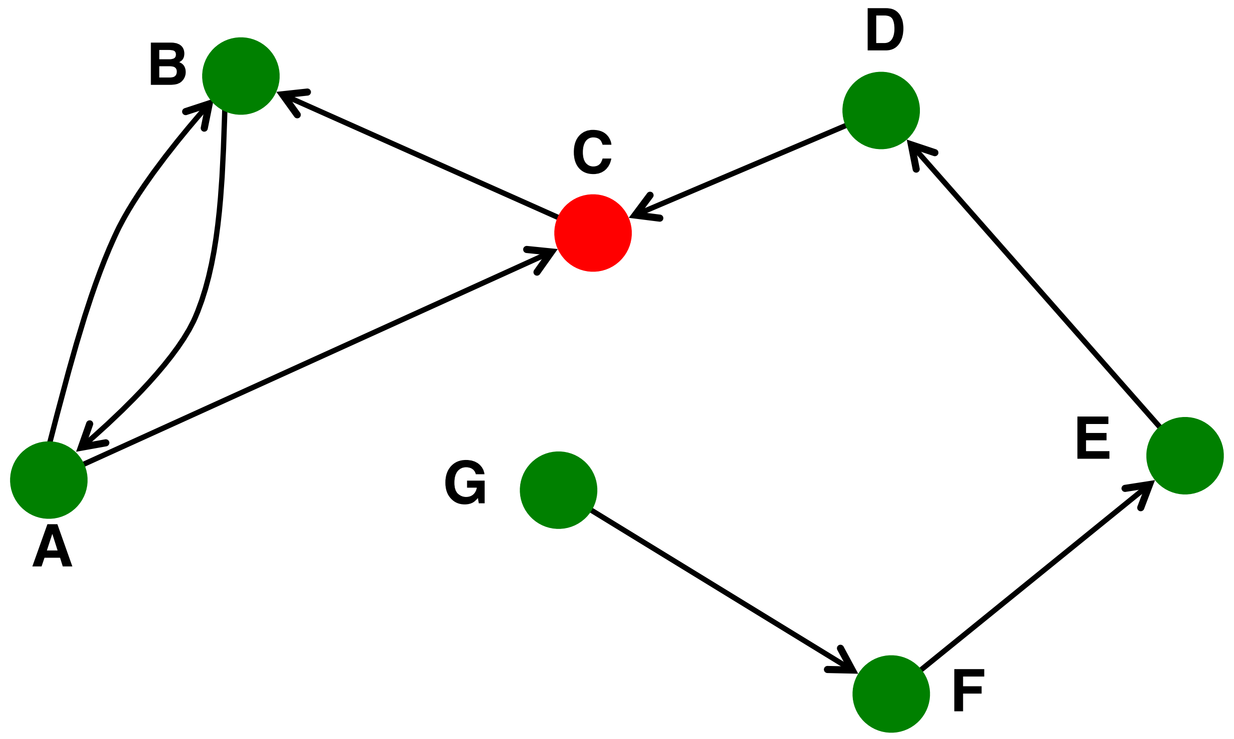


Node degree, k_i: the number of edges adjacent to node *i*

e.g. $k_A = 4$

Avg. degree: $$\bar{k} = \langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2E}{N}$$

Directed



In directed networks we define an in-degree and out-degree.
The (total) degree of a node is the sum of in- and out-degrees.

$$k_C^{in} = 2 \quad k_C^{out} = 1 \quad k_C = 3$$

Source: Node with $k^{in} = 0$
Sink: Node with $k^{out} = 0$

$$\overline{k^{in}} = \overline{k^{out}}$$

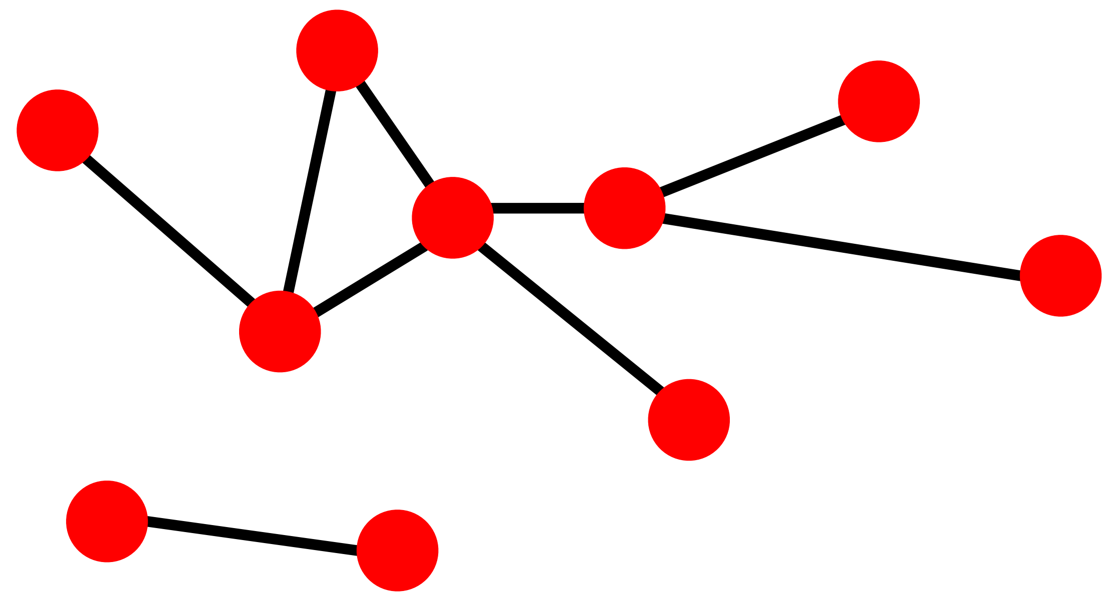
Connectivity: Degree Distribution

Degree distribution $P(k)$: Probability that a randomly chosen node has degree k

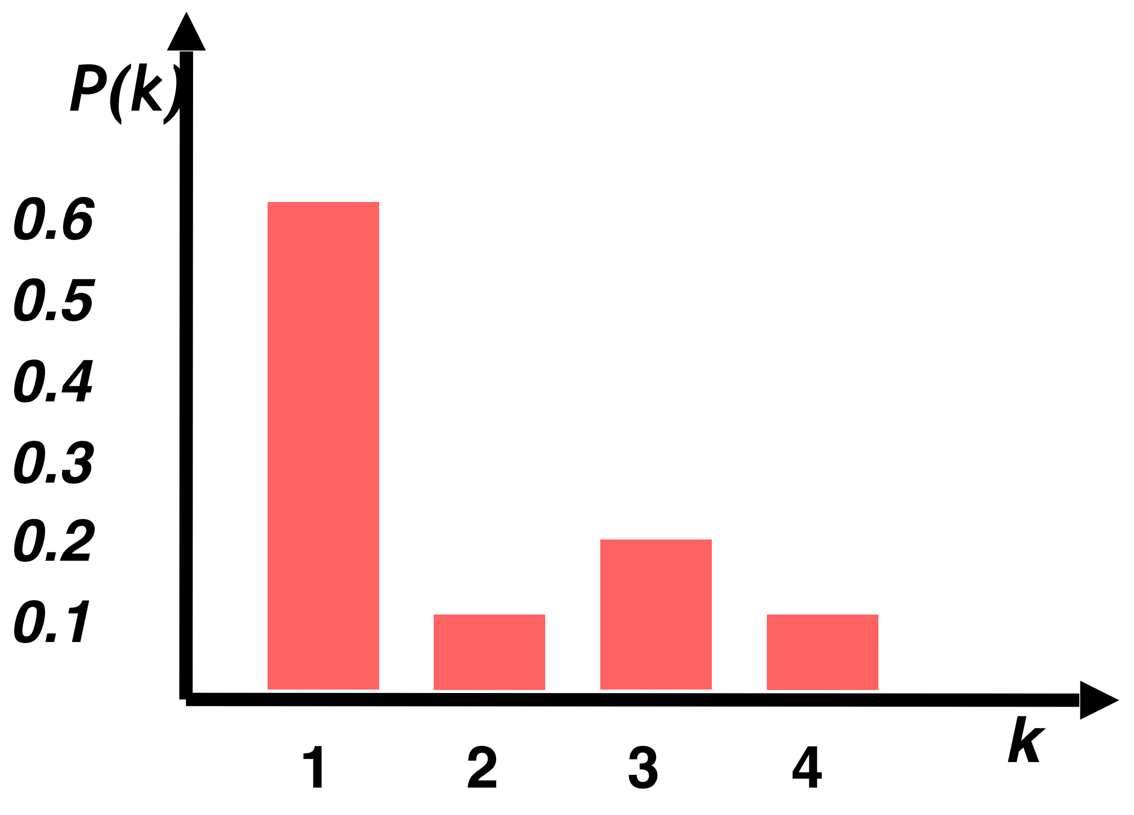
N_k = # nodes with degree k

Normalized histogram:

$P(k) = N_k / N$ → plot



→



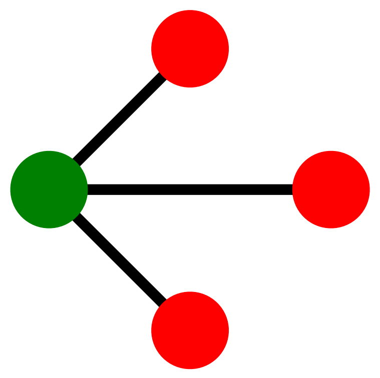
Connectivity: Clustering Coefficient

What's the probability that a random pair of your friends are connected?

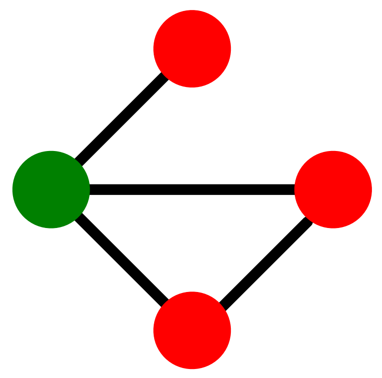
$$C_i \in [0, 1]$$

$$C_i = \frac{e_i}{\binom{k_i}{2}} = \frac{e_i}{k_i(k_i - 1)/2} = \frac{2e_i}{k_i(k_i - 1)}$$

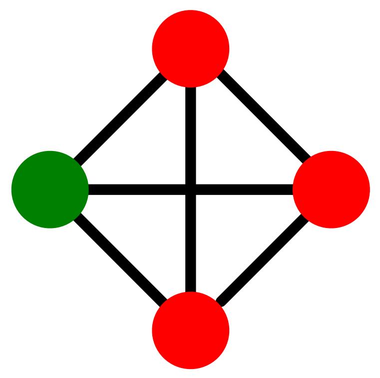
where e_i is the number of edges between the neighbors of node I and k_i is the degree of node I



$C_i = 0$



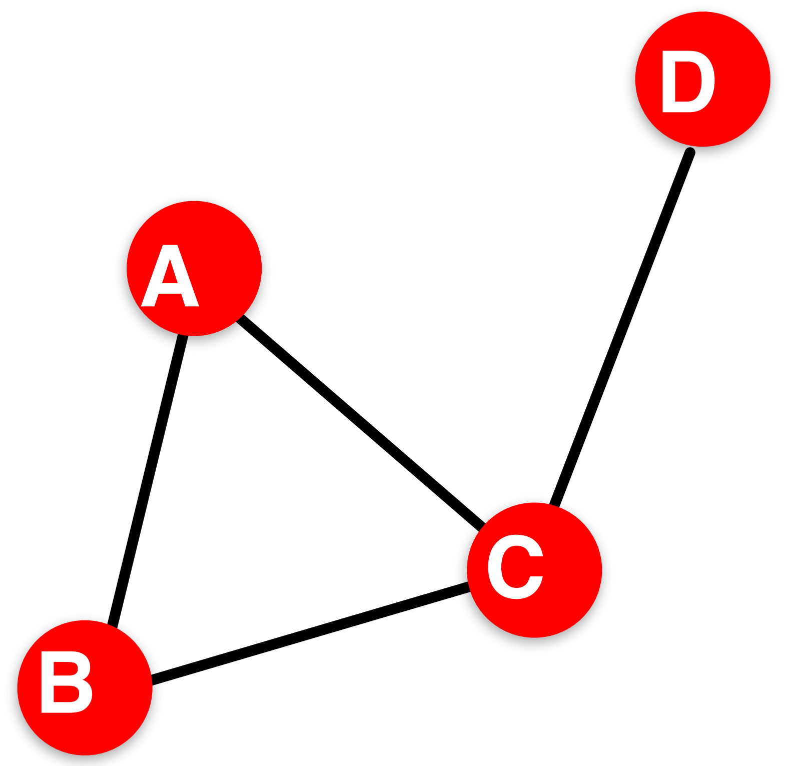
$C_i = 1/3$



$C_i = 1$

Average clustering coefficient: $C = \frac{1}{N} \sum_{i}^{N} C_i$

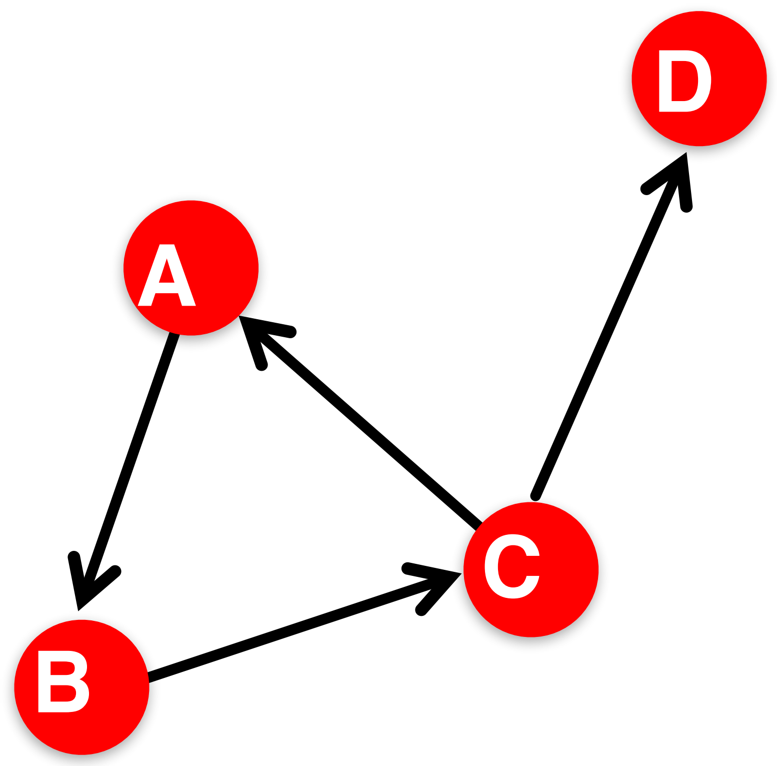
Distance: definition



$h_{B,D} = 2$

Distance (shortest path, geodesic) between a pair of nodes is defined as the number of edges along the shortest path connecting the nodes

*If the two nodes are disconnected, the distance is usually defined as infinite



$h_{B,C} = 1$, $h_{C,B} = 2$

In **directed graphs** paths need to follow the direction of the arrows

Consequence: Distance is **not symmetric**: $h_{A,C} \neq h_{C,A}$

Distance: Graph-level measures

- **Diameter:** the maximum (shortest path) distance between any pair of nodes in a graph

- **Average path length** for a connected graph (component) or a strongly connected (component of a) directed graph

$$\bar{h} = \frac{1}{2E_{max}} \sum_{i, j \neq i} h_{ij}$$

where h_{ij} is the distance from node i to node j,
And Emax is the maximum number of edges (=n*(n-1)/2)

  - Many times we compute the average only over the connected pairs of nodes (that is, we ignore “infinite” length paths)

Simplest Model of Graphs

Erdös-Renyi Random Graphs [Erdös-Renyi, ‘60]

$G_{n,p}$: undirected graph on n nodes and each edge (u,v) appears i.i.d. with probability p

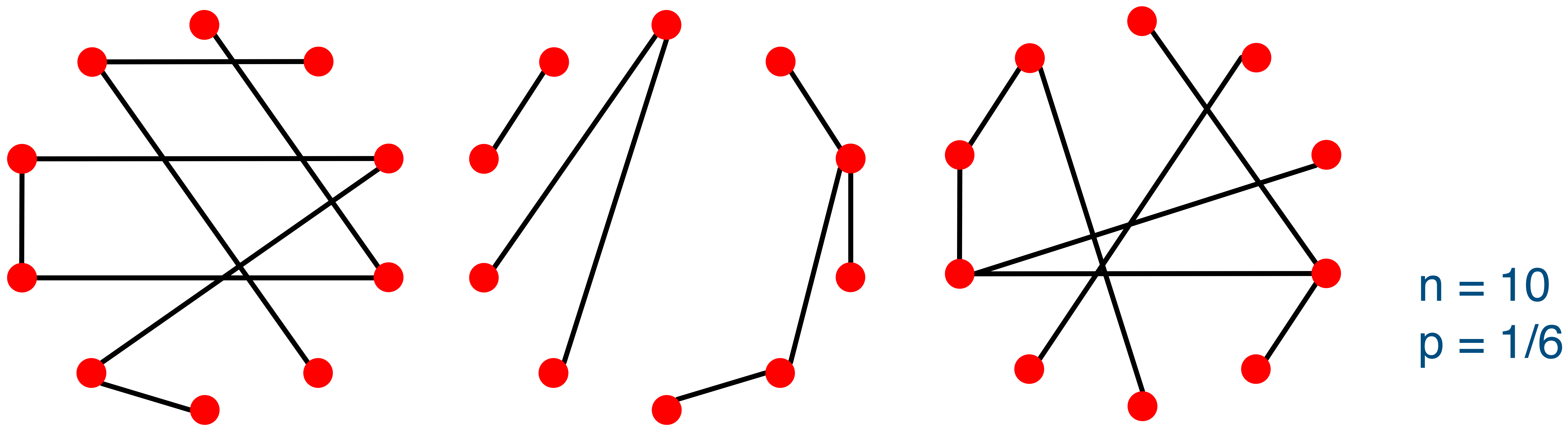
Simplest random model you can think of

Random Graph Model

n and *p* do not uniquely determine the graph!

The graph is a result of a random process

We can have many different realizations given the same *n* and *p*



Degree Distribution

Fact: Degree distribution of G_{np} is <u>Binomial</u>.

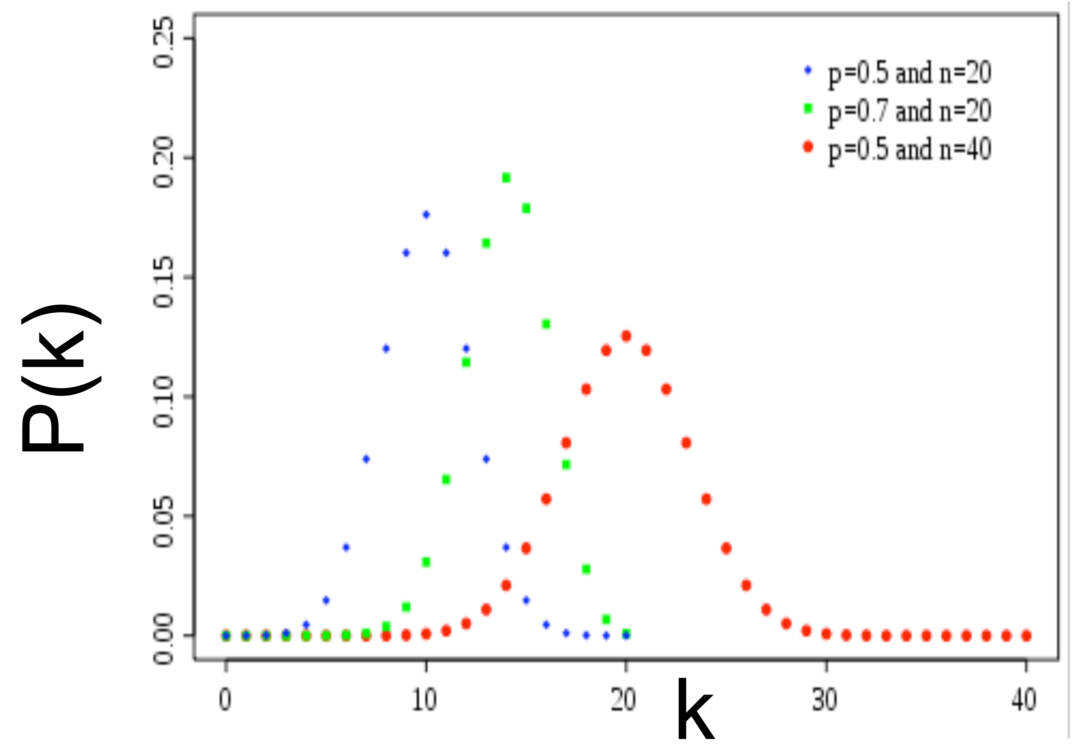
Let $P(k)$ denote a fraction of nodes with degree k:

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

Select k nodes out of $n-1$

Probability of having k edges

Probability of missing the rest of the $n-1-k$ edges

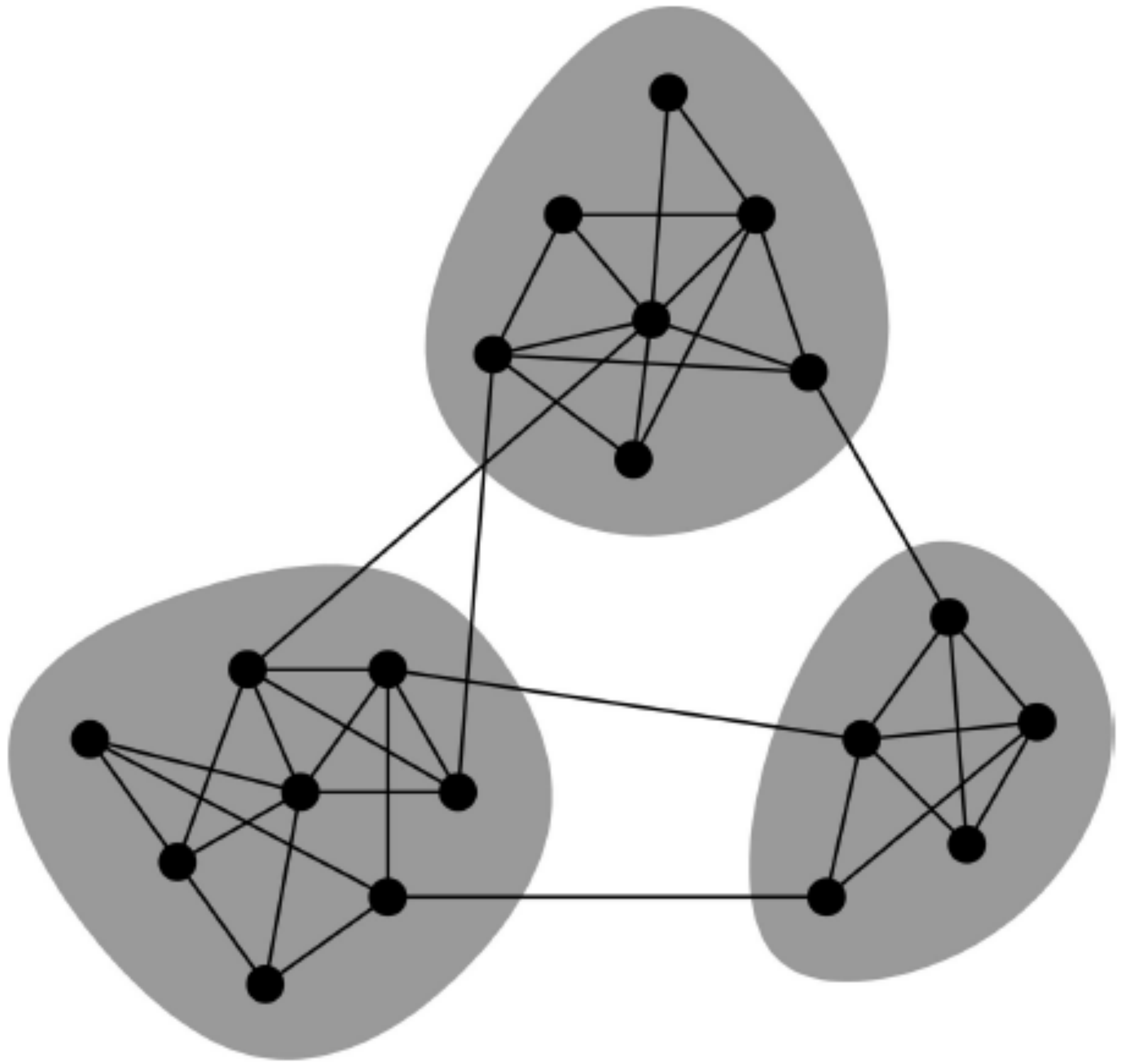


$$\bar{k} = p(n-1)$$

Lecture 3

Networks & Communities

We often think of networks "looking" like this:

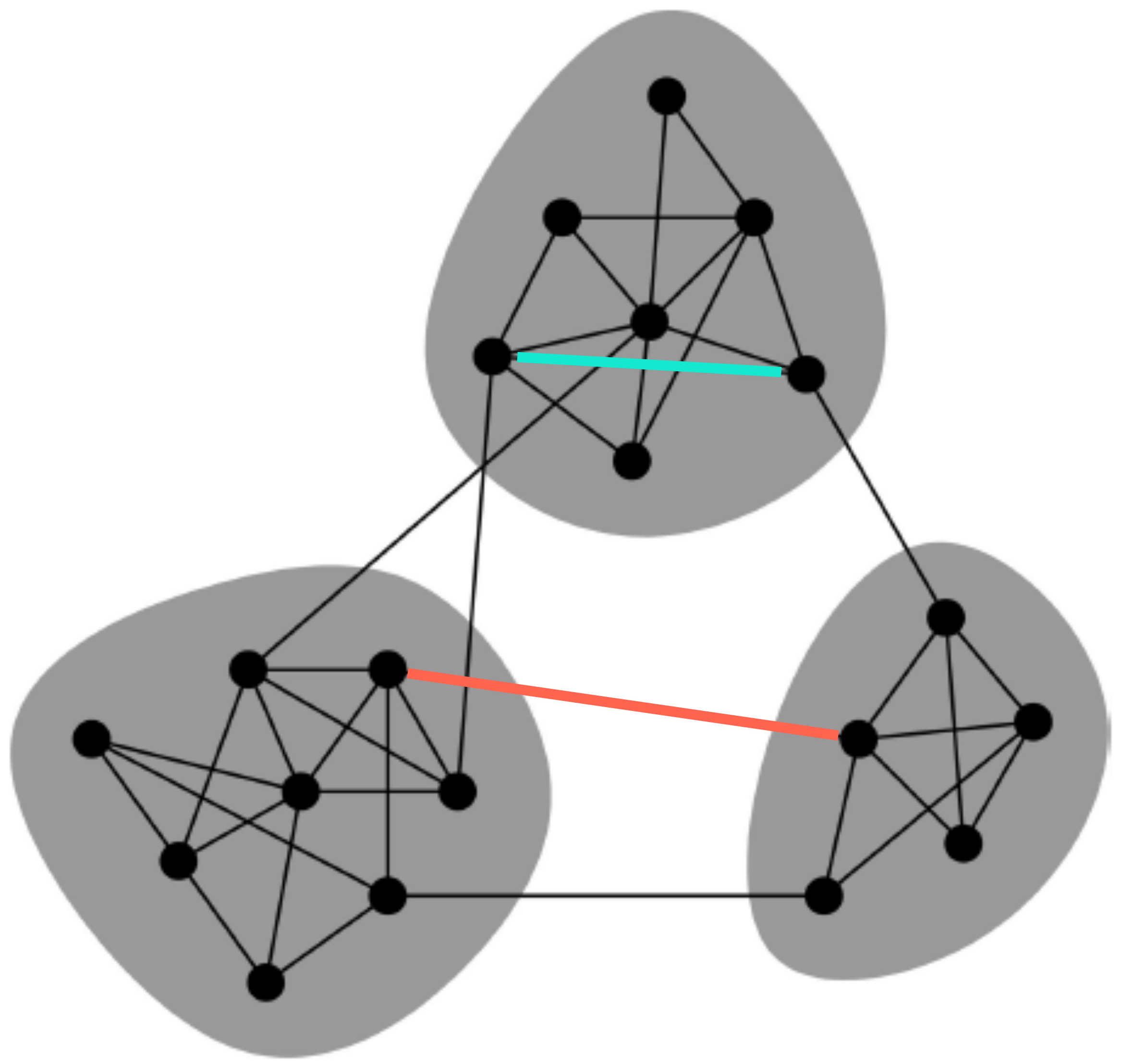


What can lead to such a conceptual picture?

Granovetter's Answer

Two perspectives on friendships:

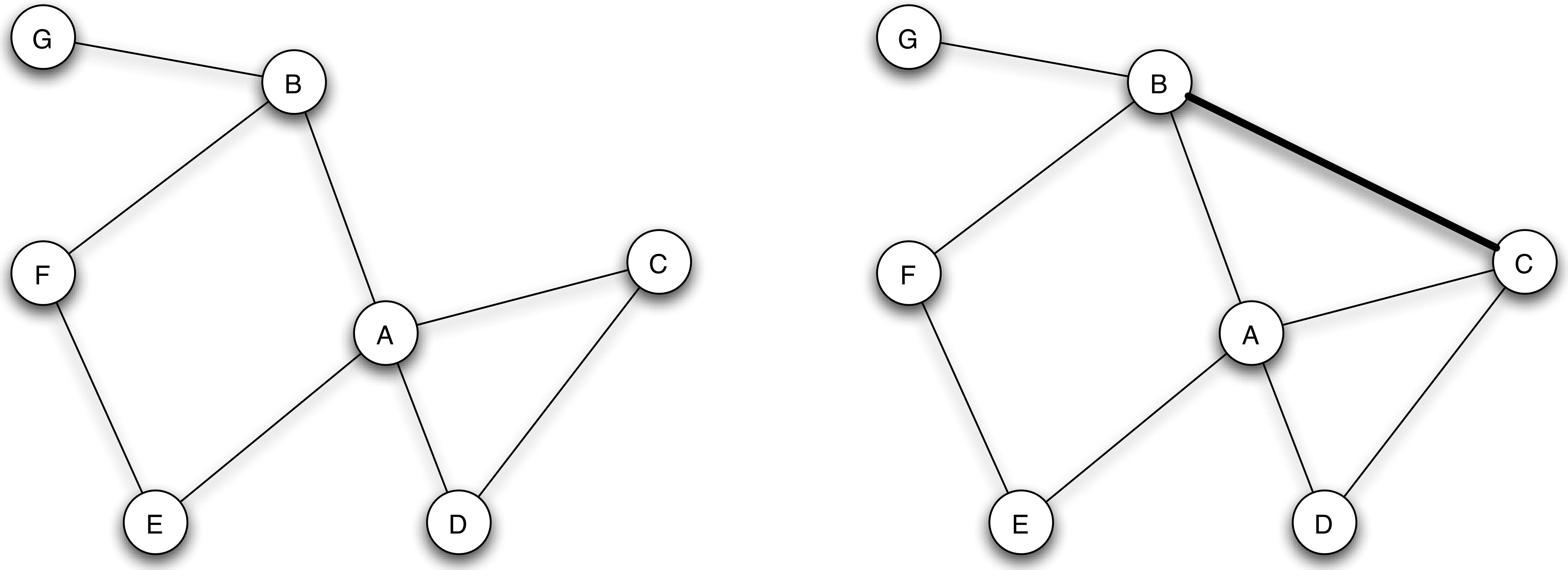
Structural: Friendships span different parts of the network



The two highlighted edges are structurally different: one spans two different "communities" and the other is inside a community

Interpersonal: Friendship between two people vary in strength, you can be close or not so close to someone

Triadic closure



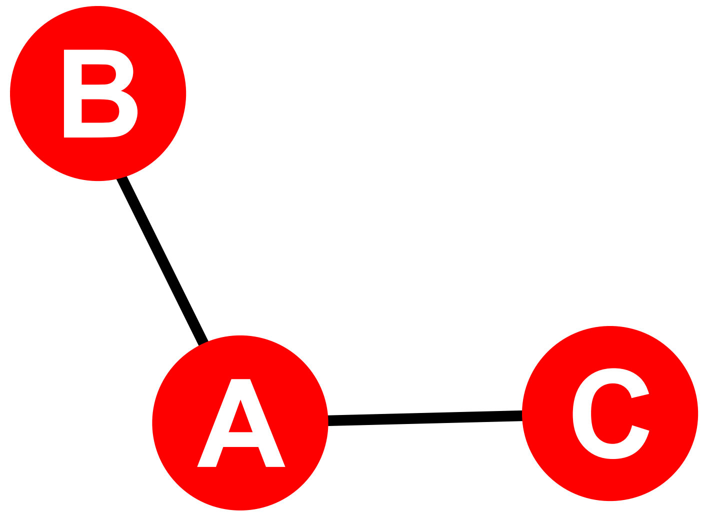
Informally: If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future.

Triadic Closure

Triadic closure == High clustering coefficient

Reasons for triadic closure:
If **B** and **C** have a friend **A** in common:

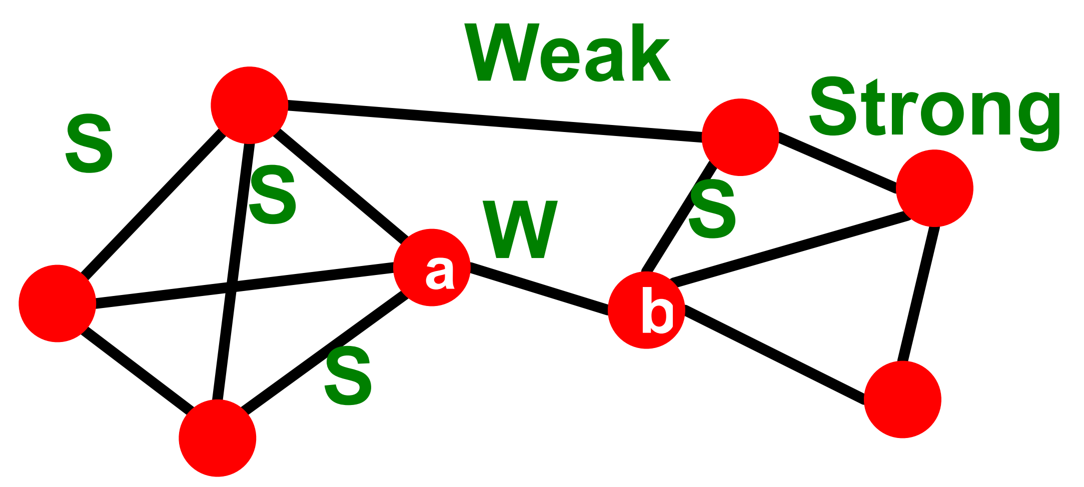
- **B** is **more likely to meet C**
  (both spend time with **A**)
- **B** and **C** **trust each other more**
  (they have a friend in common)
- **A** has an **incentive to bring B and C together**
  (easier for **A** to maintain two disjoint relationships)



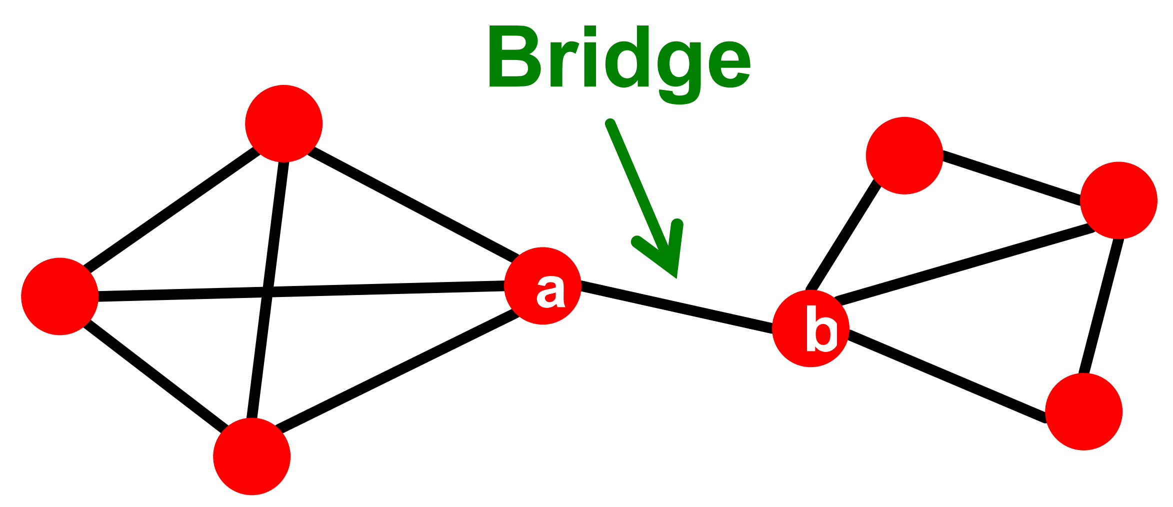
Granovetter's Explanation

Granovetter makes a connection between the social and structural roles of an edge

- **First point: Structure**
  - Structurally embedded edges are also socially strong
  - Long-range edges spanning different parts of the network are socially weak
- **Second point: Information**
  - Long-range edges allow you to gather information from different parts of the network and get a job
  - Structurally embedded edges are heavily redundant in terms of information access



Network Vocabulary: Span and Bridges



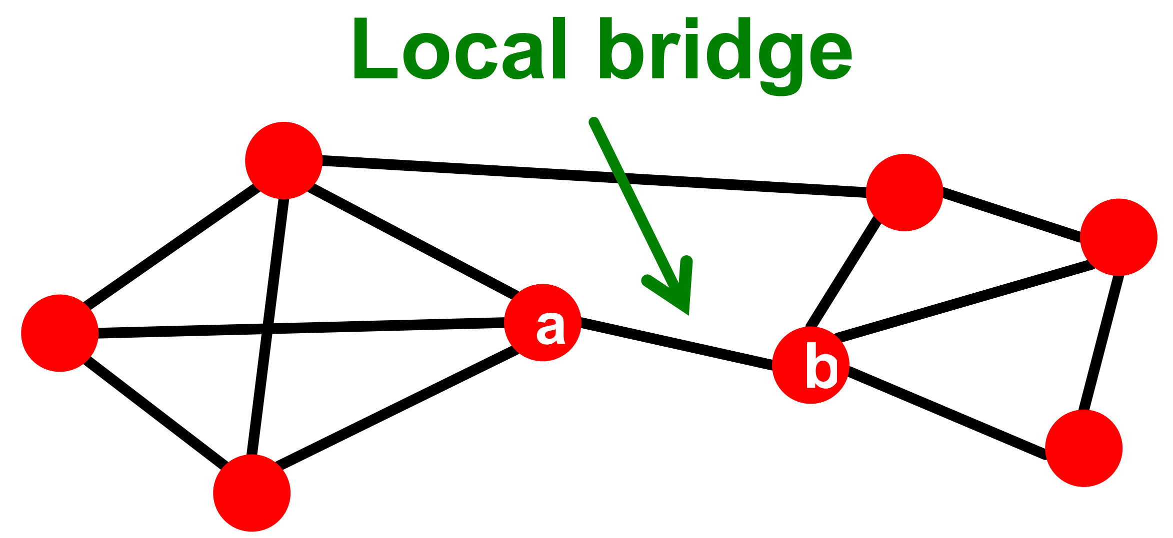
Define: Span

The **Span** of an edge is the distance of the edge endpoints if the edge is deleted.

Define: Bridge edge

If removed, it disconnects the graph

Span of a bridge edge = ∞



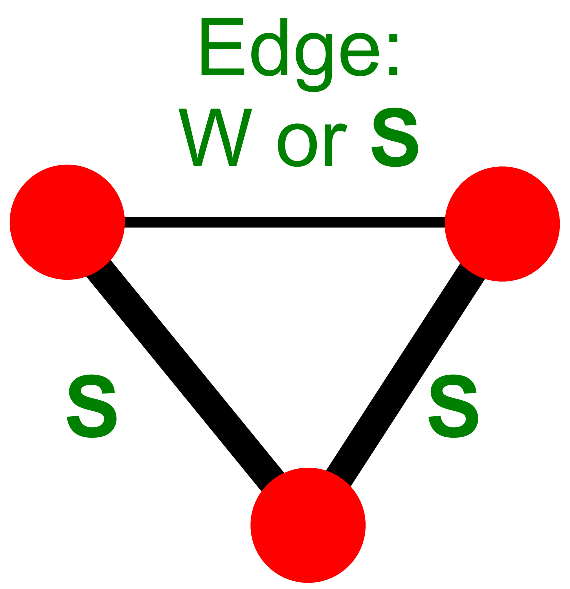
Define: Local bridge

Edge of **Span > 2**

(any edge that doesn't close a triangle)

Idea: Local bridges with long span are like real bridges

Granovetter's Explanation



Model: Two types of edges:

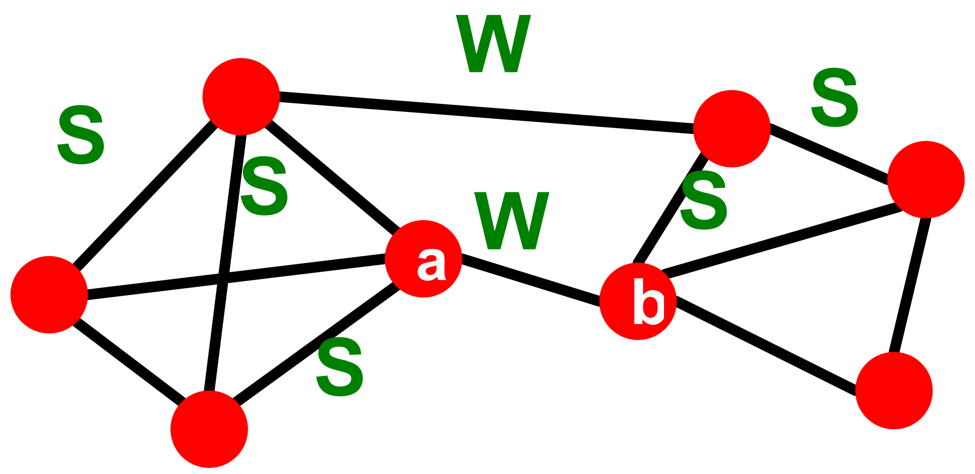
Strong (friend), **Weak** (acquaintance)

Model: Strong Triadic Closure property:

Two strong ties imply a third edge

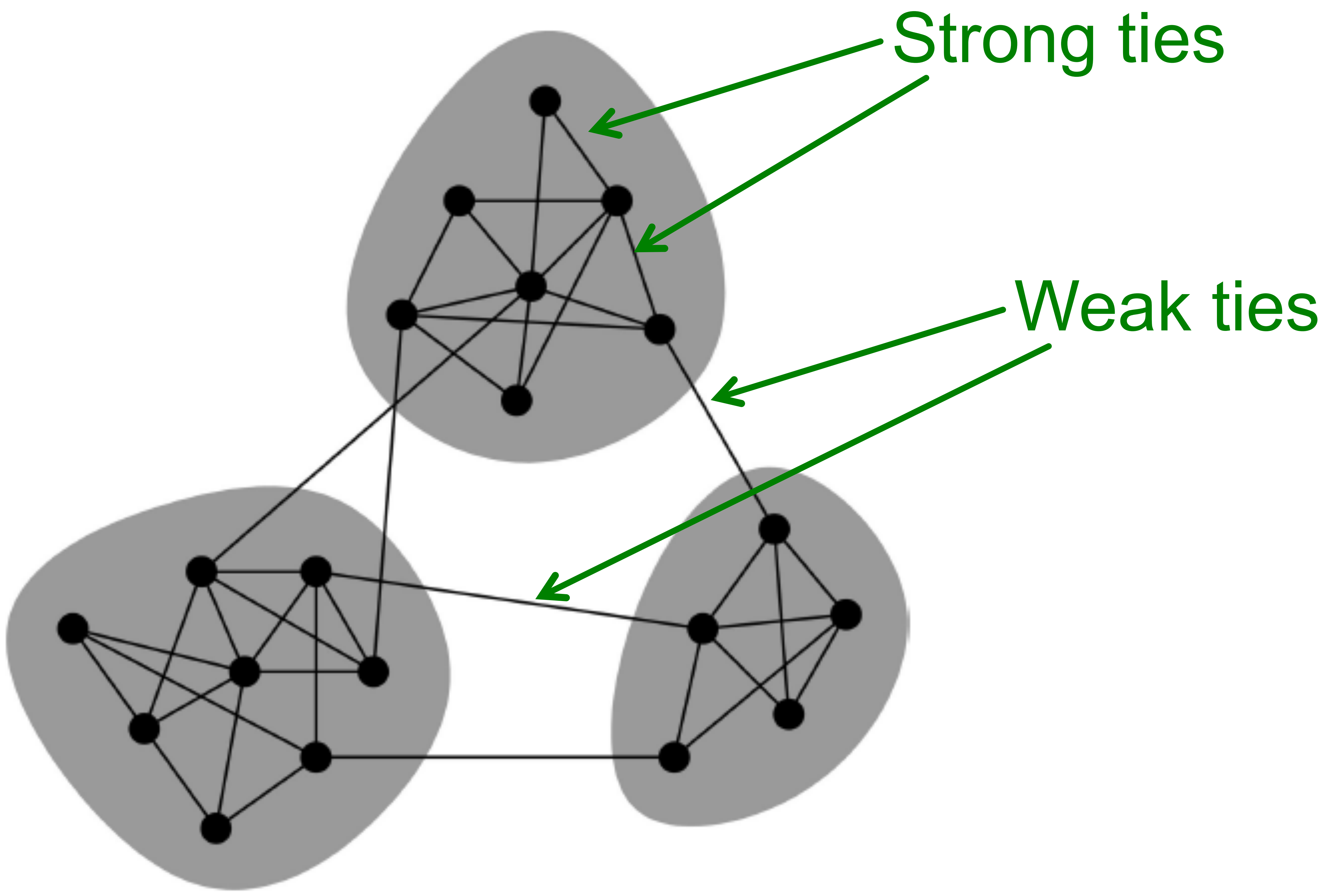
If node A has strong ties to both nodes B and C, then there must be an edge (strong or weak) between B and C

Fact: If strong triadic closure is satisfied then **local bridges are weak ties!**

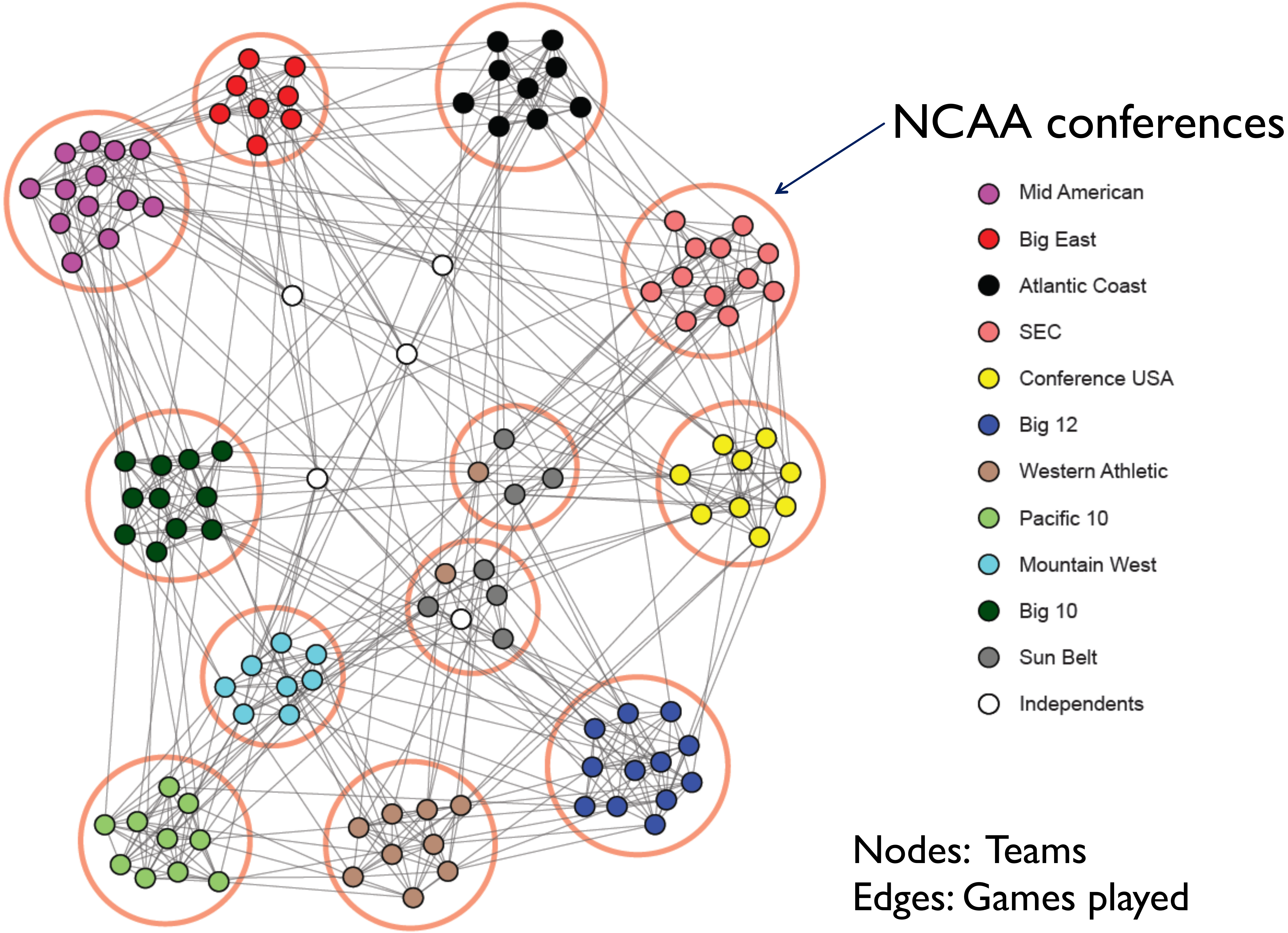


Conceptual Picture of Networks

Granovetter's theory leads to the following conceptual picture of networks



NCAA Football Network



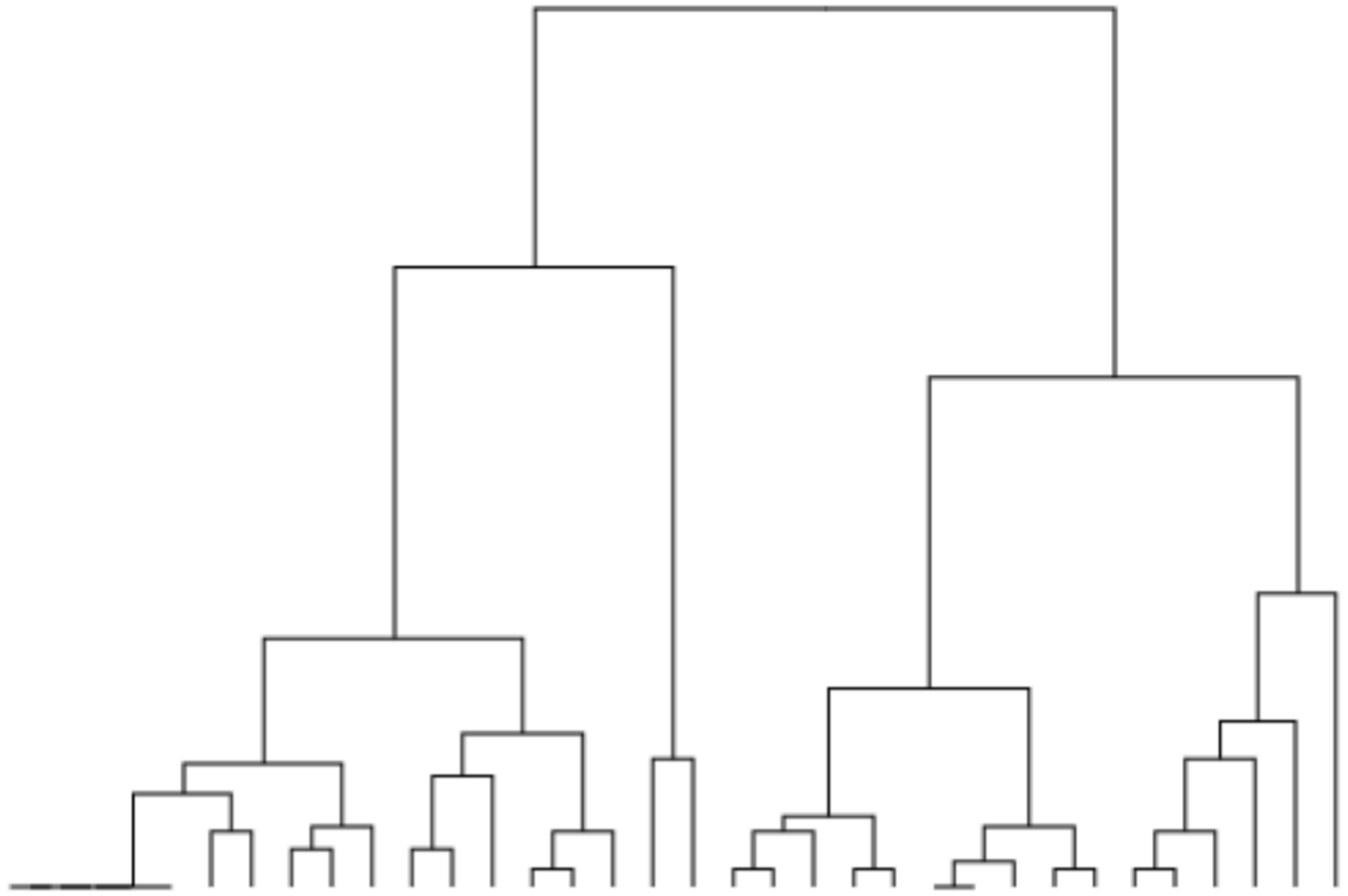
Nodes: Teams
Edges: Games played

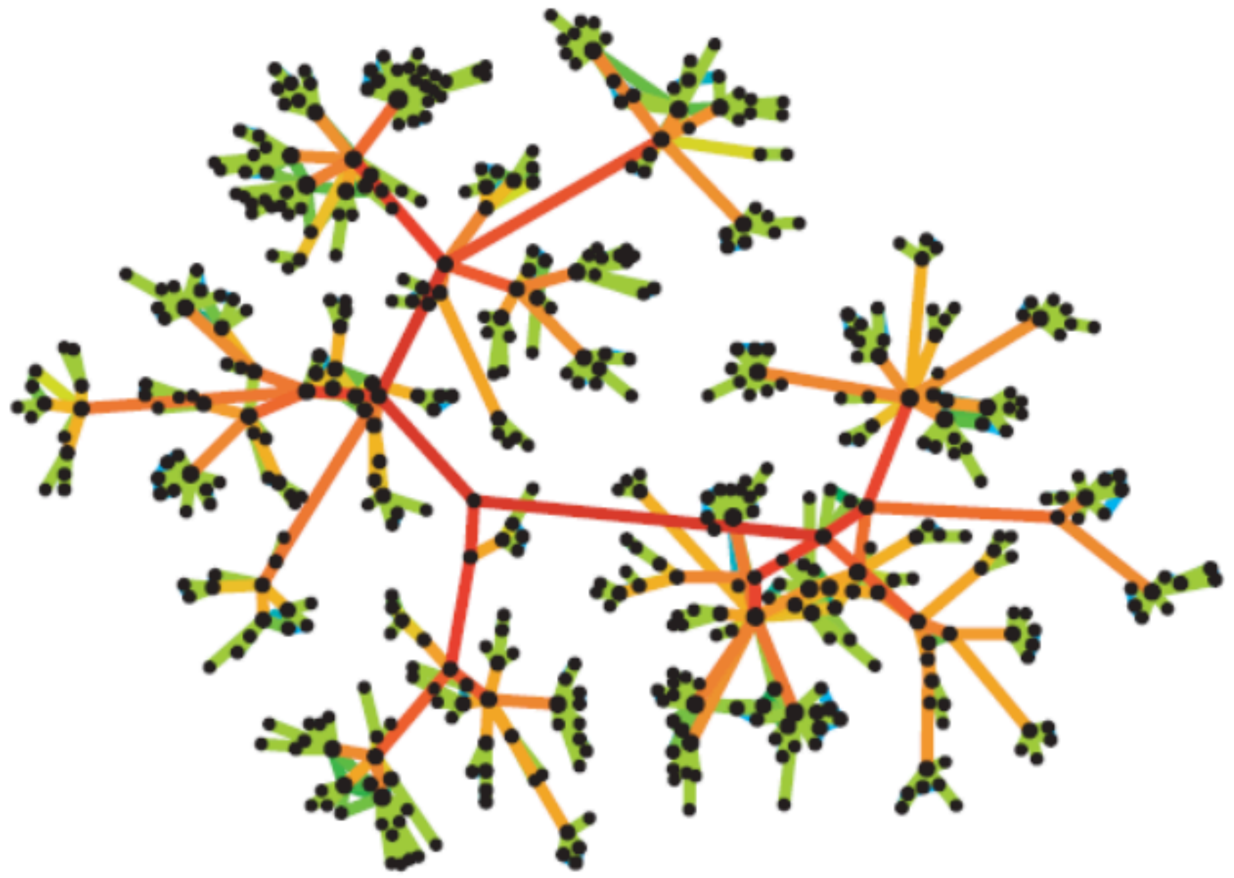
Graph Partitioning

Two general approaches:

1. Start with every node in the same cluster and break apart at "weak links" ("**divisive clustering**")

2. Start with every node in its own "community" and join communities that are close together ("**agglomerative clustering**")

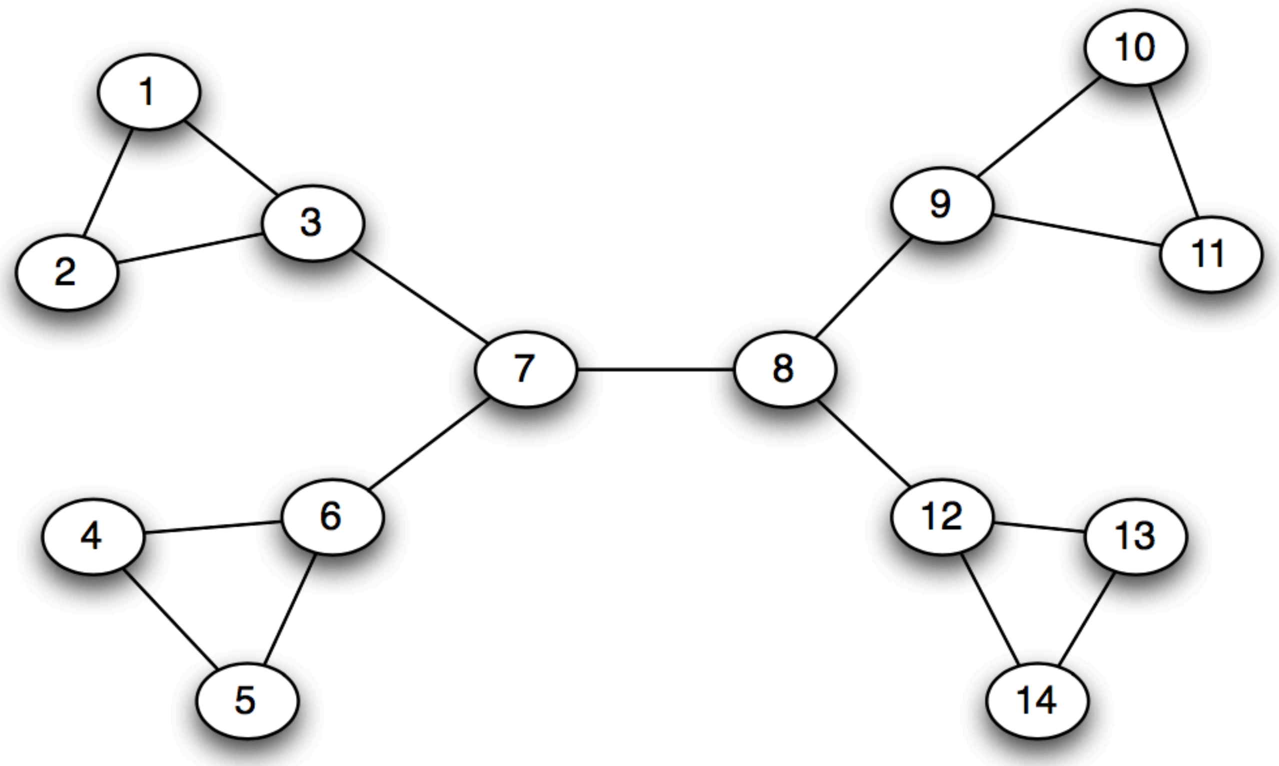




Graph Partitioning

Definition: the **betweenness** of an edge is how many (fractional) shortest paths travel through it

–For every pair of nodes A,B say there is one unit of "flow" along the edges from A to B
–Flow between A to B divides evenly among all shortest paths from A to B
–If k shortest paths, 1/k flow on each path



Girvan-Newman algorithm

Divisive hierarchical clustering based on the notion of edge **betweenness** (Number of shortest paths passing through an edge)

Girvan-Newman Algorithm (on undirected unweighted networks)**:**

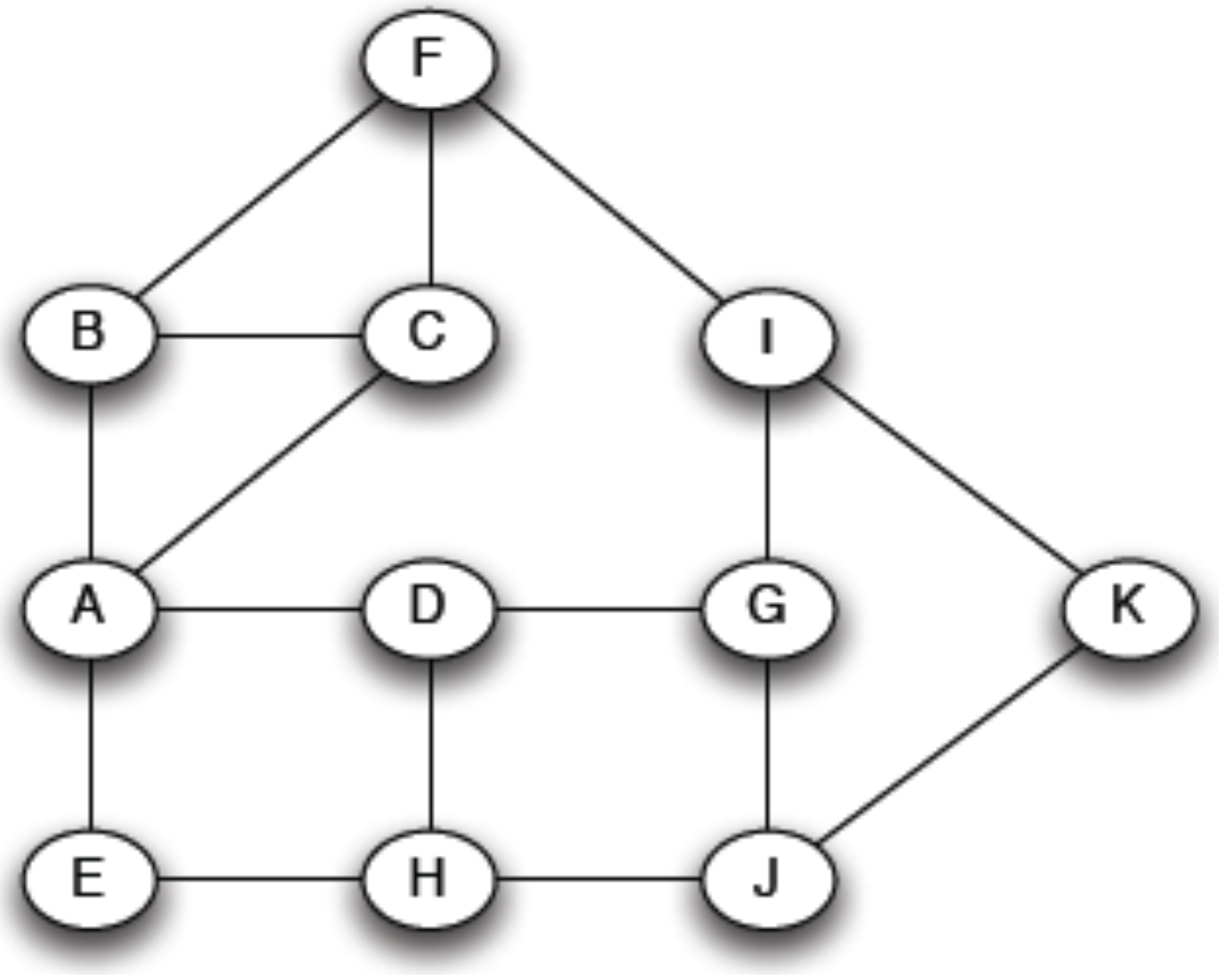
Repeat until no edges are left:

- –(Re)calculate betweenness of every edge
- –Remove edges with highest betweenness (if ties, remove all edges tied for highest)
- –Connected components are communities

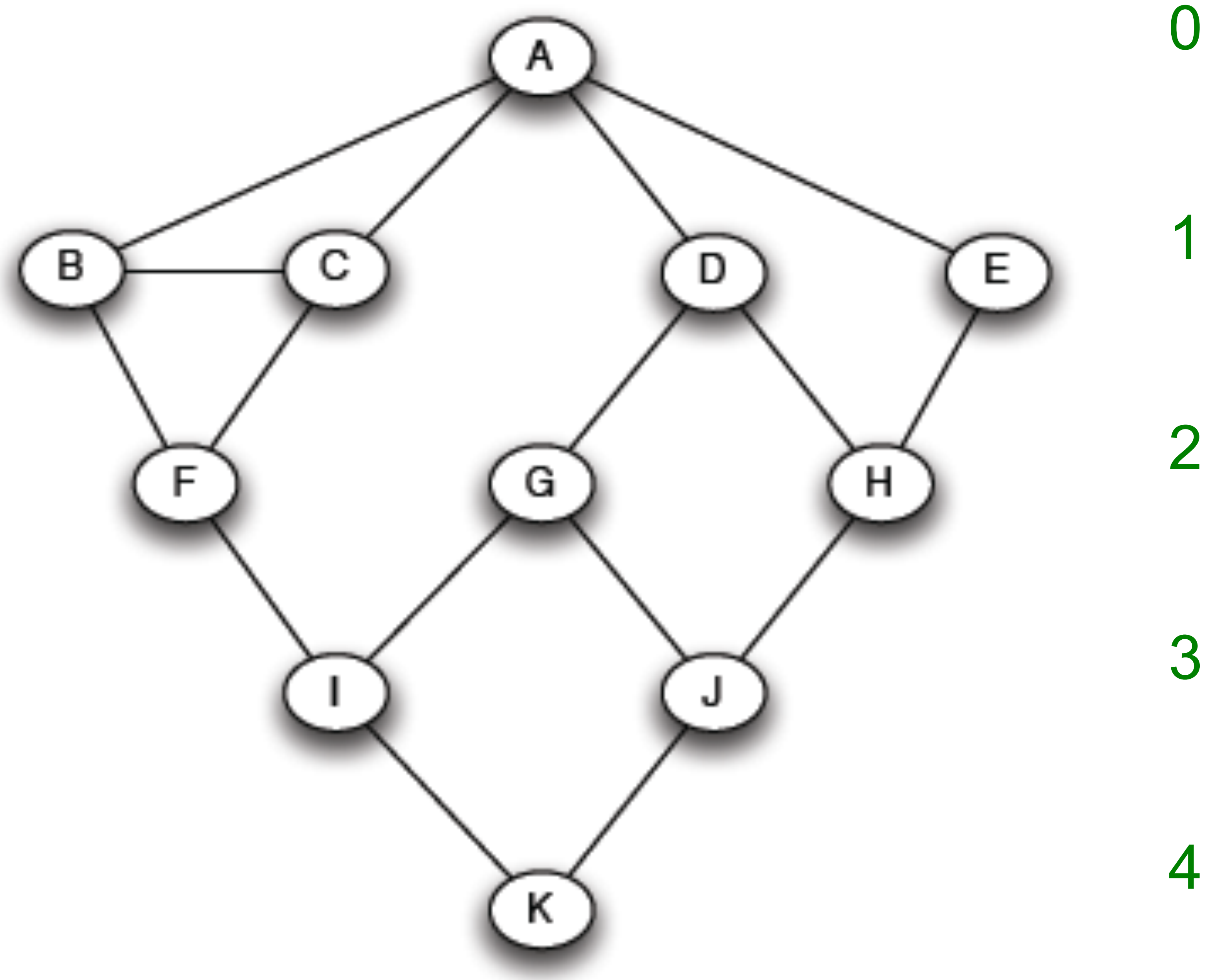
Gives a hierarchical decomposition of the network

How to Compute Betweenness?

Want to compute betweenness of paths starting at node A



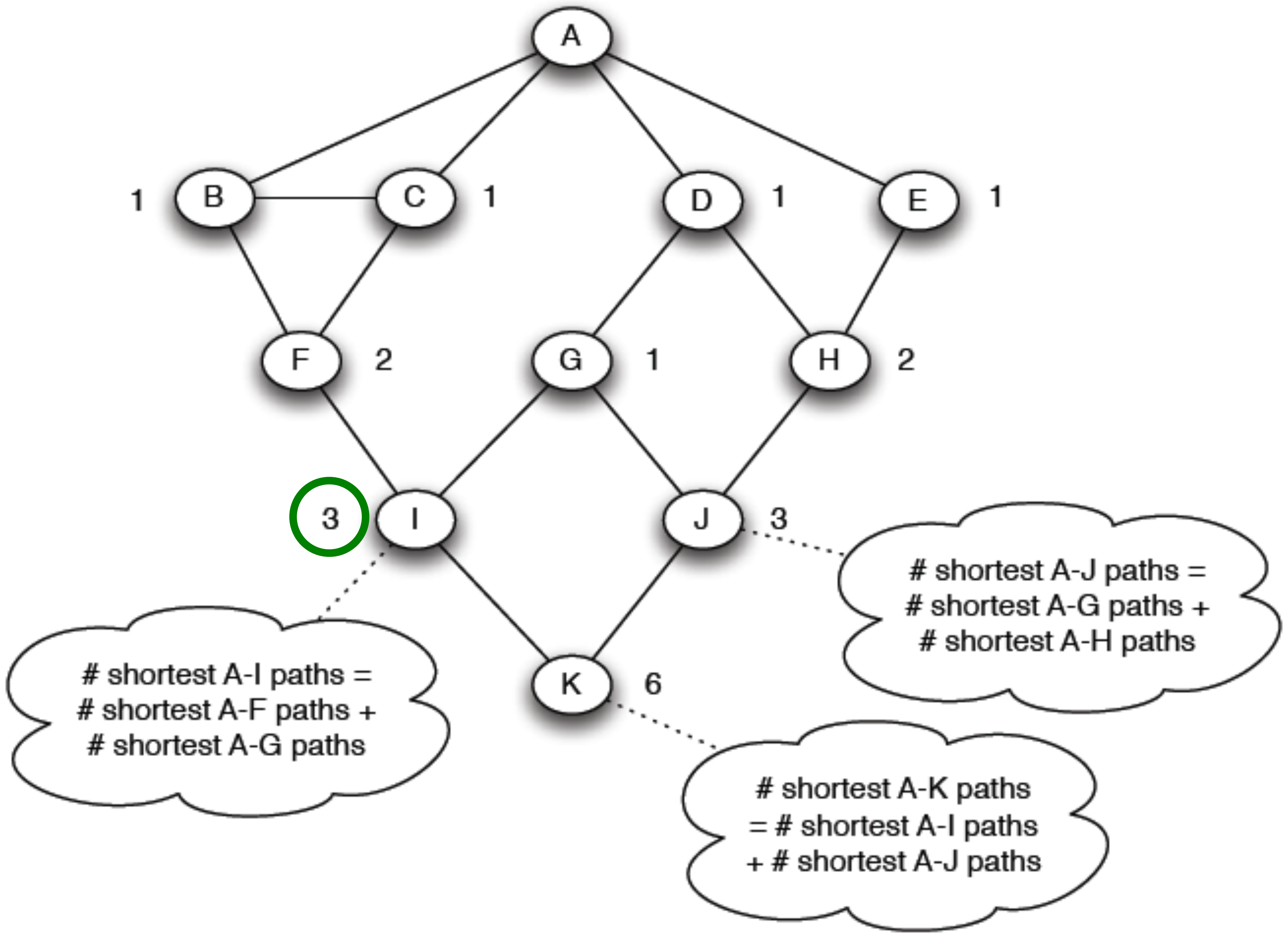
BFS starting from A:



Recall BFS goes layer-by-layer

How to Compute Betweenness?

Count the number of shortest paths from A to all other nodes in the graph:



Work downwards

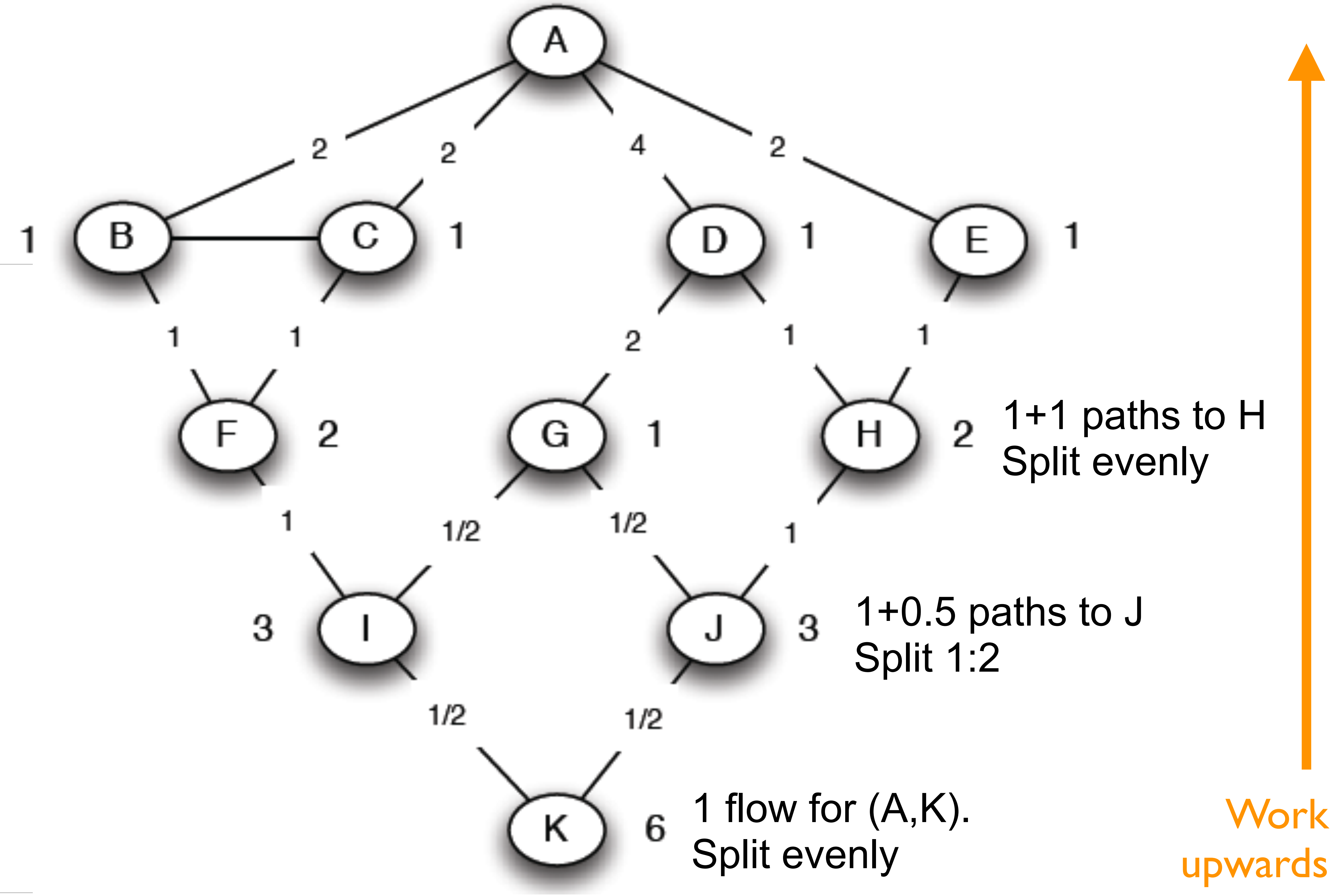
How to Compute Betweenness?

How much flow goes from A to other nodes?

Compute betweenness by working up the tree: If there are multiple paths count them fractionally

The algorithm:

- Add edge **flows**:
  - -- node flow = 1+∑child edges
  - -- split the flow up based on the parent value
- Repeat the BFS procedure for each starting node *U*



Lecture 4

Signed Networks

Networks with positive and negative relationships

Consider an undirected complete graph
Label each edge as either:

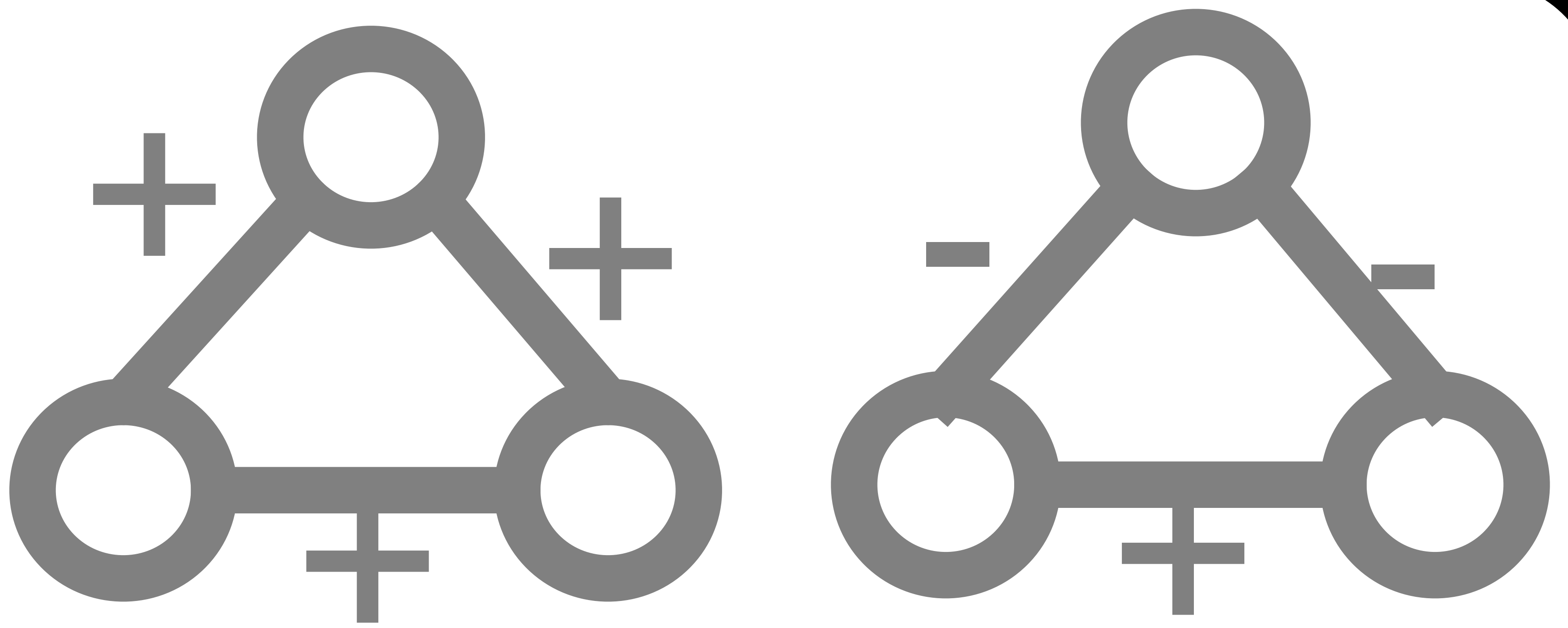
- **Positive**: friendship, trust, positive sentiment, …
- **Negative**: enemy, distrust, negative sentiment, …

Theory of Structural Balance

Start with the intuition [Heider '46]:

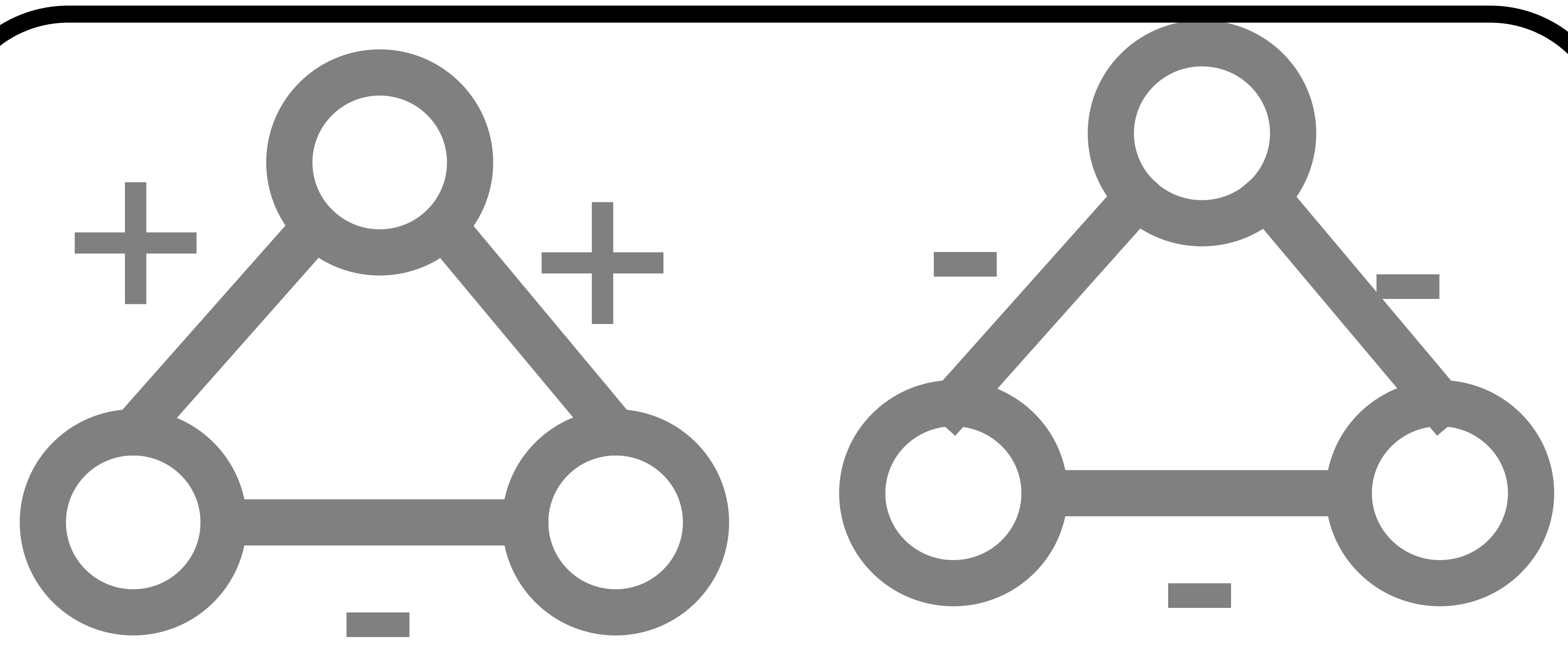
- Friend of my friend is my friend
- Enemy of enemy is my friend
- Enemy of friend is my enemy

Look at connected triples of nodes:



Balanced

Consistent with "friend of a friend" or "enemy of the enemy" intuition



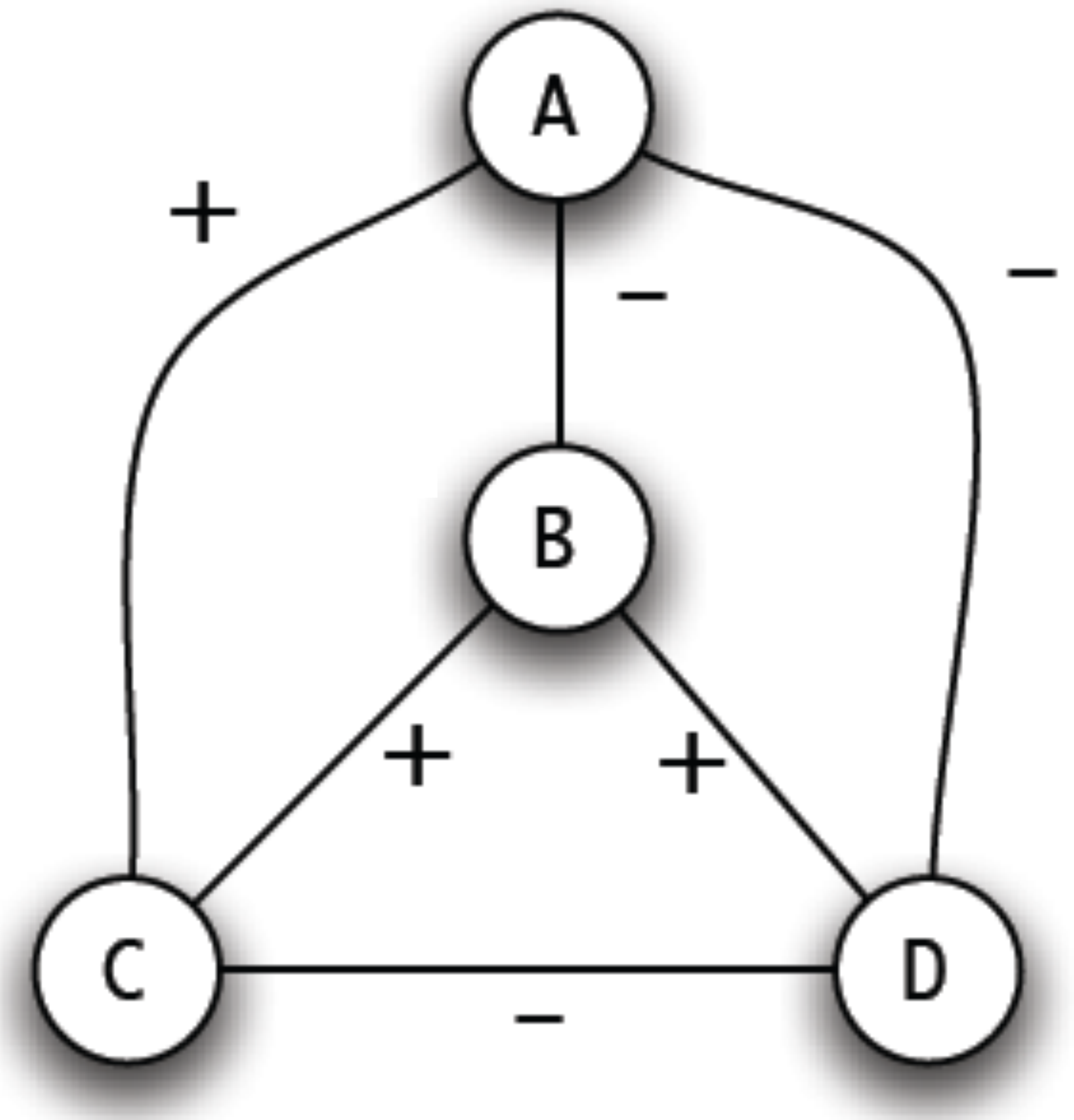
Unbalanced

Inconsistent with the "friend of a friend" or "enemy of the enemy" intuition

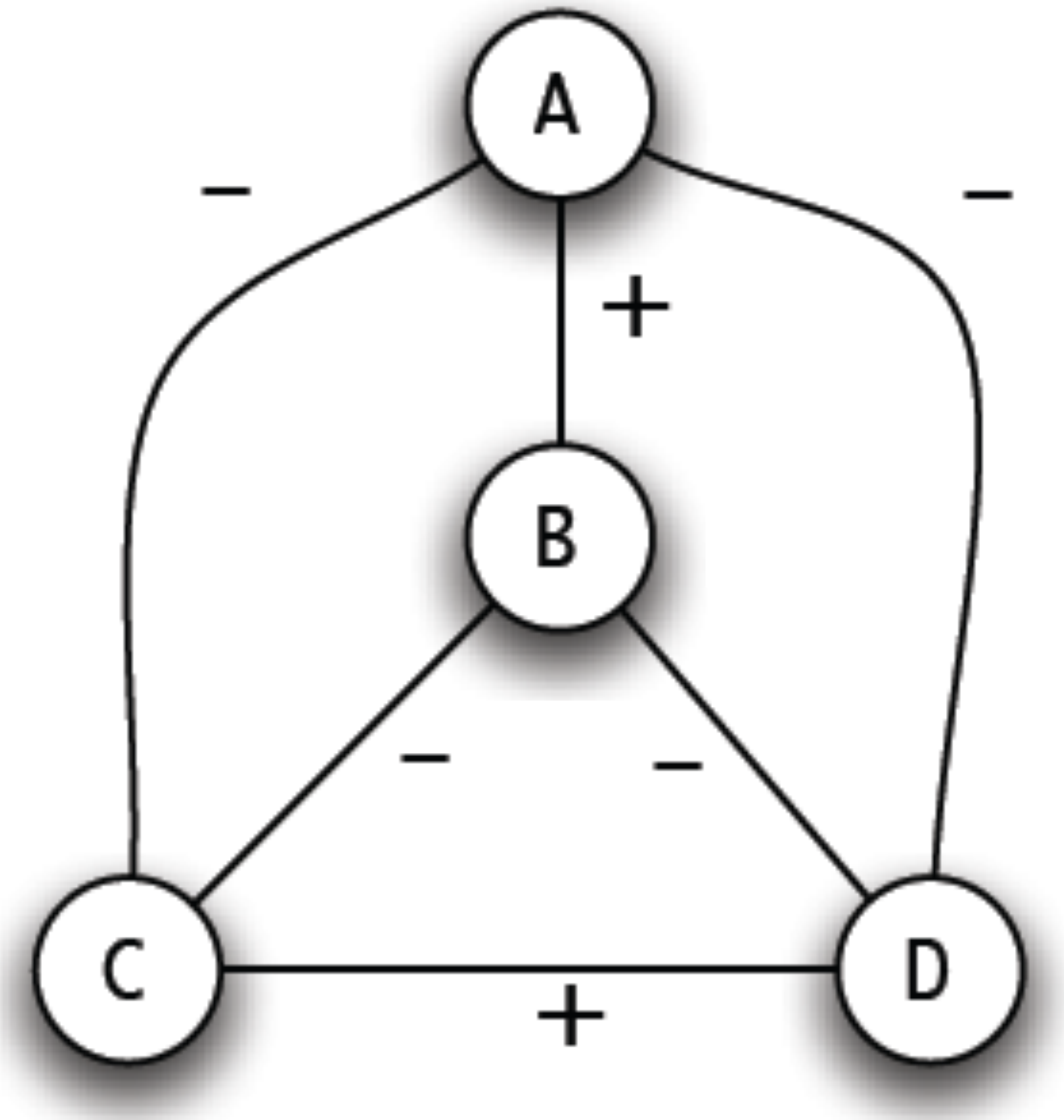
Balanced/Unbalanced Networks

Define: A complete graph is *balanced* if every connected triple of nodes has:

All 3 edges labeled + **or** Exactly 1 edge labeled +



Unbalanced



Balanced

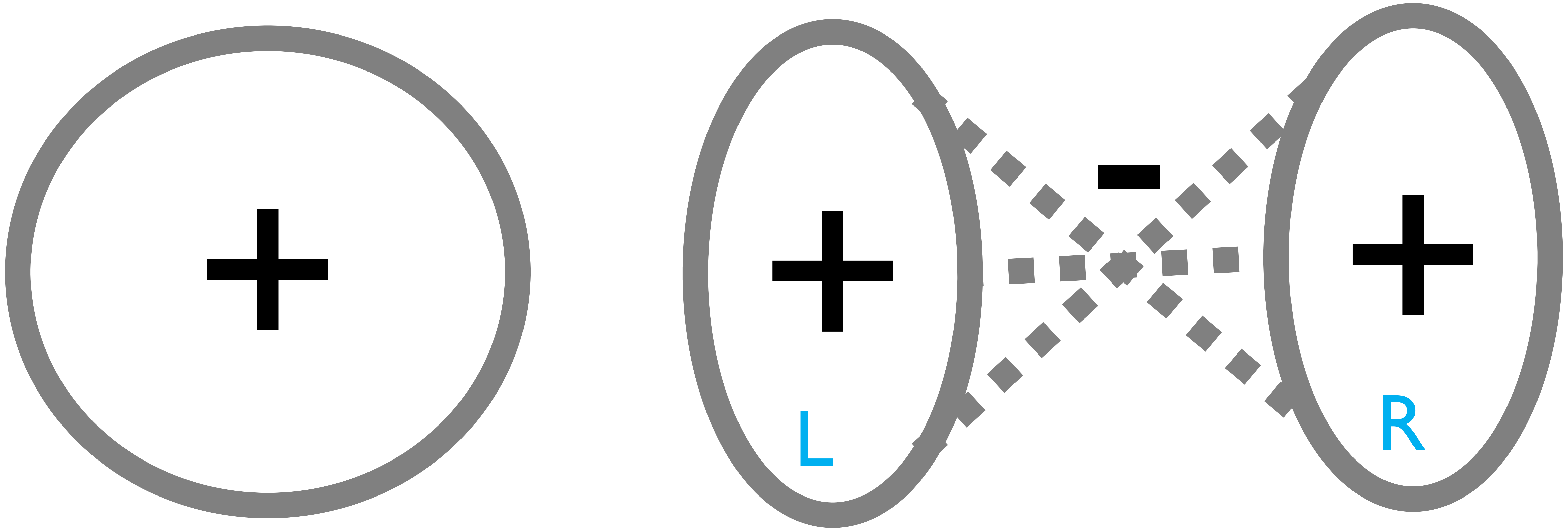
Local Balance → Global Factions

The Balance Theorem: Balance implies global coalitions
[Cartwright-Harary]

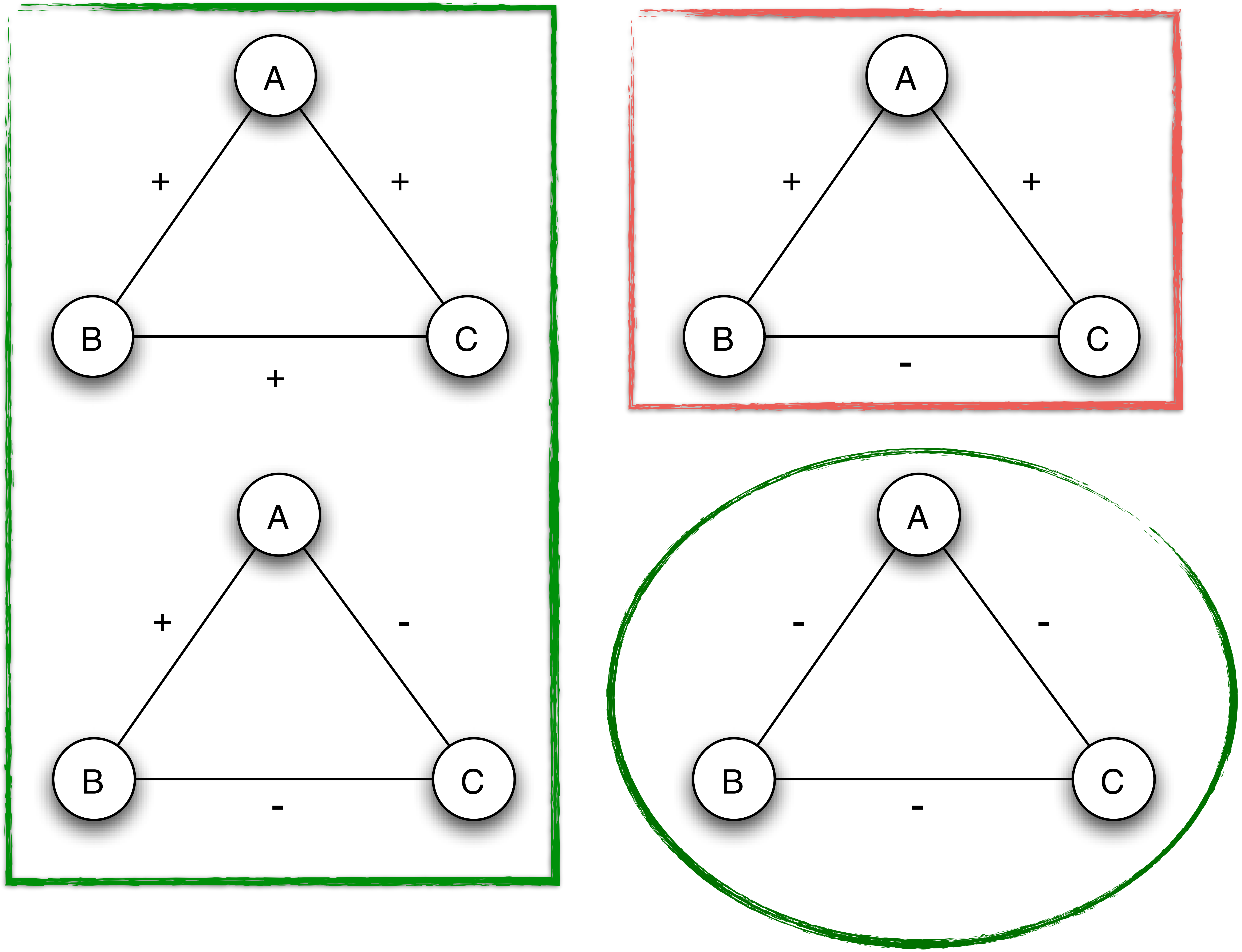
If **all triangles are balanced**, then either:

A) The network contains only positive edges, or

B) The network can be split into two factions: Nodes can be split into 2 sets where negative edges only point between the sets



Structural Balance



What if we allow three mutual enemies?

Weak Structural Balance → Many Global Factions

Define: A complete network is *weakly balanced* if there is no triangle with exactly 2 positive edges and 1 negative edge.

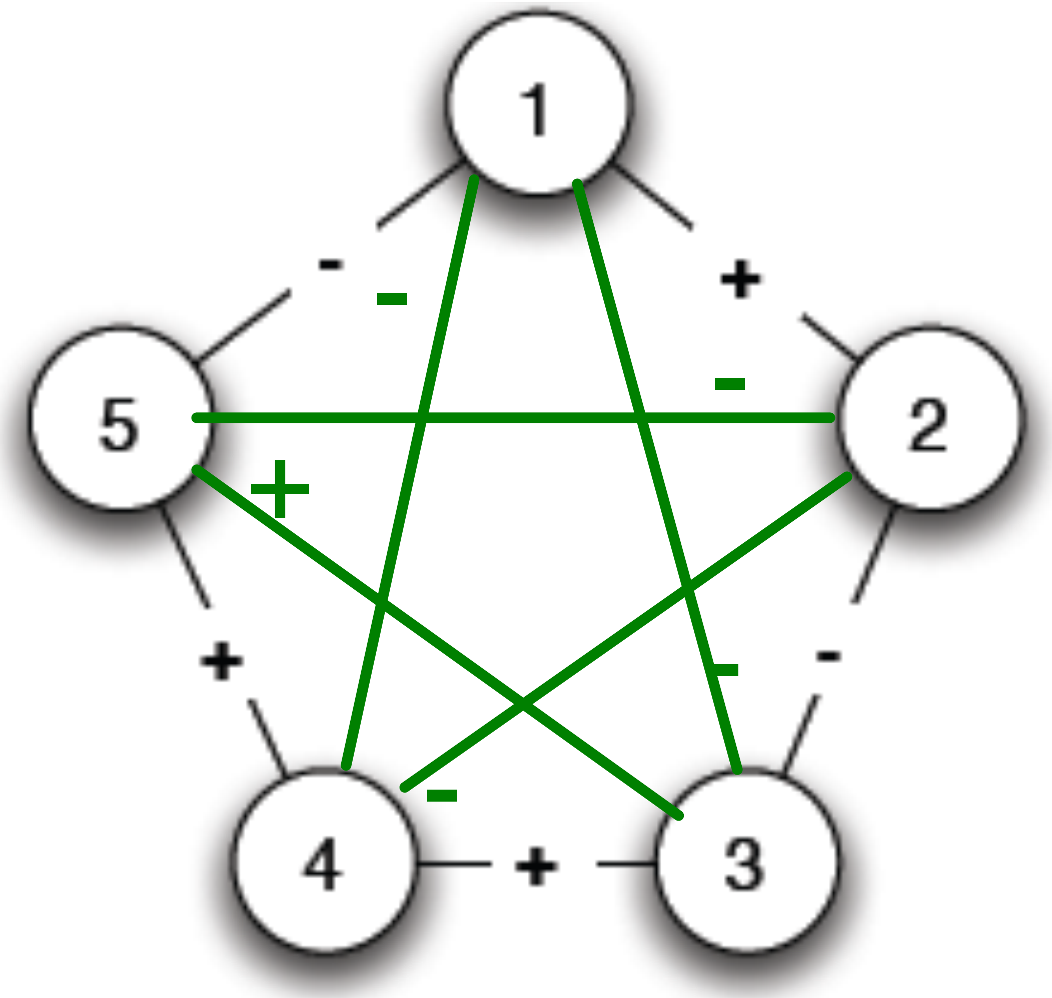
Characterization of Weakly Balanced Networks:
If a labeled complete graph is weakly balanced, then its nodes can be **partitioned**

(divided into groups such that two nodes belonging to the same group are friends, and every two nodes belonging to different groups are enemies)

Global picture: same thing as before, but with many factions, not necessarily two

Balance in General Networks

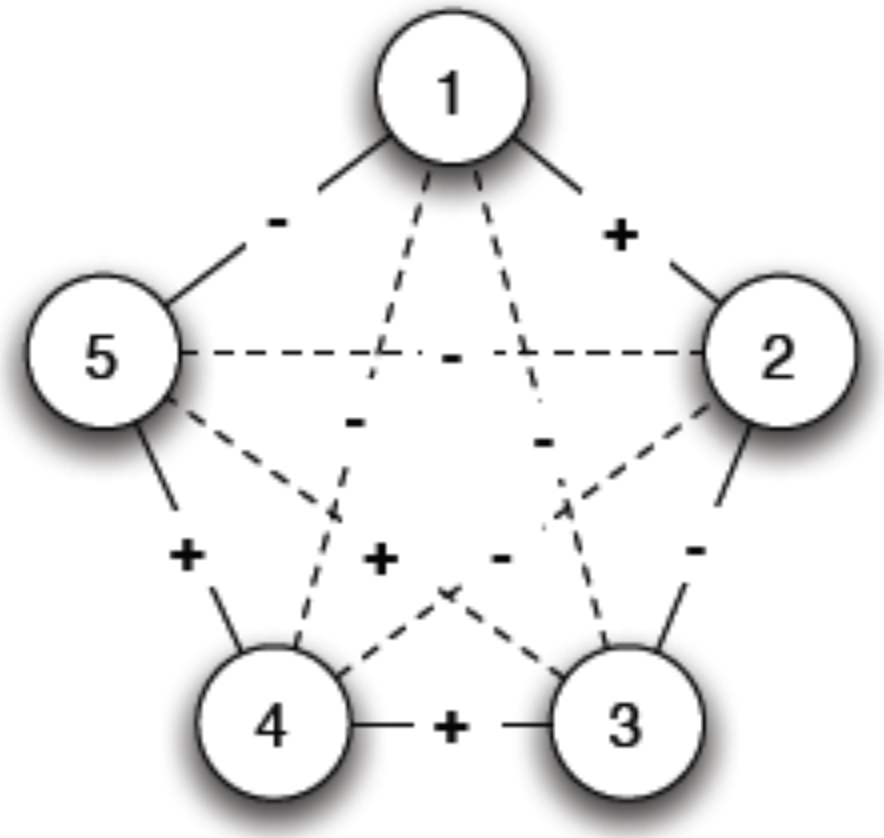
So far we talked about complete graphs



Balanced?

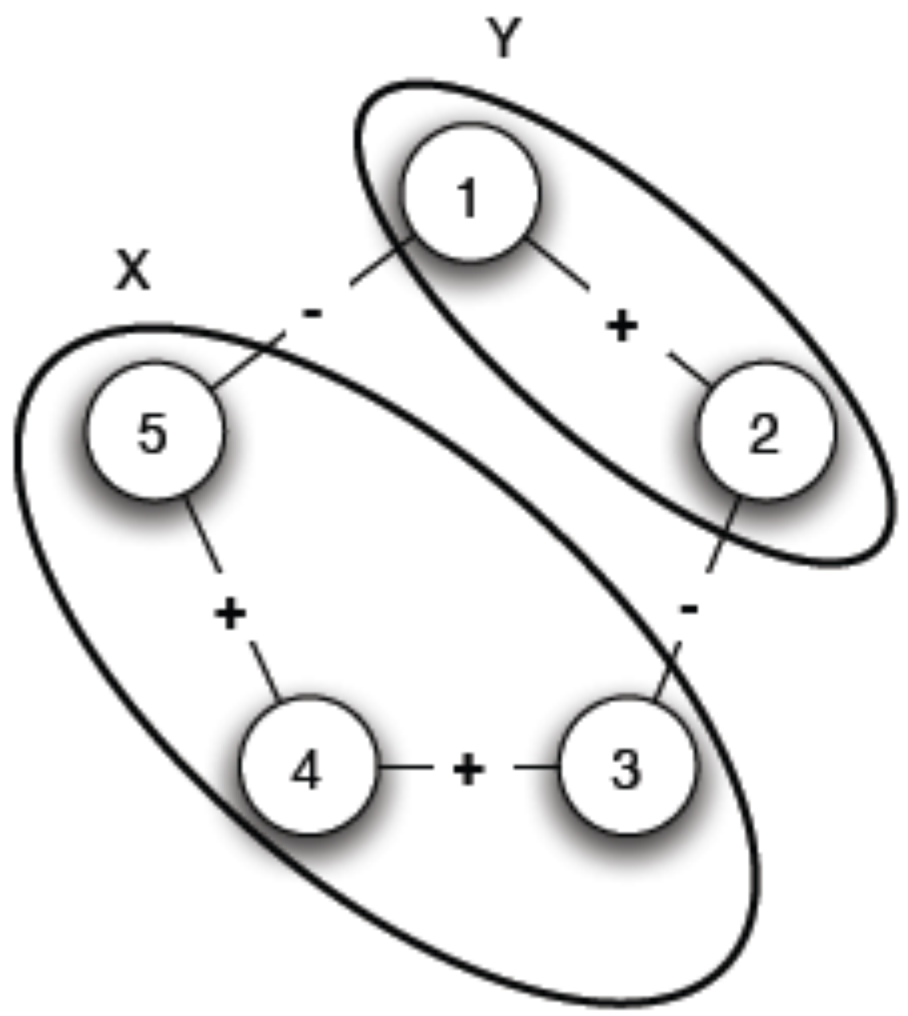
Def 1: Local view

Fill in the missing edges to achieve balance



Def 2: Global view

Divide the graph into two coalitions

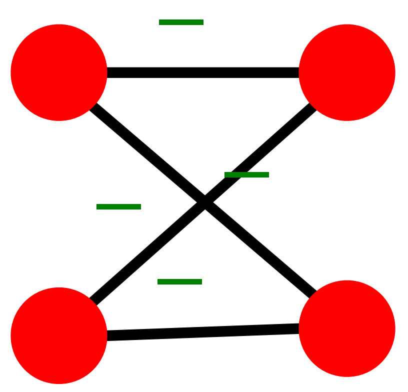


The 2 definitions are equivalent!

Is a Signed Network Balanced?

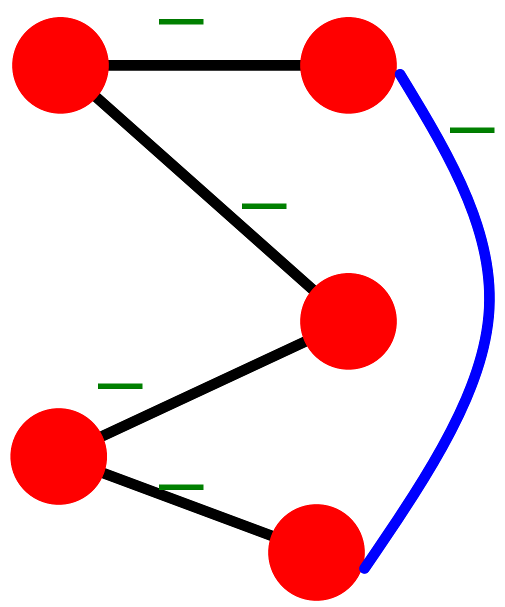
Theorem: Graph is **balanced** if and only if it contains **no cycle with an odd number of negative** edges [Harary 1953, 1956]

Proof by algorithm: We proved this by actually constructing an algorithm that either outputs a division into coalitions or a cycle with odd number of negative edges



Even length cycle

Because these are the only two outcomes, this **proves the claim**



Odd length cycle

Is a Signed Network Balanced?

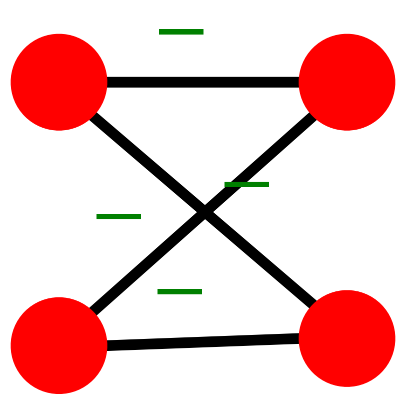
Signed graph algorithm:

Step 1: Find connected components on + edges and for each component create a super-node

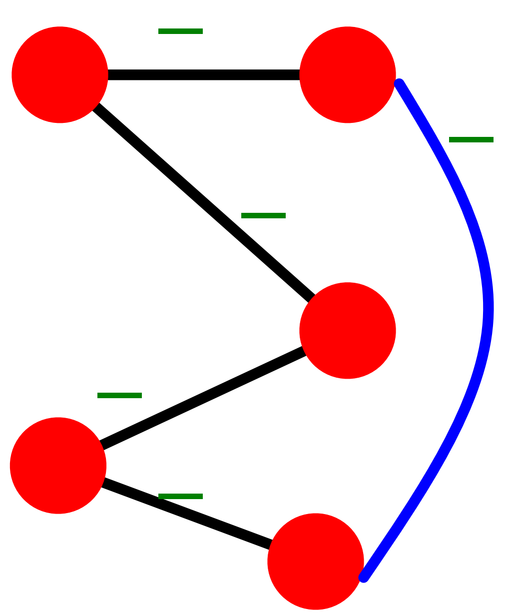
- Since nodes connected by a + edge must be in same coalition
- If any – edge in the super node, done (cycle with 1 negative edge)

Step 2: Connect components A and B if there is a negative edge between the members

- Note there are only negative edges pointing out of a super-node (otherwise should've connected the two super-nodes that have a positive edge)



Even length cycle



Odd length cycle

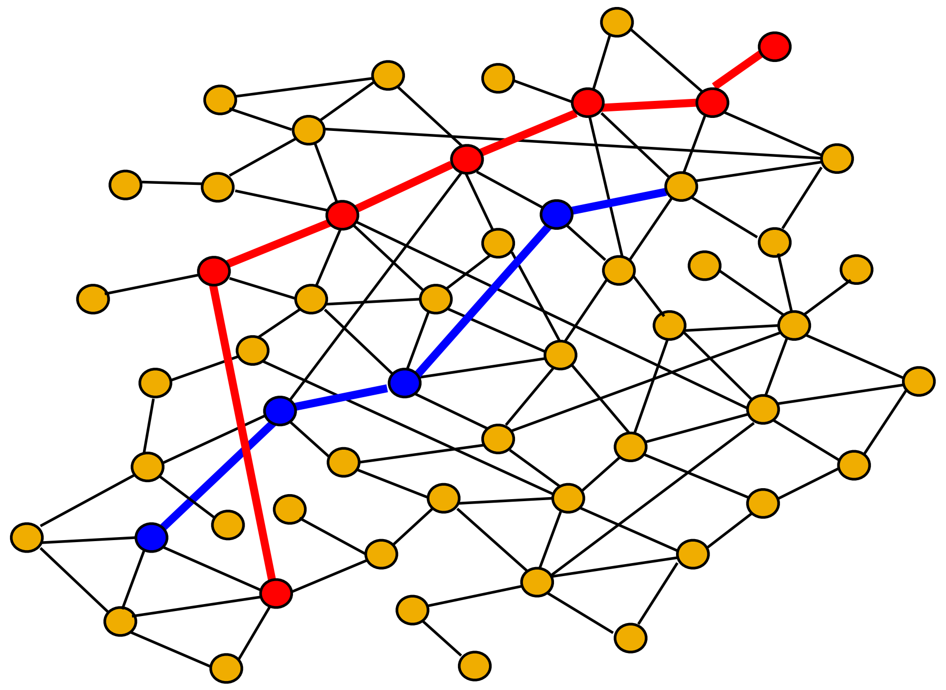
Lecture 5

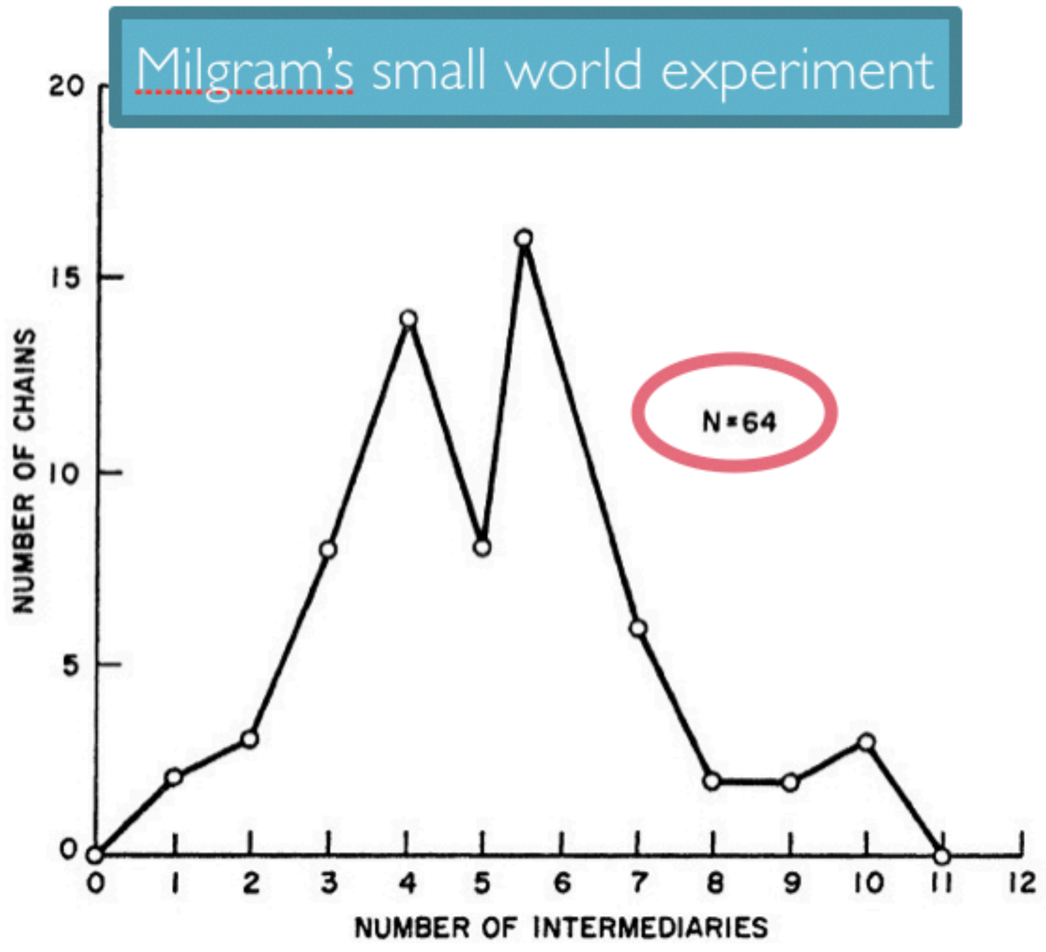
How long is the typical shortest path?

Milgram devised a clever experiment

–Picked ~300 people in Omaha, Nebraska and Wichita, Kansas
–Asked each person to try get a letter to a particular person in Boston (a stockbroker), but they could only send it to someone they know on a first-name basis
–The friends then send to their friends, etc.

64 chains completed, 6.2 steps on average





6 Degrees: Should We Be Surprised?

Assume each human is connected to 100 other people

Then:

Step 1: reach 100 people

Step 2: reach 100*100 = 10,000 people

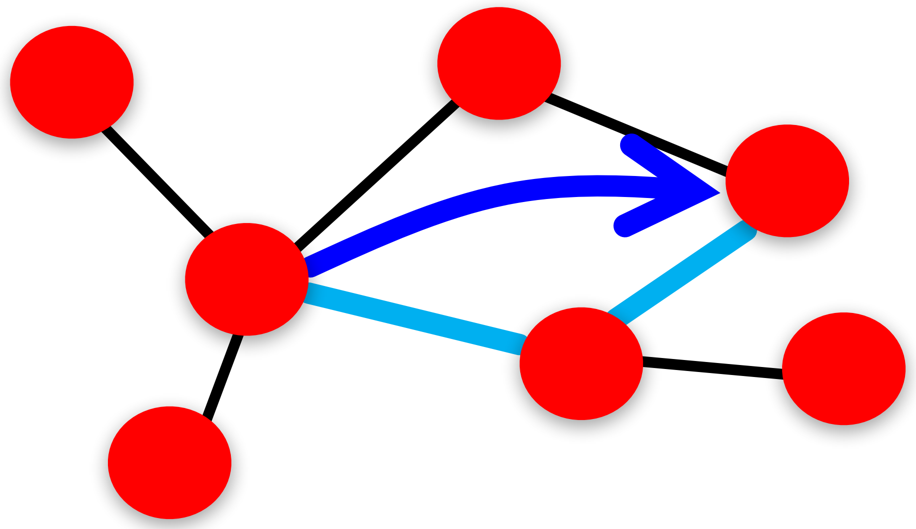
Step 3: reach 100*100*100 = 1,000,000 people

Step 4: reach 100*100*100*100 = 100M people

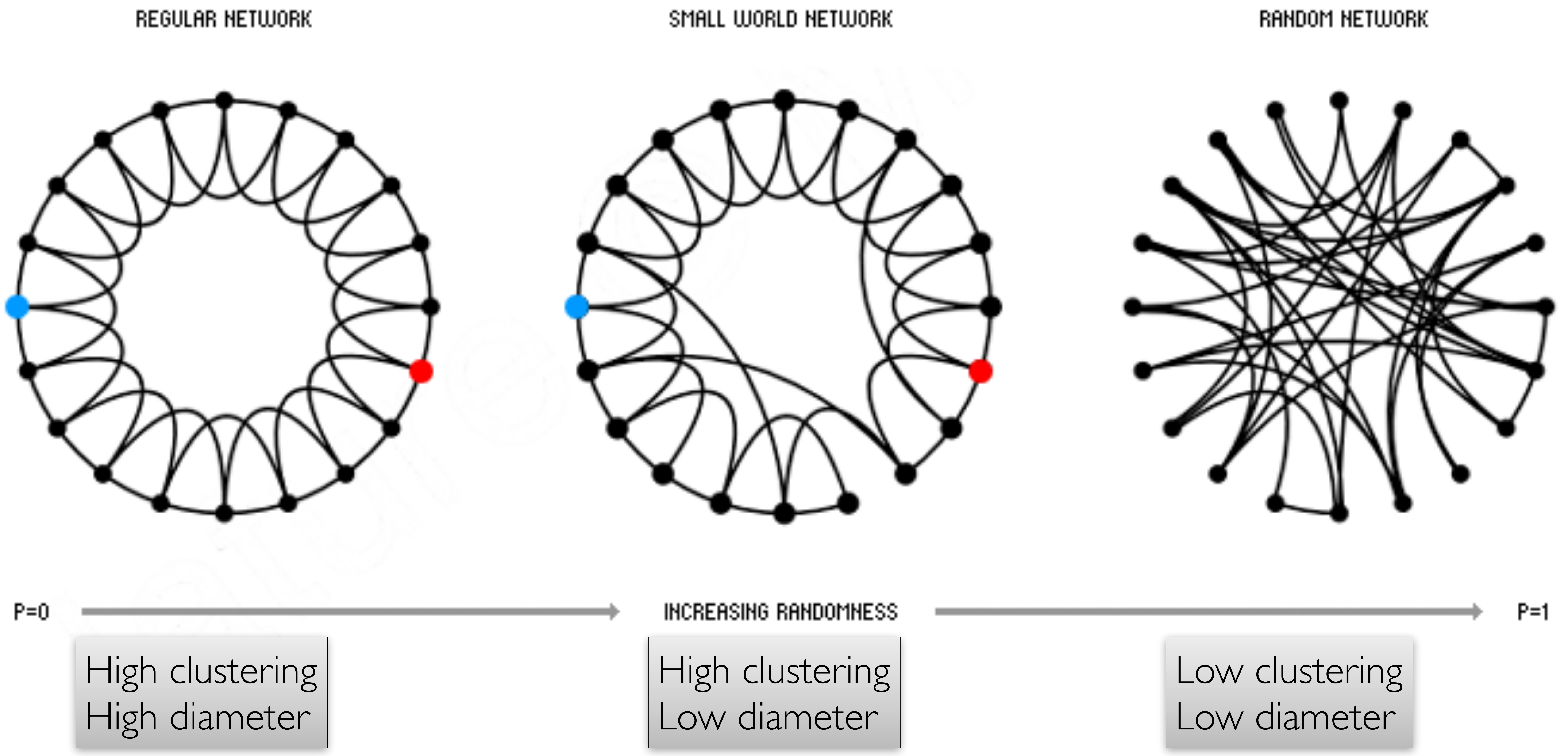
In 5 steps we can reach 10 billion people

What's wrong here?

Triadic closure: 92% of new FB friendships are to a friend-of-a-friend [Backstom-Leskovec '11]



The Small-World Model

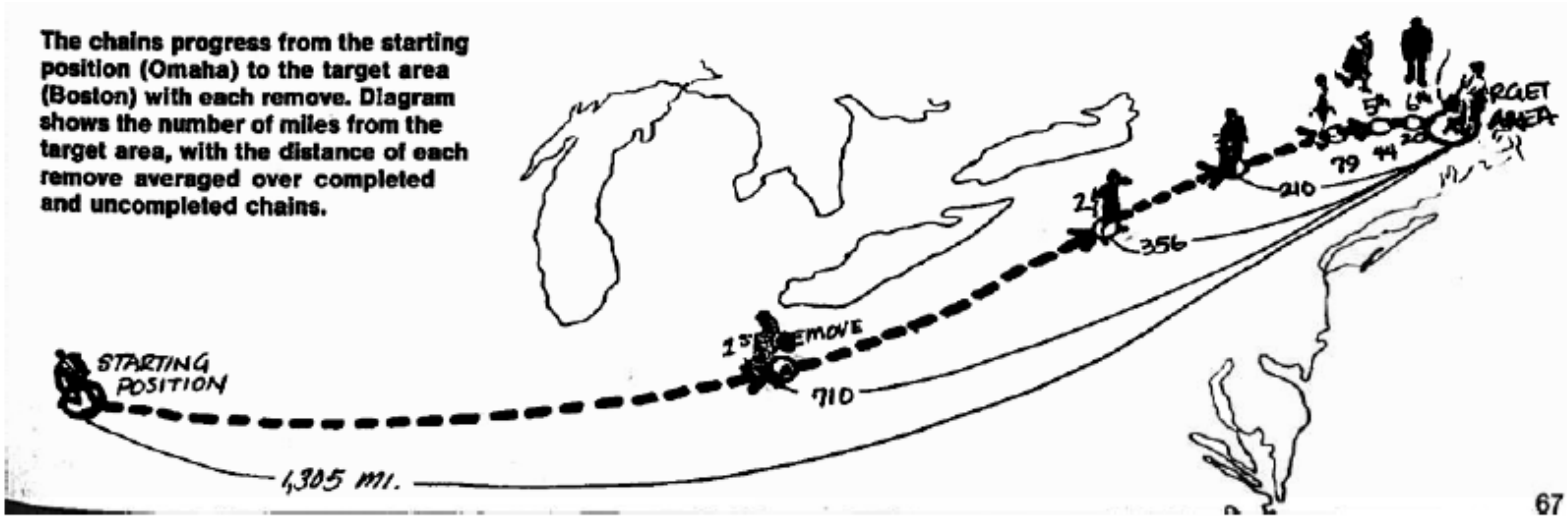


Rewiring allows us to “interpolate” between
a regular lattice and a random graph

How to Navigate a Network?

"The geographic movement of the [message] from Nebraska to Massachusetts is striking. There is a progressive closing in on the target area as each new person is added to the chain"

S.Milgram 'The small world problem', Psychology Today, 1967

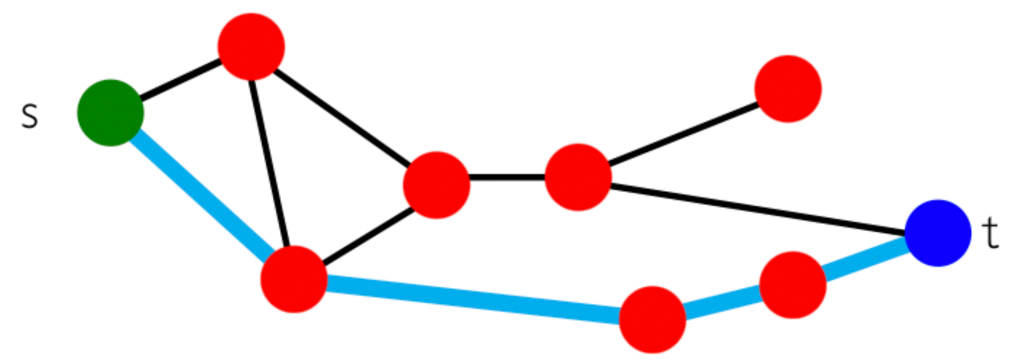


Decentralized Search

The setting:
– Nodes live in a regular lattice, just as in Watts-Strogatz
– Each node has an "address"/location in the grid
– Node ***s*** is trying to route a message to ***t***
– ***s*** only knows locations of its friends and location of the target ***t***
– ***s*** does not know random links of anyone else but itself

Geographic Navigation: nodes will act greedily with respect to geography: always pass the message to their neighbour who is geographically closest to ***t*** (what else can they do?)

Search time T: Number of steps it takes to reach ***t***



What is success?

We know these graphs have diameter O(log n), so paths are logarithmic in shortest-path length

We will say a graph is **searchable** if the decentralised search time T is polynomial in the path lengths

But it's **not searchable** if T is exponential in the path lengths

Searchable	Not searchable
Search time T:	Search time T:
$O((\log n)^{\beta})$	$O(n^{\alpha})$

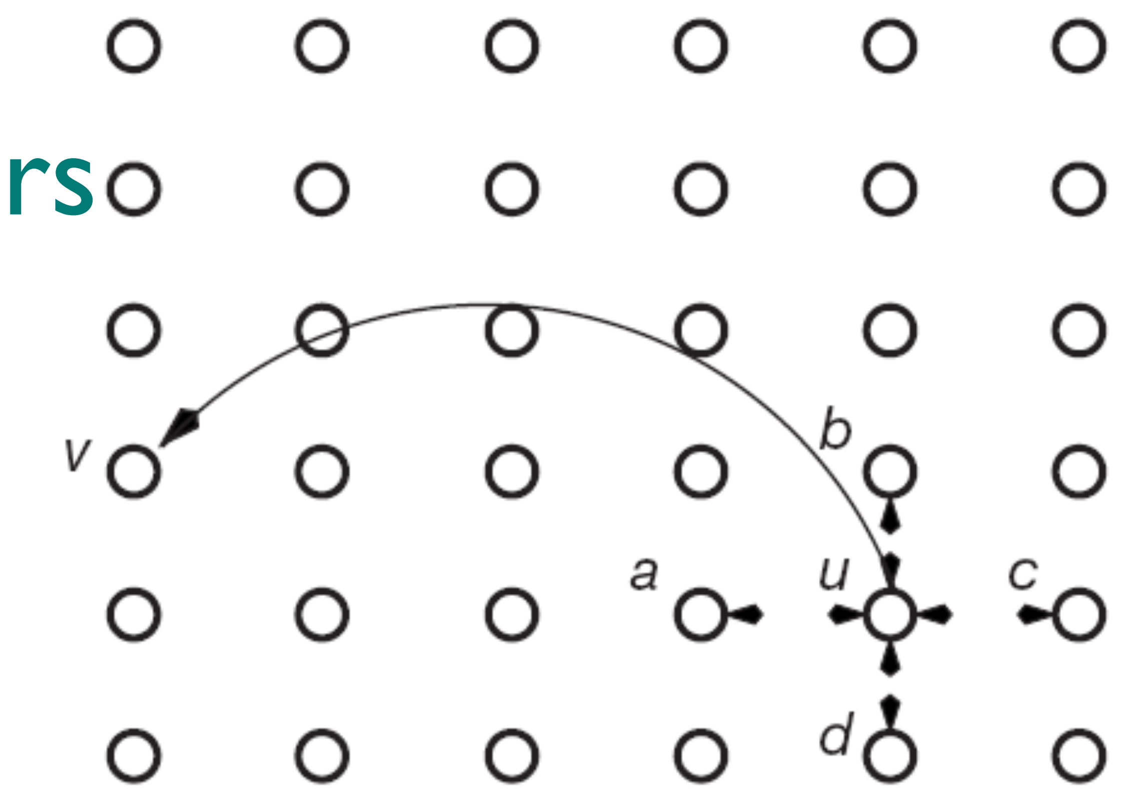
Kleinberg's Model

Kleinberg's Model [Kleinberg, Nature '01]

Nodes still live in a grid, and know their neighbours

Each node has one random "long-range" link

Key difference: the link isn't uniformly at random anymore, it follows geography



Prob. of long link to node *v*:

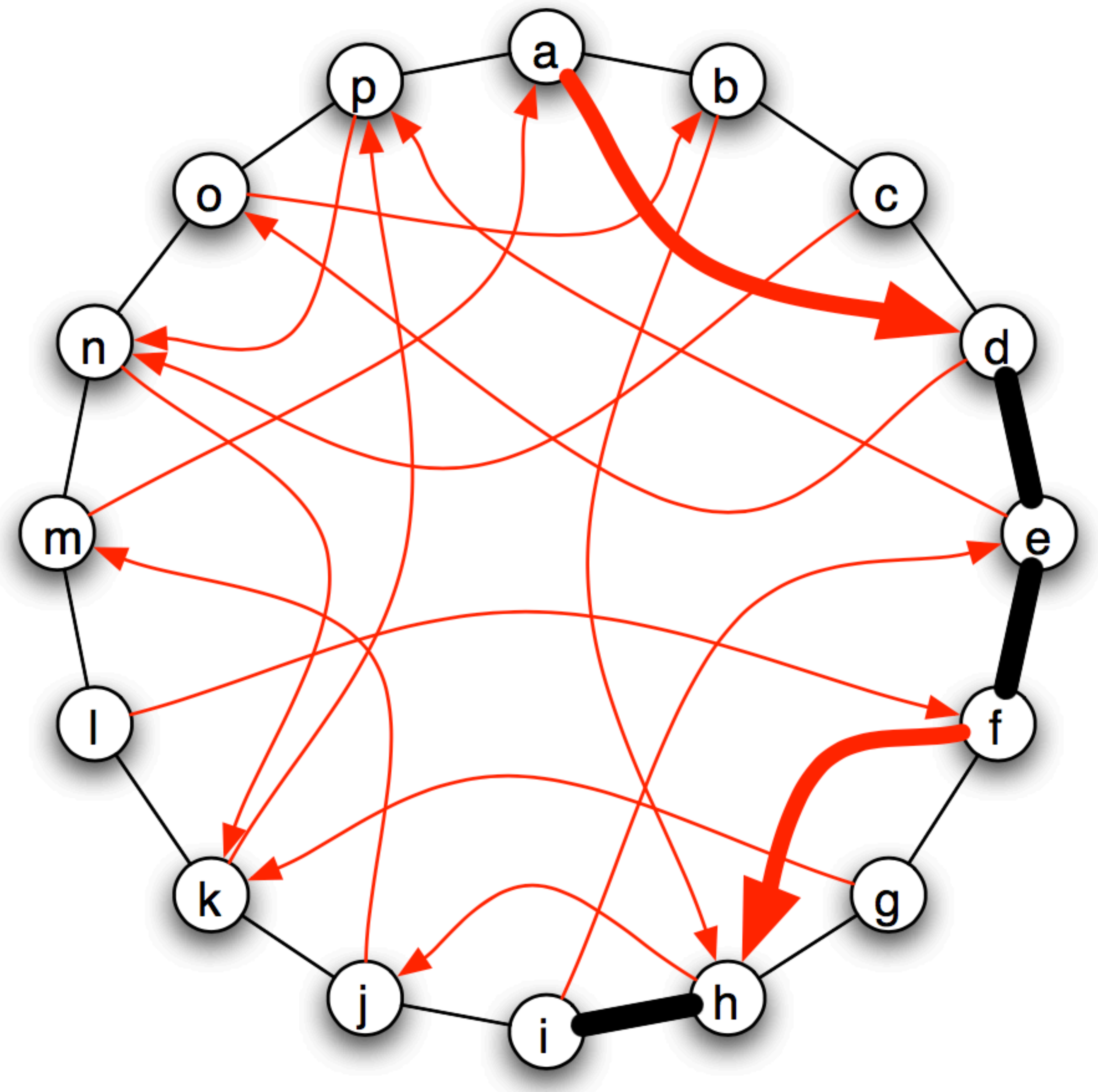
$$P(u \rightarrow v) \sim d(u,v)^{-\alpha}$$

$d(u,v)$ … grid distance between *u* and *v* (address distance, not shortest path)

α … tunable parameter ≥ 0

Kleinberg's Model in 1-Dimension

Myopic search in general doesn't find the shortest path!



Kleinberg's Model in 1-Dimension

We analyze 1-dimensional case:

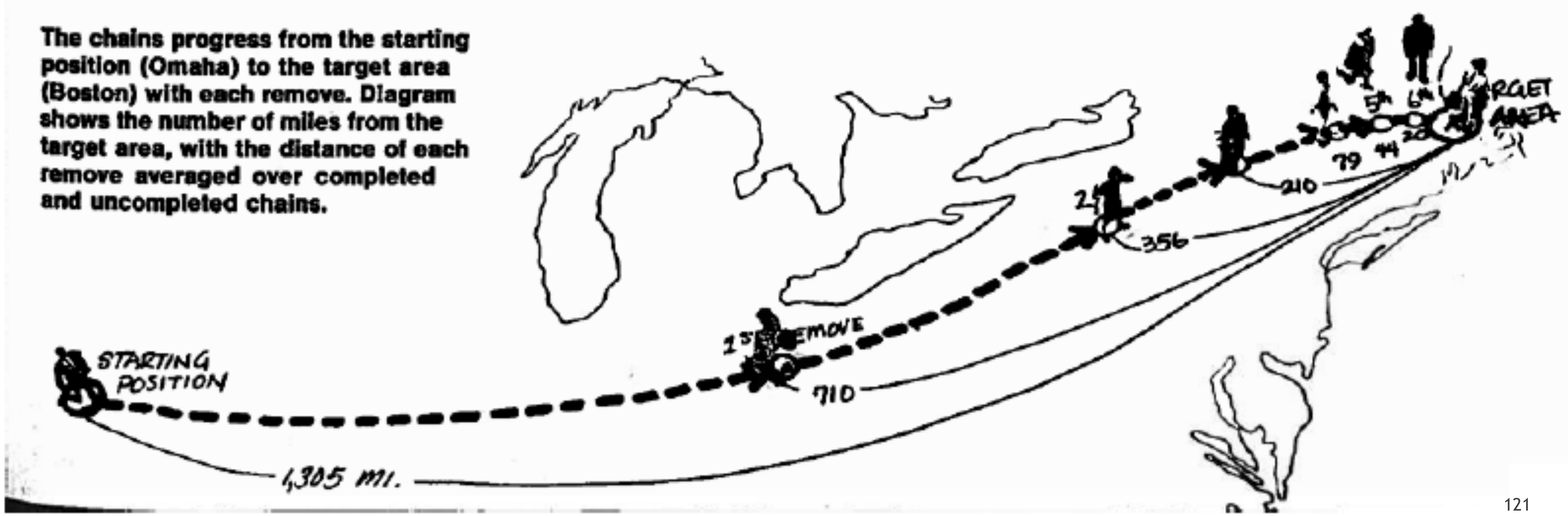
Claim: For $\alpha = 1$ we can get from *s* to *t* in O(log(n)²) steps in expectation

$$P(u \to v) \sim d(u,v)^{-\alpha} = 1/d(u,v)$$

Proof strategy:

Argue it takes O(log n) to halve the distance

O(log n) halving steps to get to target



Lecture 6

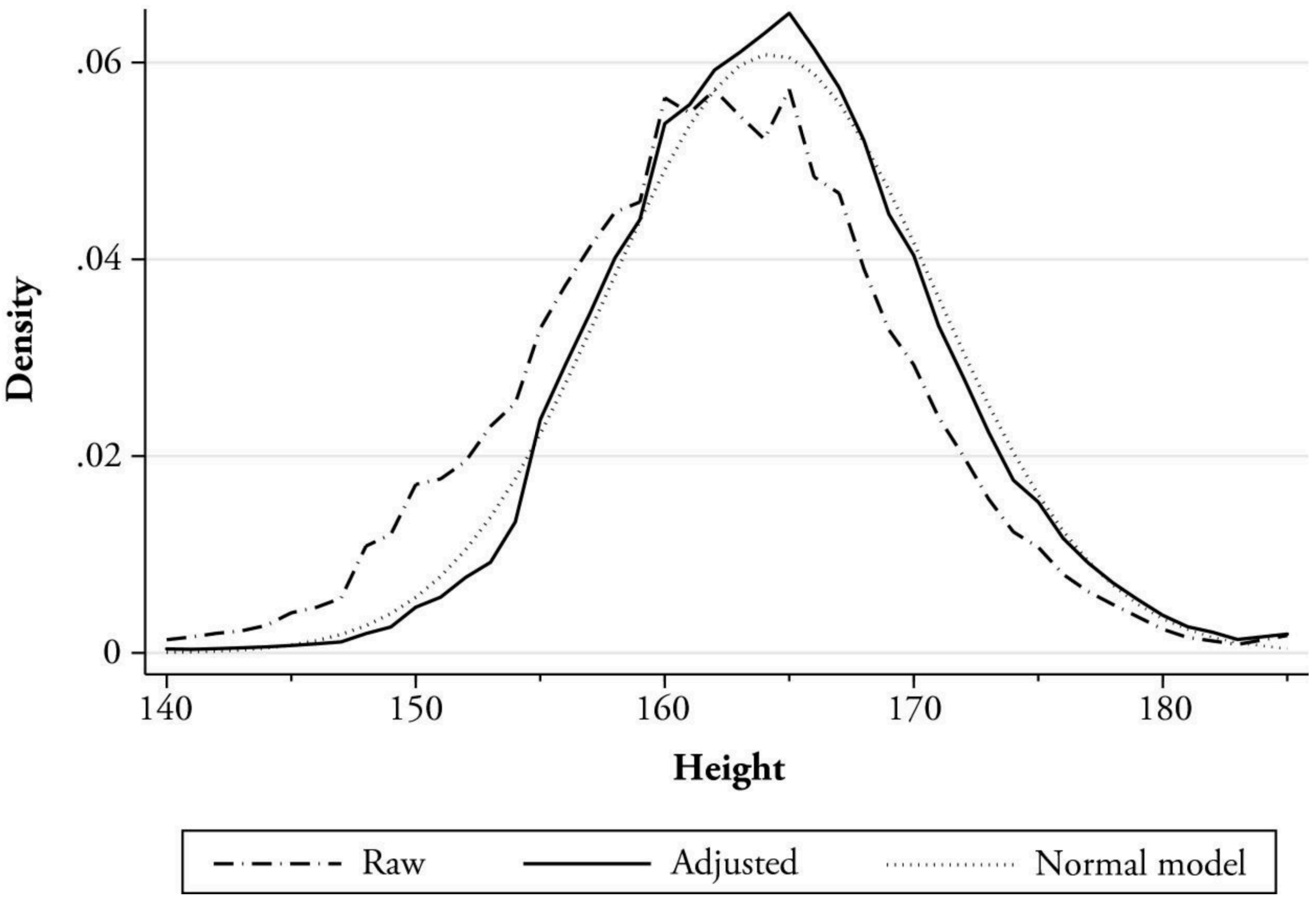
How is popularity distributed?

A deeper look at one of our central questions: how connected are people? How many people do people tend to know?

Most know some, and some know a ton

How is popularity *distributed* in the population?

A guess



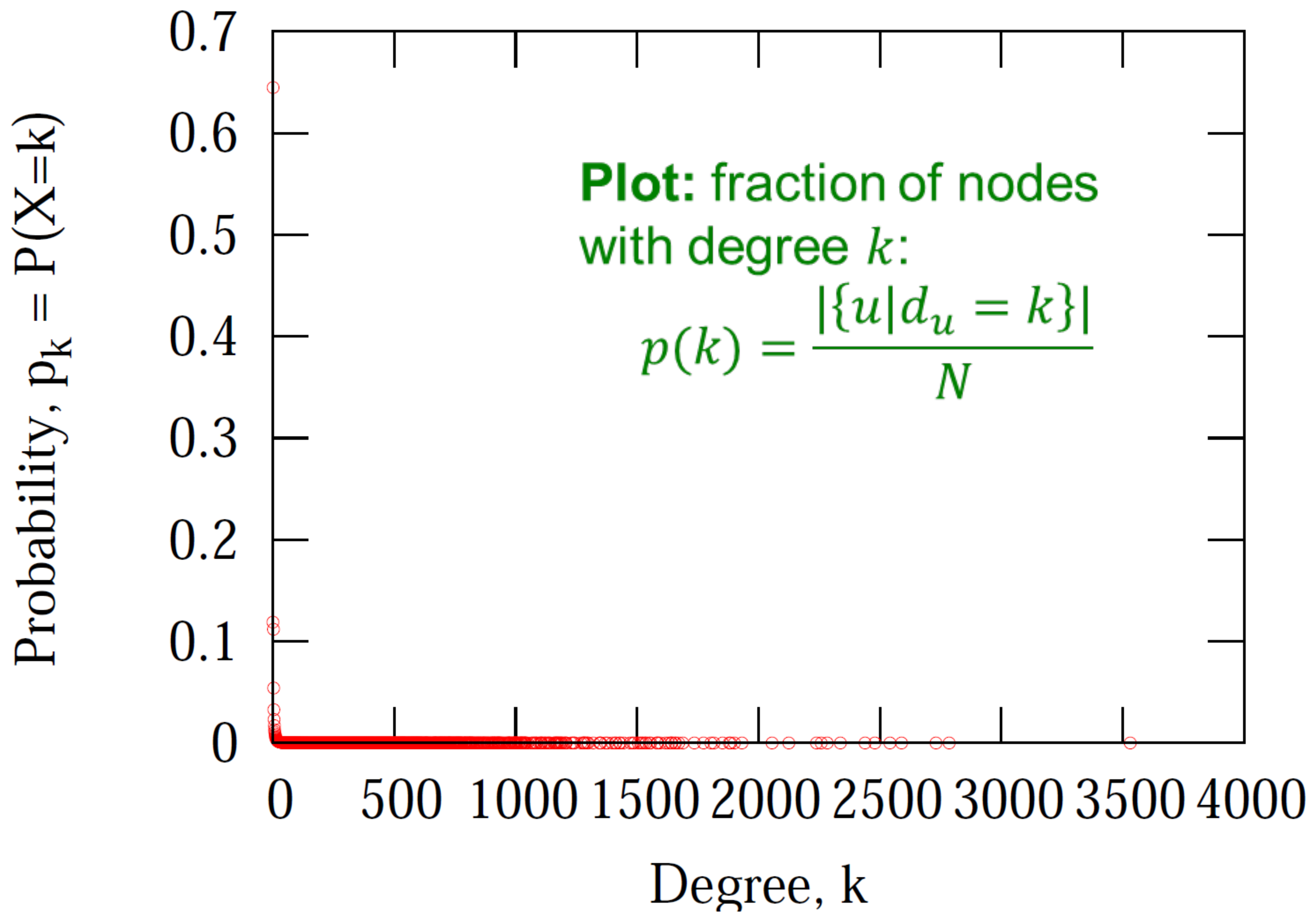
From "Height and the Normal Distribution: Evidence from Italian Military Data"

Heights of males in the Italian army

Most values are clustered around a typical value

Node Degrees in Networks

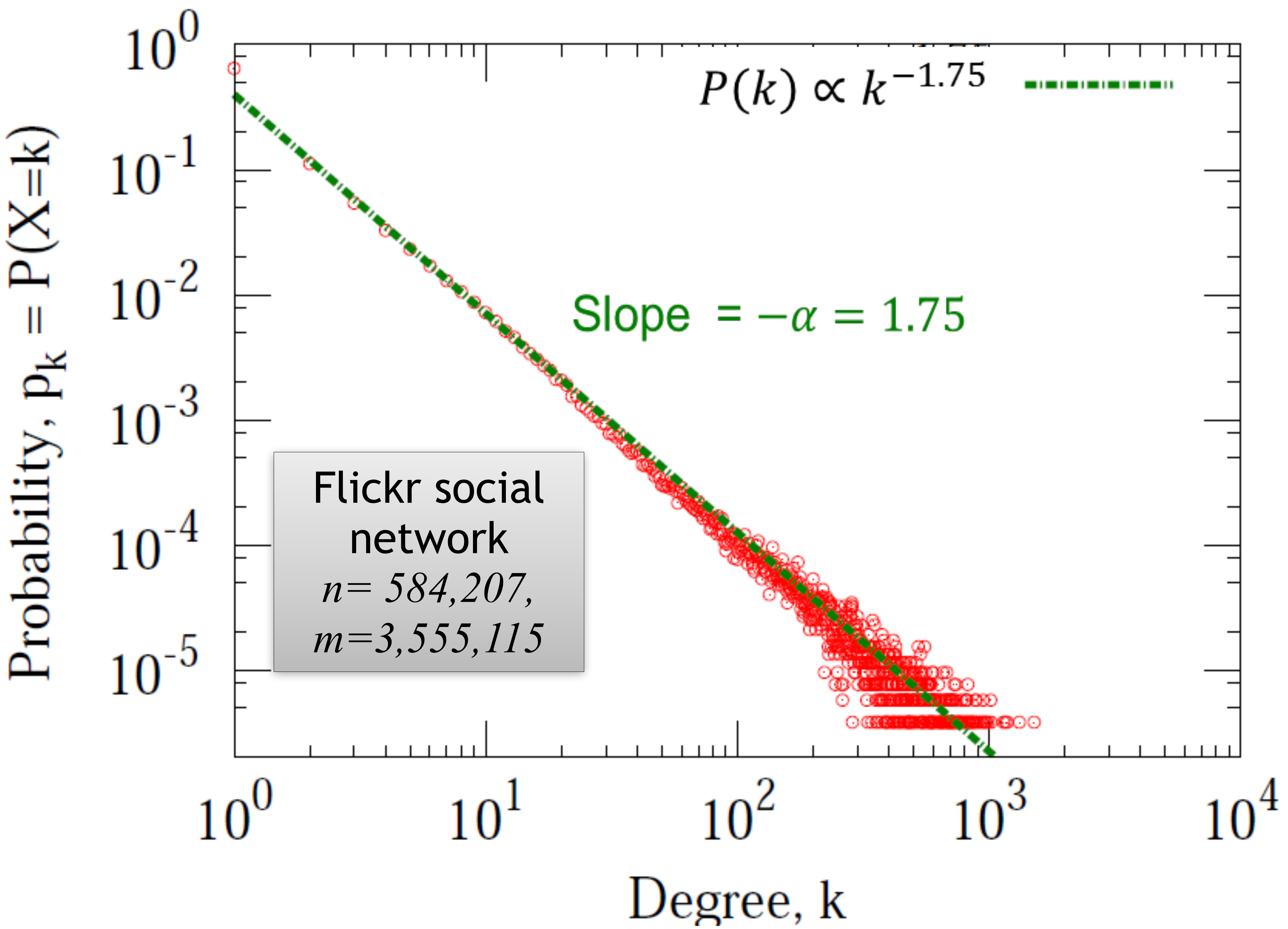
Take a network, plot a histogram of P(k) vs. k



Flickr social network
n= 584,207,
m=3,555,115

Node Degrees in Networks

Plot the same data on *log-log* scale:



The Power Law Distribution

The main heavy-tailed distribution we will consider is the **power law**:

$$p(x) \propto x^{-\alpha}$$

For example, Newton's law of universal gravitation follows an "inverse-square law", e.g. a power law:

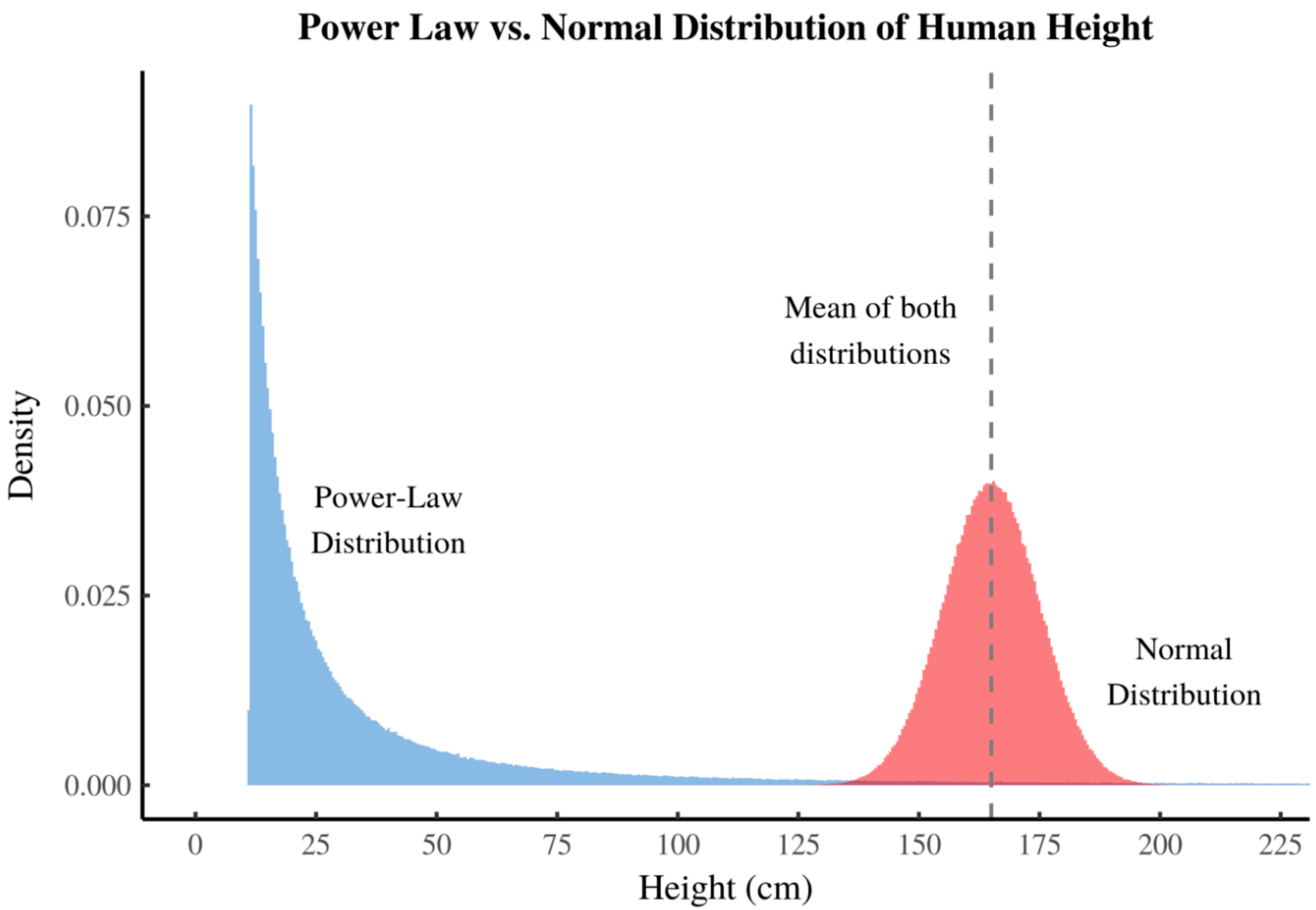
$$F(r) = G\frac{m_1 m_2}{r^2}$$

Where the distance r is the quantity that is changing

To make it an actual distribution, include a normalizing constant c

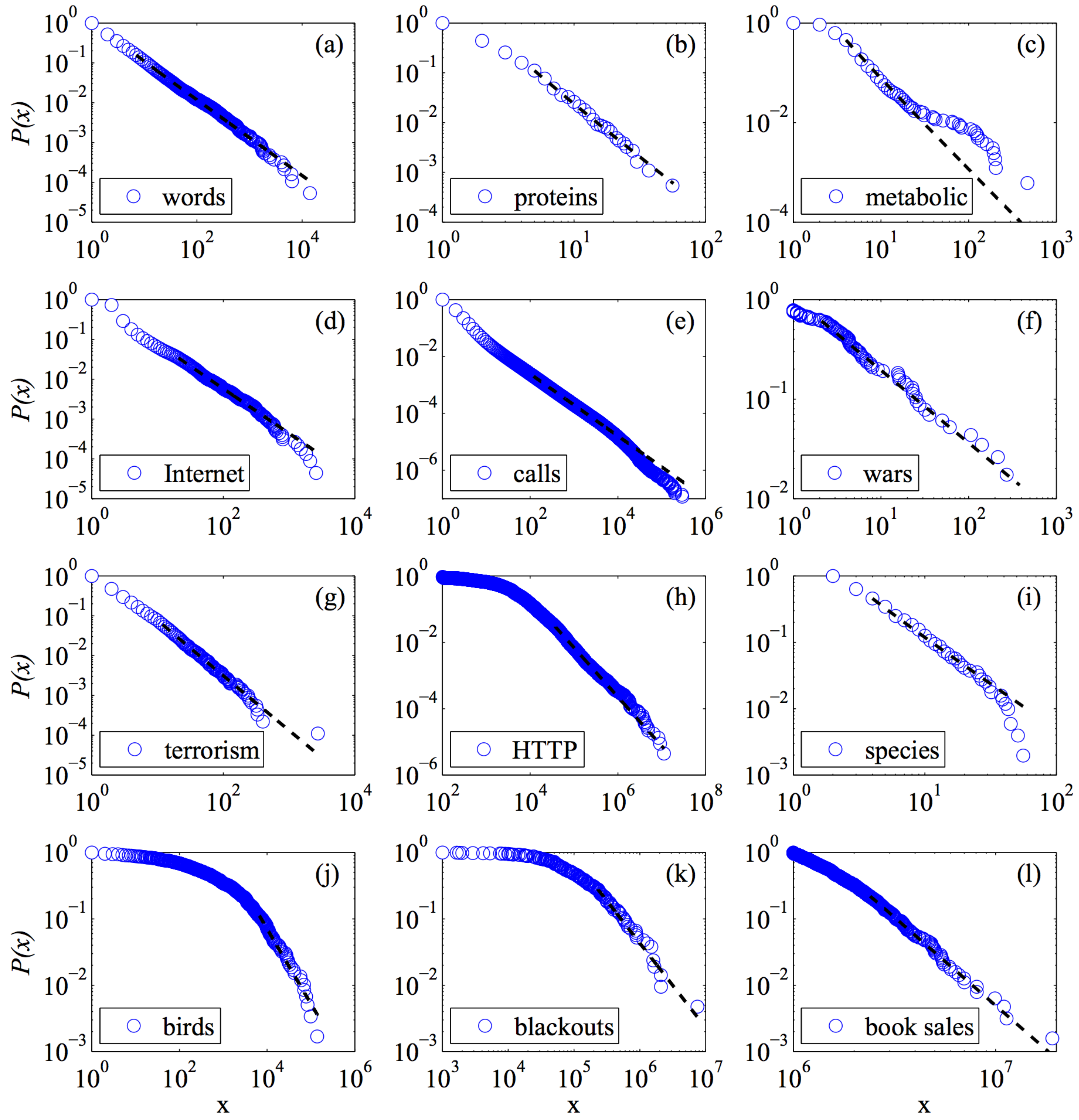
$$p(x) = cx^{-\alpha}$$

Height as a Power Law

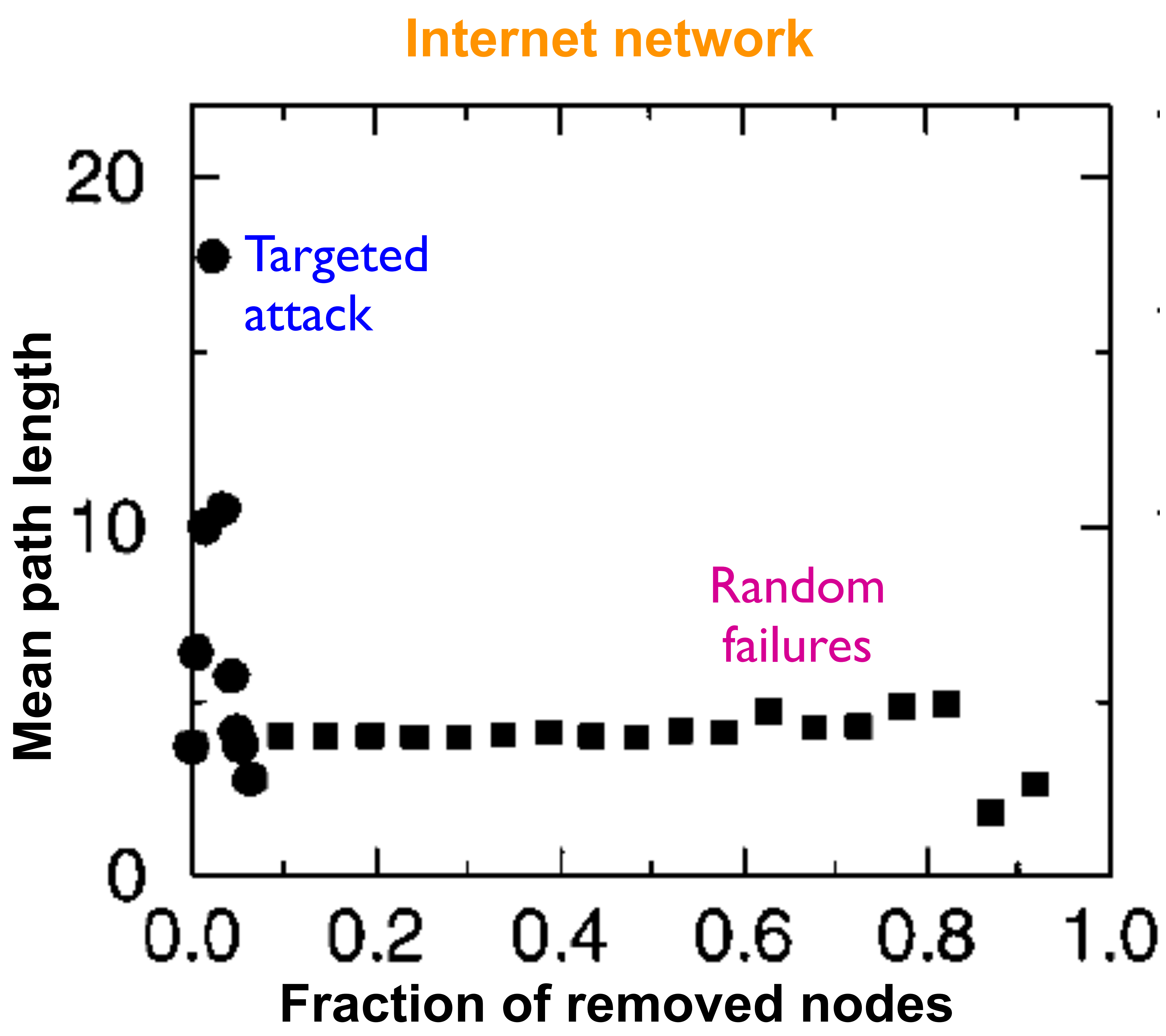


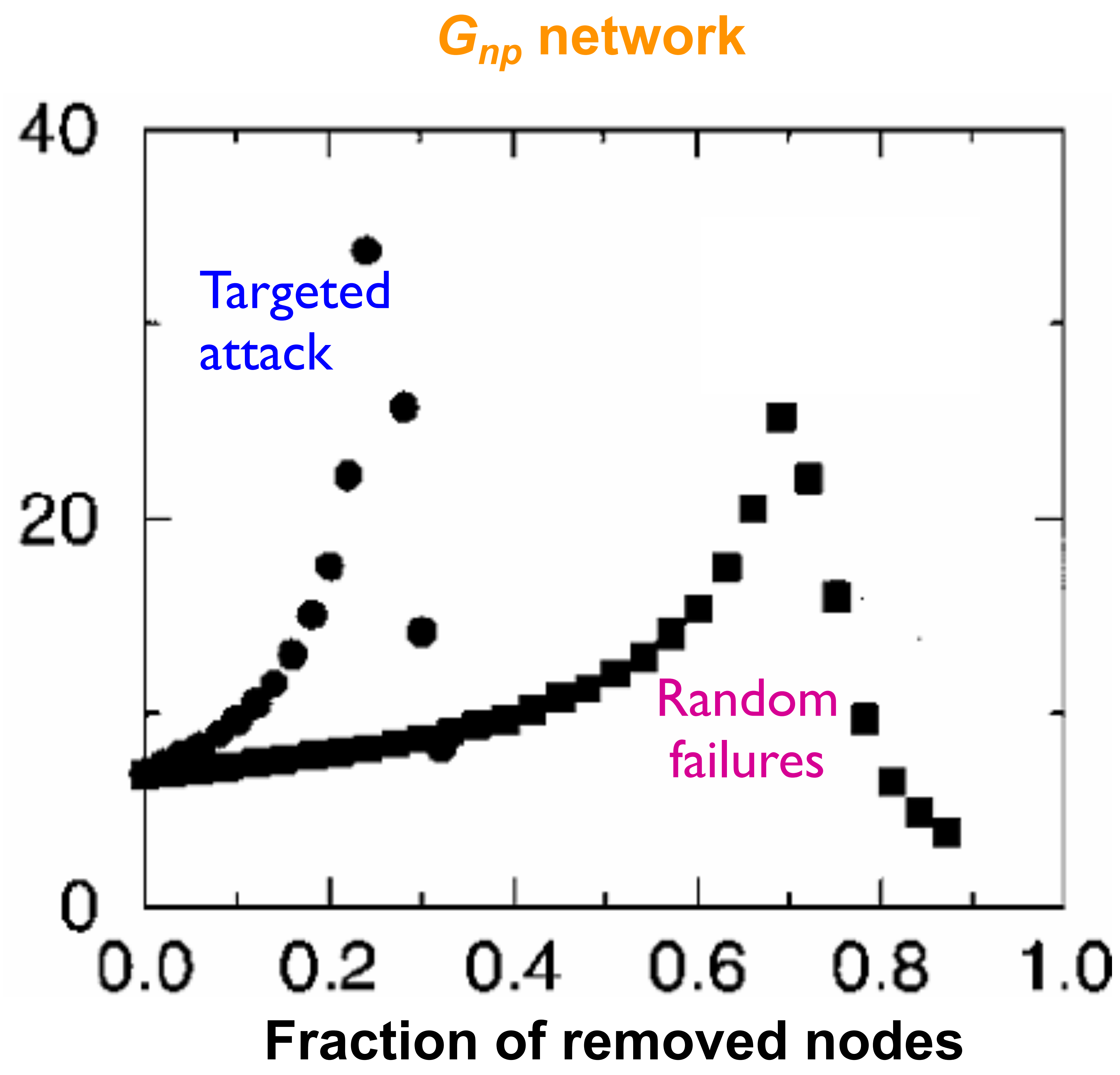
Why is the mean of the power law so far out?

Power laws are *everywhere*



Network Resilience





Real networks are resilient to random failures

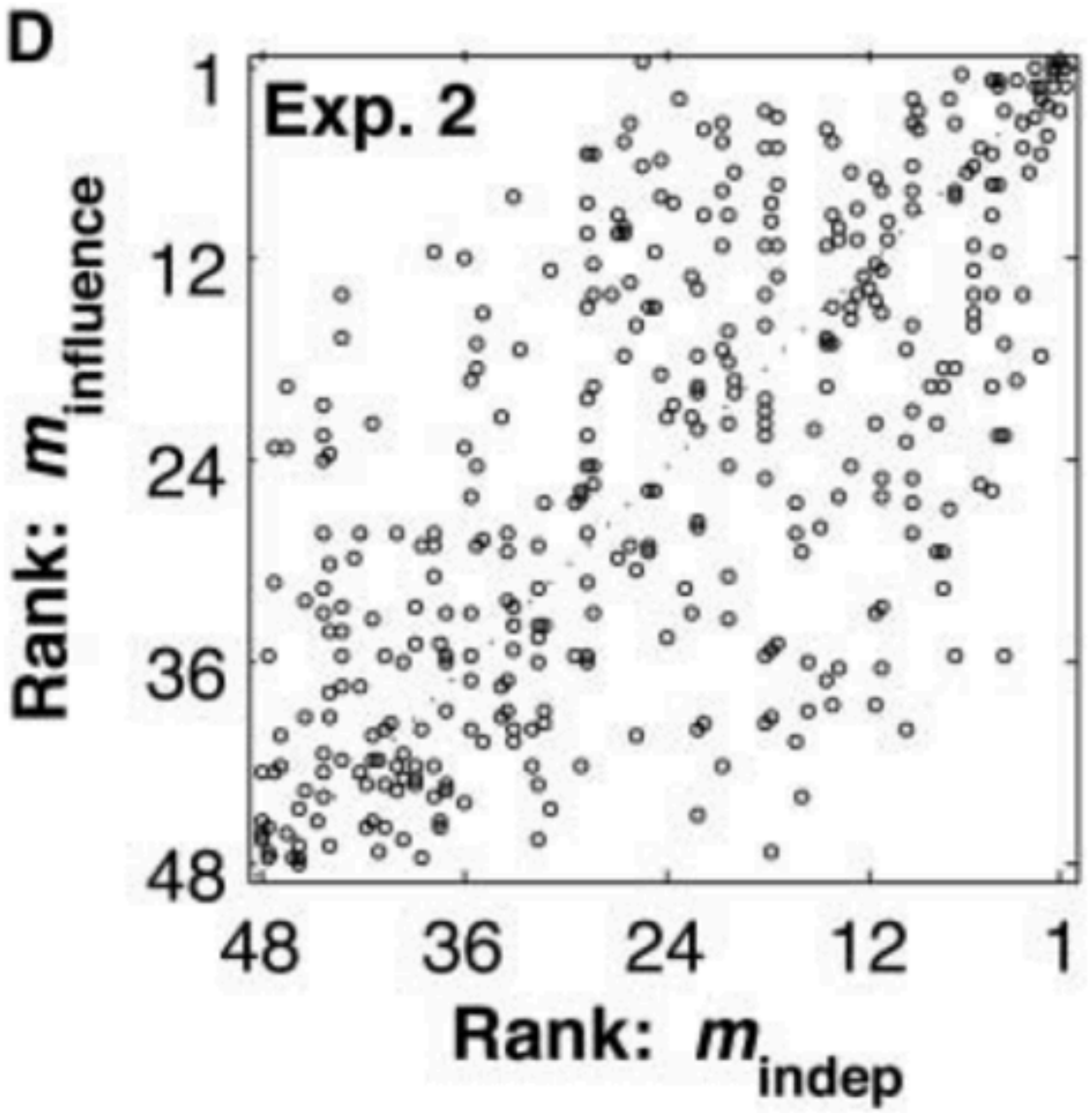
G_{np} has better resilience to targeted attacks

Need to remove all pages of degree **>5** to disconnect the Web

But this is a very small fraction of all web pages

MusicLab:

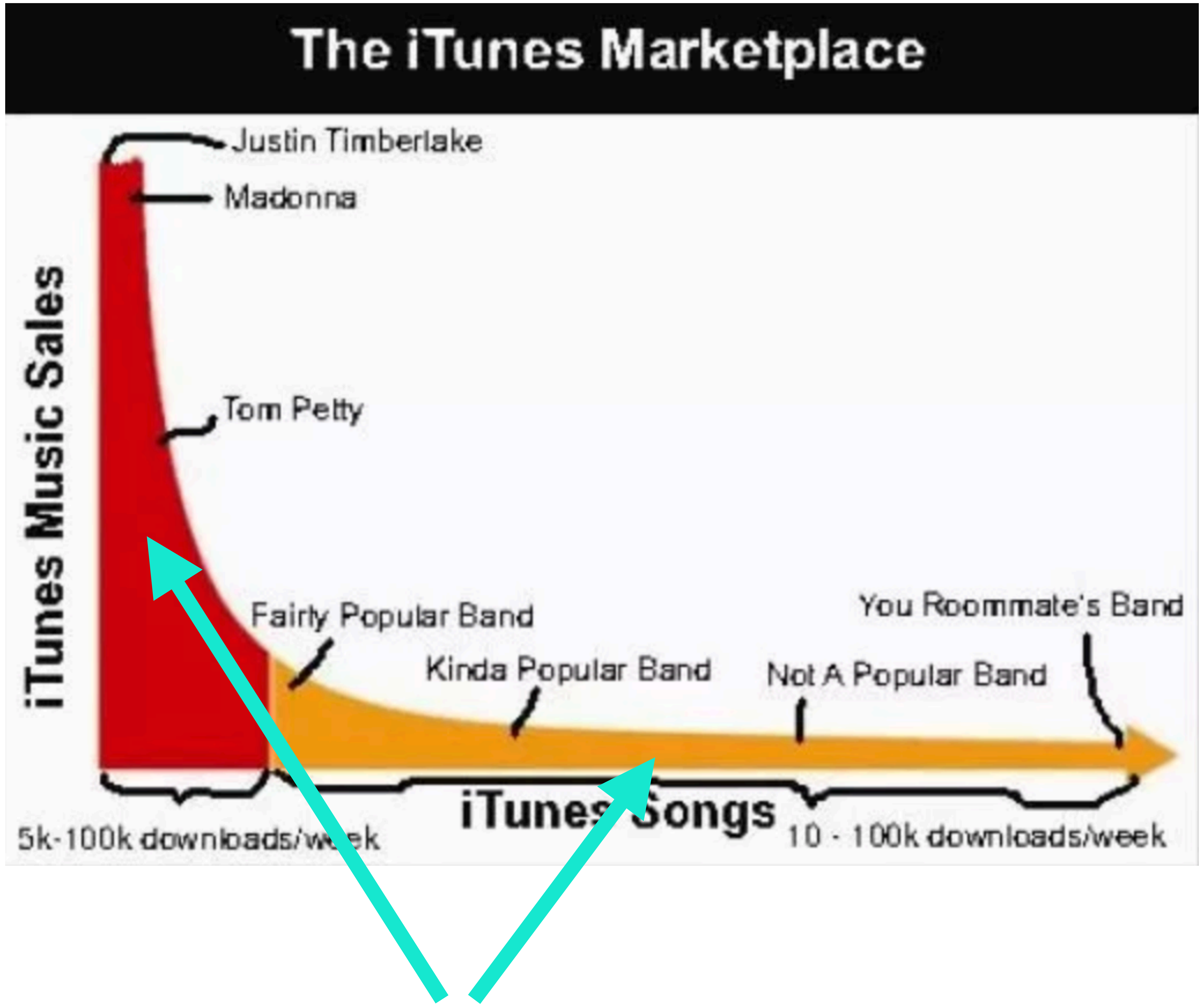
success



"quality"

Success is inherently unpredictable from quality

MusicLab:



Who ends up here is pretty **random**!

Rich Get Richer

Example in networks: new nodes are more likely to link to nodes that already have high degree

Herbert Simon's result:

- Power-laws arise from "**Rich get richer**" (**cumulative advantage**)

Examples [Price '65]

- Citations: New citations to a paper are proportional to the number it already has
- Herding: If a lot of people cite a paper, then it must be good, and therefore I should cite it too

The Model Gives Power-Laws

Claim: The described model generates networks where the fraction of nodes with in-degree *k* scales as:

$$P(d_i = k) \propto k^{-(1+\frac{1}{q})}$$

where q=1-p

So we get power-law degree distribution with exponent:

$$\alpha = 1 + \frac{1}{q} = 1 + \frac{1}{1-p}$$

Lecture 7

How to Organize the Web?

How do you organize the Web?

First try: Human curation

Web directories

Yahoo, DMOZ, LookSmart

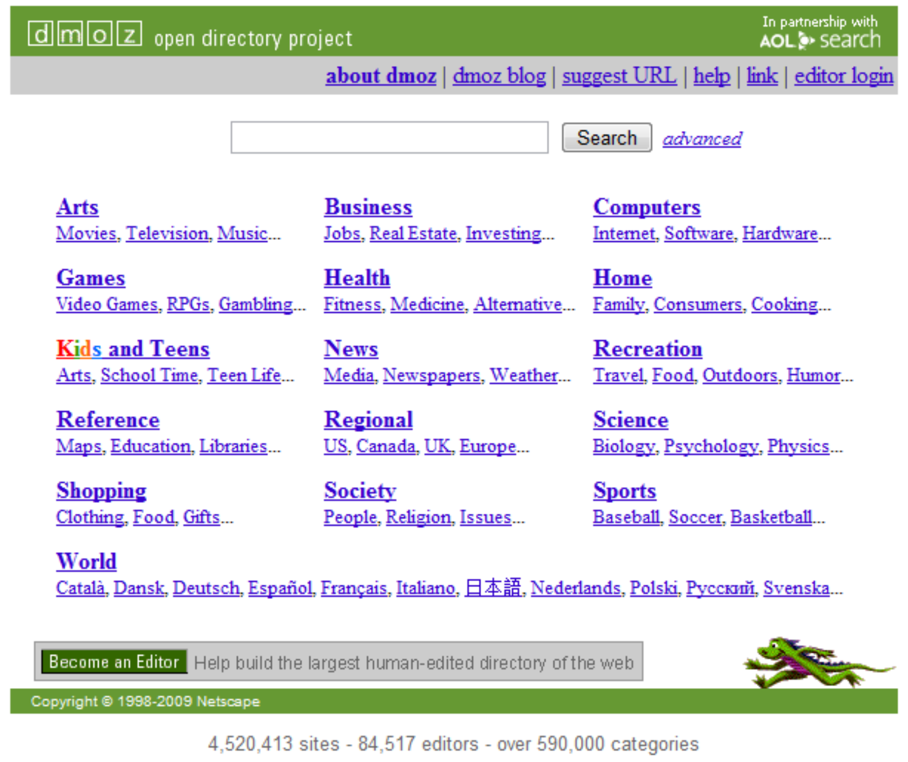
Second try: Web Search

Information Retrieval attempts to find relevant docs in a small and trusted set

Newspaper articles, Patents, etc.

But: The Web is huge, full of untrusted documents, random things, web spam, etc.

So we need a good way to rank webpages!



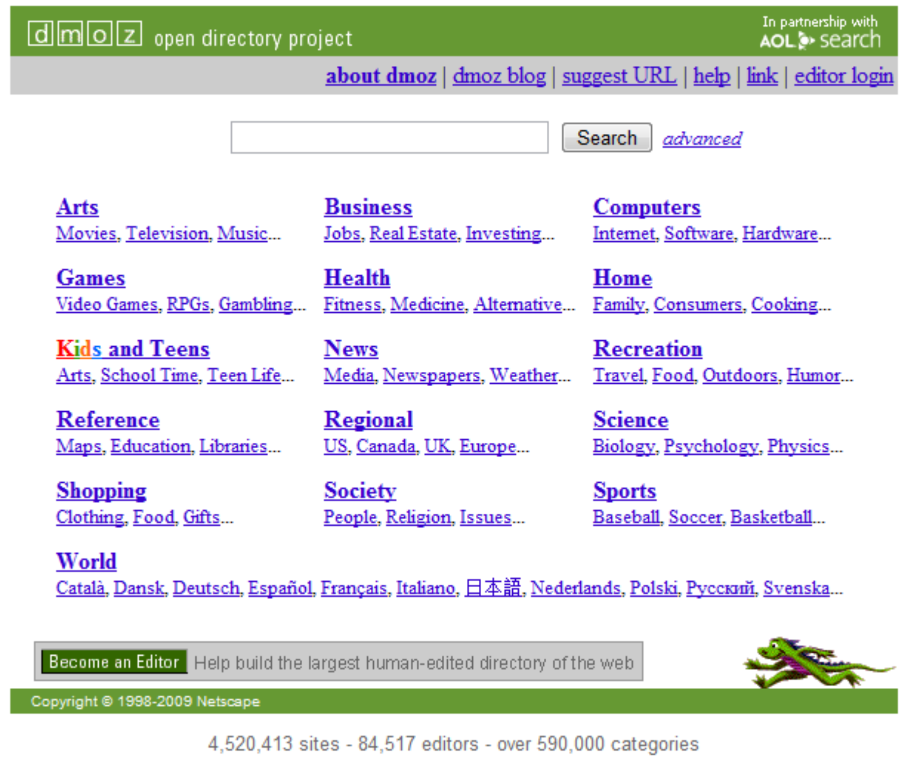
Idea: links as votes!

If I link to you, that's usually a good thing

1. Model the Web as a directed graph

2. Use the link structure to compute importance values of webpages

3. Use these importance values for ranking



Hubs and Authorities

Each page has a hub score h_i and an authority score a_i

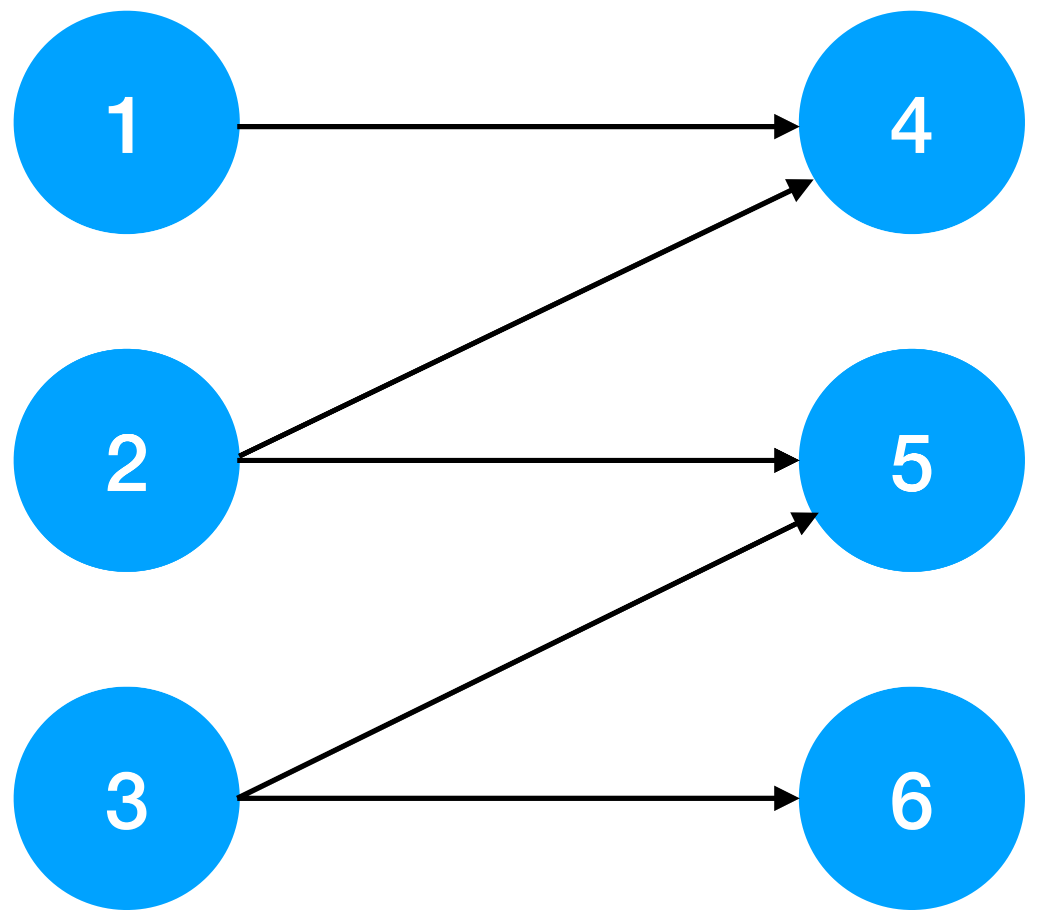
HITS algorithm:

1. Initialize all scores to 1
2. Perform a sequence of hub-authority updates:
   — First apply Authority Update Rule
   — Then apply Hub Update Rule
3. Normalize (divide authority scores by sum over a_i's and same for hubs)

(We normalize since the numbers get very big,
and we only care about the relative sizes)

Hubs and Authorities: Example

Apply 2 rounds of hub and authority update steps on the graph below:



Node	$h^{<0>}$	$a^{<1>}$	$h^{<1>}$	$a^{<2>}$	$h^{<2>}$	…	$a^{<*>}$	$h^{<*>}$
1	1	0	2/9	0	6/29	…	0	0.198
2	1	0	4/9	0	13/29	…	0	0.445
3	1	0	3/9	0	10/29	…	0	0.357
4	1	2/5	0	6/16	0	…	0.357	0
5	1	2/5	0	7/16	0	…	0.445	0
6	1	1/5	0	3/16	0	…	0.198	0

Note: in this example, values are very close to convergence after only 2 steps

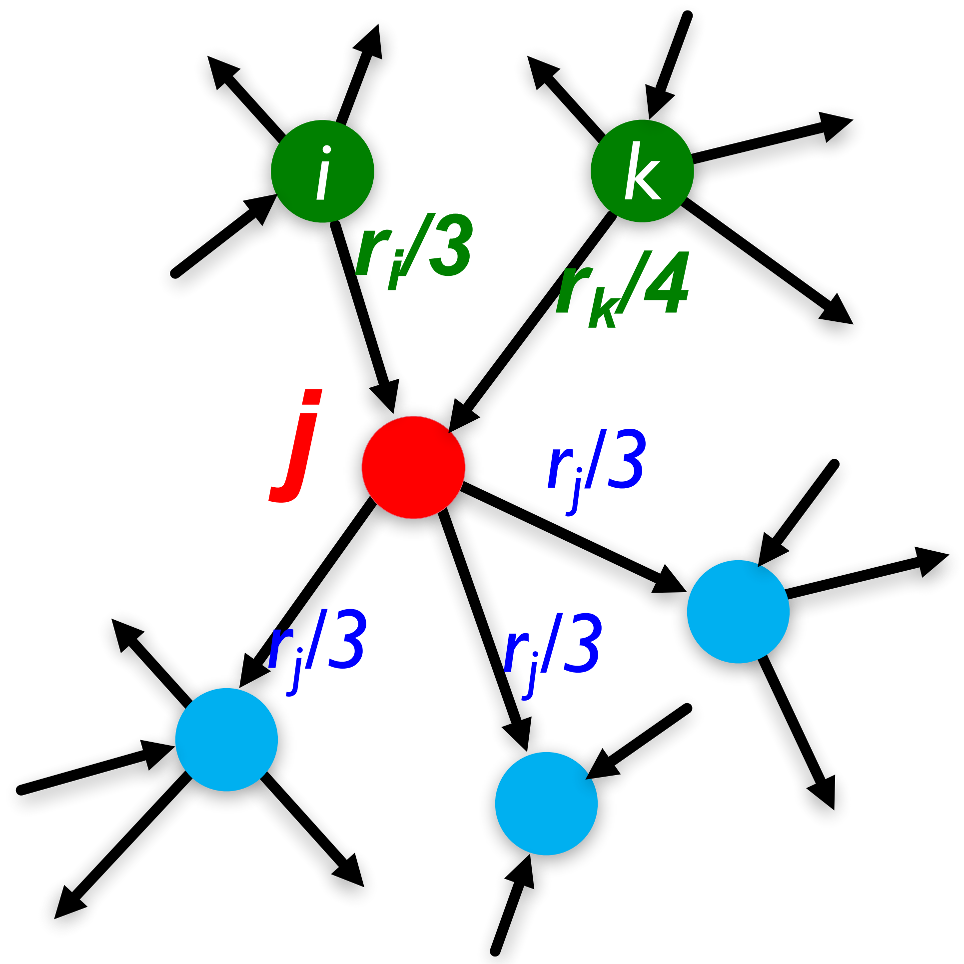
PageRank: The "Flow" Model

A "vote" from an important page is worth more:

Each link's vote is proportional to the **importance** of its source page

If page i with importance r_i has d_i out-links, each link gets r_i / d_i votes

Page j's own importance r_j is the sum of the votes on its in-links



$$r_j = r_i/3 + r_k/4$$

Mental Model: PageRank as a Fluid

Think of PageRank as a "fluid" that circulates around the network, passing from node to node and pooling at the most important ones

PageRank Algorithm:

1. Initialize all nodes with 1/n PageRank
2. Perform k PageRank updates:

Basic PageRank Update Rule: Each page divides its current PageRank equally across its outgoing links. New PageRank is the sum of PR you receive.

Page j's PageRank Update equation: $$r_j = \sum_{i \to j} \frac{r_i}{d_i}$$

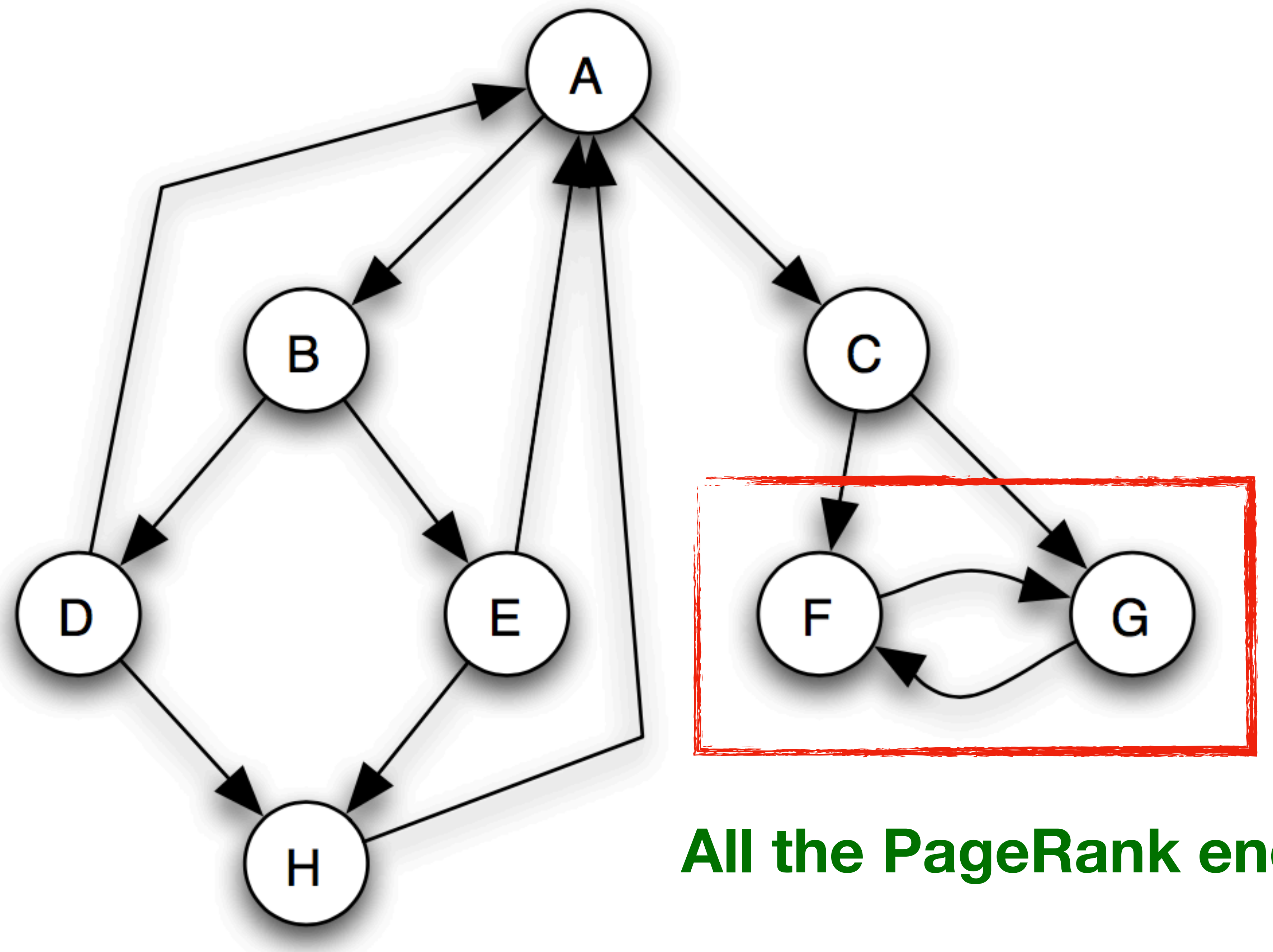
Where d_i = out-degree of node i

PageRank: A Problem

In real graph structures, PageRank can pool in the wrong places

Consider a slightly different graph:

What happens?



All the PageRank ends up here!

PageRank: A Solution

Scaled PageRank: only divide a fraction *s* of the PageRank among outgoing links
The rest gets spread evenly over all nodes

In effect we create a complete graph

Scaled PageRank Update Rule: First apply Basic PageRank Update Rule, scale down the values by *s*, then divide the residual *1-s* units of PageRank equally: (1-s)/n to each.

PageRank: Random Surfer

Claim: The probability of being at page X after k steps of this random walk is equal to the PageRank of X after k applications of the Basic PageRank Update rule.

The Random Walk: Walker chooses a starting node at random, then at each step picks one of the out-links of its current node uniformly at random.

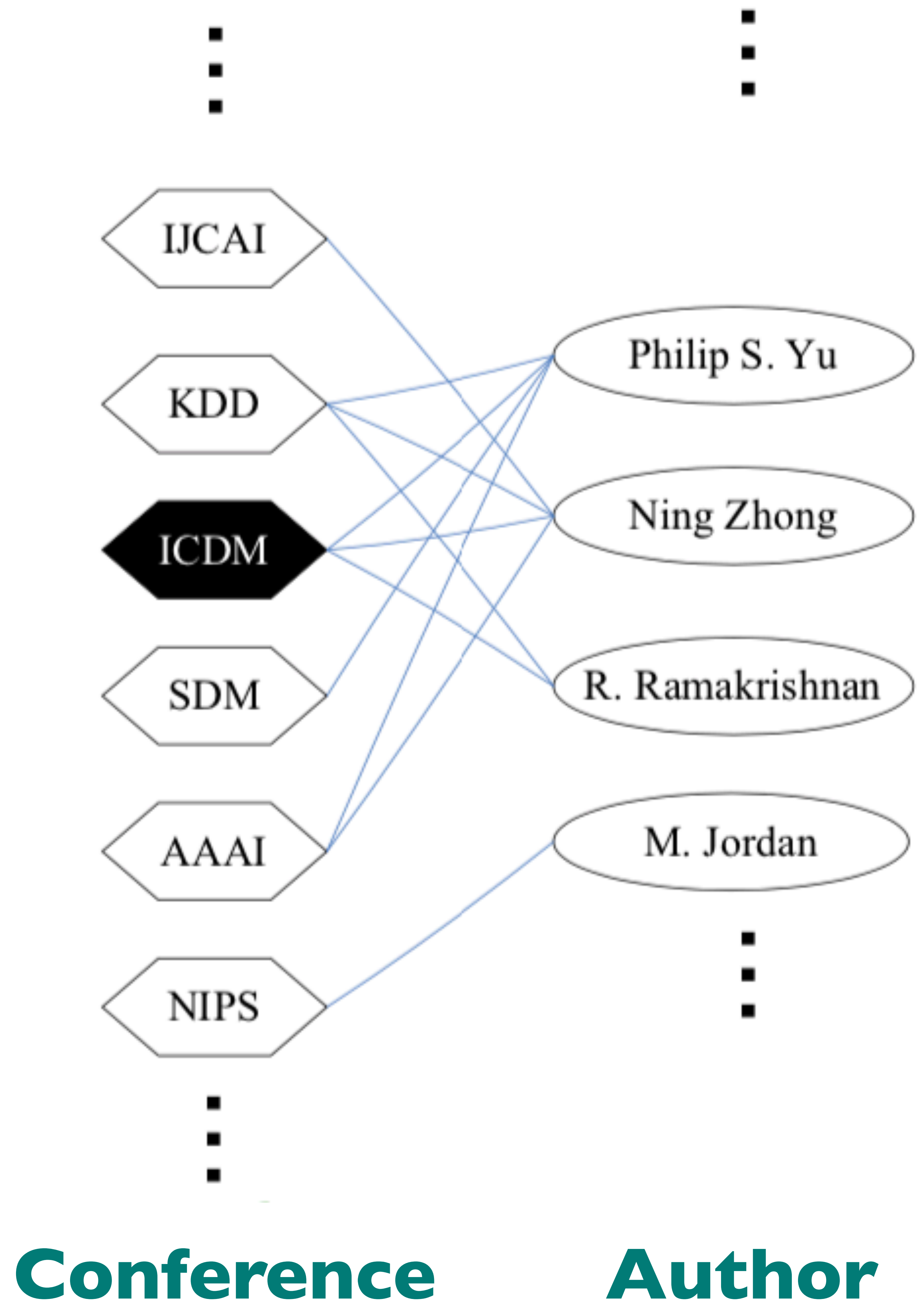
Personalized PageRank

Goal: Evaluate pages not just by popularity or global importance, but by how close they are to a given topic

Solution: change teleportation vector!

Teleporting can go to:

- Any page with equal prob. (normal PageRank)
- A topic-specific set of "relevant" pages
- A single page/node (random walk with restarts)

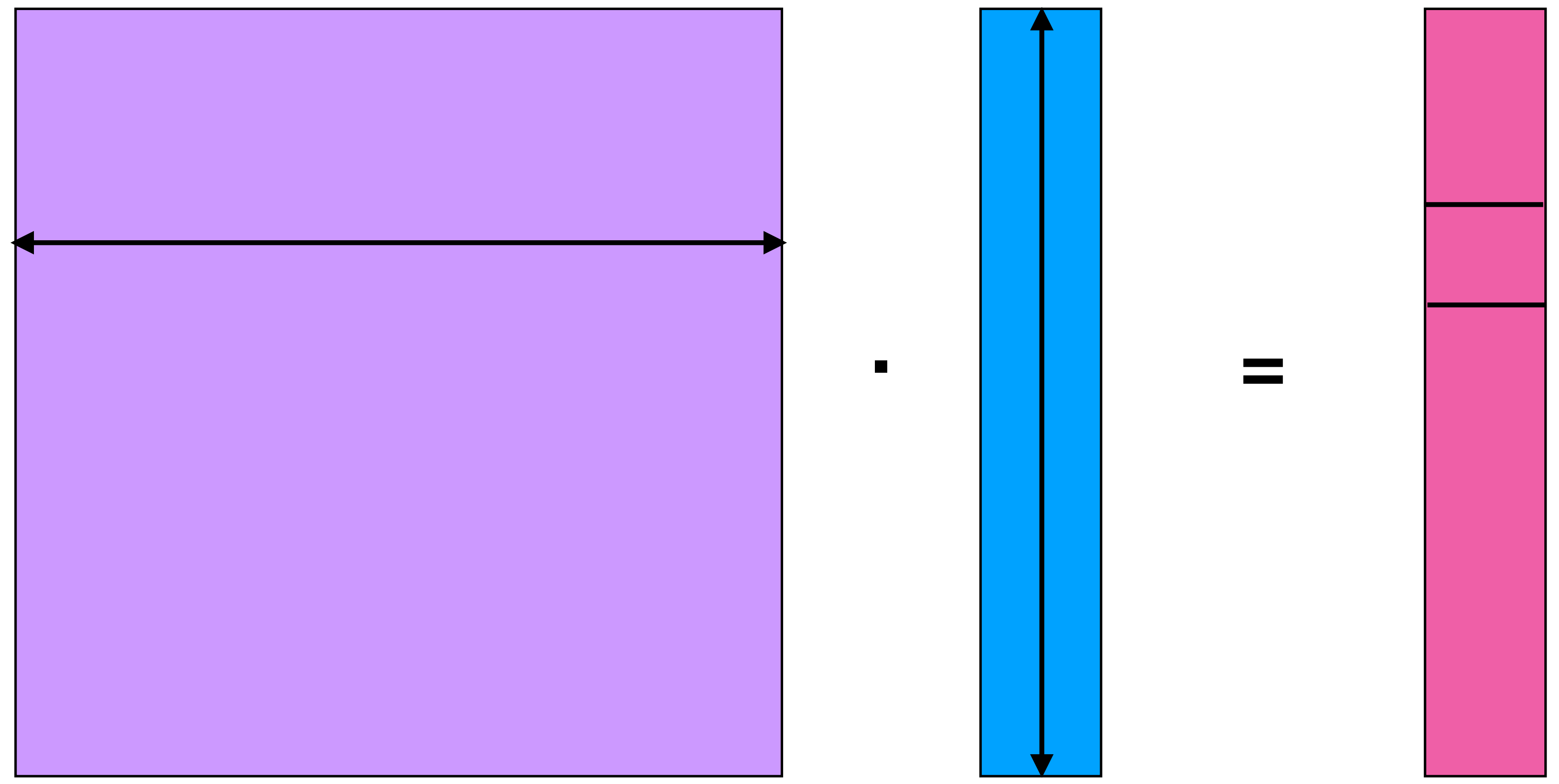


Update Rules as Matrix-Vector Multiplication

Recall Hub Update Rule:

$$h_i \leftarrow M_{i1}a_1 + M_{i2}a_2 + \ldots + M_{in}a_n$$

This corresponds exactly to the simple matrix-vector multiplication $h \leftarrow Ma$

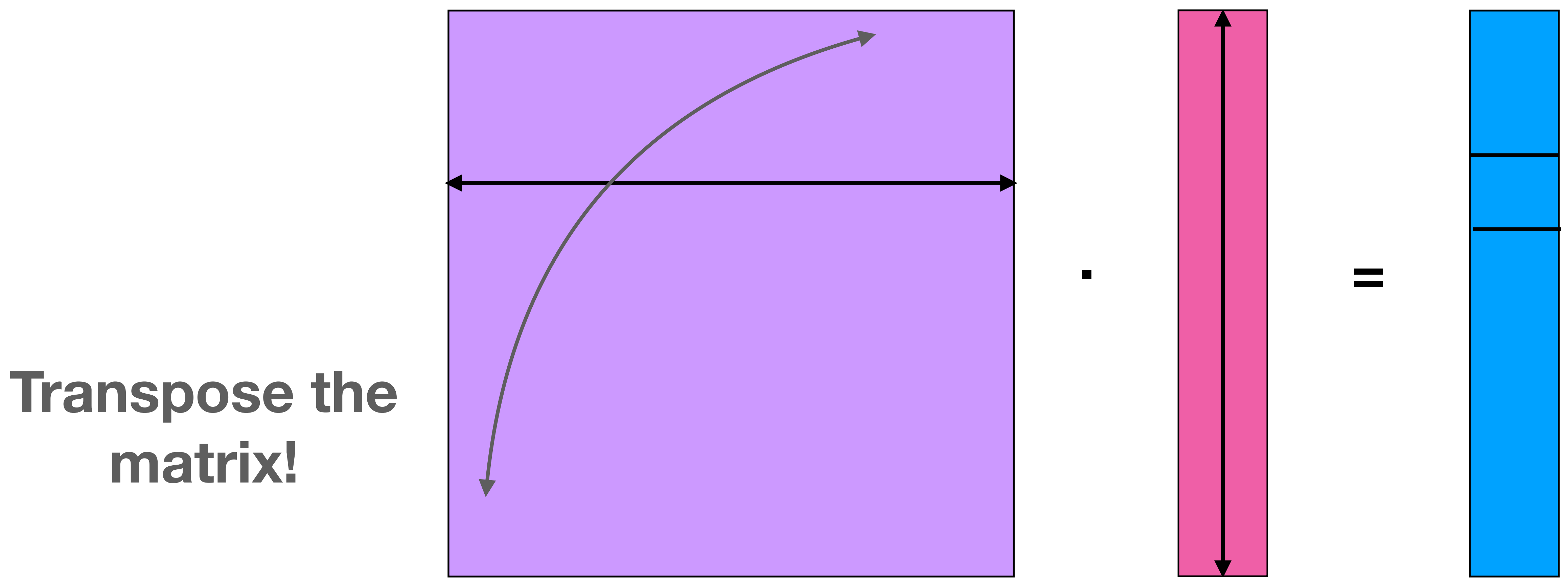


Update Rules as Matrix-Vector Multiplication

Authority update rule is similar

$$a_i \leftarrow M_{1i}h_1 + M_{2i}h_2 + \ldots + M_{ni}h_n$$

This corresponds exactly to the simple matrix-vector multiplication $a \leftarrow M^T h$



Convergence

Recall your eigenvectors and eigenvalues:

$$Av = \lambda v$$

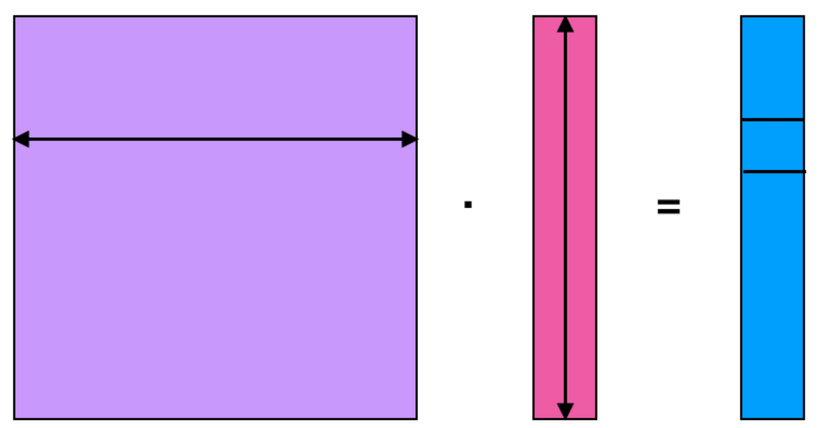
v is an eigenvector of A, with corresponding eigenvalue lambda

At convergence, performing additional hub-authority steps won't change anything

Thus Hubs and Authorities converges to the leading eigenvector of MM^T and M^TM!

$$(MM^T)h^{\langle * \rangle} = c \cdot h^{\langle * \rangle}$$

eigenvector eigenvalue

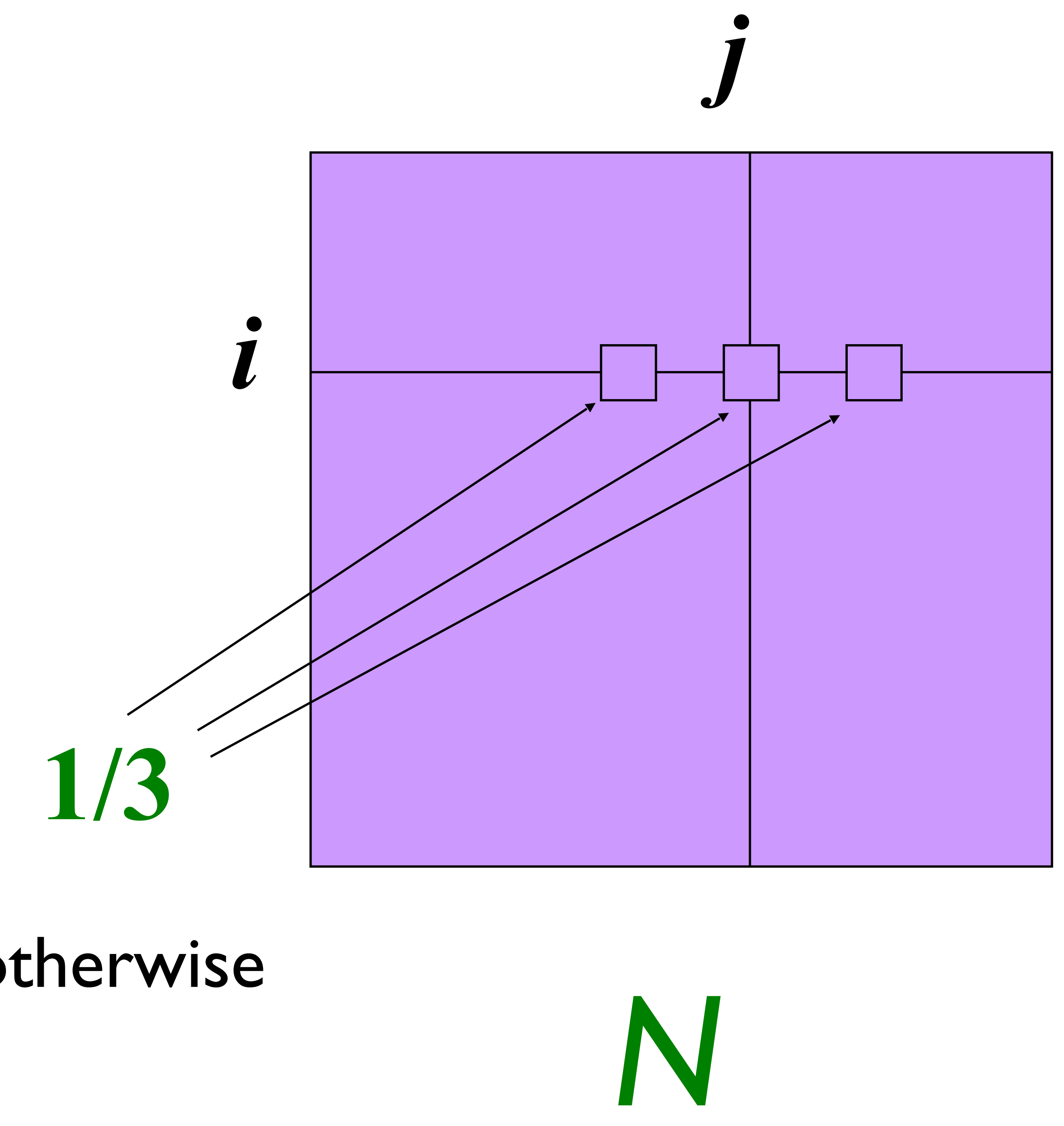


(Full details in the reading)

PageRank Spectral Analysis

Recall the Basic PageRank Update Rule:

$$r_j^{\langle k+1 \rangle} = \sum_{i \to j} \frac{r_i^{\langle k \rangle}}{d_i}$$



Define a new matrix N: $N_{ij} = \frac{1}{d_i}$ for edges i->j, 0 otherwise

where page i has
d_i out-links

$$r^{\langle k+1 \rangle} = N_{1i} r_1^{\langle k \rangle} + N_{2i} r_2^{\langle k \rangle} + \cdots N_{ni} r_n^{\langle k \rangle}$$

$$r^{\langle k+1 \rangle} = N^T r^{\langle k \rangle}$$

Similarly, PageRank converges to the leading eigenvector of N^T

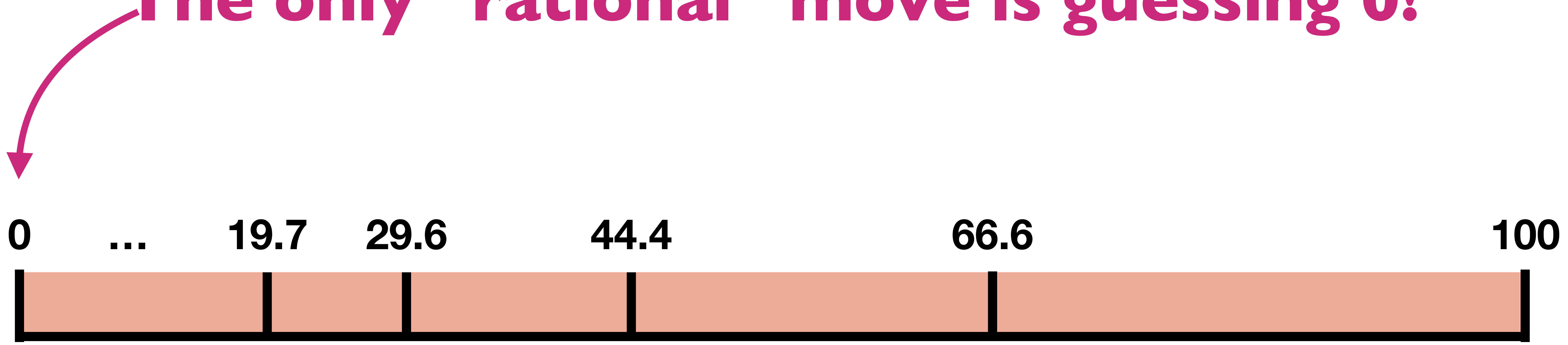
Lecture 8

What is “rational” play?

Repeat!

44.4 is the new 66.6, and so on

The only “rational” move is guessing 0!



(of course, in real life not everyone is rational)

Exam-Presentation Game

What should you do?

If you knew your partner would study for the **exam**, what should you do?

You'd choose **exam (88 > 86)**

If you knew your partner would work on the **presentation**, what should you do?

You'd choose **exam (92 > 90)**

No matter what, you should choose exam!

		Your Partner: *Presentation*	Your Partner: *Exam*
You	*Presentation*	90, 90	86, 92
You	*Exam*	92, 86	88, 88

Basic Definitions

A game G is a tuple (P,S,O):

P = set of Players

S = a set of strategies for every player

O = for every outcome (where every player is picking one strategy), a payoff for each player

Payoff matrix summarizes all of these (each dimension is a player, every row/column/etc is a strategy for one player, every cell expresses payoffs for each player)

Underlying Assumptions

Payoffs summarize **everything** a player cares about

Every player knows everything about the structure of the game: who the players are, the strategies available to everyone, payoffs for each player and strategy

Every player is **rational**: wants to maximize payoff and succeeds in doing so

		Your Partner: *Presentation*	Your Partner: *Exam*
You	*Presentation*	$90, 90$	$86, 92$
	Exam	$92, 86$	$88, 88$

Fundamental Concepts: Strict Dominant Strategy

A strategy that is strictly better than all other options, regardless of what other players do

Exam is a **strictly dominant strategy** for both players

Sadly, (90,90) is not achievable with rational play

Even if you could commit to preparing for the presentation, your partner would still be better off studying for the final

		Your Partner	
		Presentation	*Exam*
You	*Presentation*	90, 90	86, 92
	Exam	92, 86	88, 88

Fundamental Concepts: Best Response

Let's define some more of the fundamental concepts we just used

Strategy **S** by P_1 is a **best response** to strategy **T** by P_2 if the payoff from **S** as at least as good as anyone other strategy against **T**

$P_1(S,T) \geq P_1(S',T)$ for all other S' by P_1

It's a **strict best response** if:

$P_1(S,T) > P_1(S',T)$ for all other S' by P_1

		Suspect 2: NC	Suspect 2: C
Suspect 1	NC	$-1, -1$	$-10, 0$
Suspect 1	C	$0, -10$	$-4, -4$

S1's best response to NC is: C
S1's best response to C is: C

Fundamental Concepts: Dominant Strategy

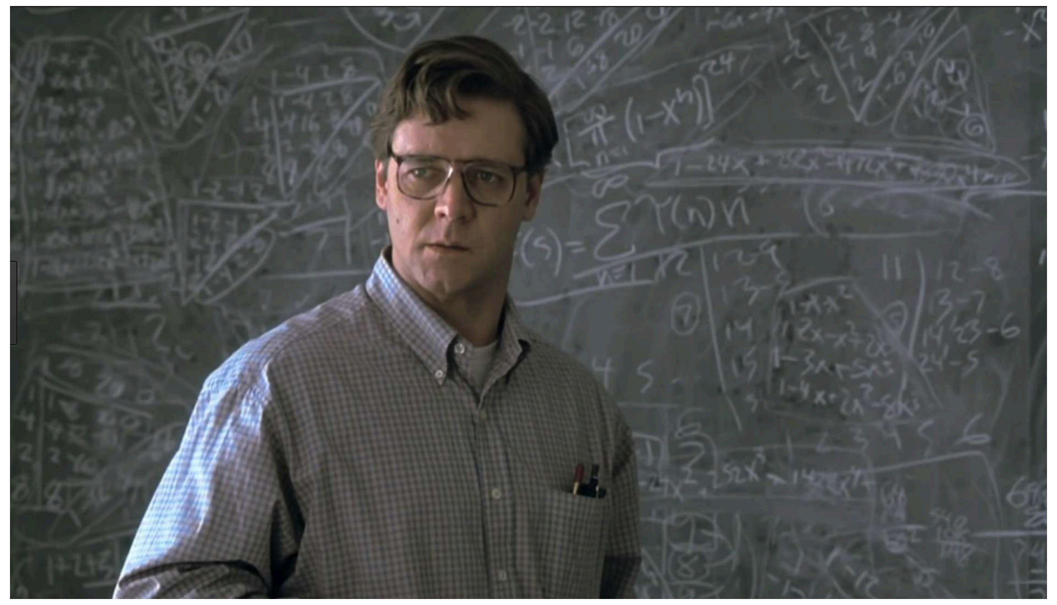
A **dominant strategy** for P_1 is a strategy that is a best response every strategy by P_2

A **strict dominant strategy** for P_1 is a strategy that is a strict best response every strategy by P_2

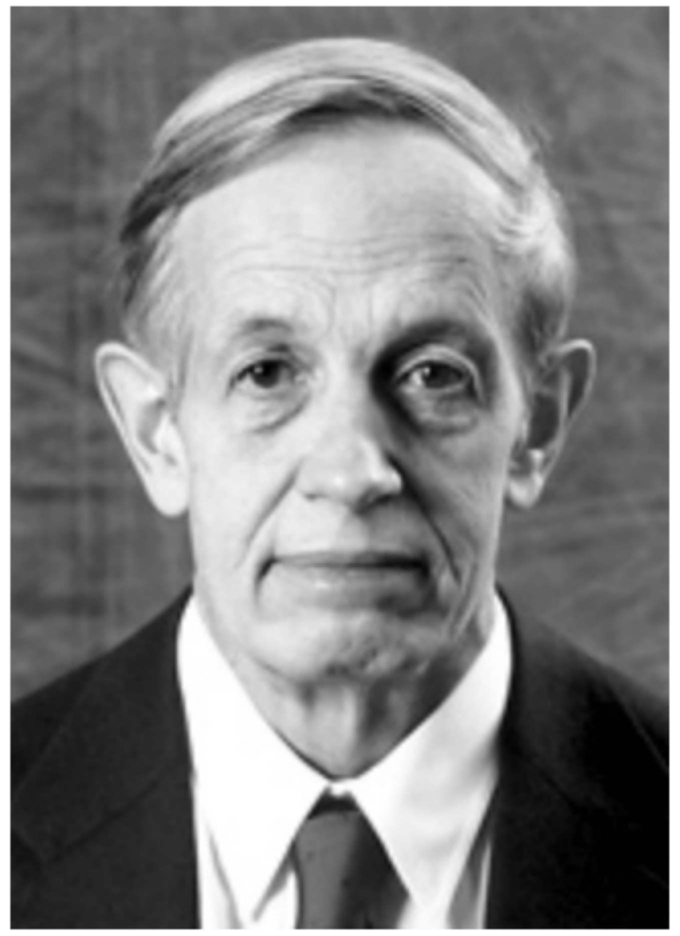
		Suspect 2	
		NC	C
Suspect 1	NC	$-1, -1$	$-10, 0$
	C	$0, -10$	$-4, -4$

(Note: In Prisoner's Dilemma, P_1 has a strict dominant strategy, so we expect P1 to play it. There can be several dominant strategies, and it'd be unclear which one to expect)

Nash Equilibrium



In 1950, John Nash proposed a **simple** and **powerful** principle for reasoning about behaviour in general games (and won the Nobel Prize for it in 1994)



Even when there are no dominant strategies, **we should expect players to use strategies that are best responses to each other**

A pair of strategies (**S**,**T**) is a Nash equilibrium if **S** is a best response to **T** and **T** is a best response to **S**

Mixed Strategies Example: Football

Players: Offense, Defense

Strategies: Run, Pass and Defend Run, Defend Pass

Payoff matrix:

		Defense	
		Defend Pass	*Defend Run*
Offense	*Pass*	$0, 0$	$10, -10$
	Run	$5, -5$	$0, 0$

Mixed Nash:
q = 2/3
p = 1/3

No Nash equilibria in this game

O's expected payoff for **Pass** when D plays p: 0*(q)+10*(1-q) = 10-10q

O's expected payoff for **Run** when D plays q: 5*(q)+0*(1-q) = 5q

Defense makes Offense indifferent when q=2/3

Lecture 9

Traffic modeled as a game

Players: Drivers 1,2,3…,4000

Strategies: A-C-B, A-D-B

Payoffs: Negative drive time

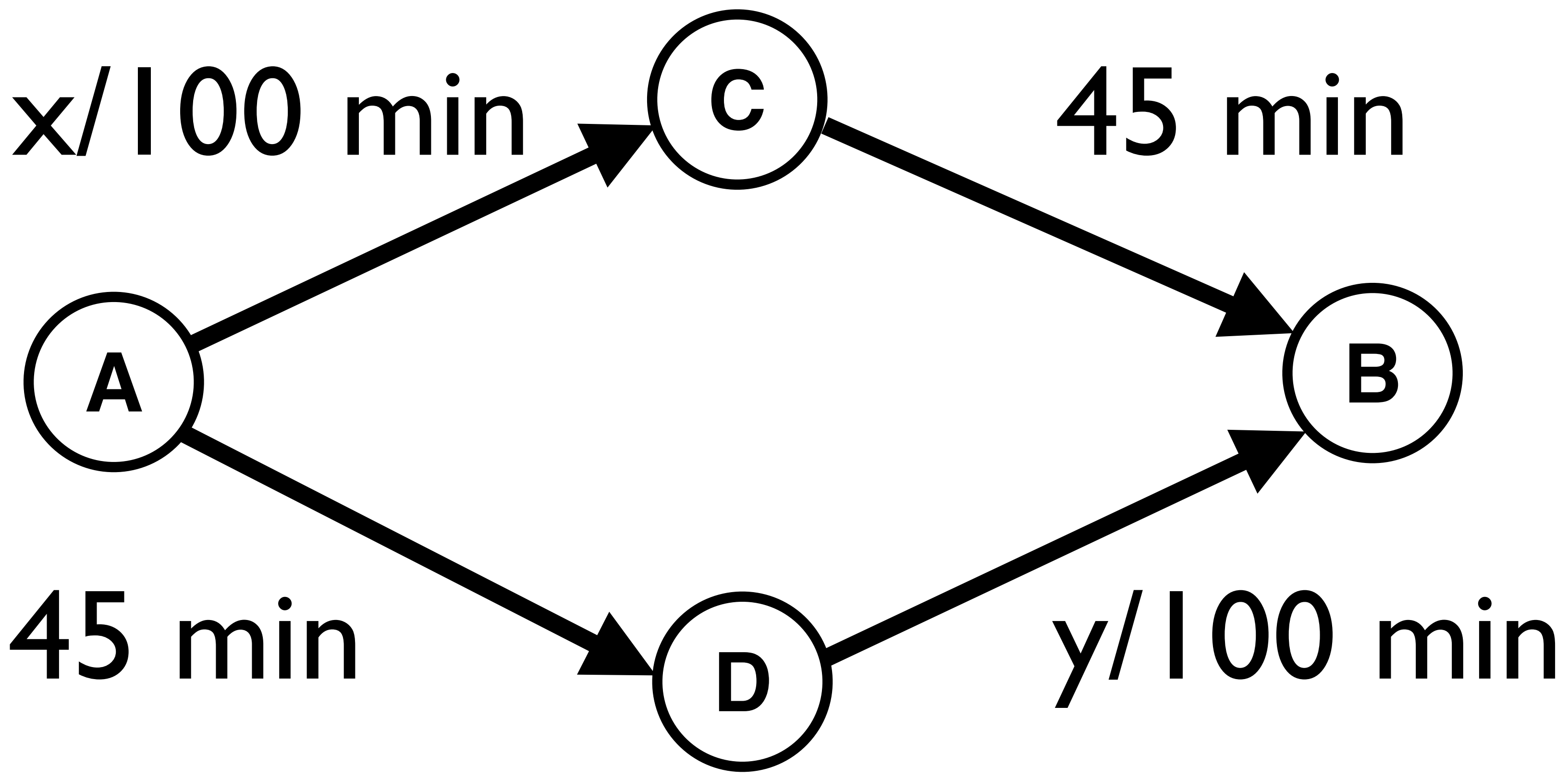
A-C-B time: -(x/100 + 45)

A-D-B time: -(45 + y/100)

You want to lower your drive time,
so we take the negative drive time as the "payoff"

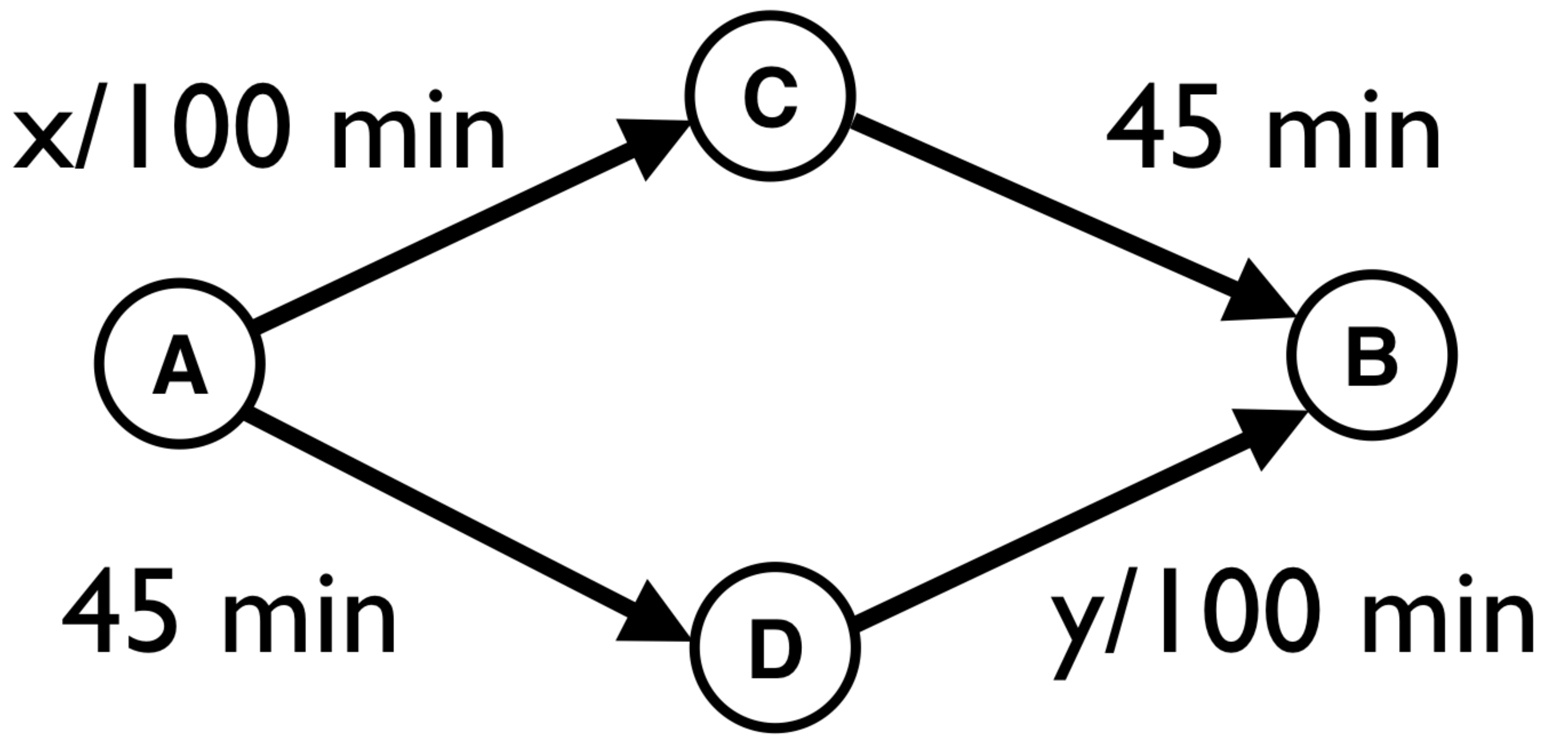
Notice that this actually describes many equilibria: any set of strategies "2000 choose top, 2000 choose bottom" is an equilibrium (players are interchangeable, so any set of 2000 can be using ACB and any set of 2000 can be using ADB)

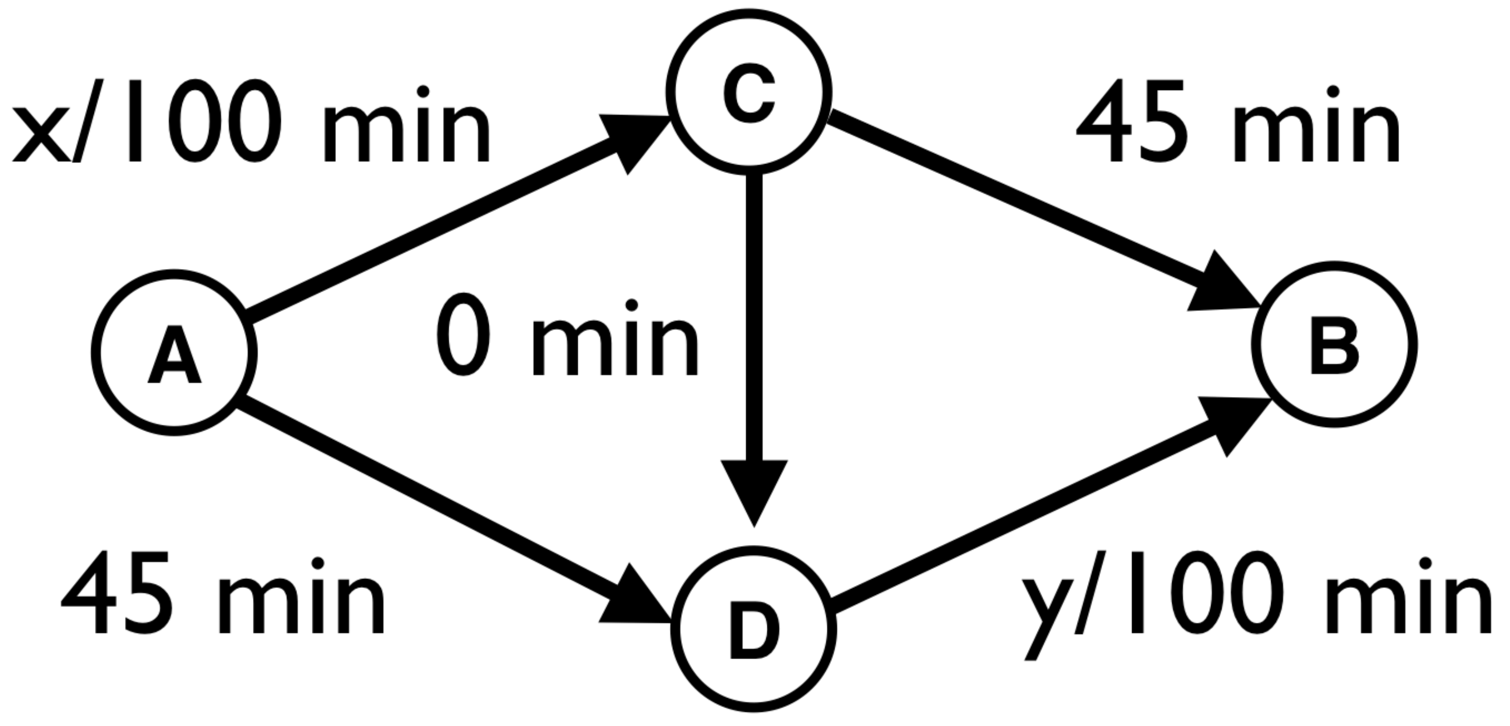
For any other set of strategies, deviation benefits someone (therefore isn't an equilibrium)



Braess's Paradox

Routing:



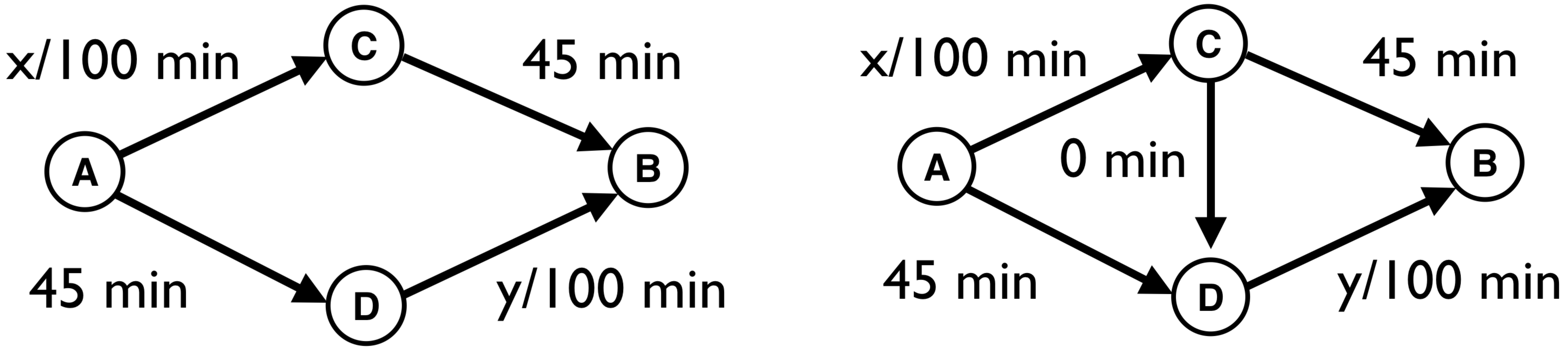


Prisoner's Dilemma:

		Suspect 2	
		NC	C
Suspect 1	NC	$-1, -1$	$-10, 0$
	C	$0, -10$	$-4, -4$

How bad can it get?

Routing:



Ratio between socially optimal and selfish routing (called the "Price of Anarchy")?

This example: 80/65 = 1.23x worse

Worst case: How bad can it get?

For selfish routing, "Price of Anarchy" = 4/3

Game Theoretic Model of Cascades

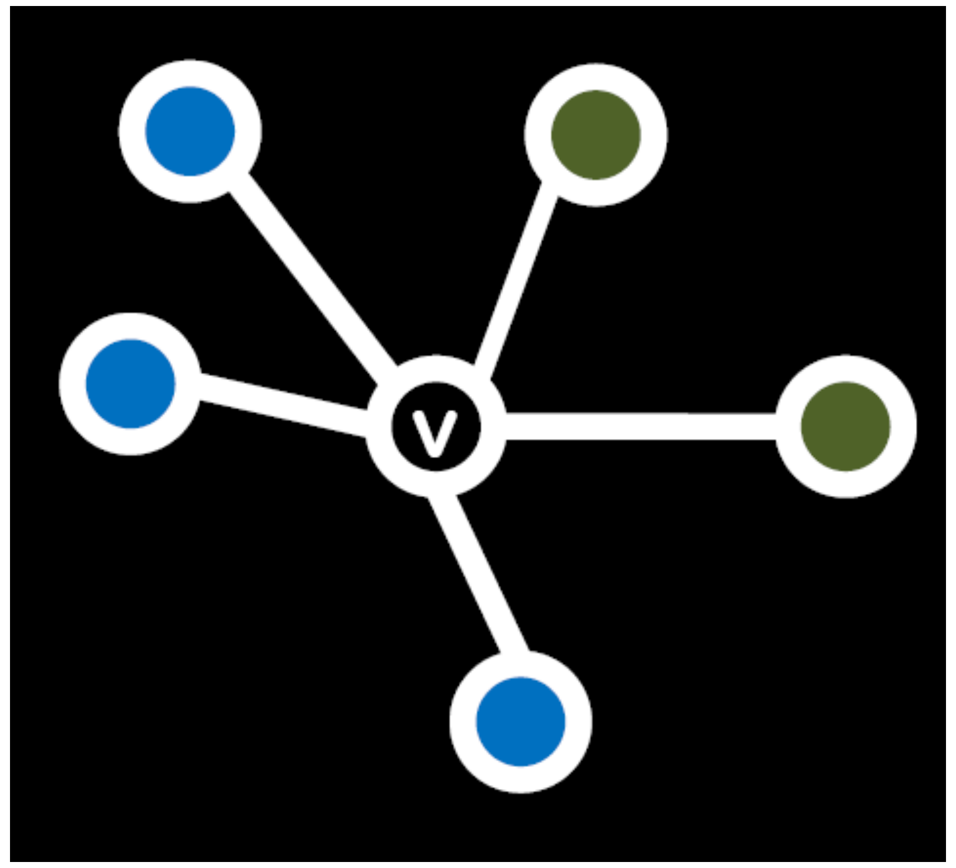
Game Theory + Social Networks can help us think about this question!

Model every friendship edge as a 2 player coordination game

2 players – each chooses technology A or B

Each person can only adopt **one** "behavior", **A** or **B**

You gain more payoff if your friend has adopted the **same** behavior as you



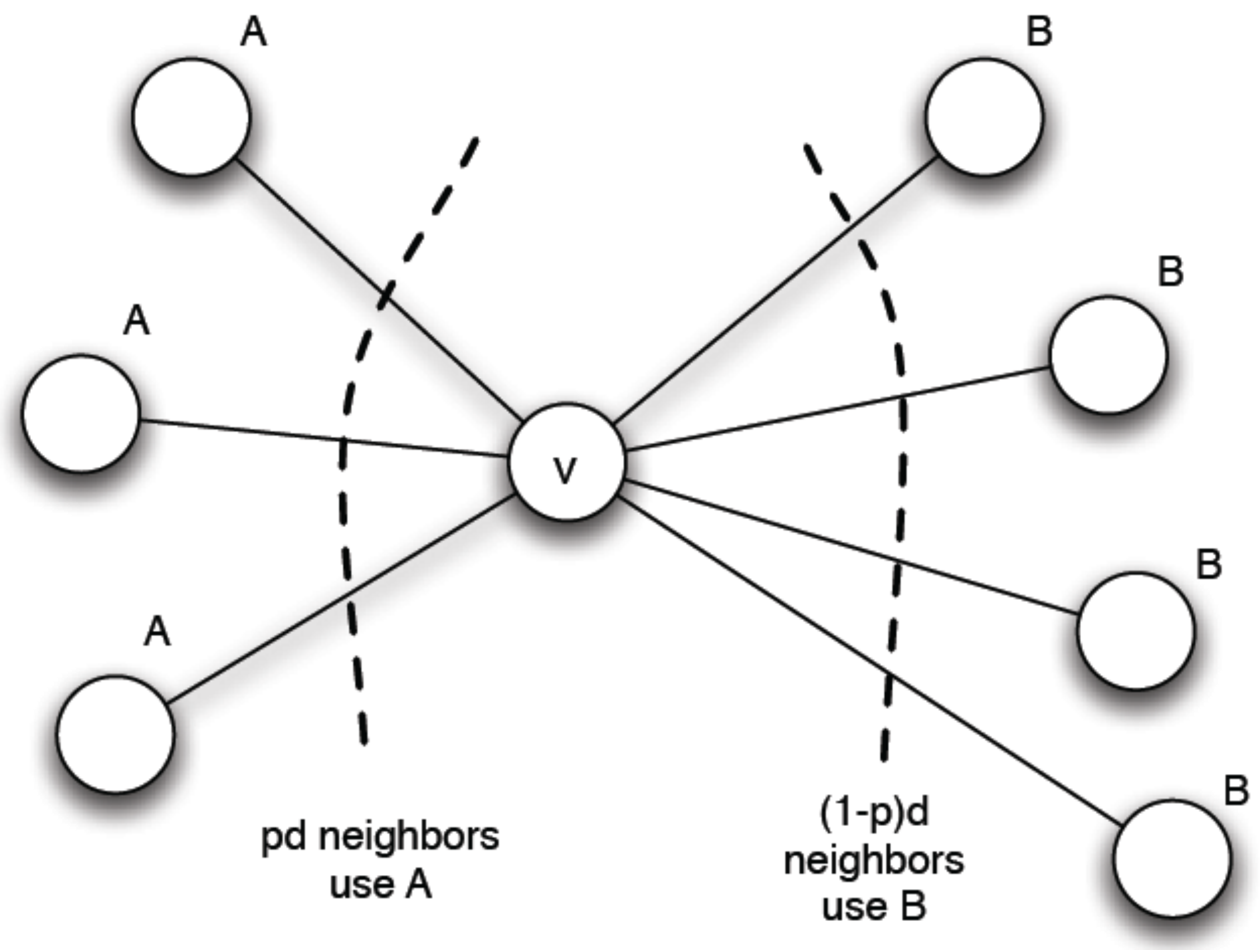
Local view of the network of node **v**

		w: A	w: B
v	A	a, a	$0, 0$
	B	$0, 0$	b, b

Calculation of Node *v*

Let **v** have ***d*** neighbours — some adopt **A** and some adopt **B**

Say fraction ***p*** of ***v***'s neighbours adopt ***A*** and **1-p** adopt **B**



$Payoff_v = a \cdot p \cdot d$ if v chooses A

$= b \cdot (1-p) \cdot d$ if v chooses B

Thus: *v* chooses A if:

$a \cdot p \cdot d > b \cdot (1-p) \cdot d$

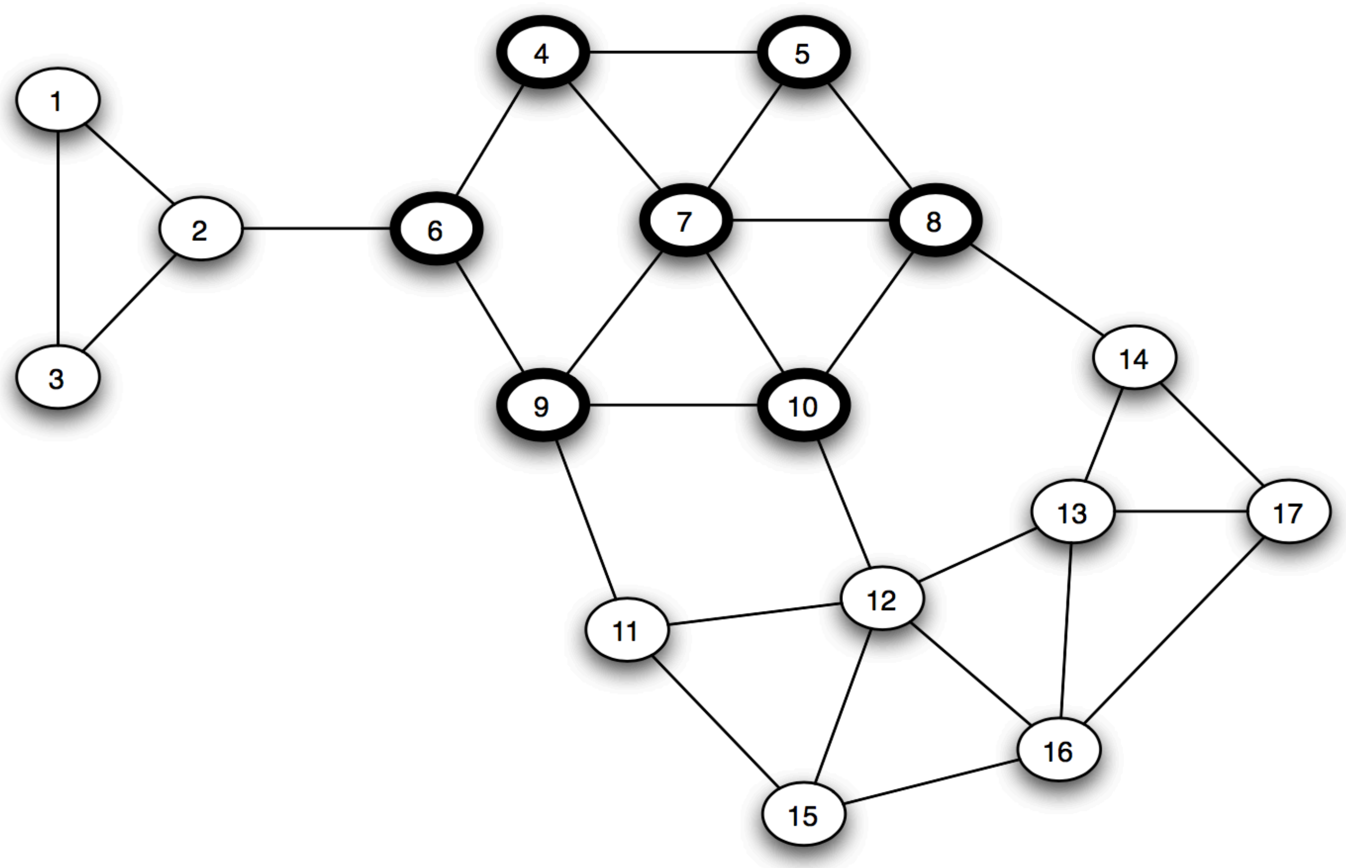
Threshold:

v chooses ***A*** if $p > \frac{b}{a+b} = q$

p… frac. v's neighbours choosing A
q… payoff threshold

Another example with a=3 and b=2

What are the impediments to spread?



Densely connected communities

- 1–3 are well-connected with each other but poorly connected to the rest of the network
- Similar story for 11–17
- **Homophily impedes diffusion**

A **cluster of density *p*** is a set of nodes such that every node in the set has at least a p fraction of its neighbours in the set

Nodes {1,2,3} are a cluster of density p = *2/3*

Nodes {11,12,13,14,15,16,17} are a cluster of density p = *2/3*

Simple Herding Model

Decision to be made (resto choice, adopt a new technology, support political position, etc)

People decide sequentially, and see all choices of those who acted earlier

Each person has some **private information** that can help guide their decision

People **can't** directly observe what others **know**, but **can** observe what they **do**

Lecture 10

Simple Herding Model

Model: n students in a classroom, urn in front

Two urns with marbles:

- "Majority-blue" urn has 2/3 blue, 1/3 red
- "Majority-red" urn has 2/3 red, 1/3 blue
- 50%/50% chance that the urn is majority blue/red

One by one, each student privately gets to look at 1 marble, put it back without showing anyone else, and guess if the urn is Majority-blue or Majority-red

Simple Herding Model

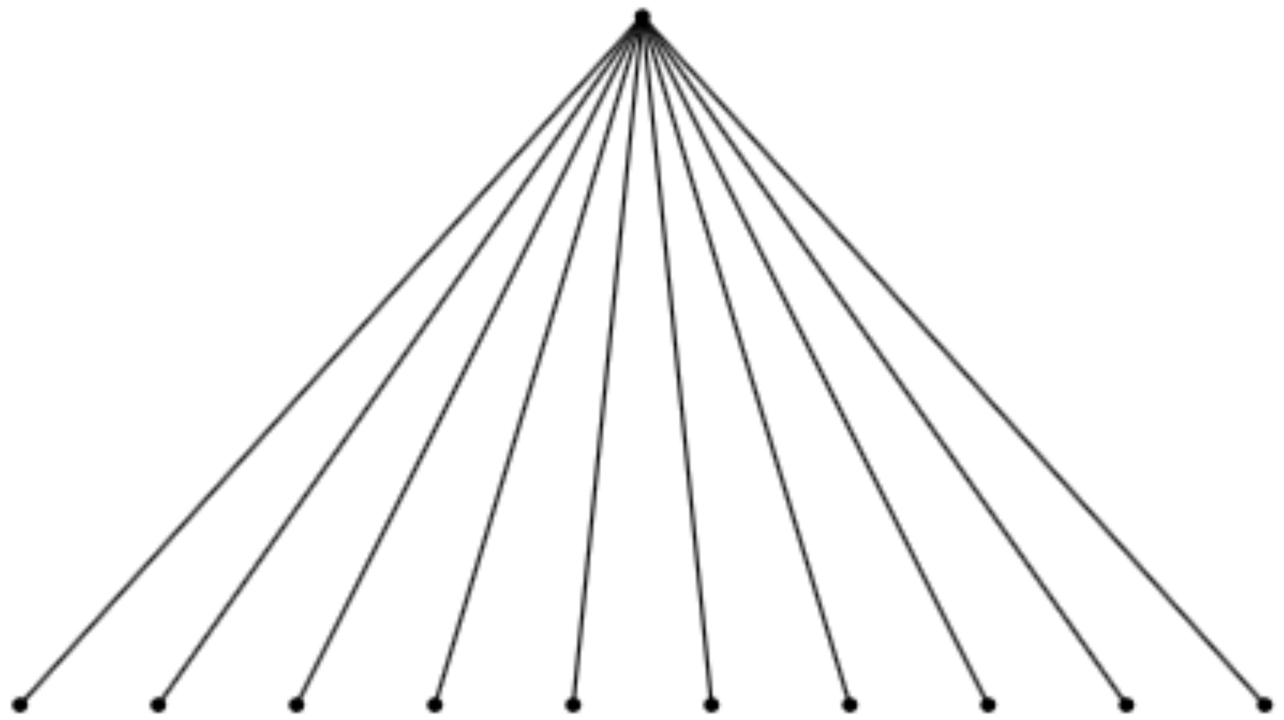
Student 1: Just guess the colour she sees
Student 2:

If same as first person, guess that colour.
But if different from first, then since he knows first guess was what first person saw, then he's indifferent between the two. Guess what he saw

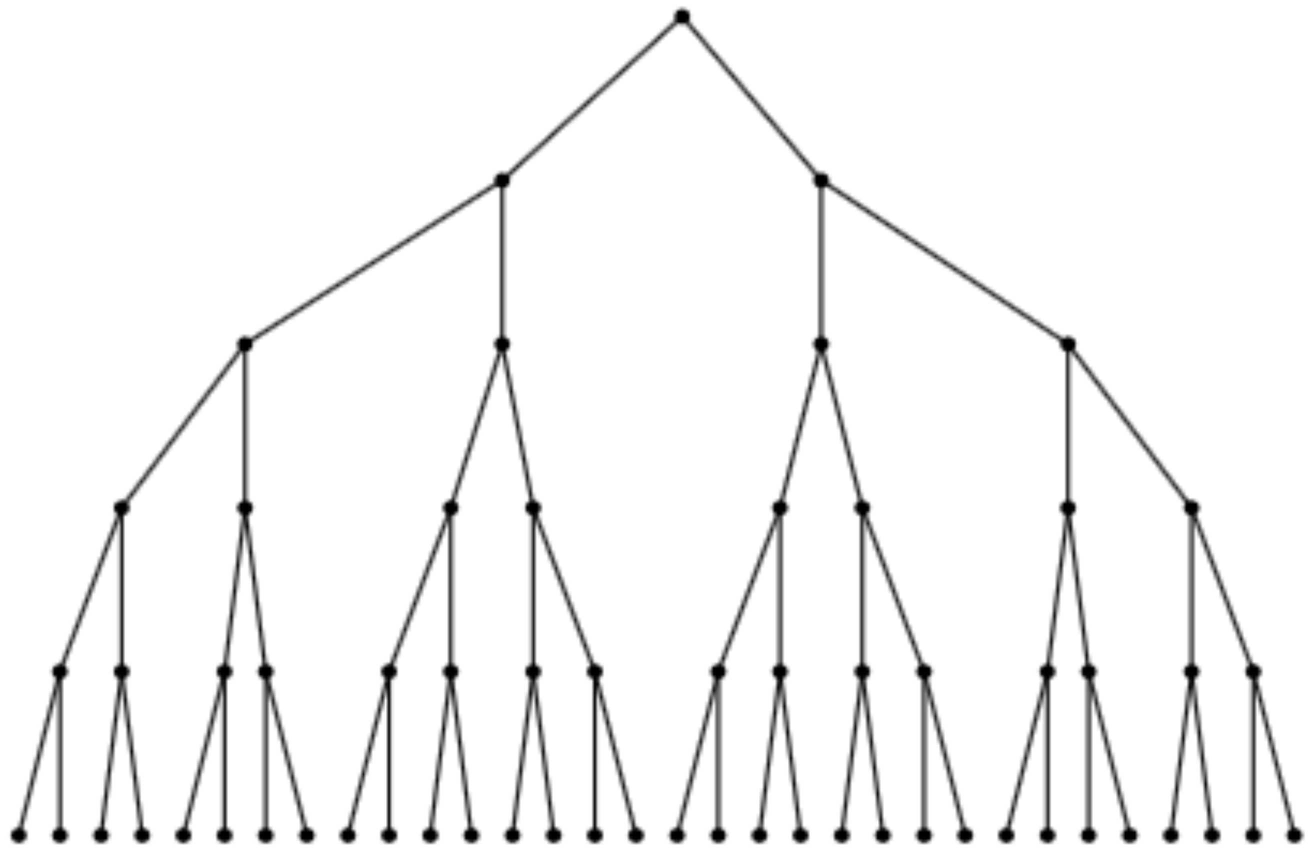
Student 3:

If first 2 are opposite colours, guess what she sees (tiebreaker)
If previous 2 are the same colour (blue) and S3 draws red, then it's like he has drawn three times and gotten two blue, so she should guess majority-blue, **despite her own private information!**

Which is it?



or



“Broadcast”

Big media (CNN, BBC, NYT, Fox)
Celebrities (Biebs, Taylor Swift)

“Viral”

Organically spreading content
Chain letters

How to measure virality?

Solution: **average path length between nodes**

$$\nu(T) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}$$

Simple average!

Originally studied in mathematical chemistry [Wiener 1947] => "Wiener index"

Not viral

Super viral

Lecture 11

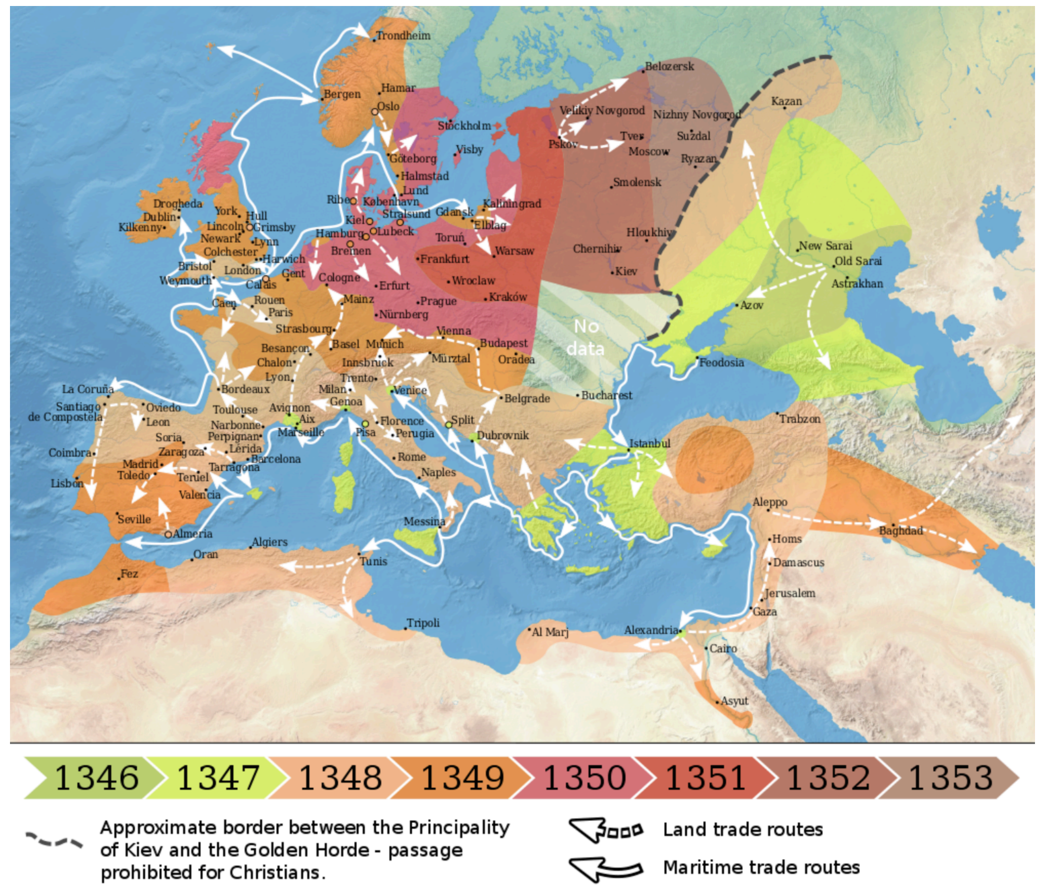
How Things Spread

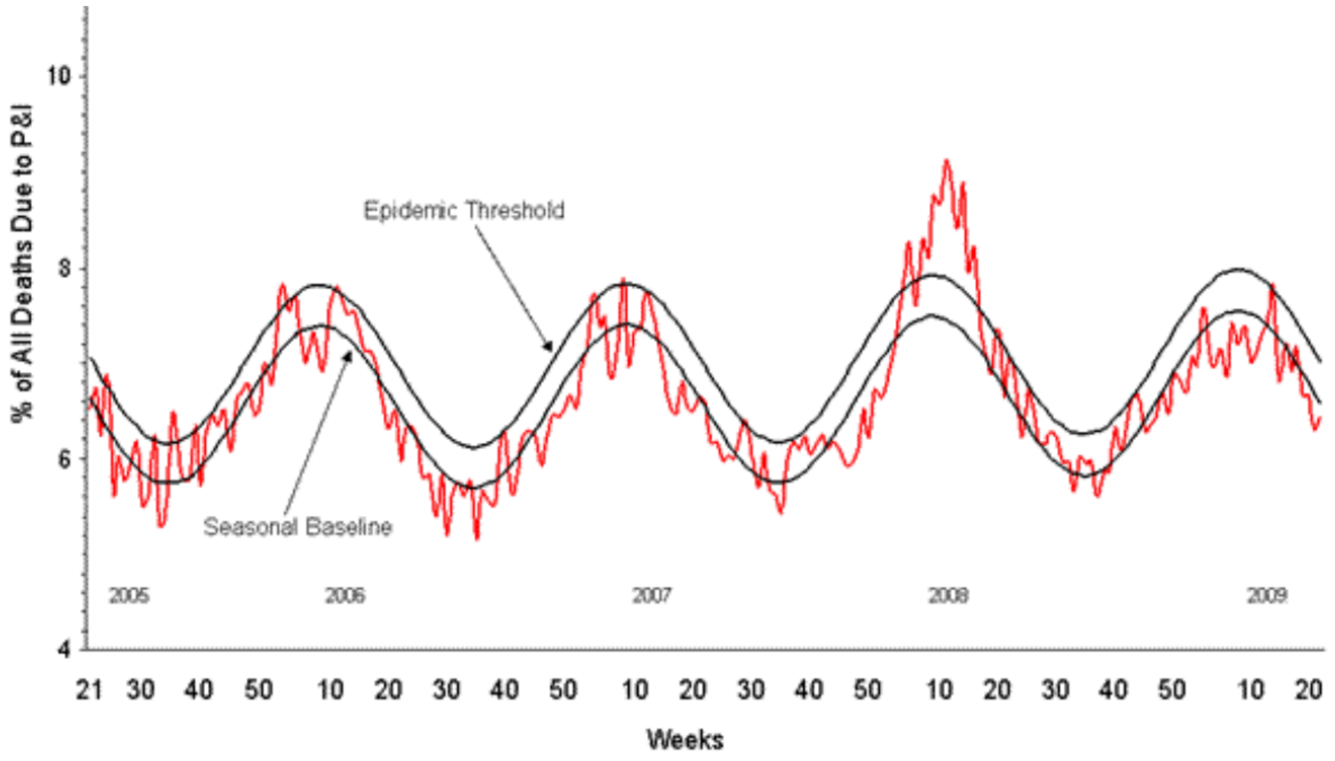
Networks define how behaviours, ideas, beliefs, diseases, etc. spread

Last class: behaviour (adoption of an innovation or technology) and information

Today:

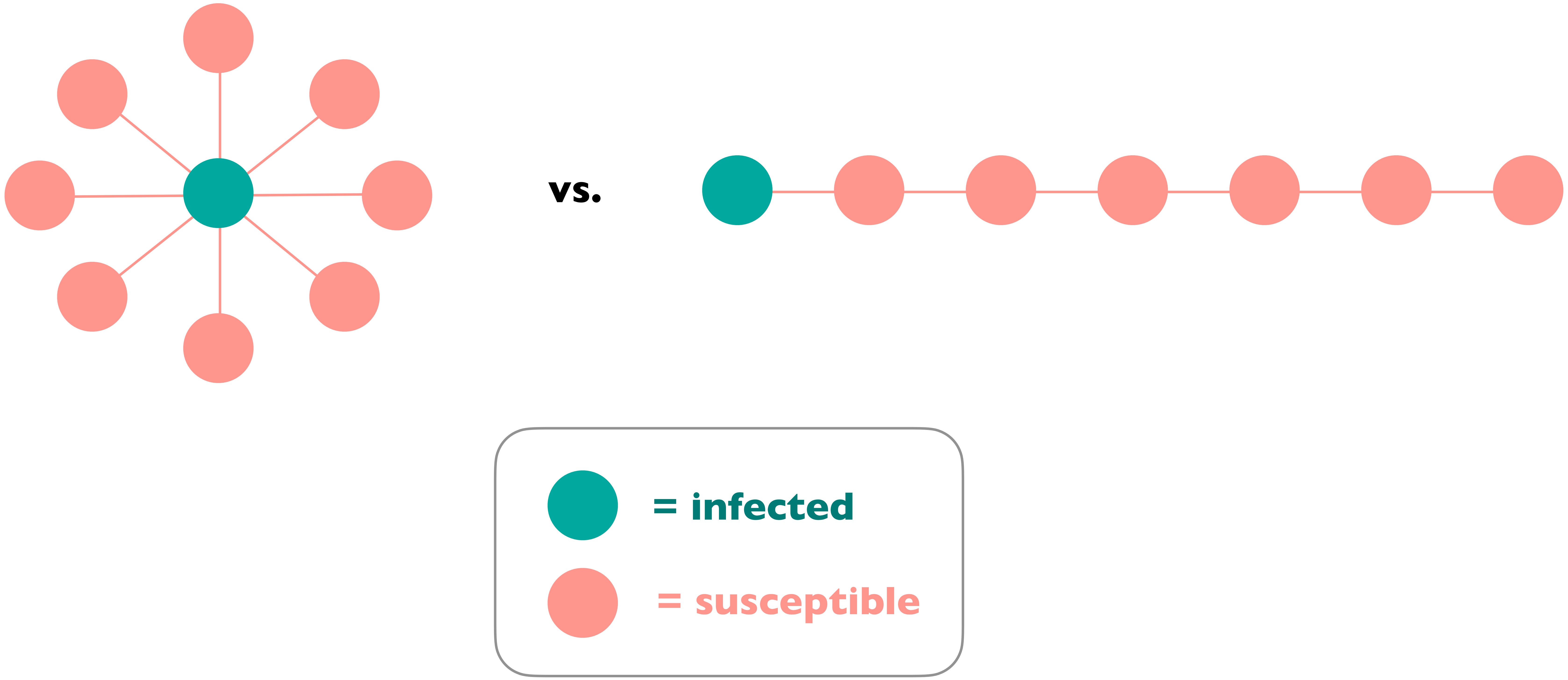
Epidemics





Epidemics

Which disease is more dangerous to the population?



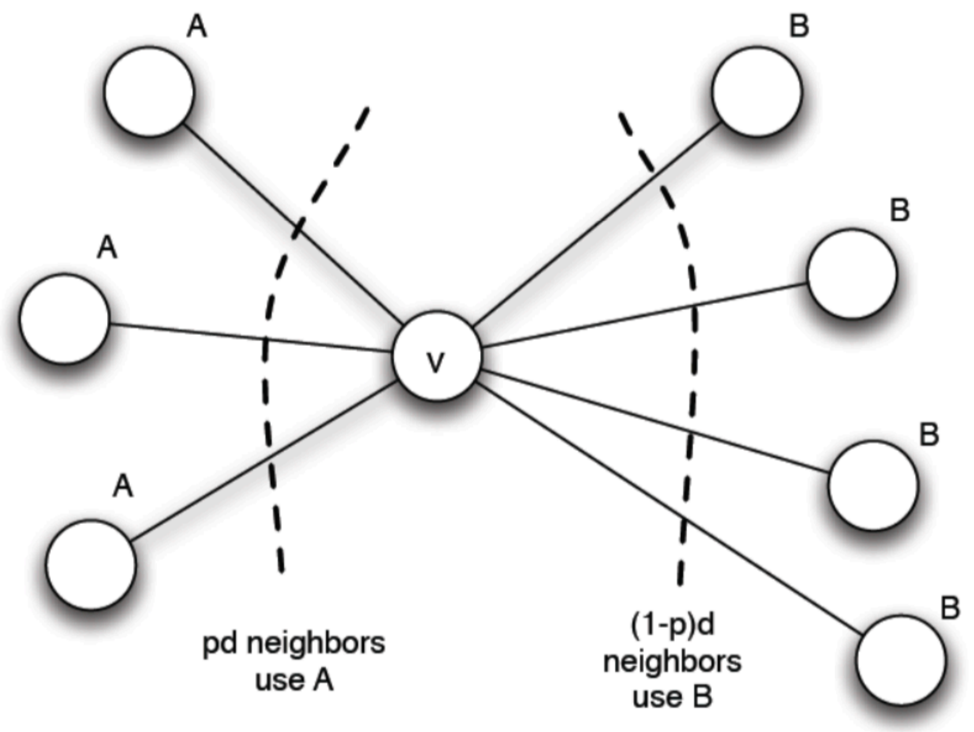
Modeling Epidemic Diffusion

Biggest difference: model transmission as **random**

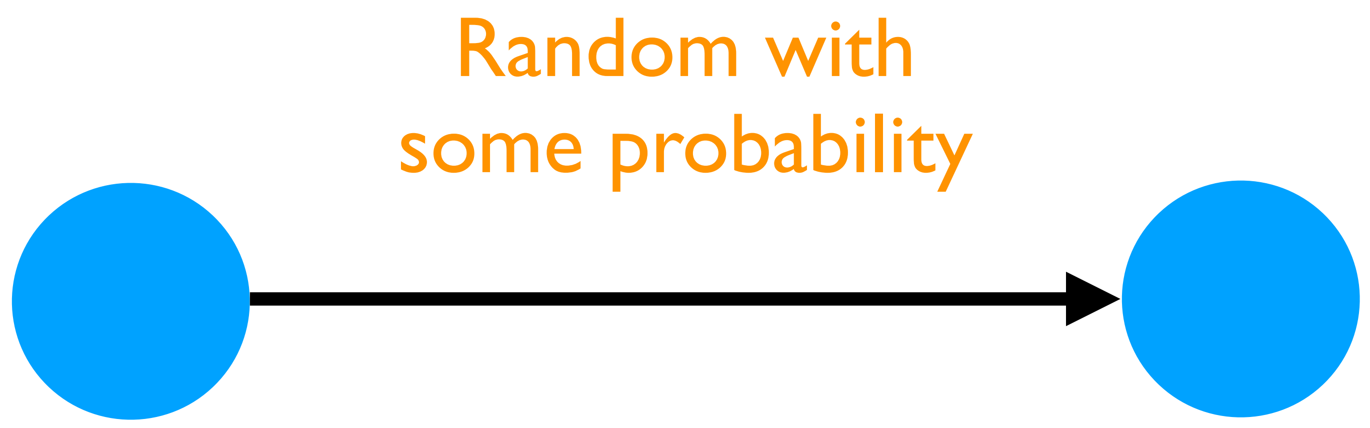
No decision-making, but also the processes by which diseases spread from one person to another are so complex and unobservable at the individual level that it's most useful to think of them as random

Use randomness to abstract away difficult biological questions about the mechanics of spread

Behaviour (last class):



Epidemics (today):



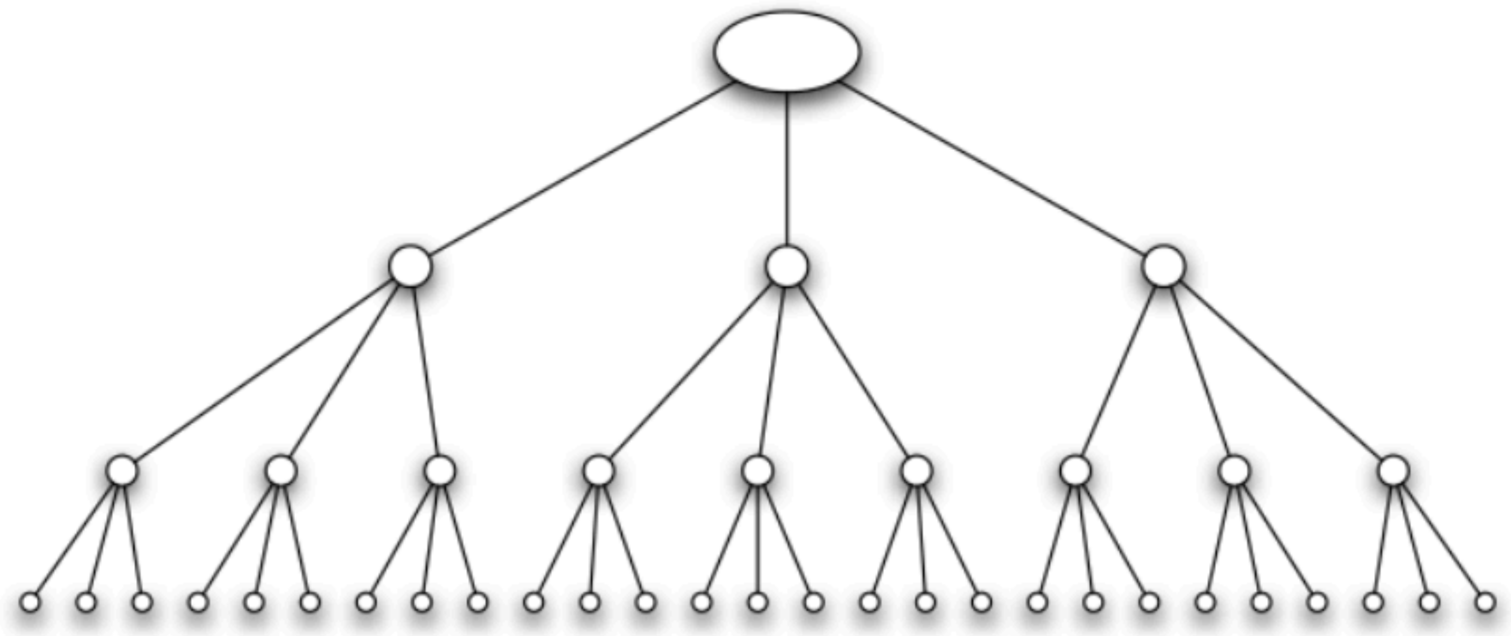
Branching Process

Model as a random process on a tree:

Wave 1: First person infected, infects each of k neighbors with independent probability p

Wave 2: For each infected person, they infect each of k neighbors with independent probability p

Wave 3+: repeat for each infected person



Here k=3

Extends infinitely below

Branching Process: R_0

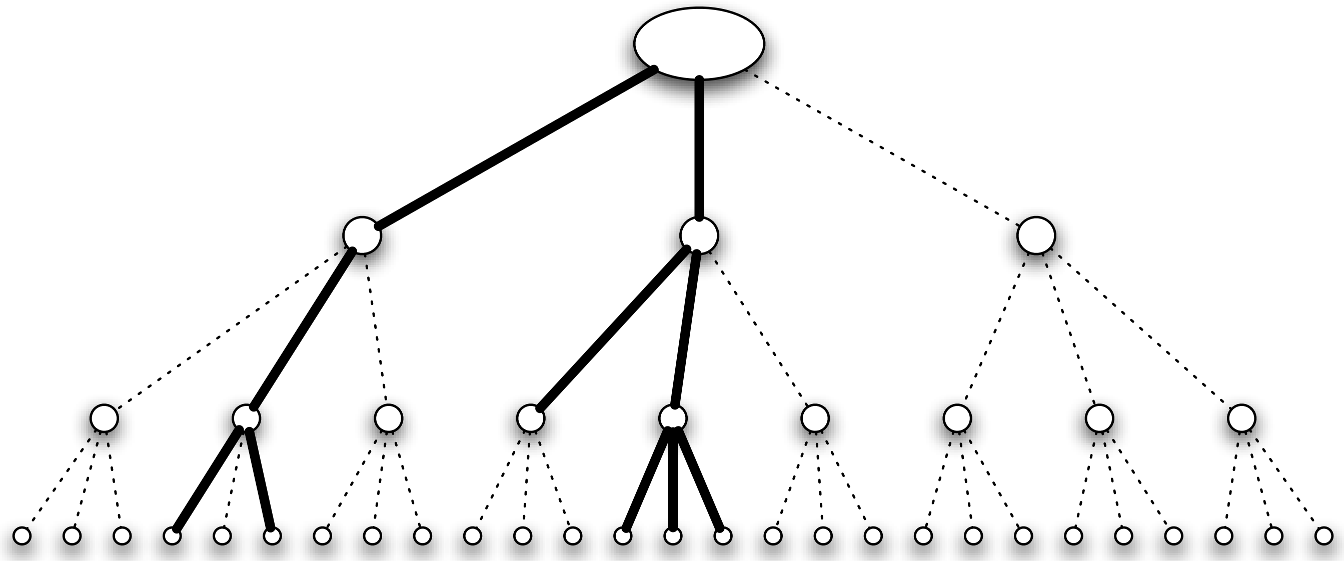
Only two possibilities in the long run: **blow up** or **die out**

How does it die out?

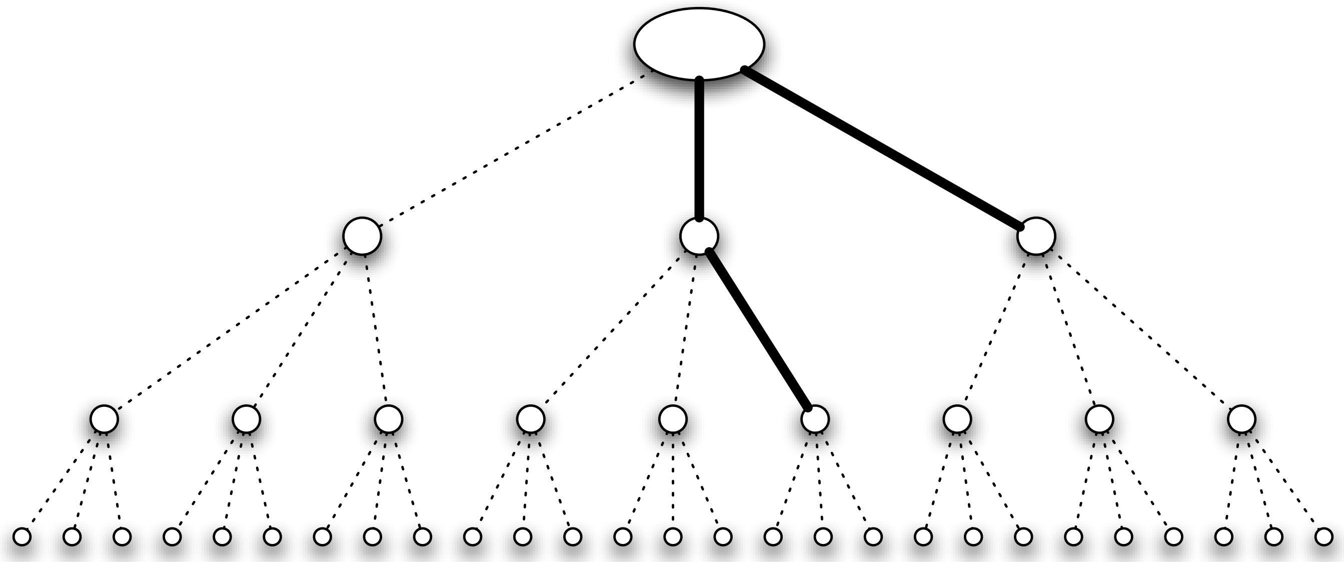
- Dies out if and only if none of the nodes on a given level are infected

Define **Basic reproductive number R_0:**
the number of expected new cases caused by an individual

$$R_0 = pk$$



(b) *With high contagion probability, the infection spreads widely*



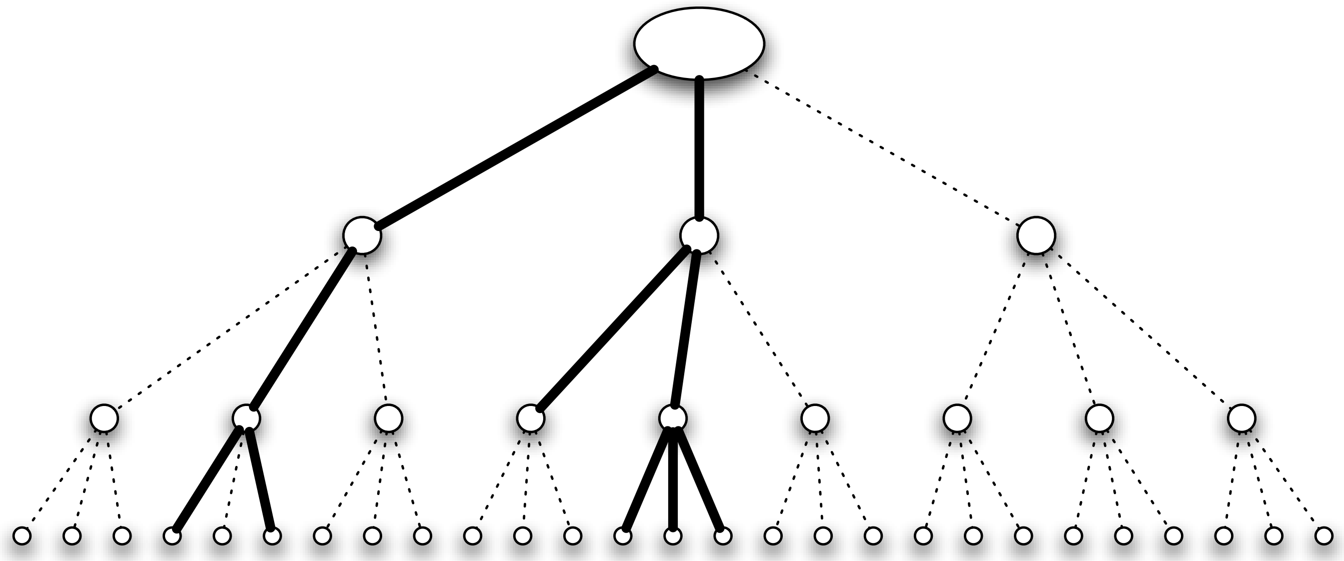
(c) *With low contagion probability, the infection is likely to die out quickly*

Branching Process: R_0

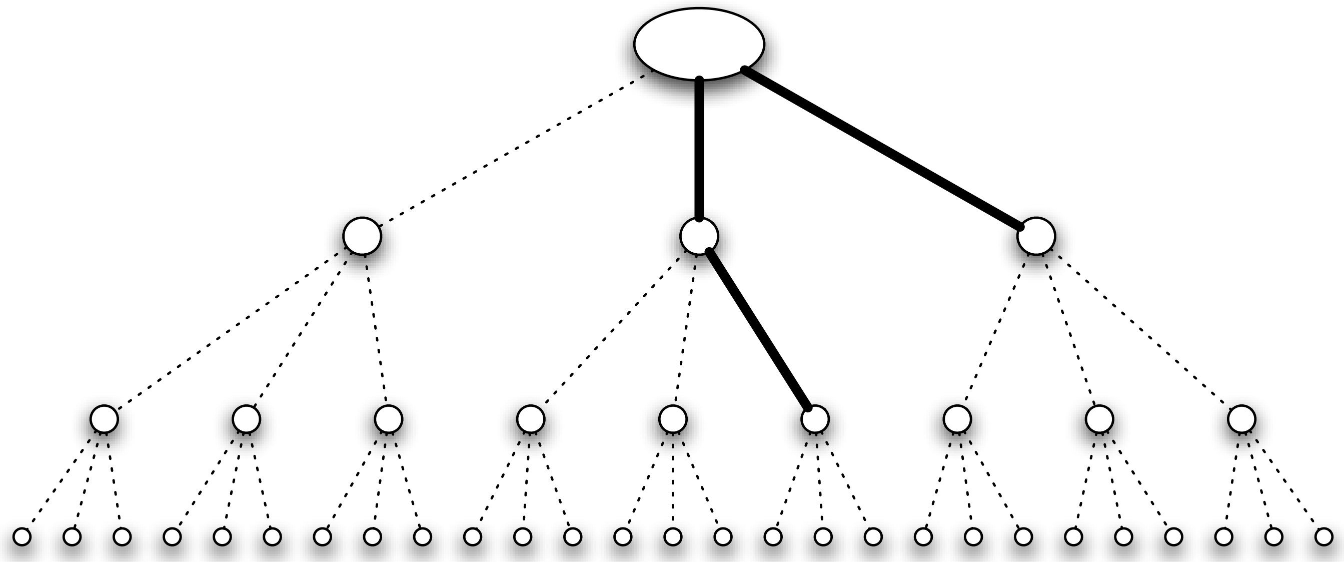
Claim: Epidemic spread in the branching process model is entirely controlled by the reproductive number R_0 :

- If $\mathbf{R_0 < 1}$ then with probability 1 the disease dies out after a finite number of steps.
- If $\mathbf{R_0 > 1}$ then with probability > 0 the disease persists by infecting at least one person in each wave.

"Go big or go home."



(b) *With high contagion probability, the infection spreads widely*



(c) *With low contagion probability, the infection is likely to die out quickly*

$R_0 = pk$

SIR Epidemic Models

S = Susceptible

I = Infectious: node is infected and infects with prob **p**

R = Removed: after $\mathbf{t_I}$ time, no longer infected or infectious

Initially some nodes in **I** state, rest in **S** state.

Each node in I state remains infected for $\mathbf{t_I}$ time steps

During each step, each node has probability **p** of infecting each susceptible neighbour

After $\mathbf{t_I}$ time steps, no longer **S** nor **I;** removed to **R**

Now: SIS Epidemic Model

S = Susceptible

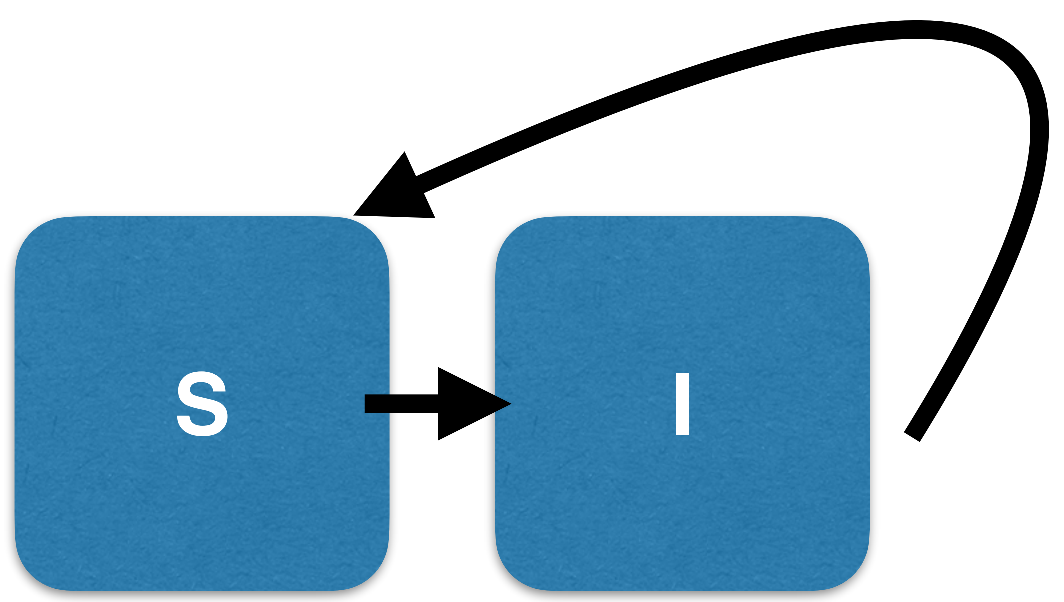
I = Infectious: node is infected and infects with prob **p**

Initially some nodes in **I** state, rest in **S** state.

Each node in I state remains infected for $\mathbf{t_I}$ time steps

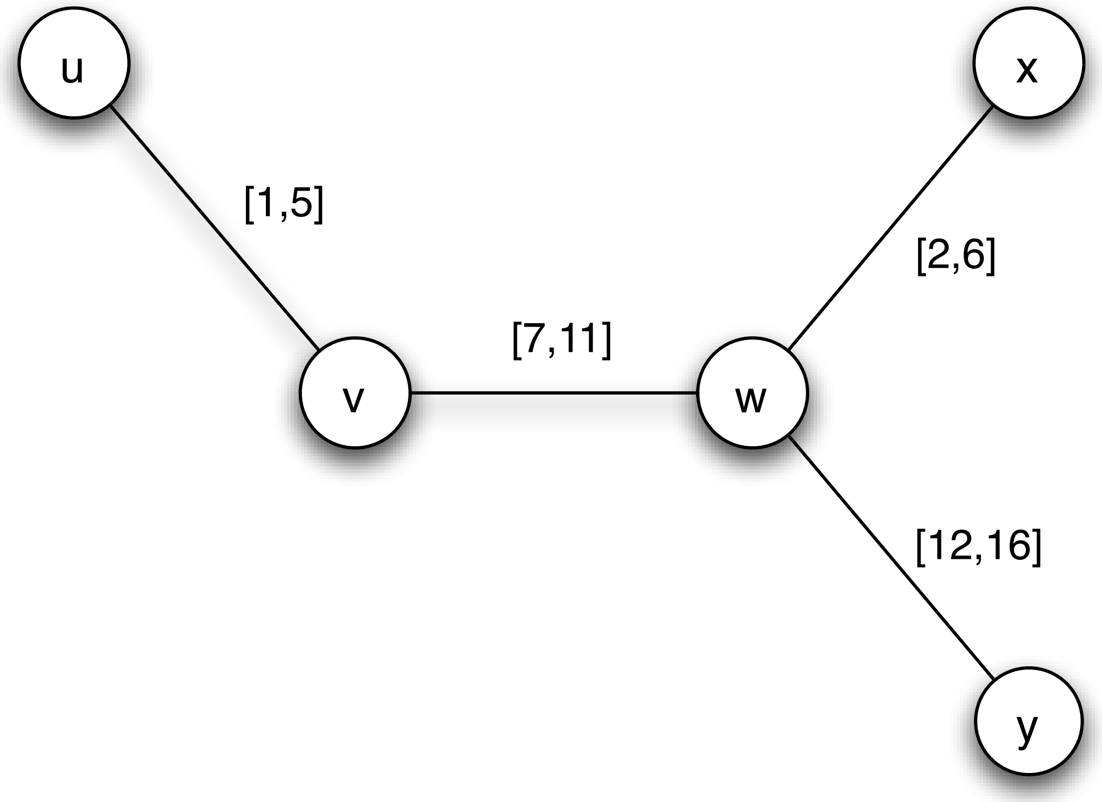
During each step, each node has probability **p** of infecting all neighbors

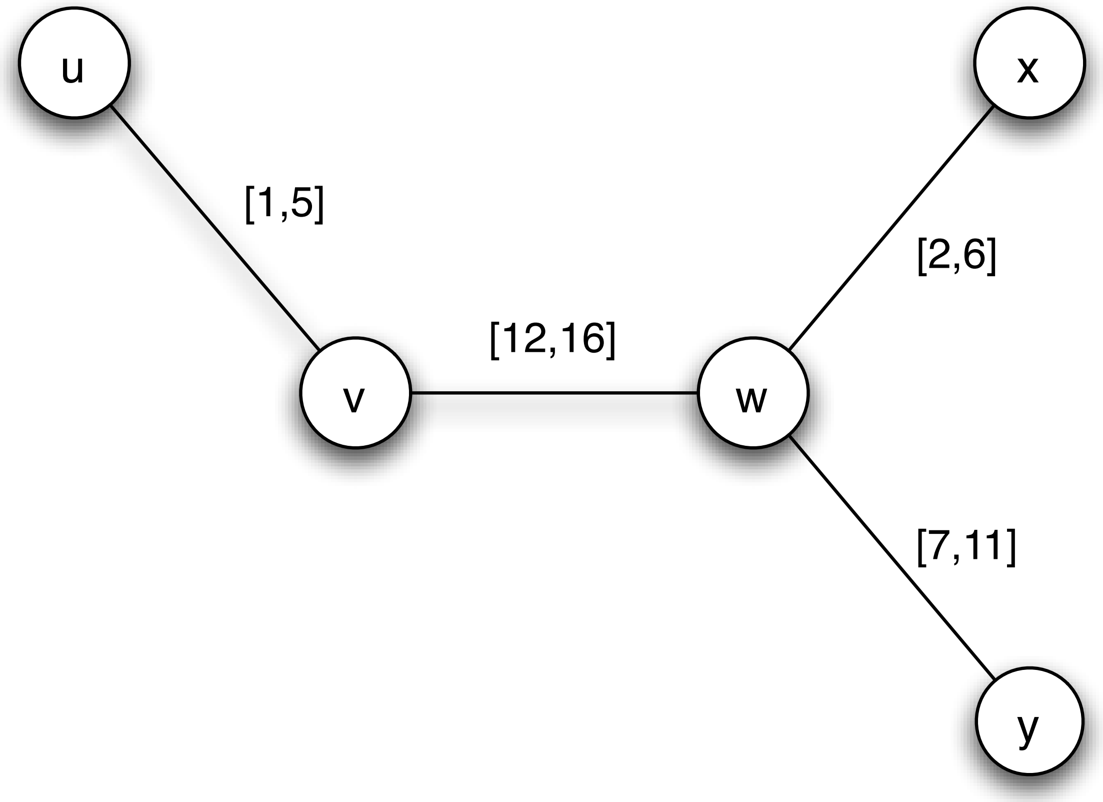
After $\mathbf{t_I}$ time steps, node **returns to S**



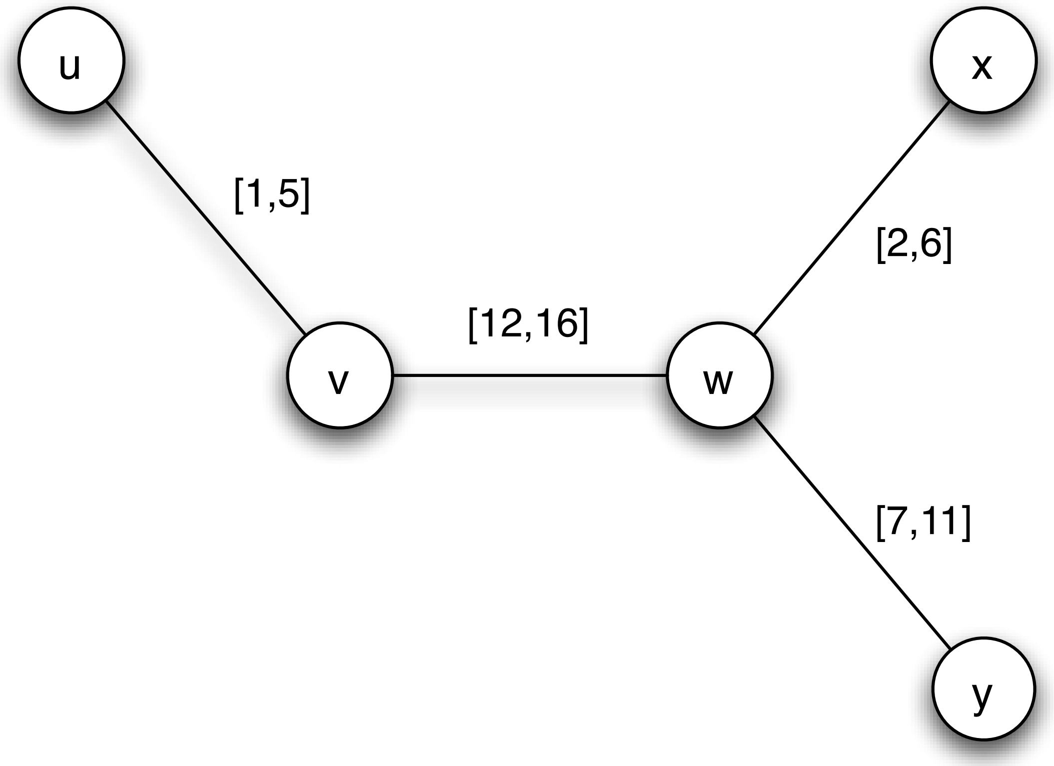
Transient Contacts & Concurrency

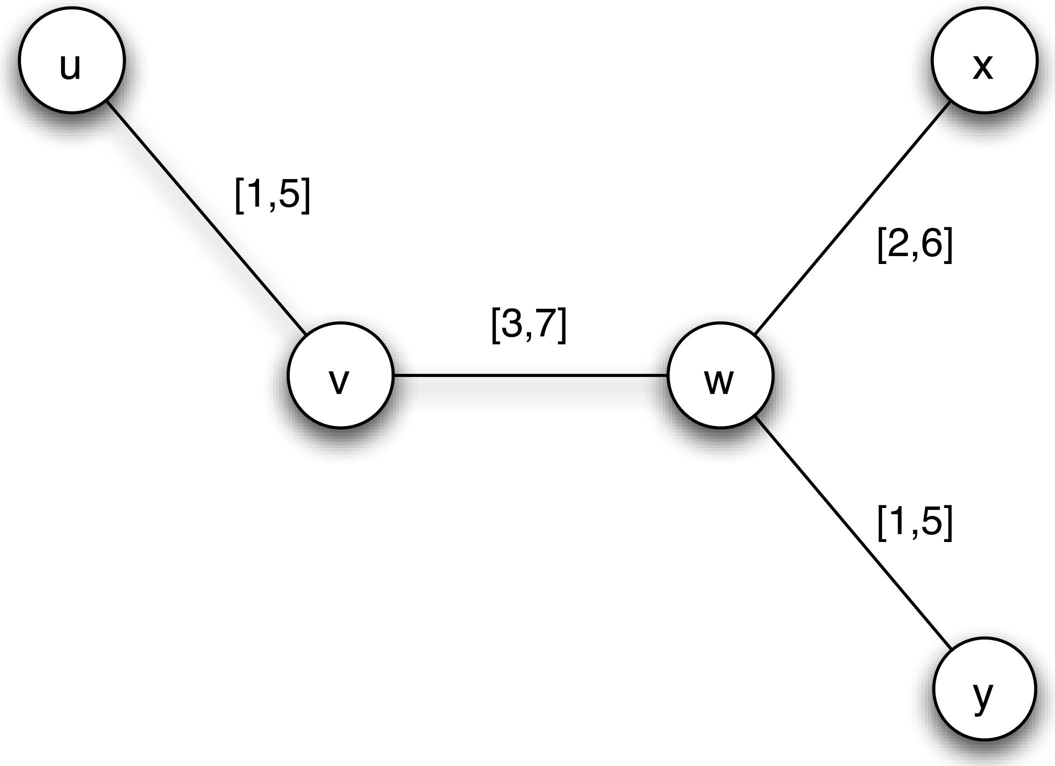
A less random model: it matters in what order contact is made in the contact network.





Concurrency: having two or more contacts at once.



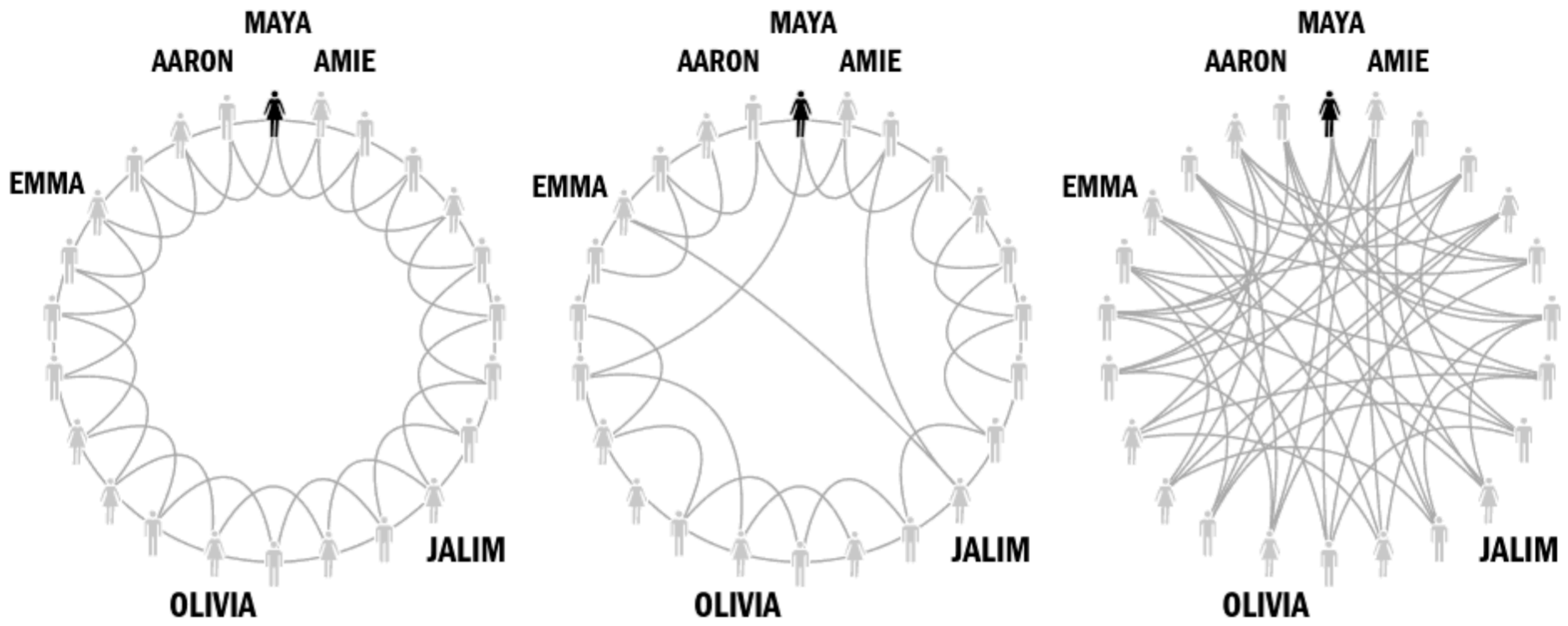


Epidemics vs. Behaviour

Simple vs. complex diffusion
Epidemics vs. behaviour

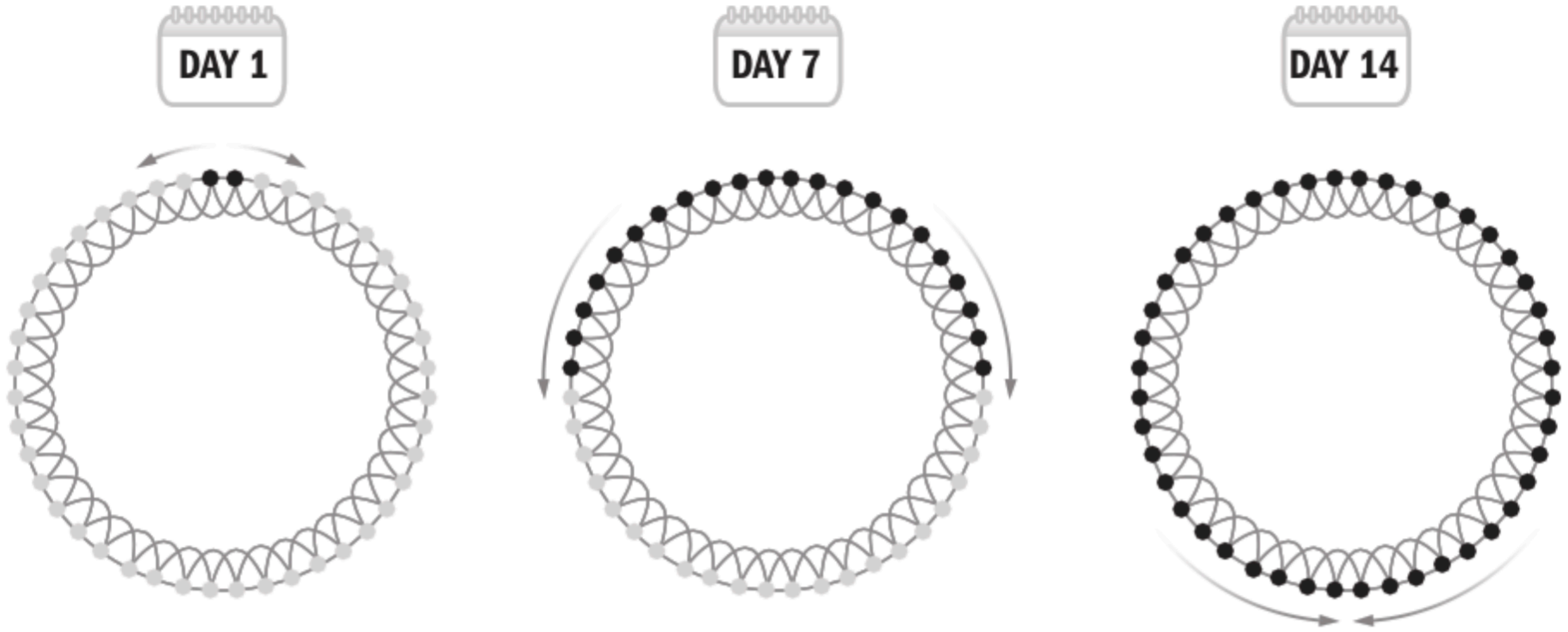
What's the difference?

Recall the small-world model



Simple Diffusion

Large world:



Small world: