Social and Information Networks

CSCC46H, Fall 2025
Lecture 11

Prof. Ashton Anderson
ashton@cs.toronto.edu

Today

A3 due next Tuesday, Dec 2 @ 11:59pm, plus no-penalty extension to next Friday, Dec 5 @ 11:59pm (late days not needed)

Today

Epidemics and Contagion

Voting

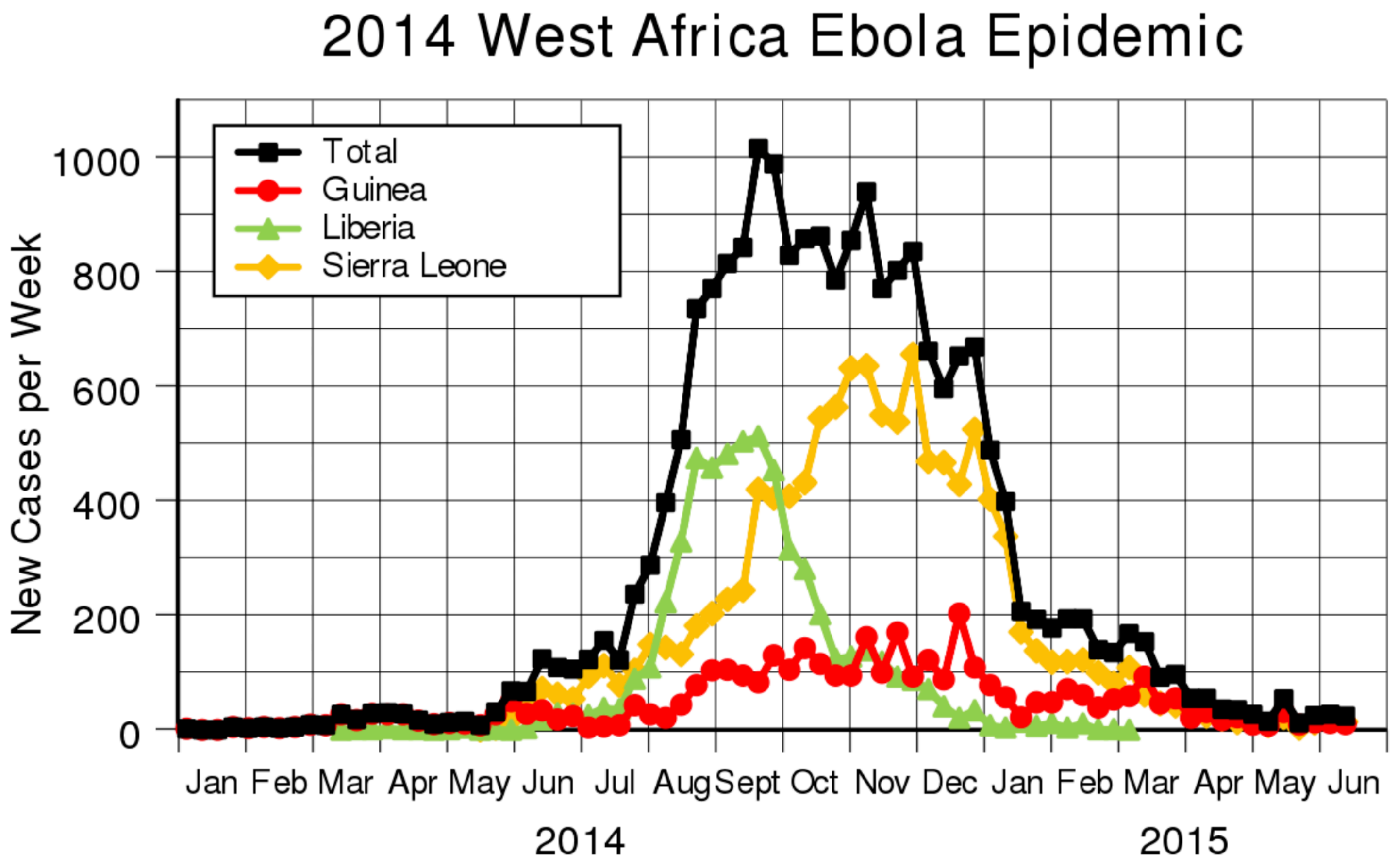
Epidemics

Epidemics

Why study epidemics in a computer science class?

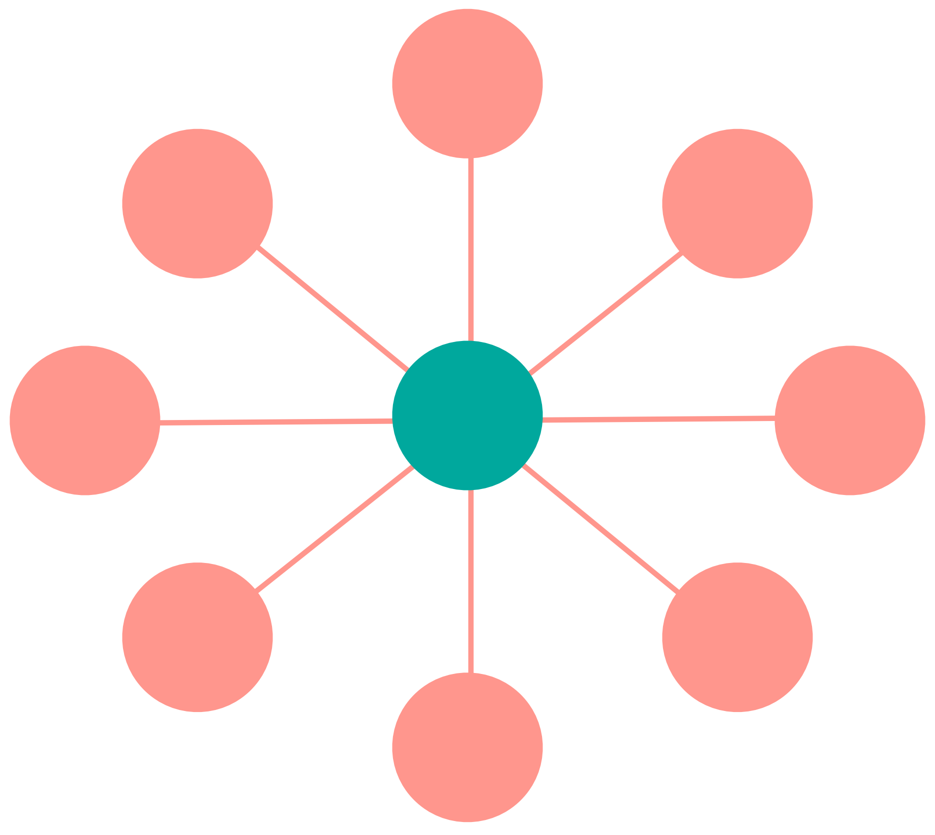
Epidemics are diseases that **travel socially**

The structure of social interaction networks determine the spread of disease

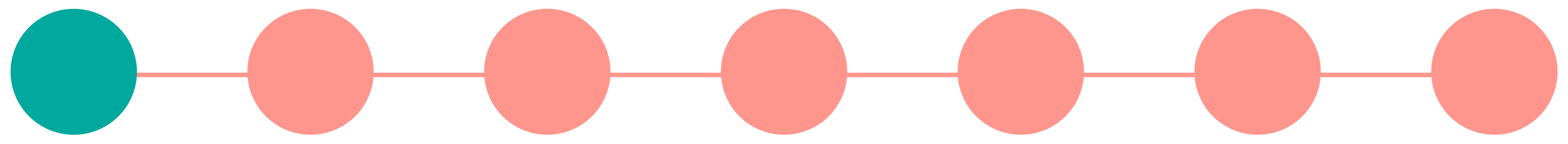


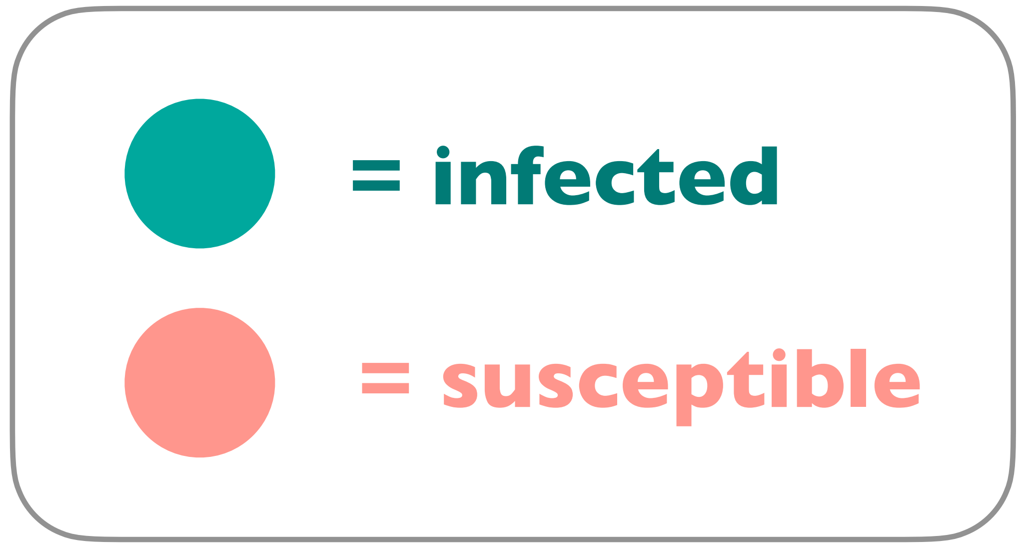
Epidemics

Which outbreak is more dangerous to the population?



vs.

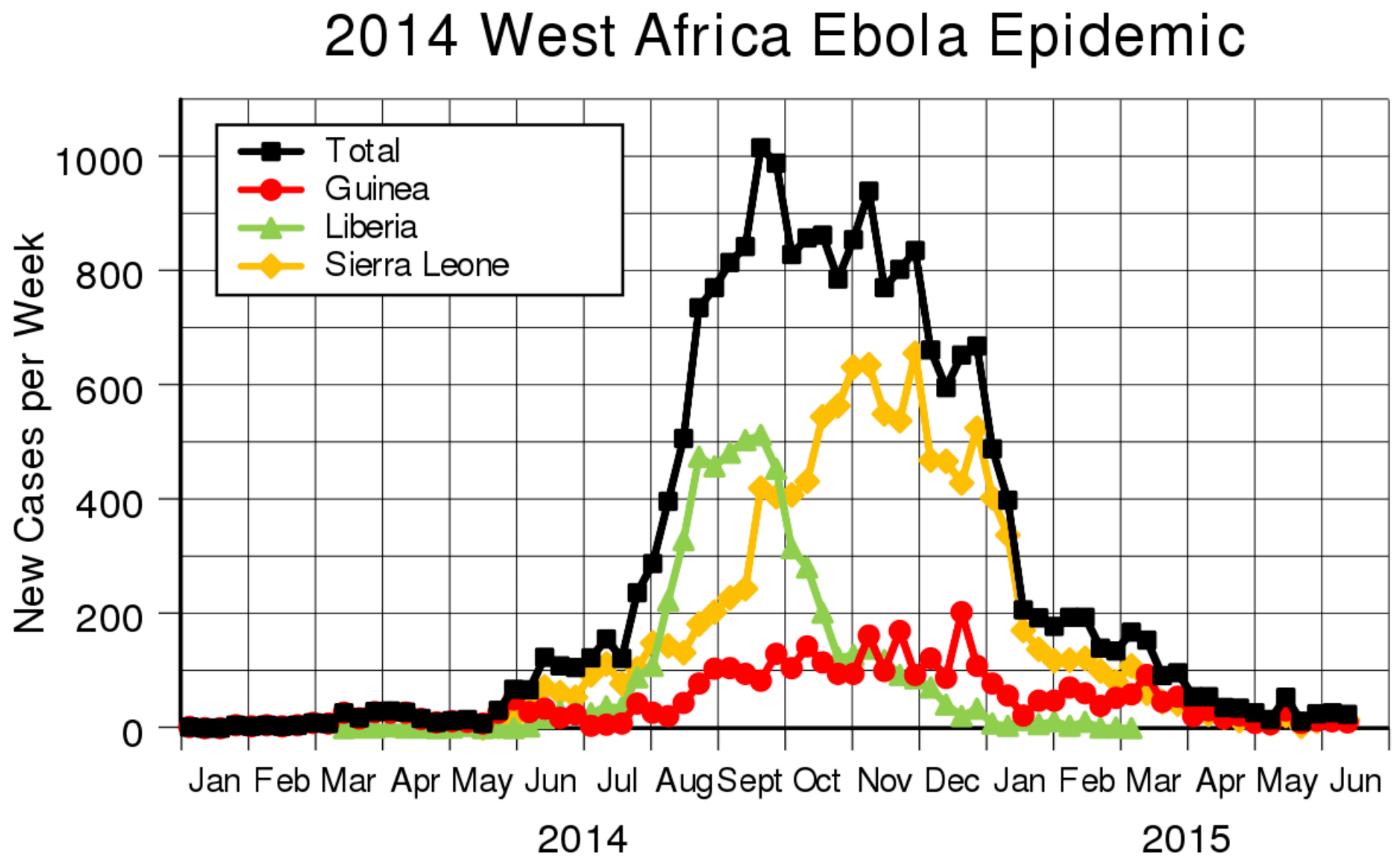




Epidemics

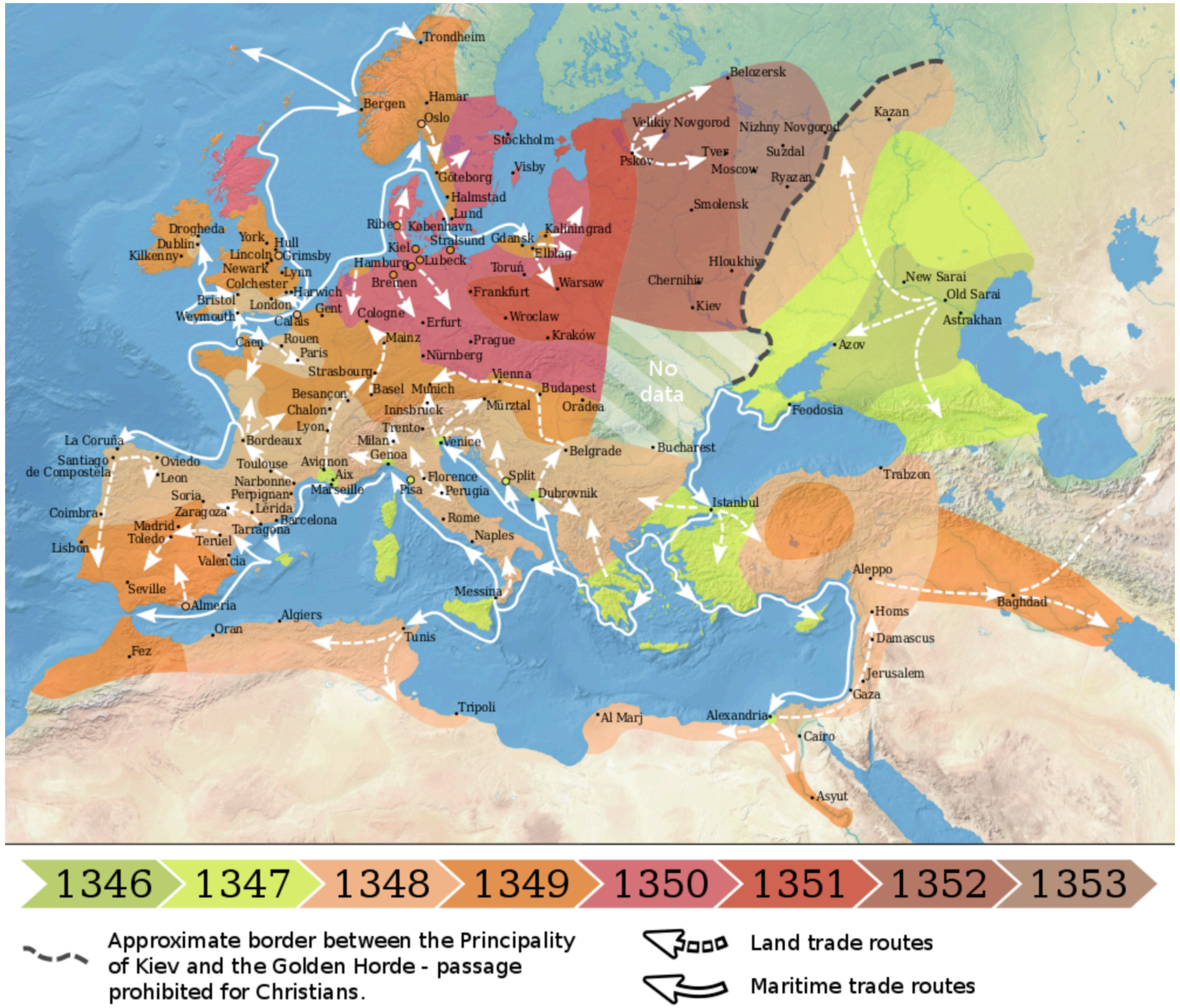
Types of epidemic diffusion:

- **Explosive spread** through a population
- **"Slow burn"** persistence over long periods of time
- Wave-like **cyclical patterns**



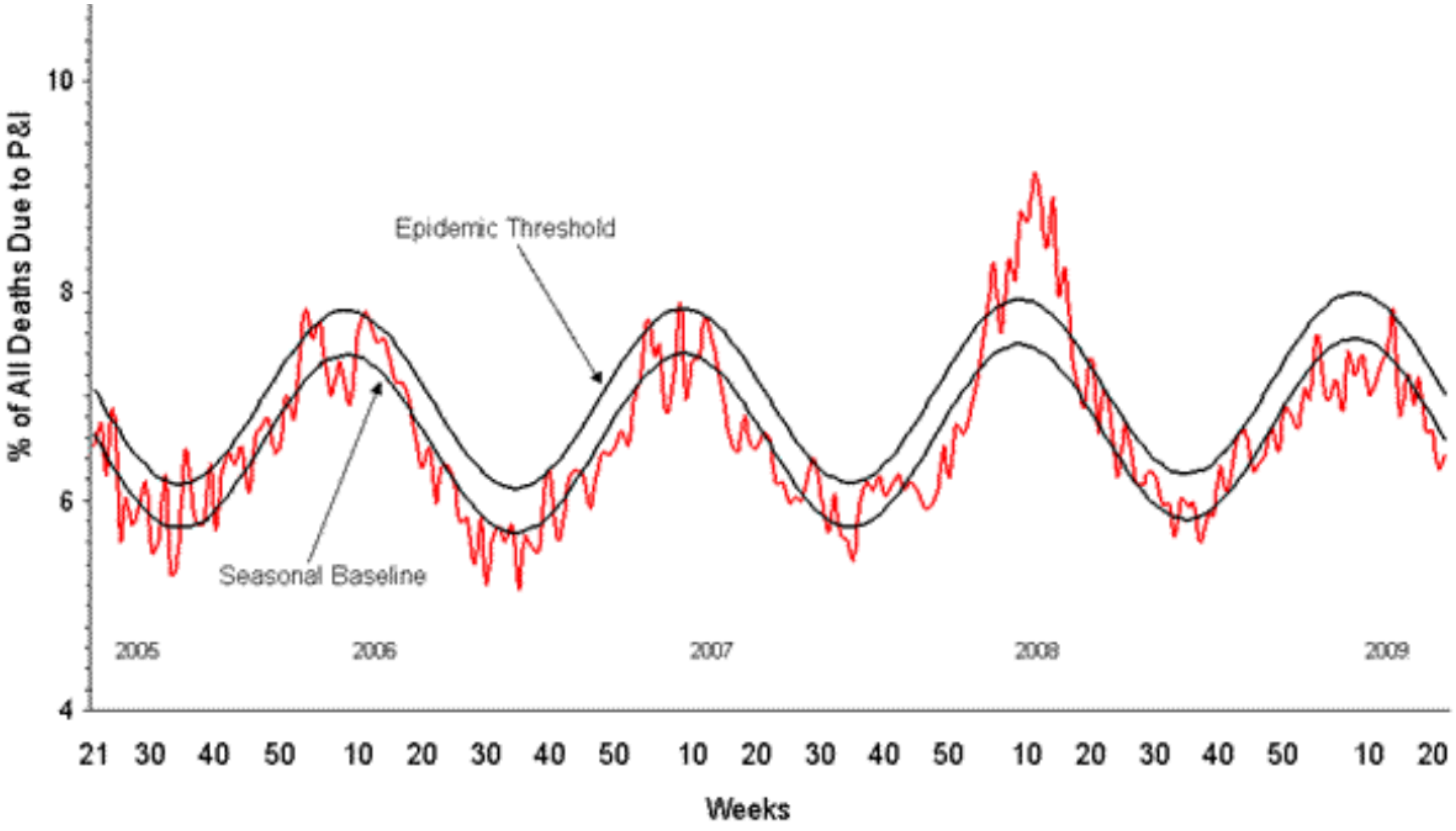
Epidemics

Explosive spread: Bubonic Plague (the "Black Death"): wiped out ~50% of the population in Europe (~150 million people) in 7 years



Epidemics

Other epidemics are cyclical

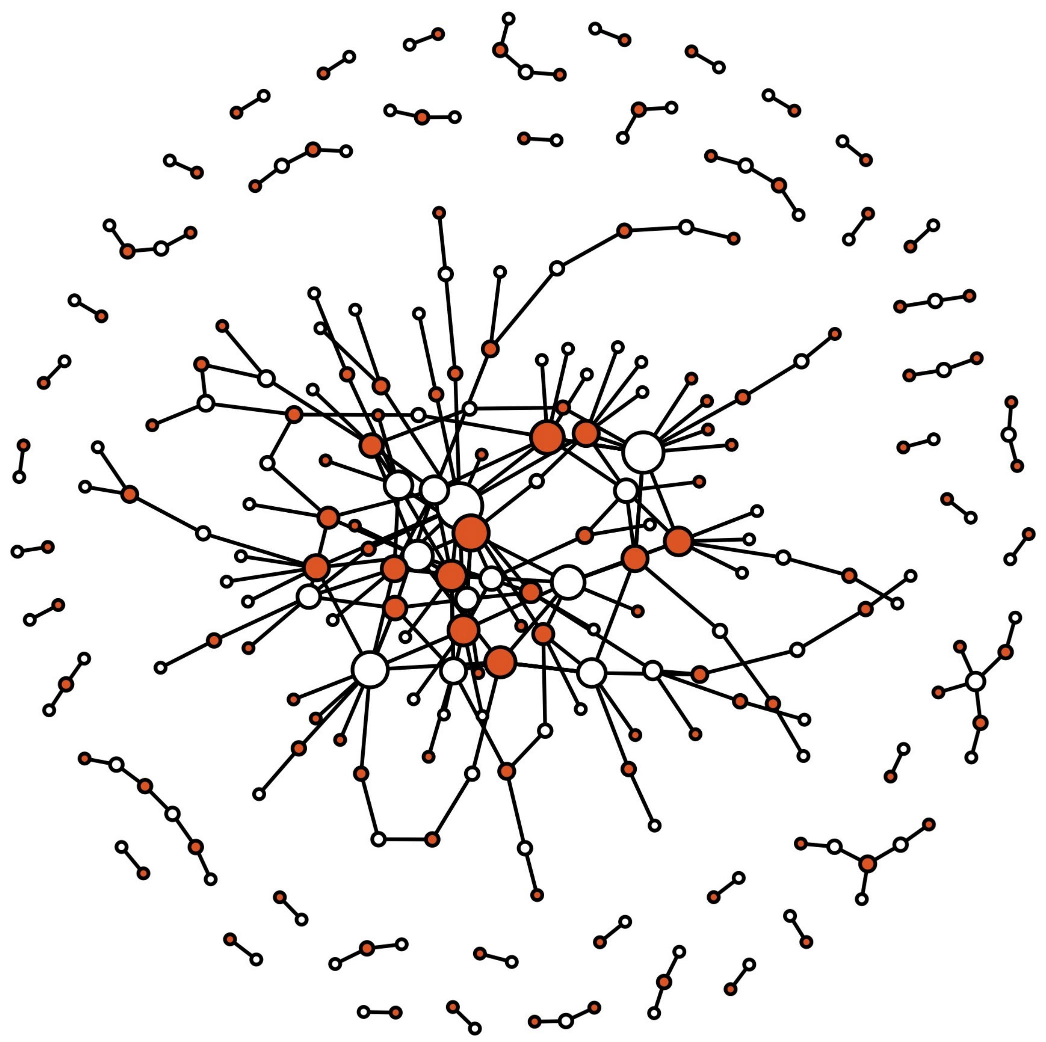


Epidemics

What determines how an epidemic might spread?

- Properties of the disease
- Structure of the network

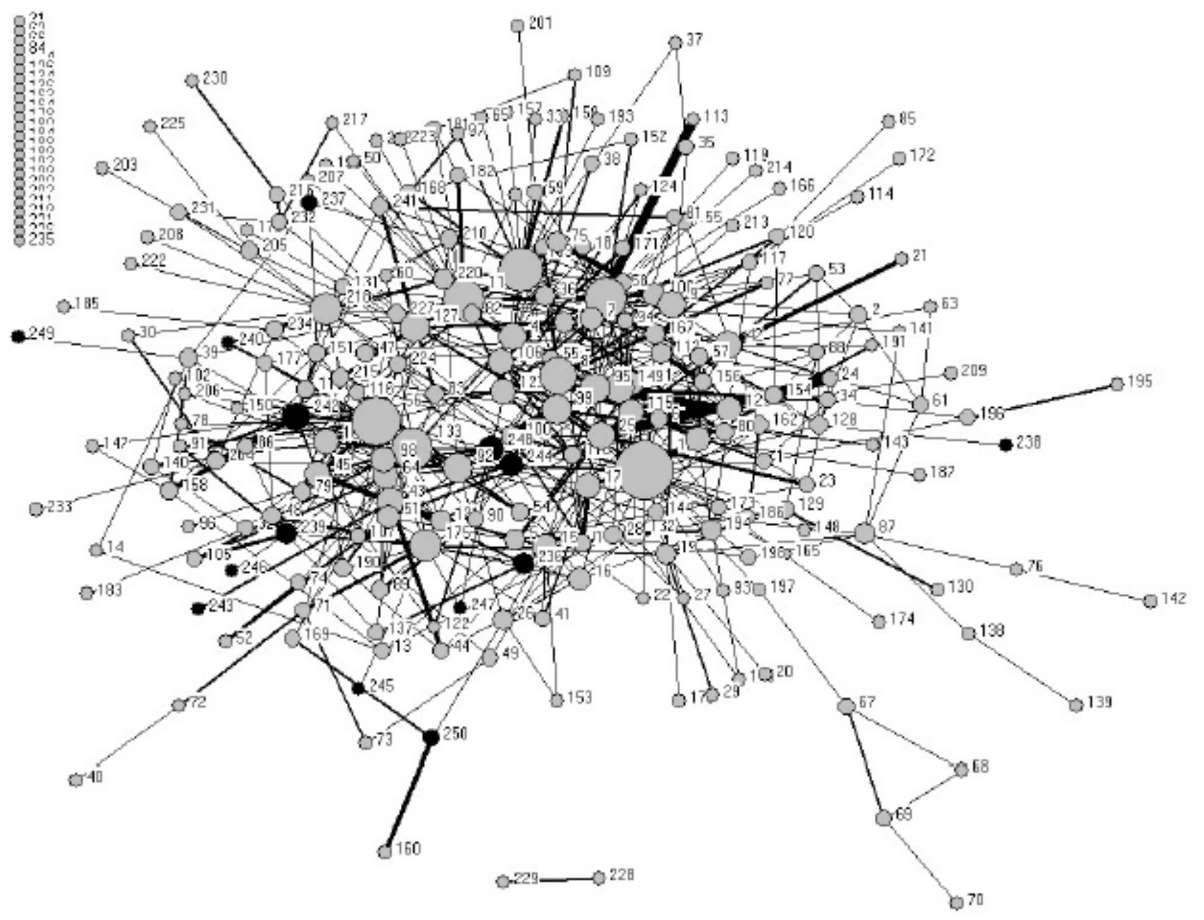
What network?



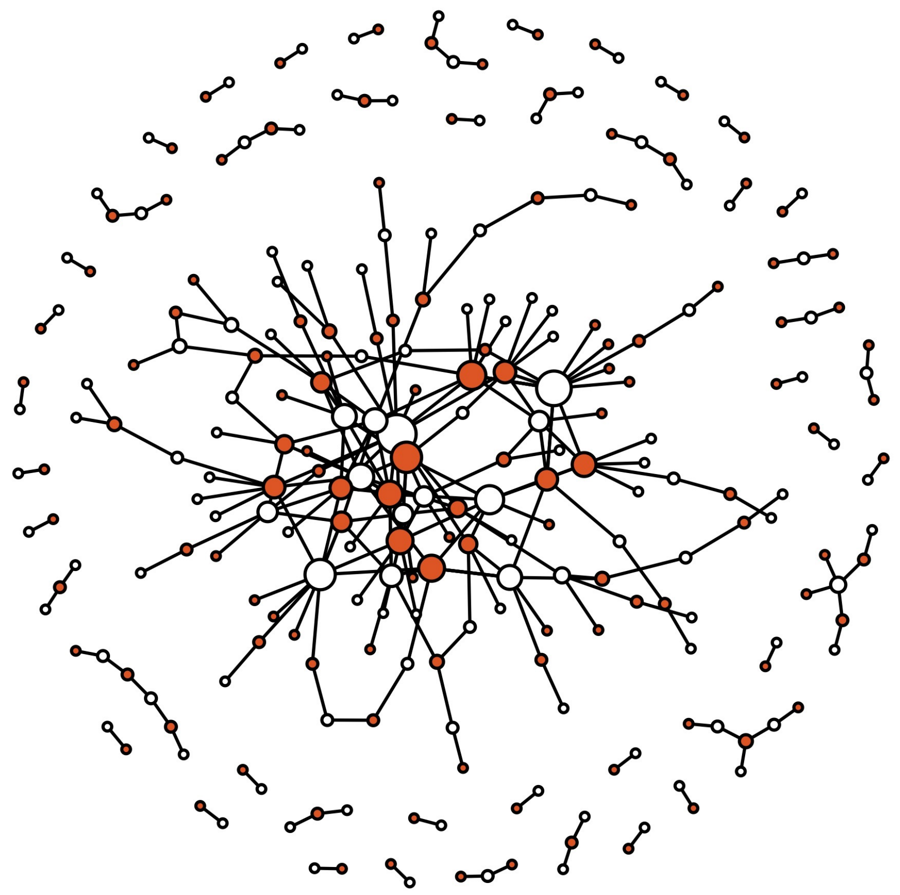
Contact Networks

Node for each person

Edge if two people come into contact with each other in a way that makes it possible for a disease to spread



UK fish farm exchanges



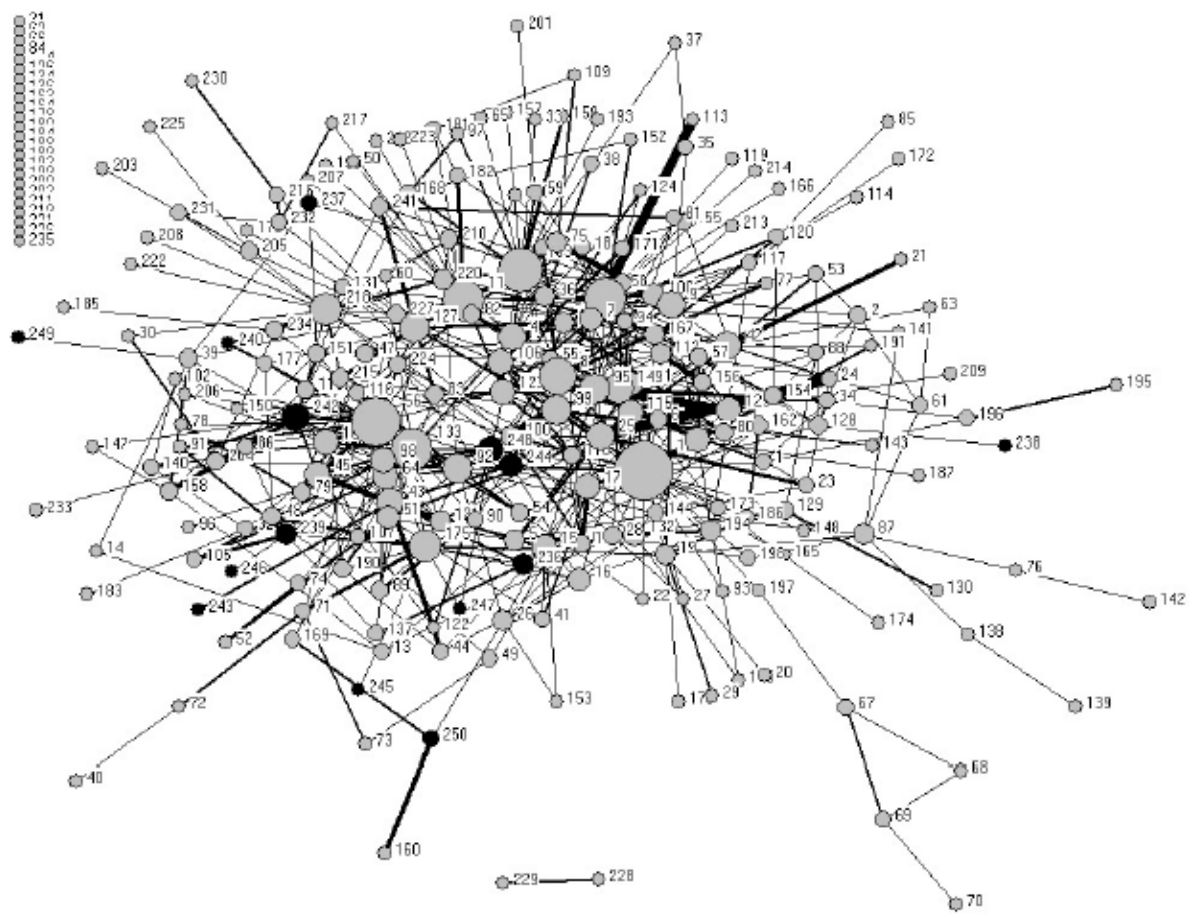
CH sexual contact network

Contact Networks

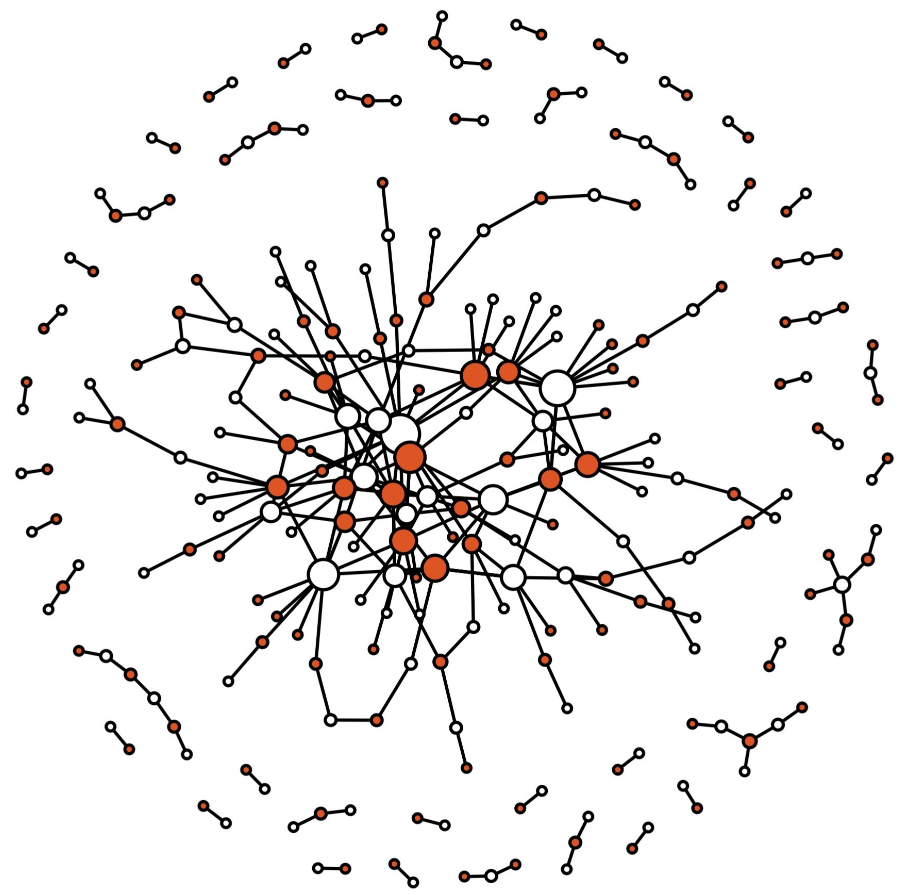
Once you've got through the laborious process of mapping out the contact network, can you use it to study any disease?

No! Definition of "contact" depends on the disease

- Airborne transmission: edge between everyone who was in the same car, etc.)
  — many edges
- Close contact / sexual transmission: sparser graph



UK fish farm exchanges

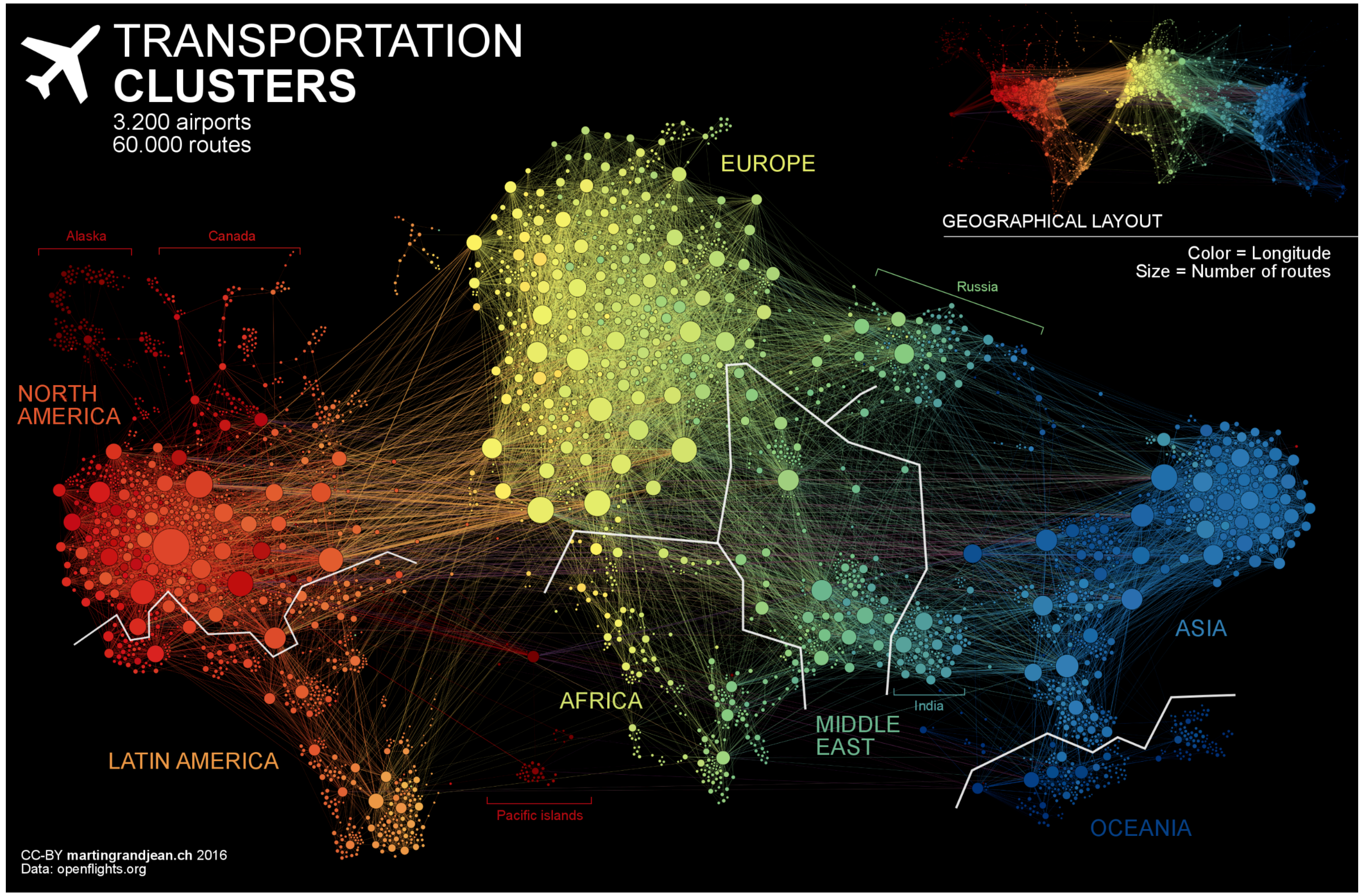


CH sexual contact network

Contact Networks

Big part of real-world epidemic research is constructing contact networks

Lots of work on travel patterns in cities, the worldwide airline network, etc. to understand how diseases can spread in today's world

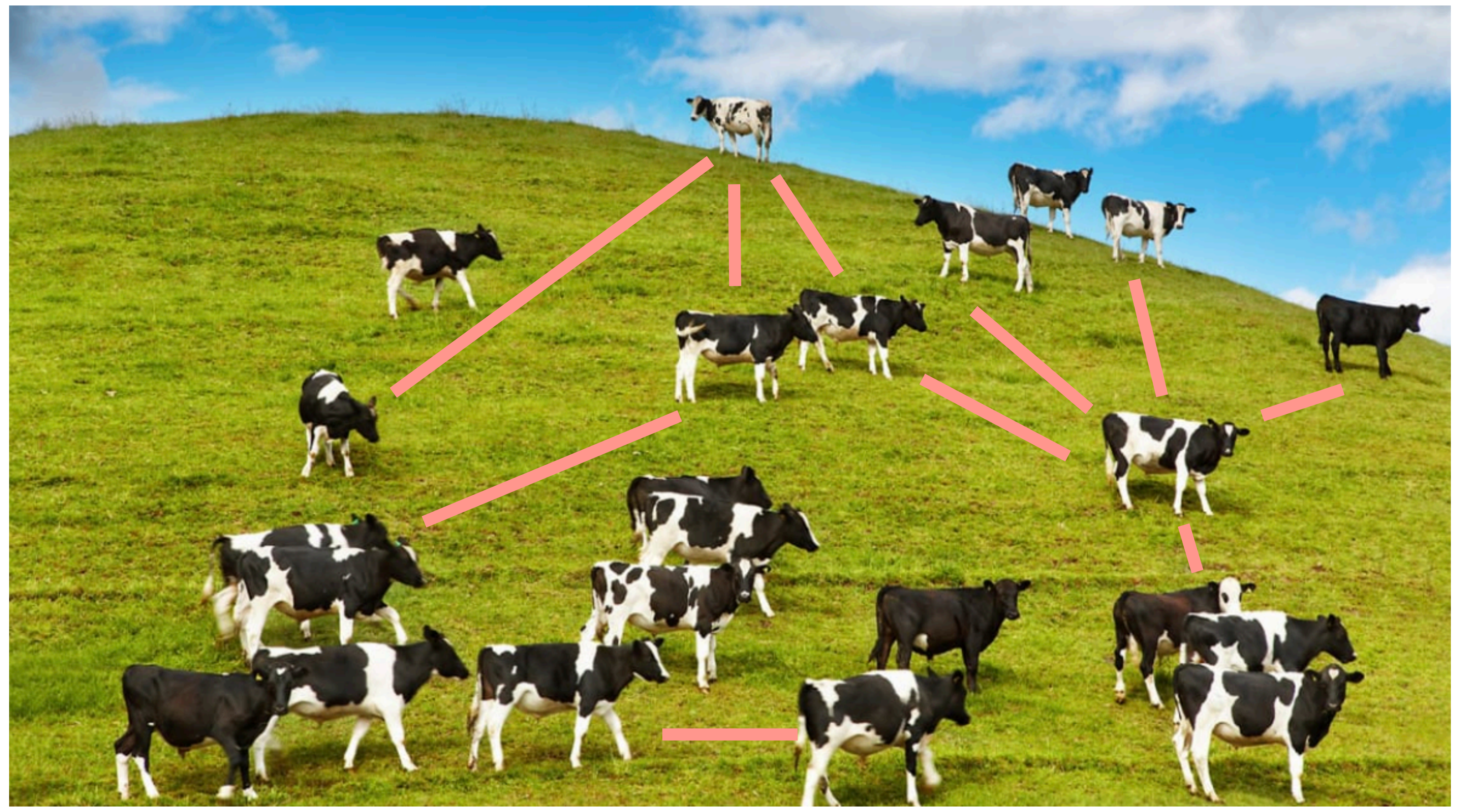


http://www.martingrandjean.ch/connected-world-air-traffic-network/

Contact Networks

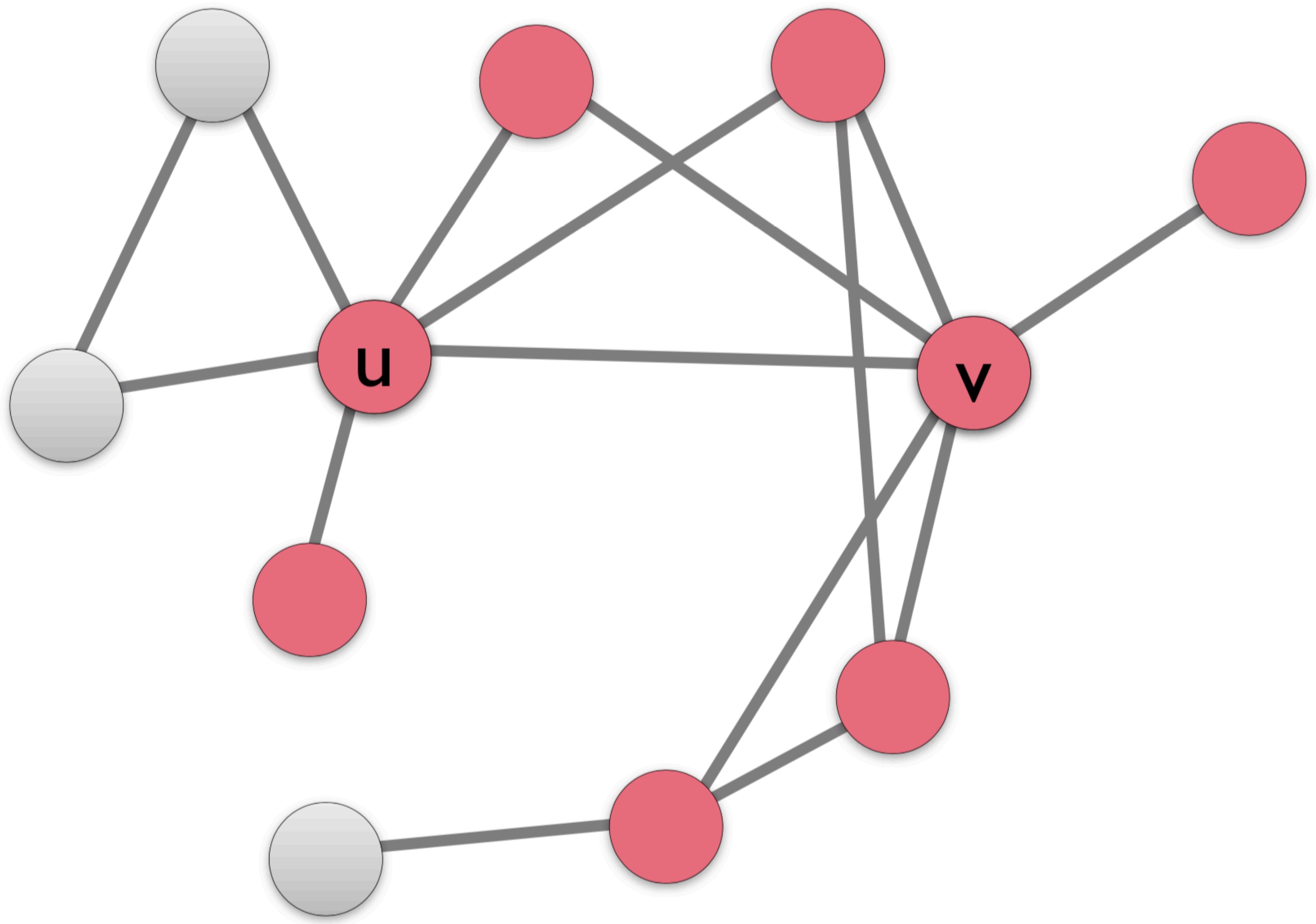
Not just human contact networks

Animal/livestock networks and plant networks

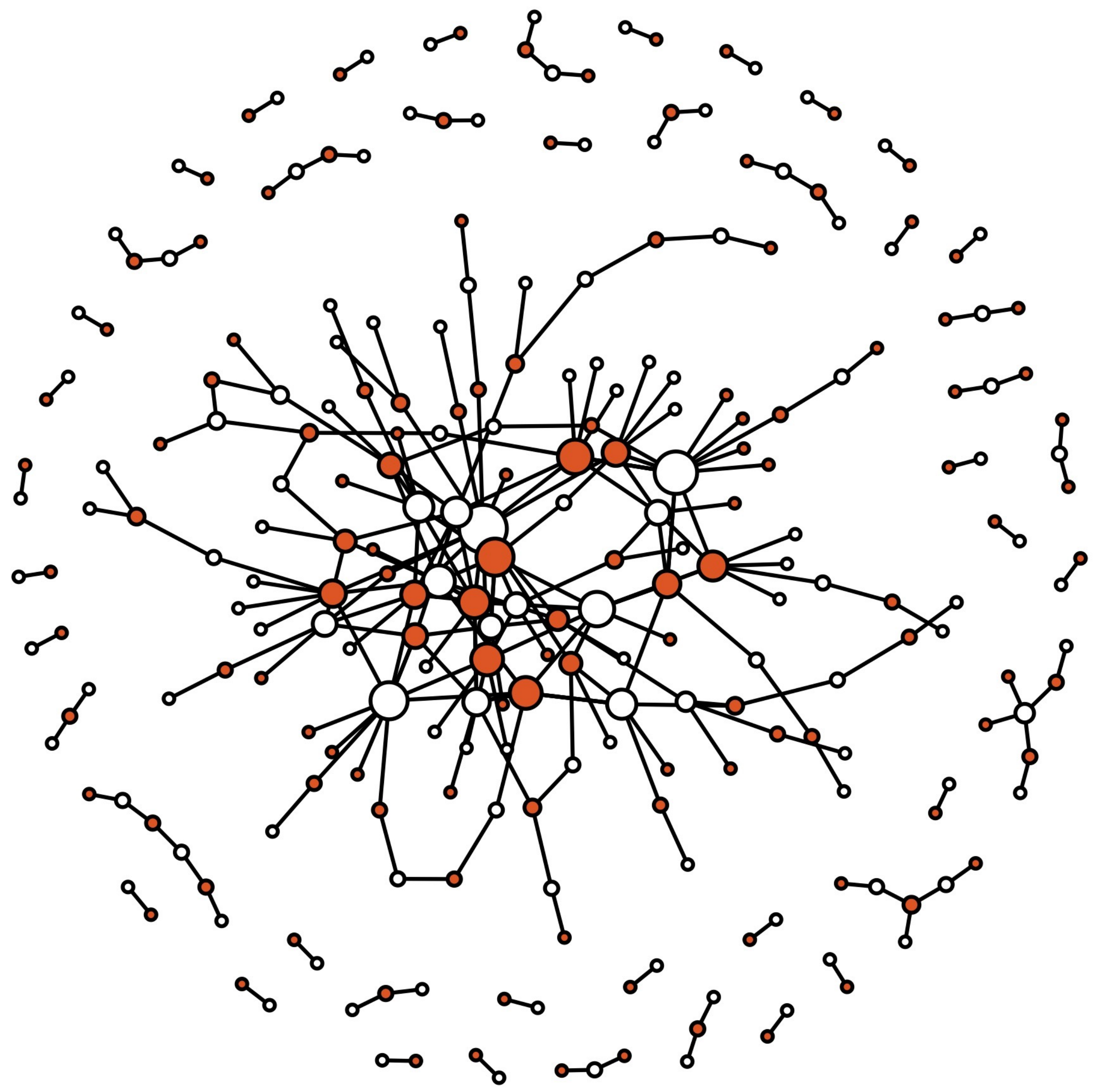


Behavioural vs. Biological Contagion

Biological/epidemic diffusion: **no decision-making!**



Decision cascade



High school contact network

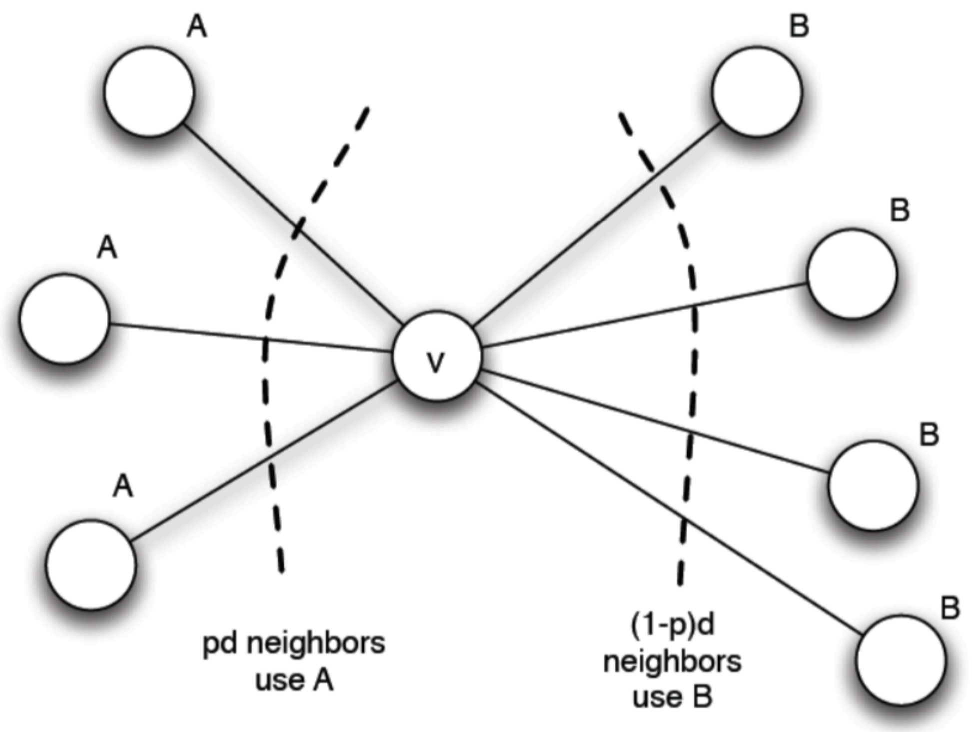
Modeling Epidemic Diffusion

Biggest difference: model transmission as **random**

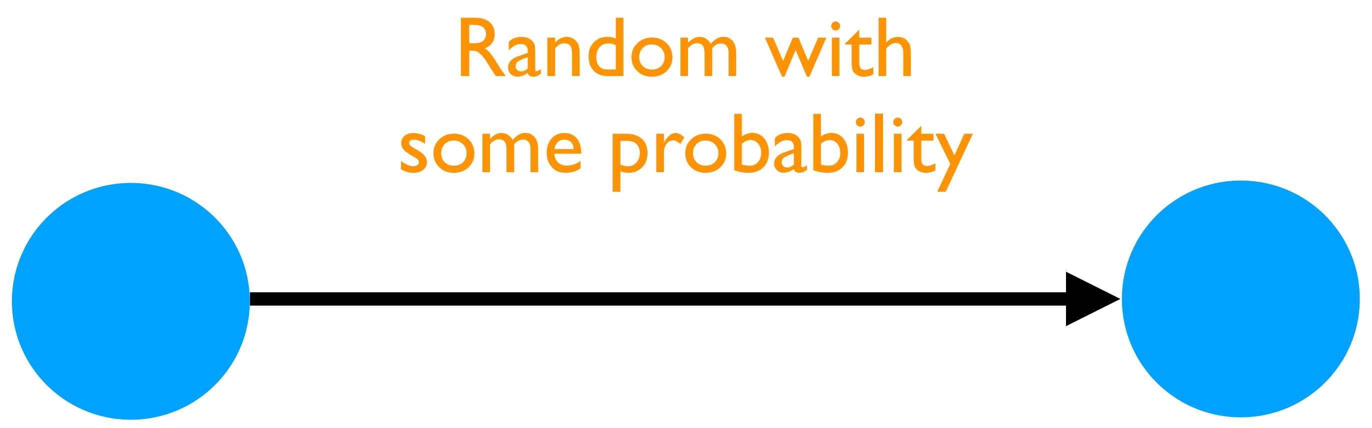
No decision-making, but also the processes by which diseases spread from one person to another are so complex and unobservable at the individual level that it's most useful to think of them as random

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

Behaviour (last class):



Epidemics (today):

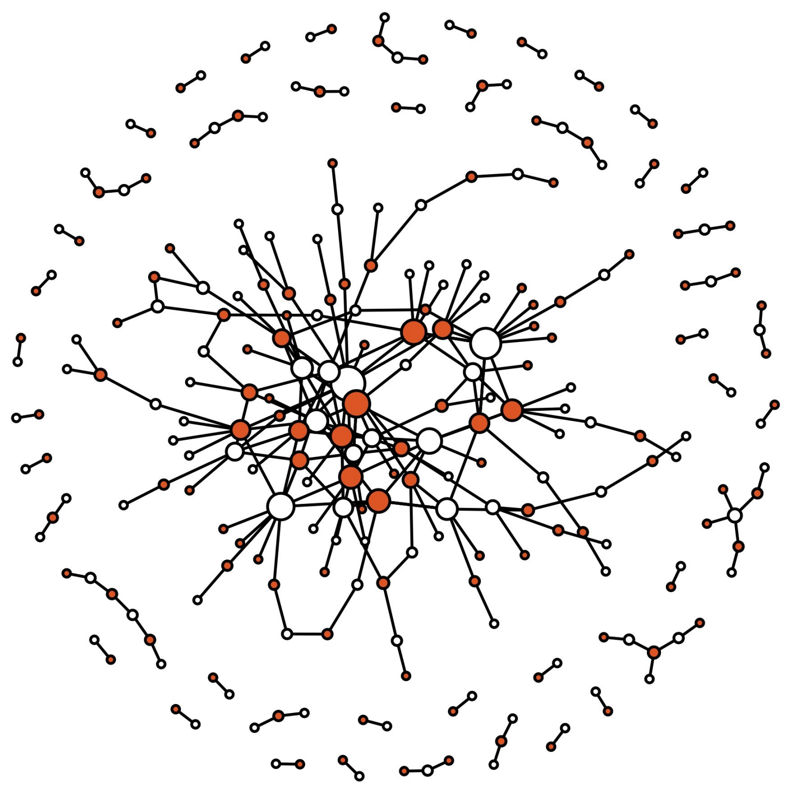


Modeling Epidemic Diffusion

Branching Process

Basic structure of epidemic diffusion:

- Someone gets infected
- Then they infect some number of people
- Those people infect others



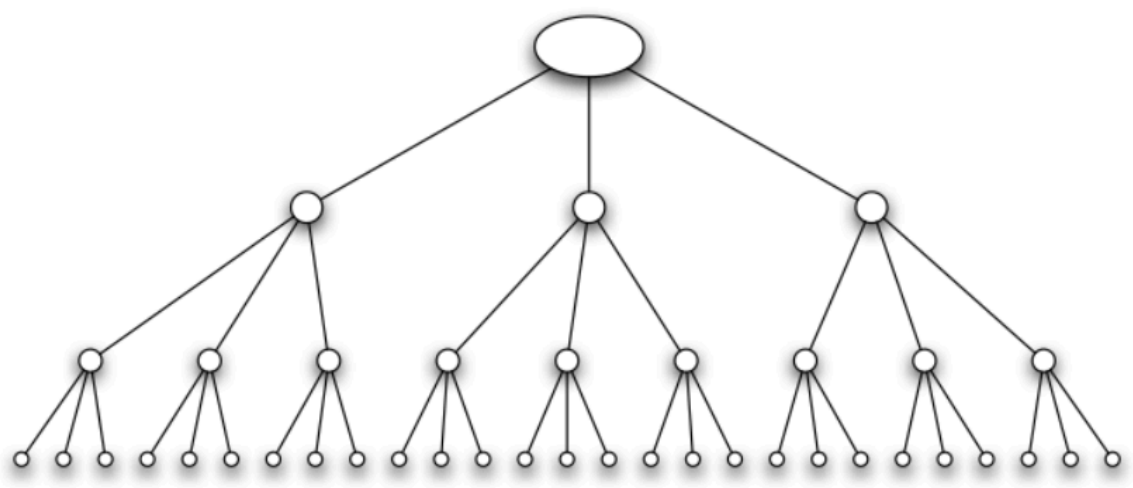
Branching Process

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



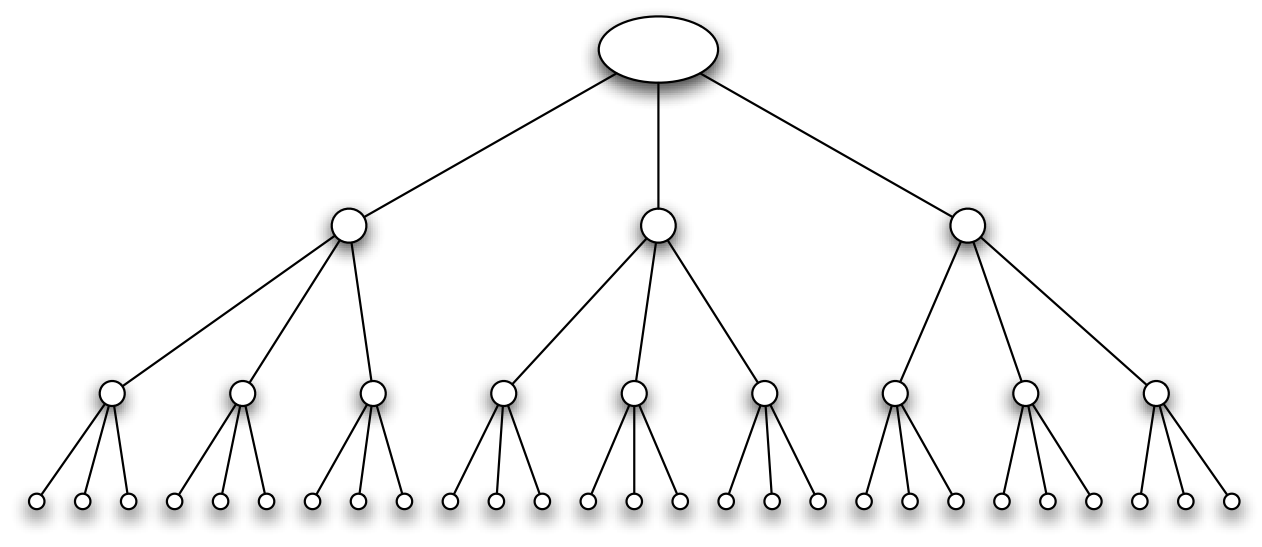
Here k=3

Extends infinitely below

Branching Process

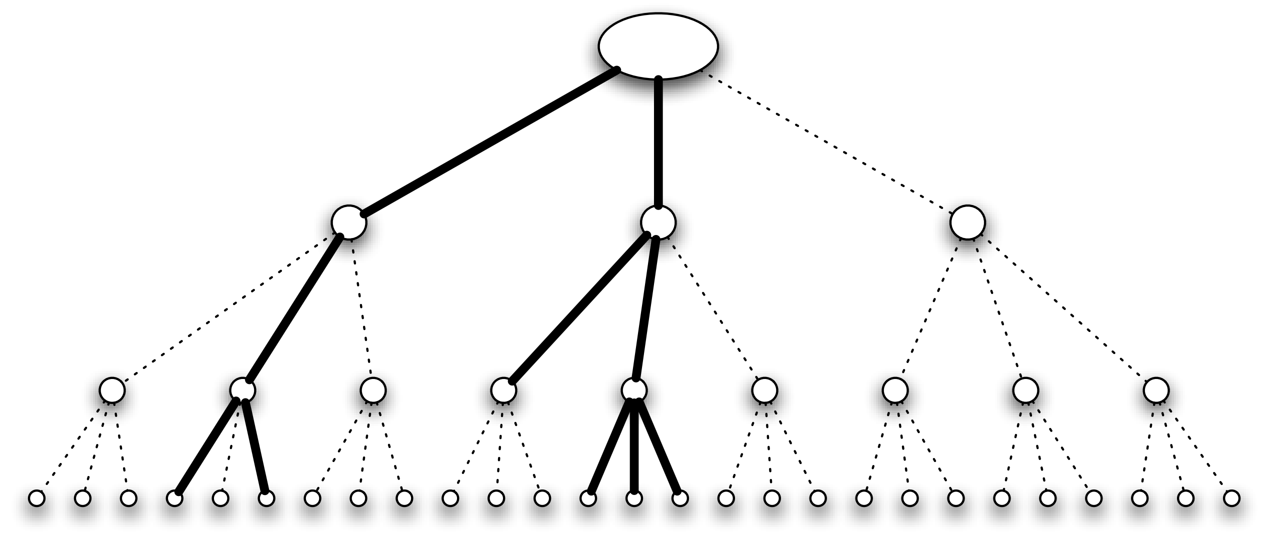
Model parameters:

k: number of individuals each person can possibly infect:



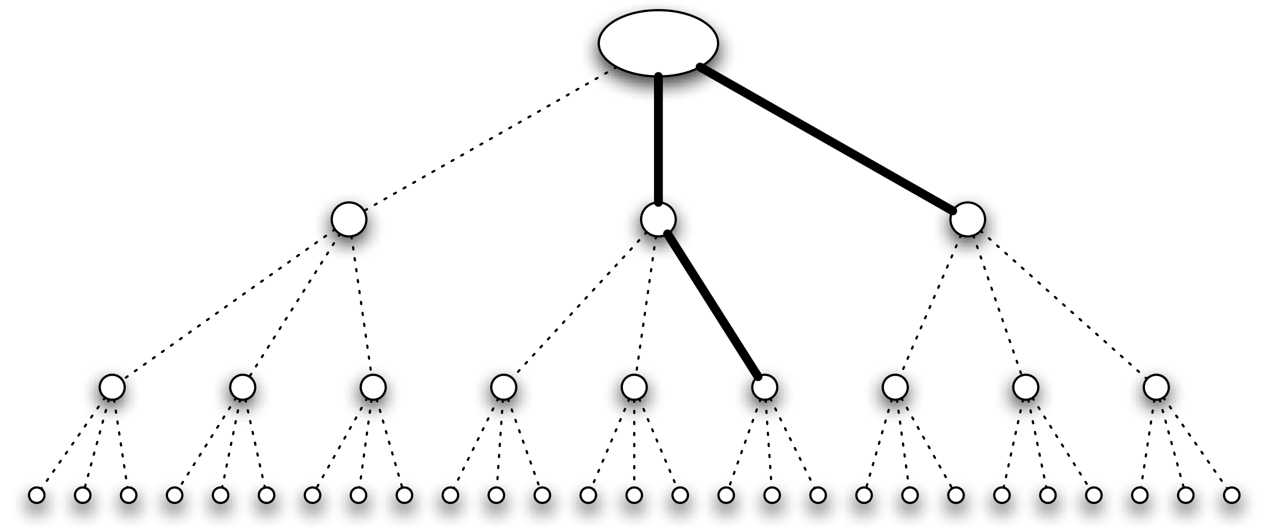
(a) *The contact network for a branching process*

Higher transmission probability *p*:



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

Lower transmission probability *p*:



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

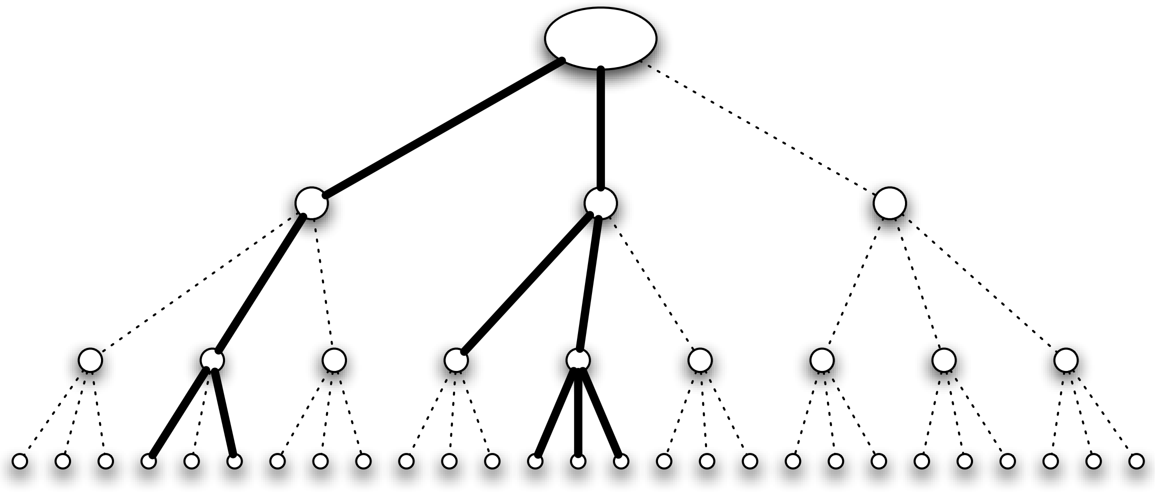
Branching Process: Outcomes

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

How does it die out?

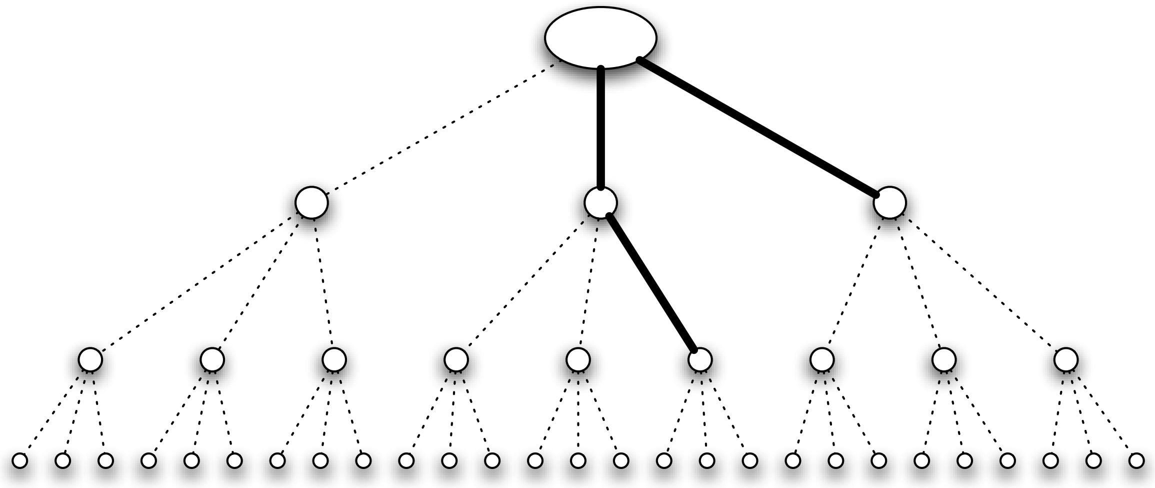
- Dies out if and only if none of the nodes on a given level are infected

Disease might blow up:



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

Disease has already died out:



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

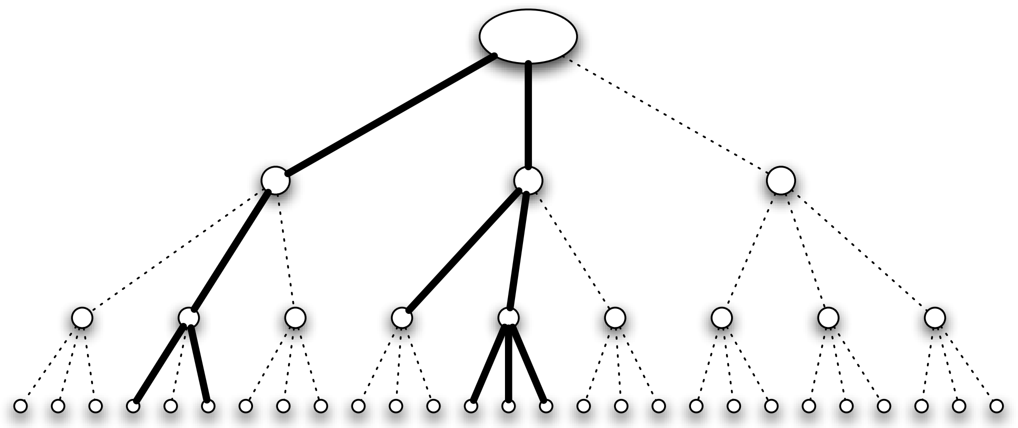
Branching Process

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

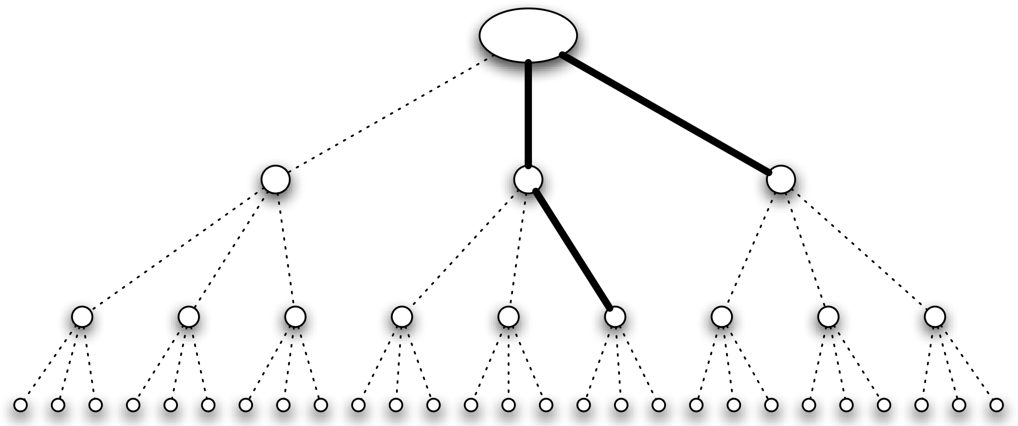
How does it die out?

- 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



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



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

Branching Process

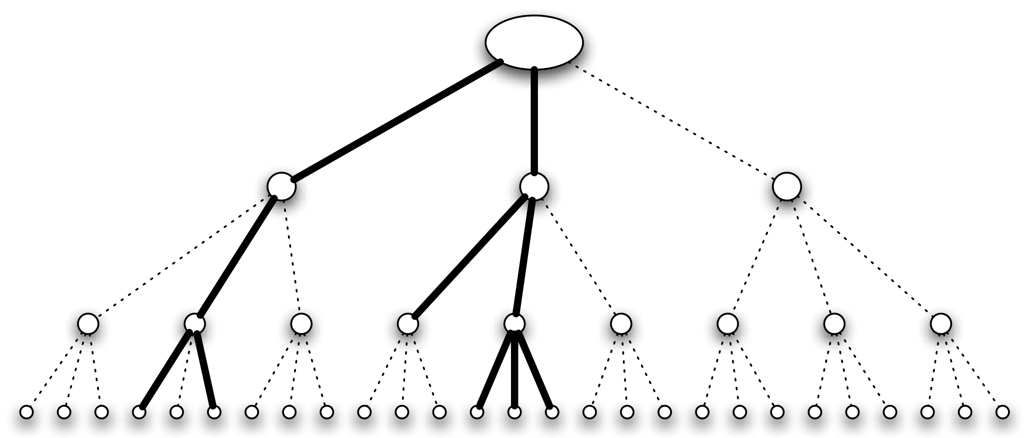
Only two possibilities in the long run: **blow up** or **die out**

How does it die out?

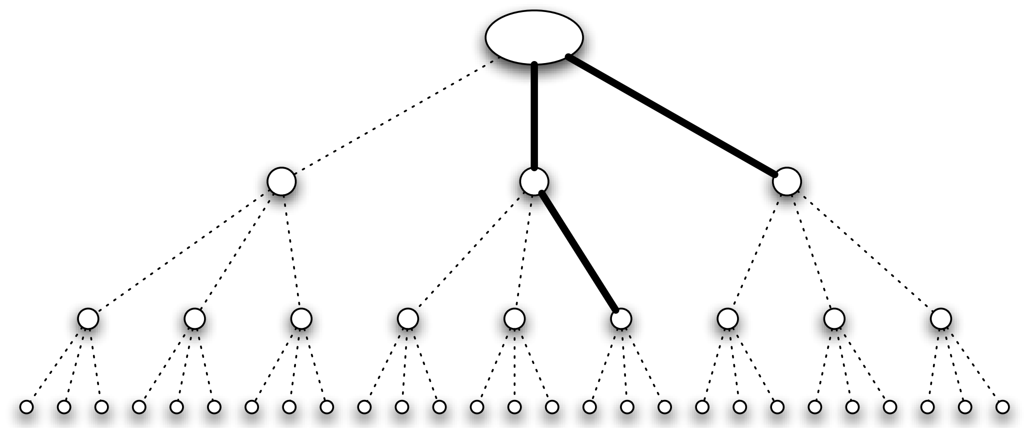
- Dies out if and only if none of the nodes on a given level are infected

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*

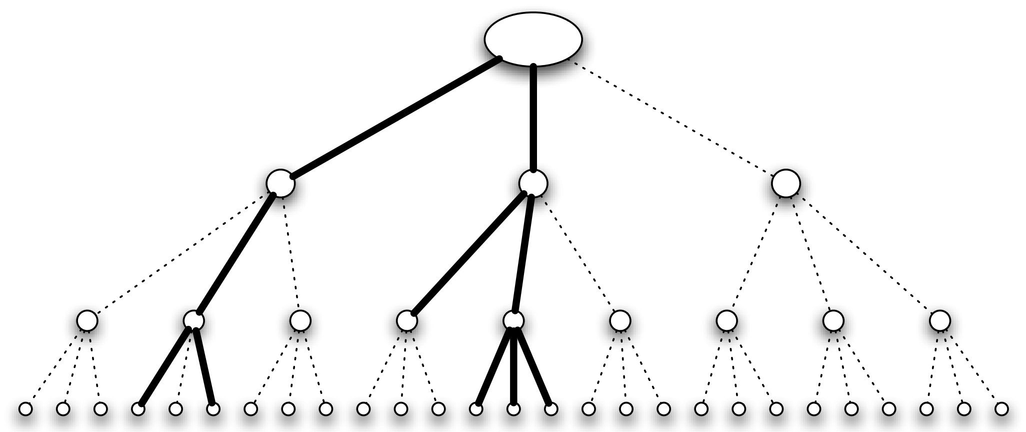
Branching Process: R_0

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

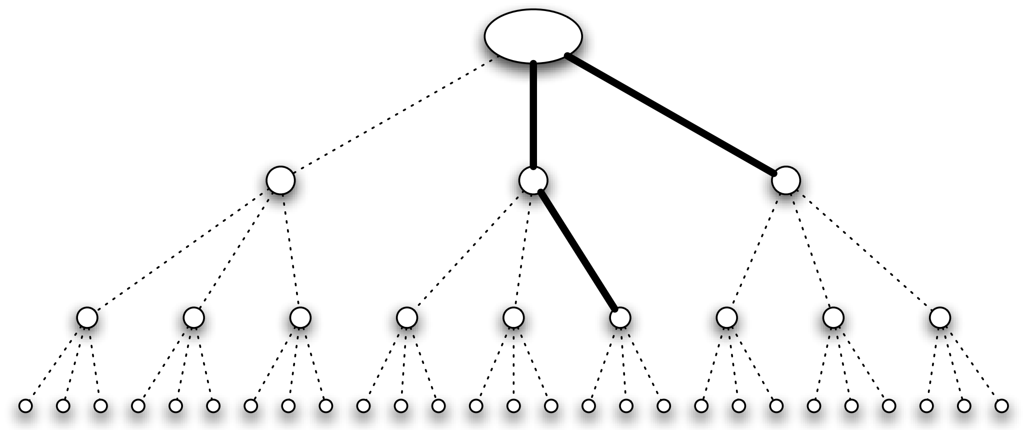
- If $\mathbf{R_0 < 1}$ then with probability 1 the disease dies out after a finite number of steps.
- If $\mathbf{R_0 > 1}$ then with probability > 0 the disease persists by infecting at least one person in each wave.

"Go big or go home."

$R_0 = pk$



(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

$R_0 = pk < 1$:

With probability 1 the disease dies out after a finite number of steps

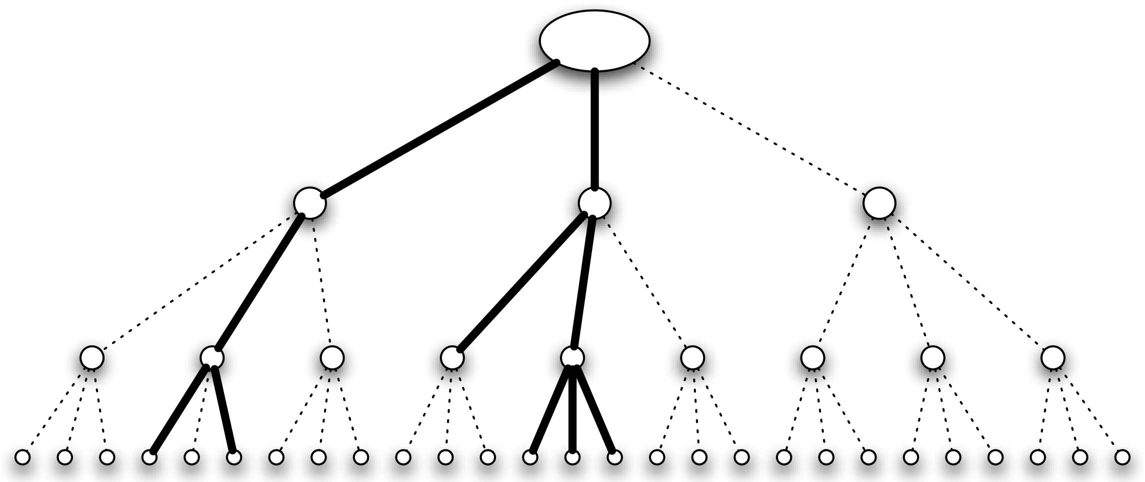
Below replacement; disease isn't able to replenish itself.

Even if it grows momentarily, it trends downward.

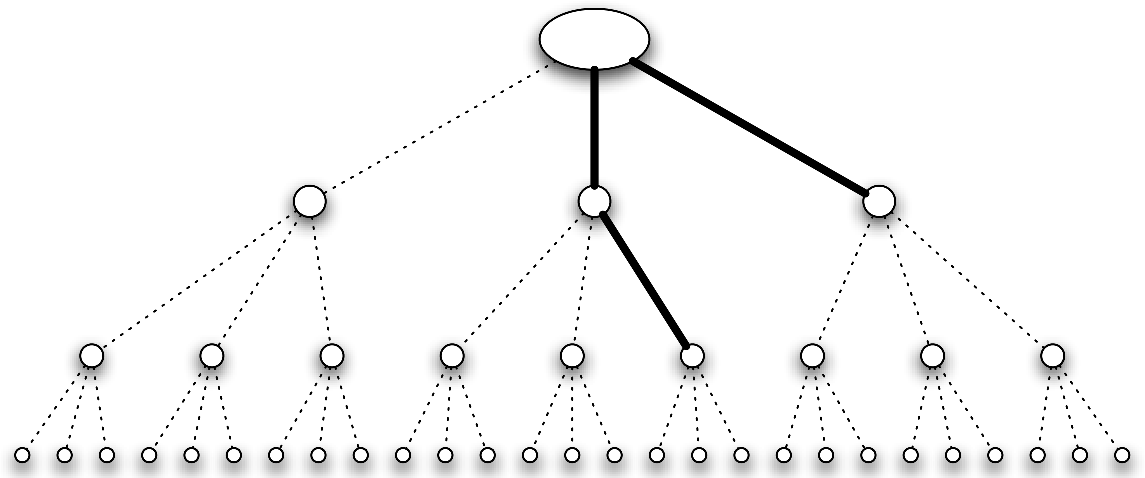
$R_0 = pk > 1$:

with probability > 0 the disease persists by infecting at least one person in each wave

Always trending upward. Could still get "unlucky" and die out, but there's a non-zero chance it runs forever.



(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

$R_0 = pk < 1$:

With probability 1 the disease dies out after a finite number of steps

Below replacement; disease isn't able to replenish itself.

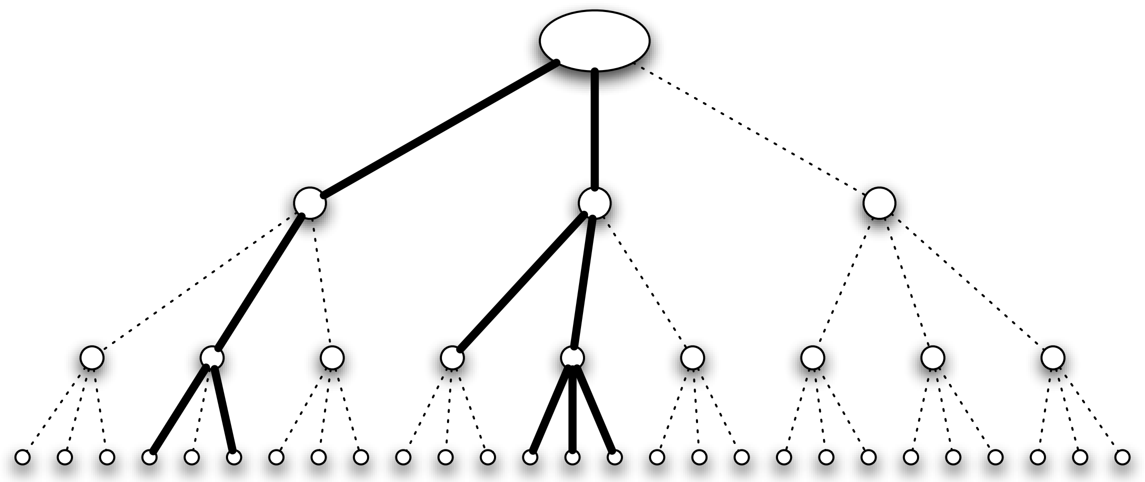
Even if it grows momentarily, it trends downward.

$R_0 = pk > 1$:

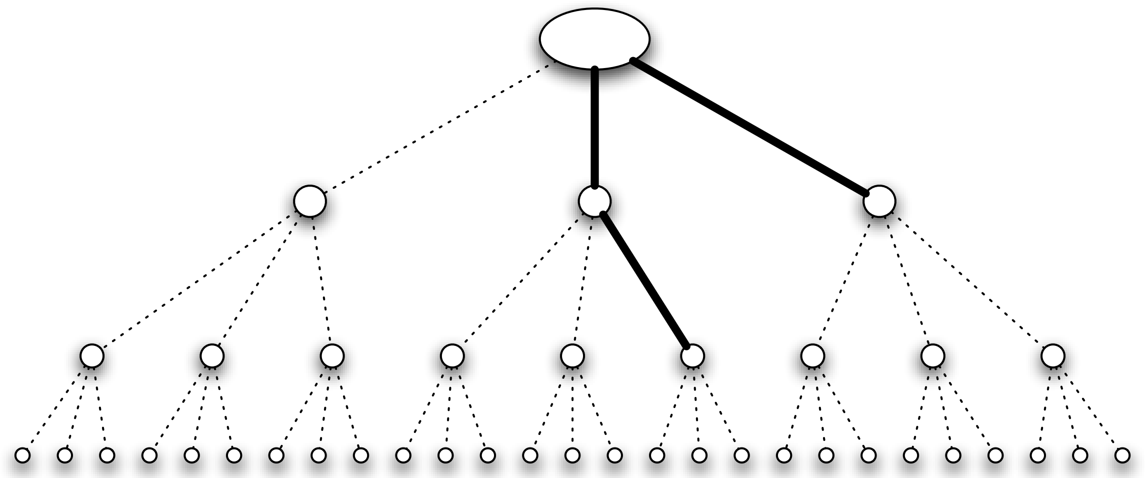
with probability > 0 the disease persists by infecting at least one person in each wave

Always trending upward. Could still get "unlucky" and die out, but there's a non-zero chance it runs forever.

What happens when **p** or **k** change near **pk=1**?



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



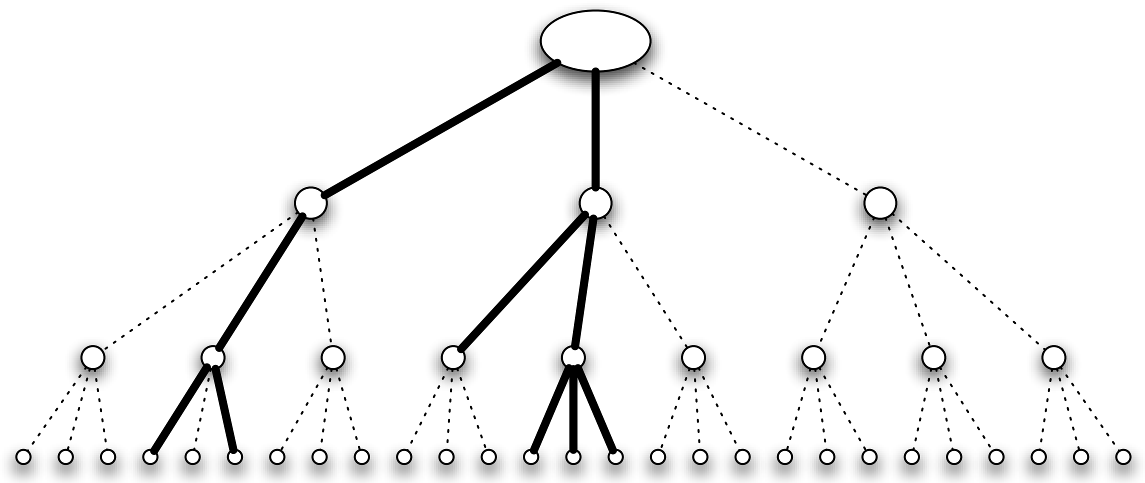
(c) *With low contagion probability, the infection is likely to die out quickly*

Sensitivity of p and k

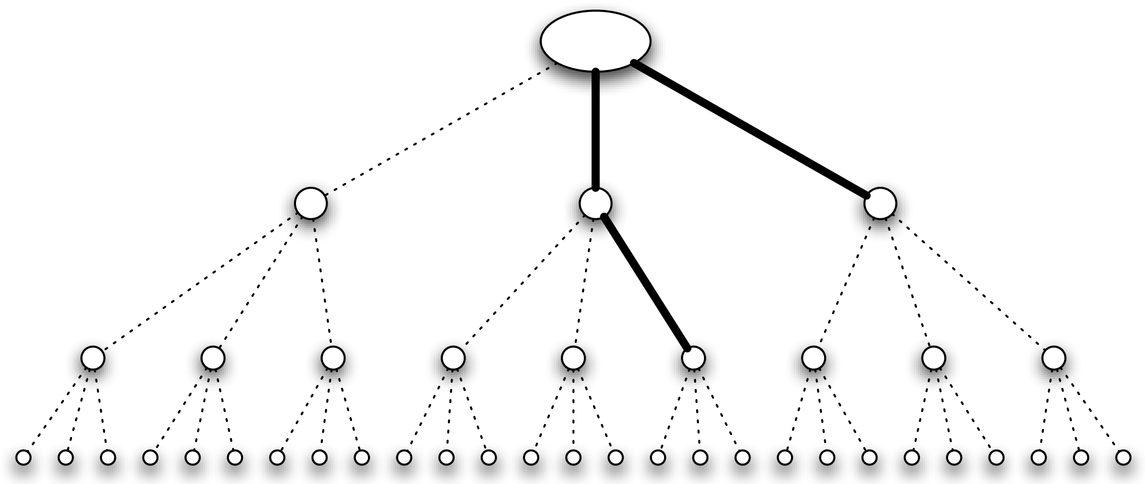
Because epidemics have a "critical threshold", it can be worth it to do a lot of work or expend resources to push **p** or **k** **down** a little bit.

Quarantine = reduce **k**

Improved sanitation = reduce **p**



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



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

Disease	Transmission	R_0
Measles	Airborne	12–18
Diphtheria	Saliva	6-7
Smallpox	Airborne droplet	5–7
Polio	Fecal-oral route	5–7
Rubella	Airborne droplet	5–7
Mumps	Airborne droplet	4–7
HIV/AIDS	Sexual contact	2–5
Pertussis	Airborne droplet	5.5[2]
SARS	Airborne droplet	2–5[3]
Influenza (1918 pandemic strain)	Airborne droplet	2–3[4]
Ebola (2014 Ebola outbreak)	Bodily fluids	1.5-2.5[5]

COVID-19: ~2
COVID Delta: ~5
COVID Omicron: ~8

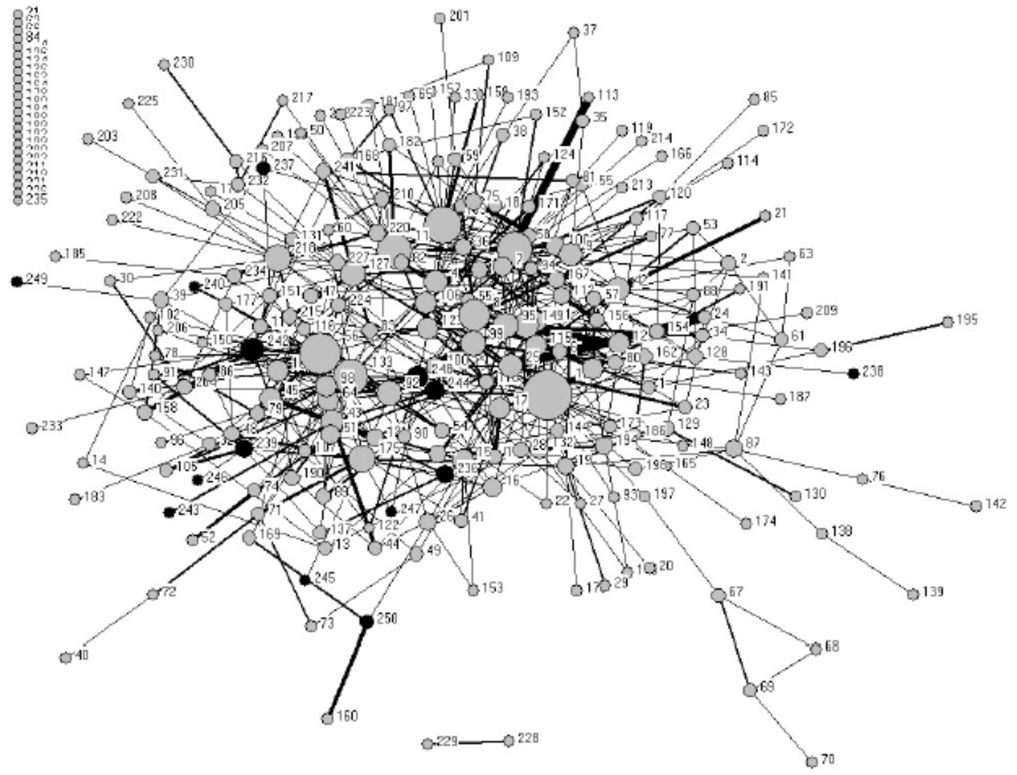
General Models of Contagion

Epidemics on General Graphs

We just studied epidemics as ideal trees

But of course real-life networks are more complicated than that

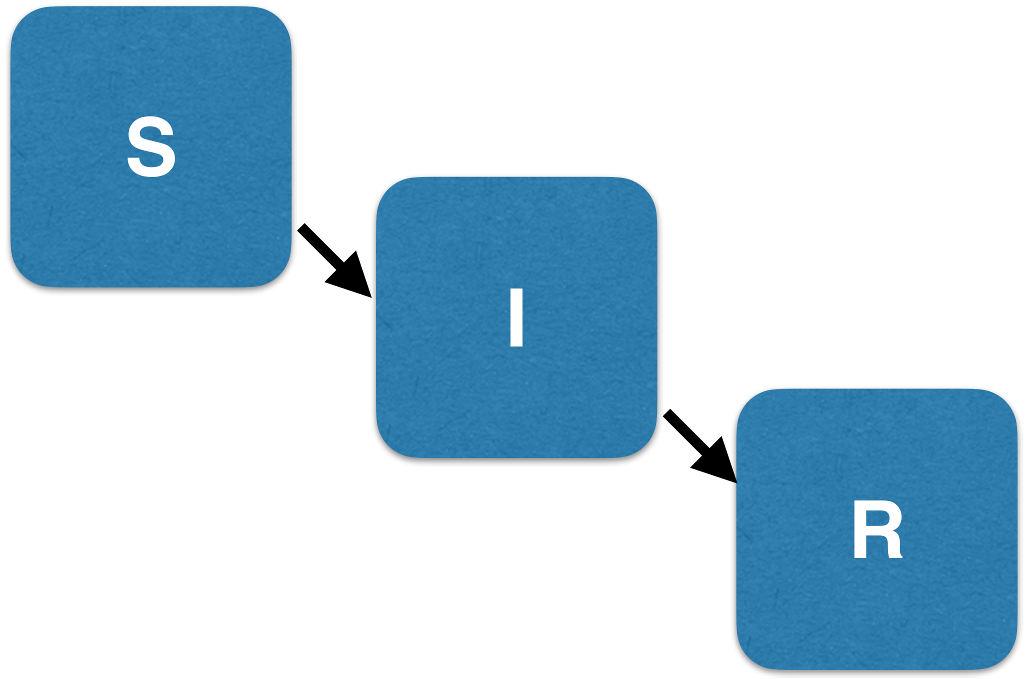
What does epidemic diffusion look like in general graphs?

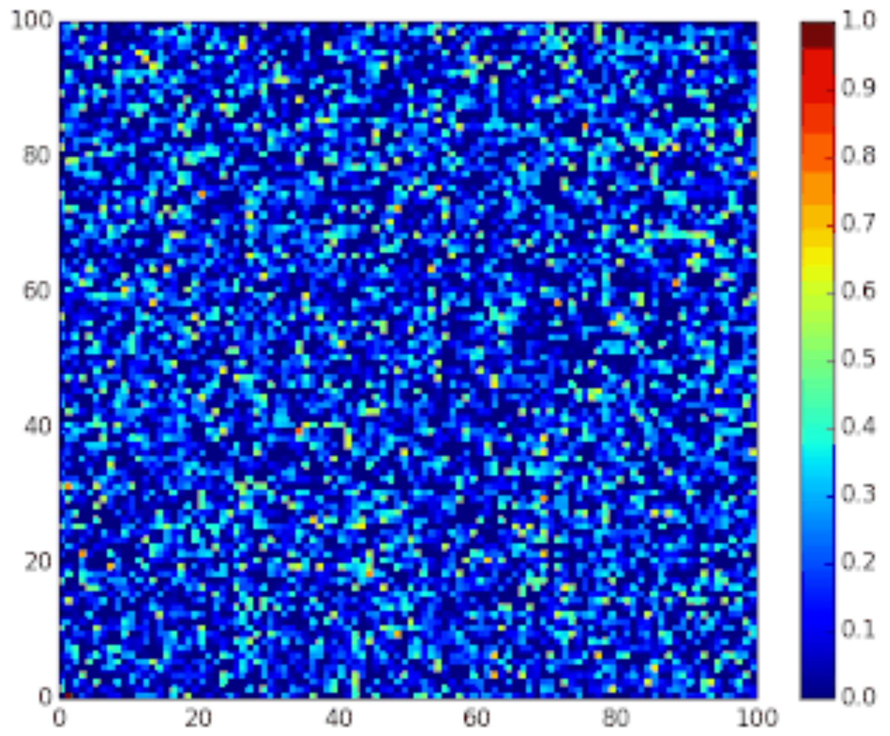


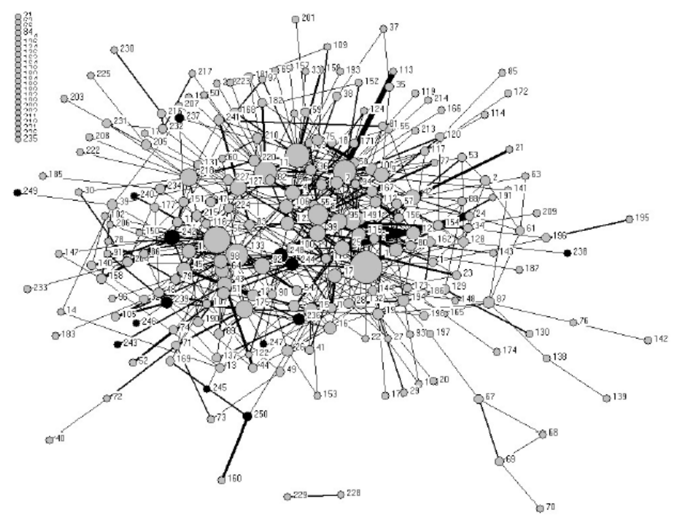
SIR Epidemic Models

Simple lifecycle model with three stages:

- **S** = Susceptible
- **I** = Infectious: node is infected and infects with prob **p**
- **R** = Removed: after **t_I** time, no longer infected or infectious







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

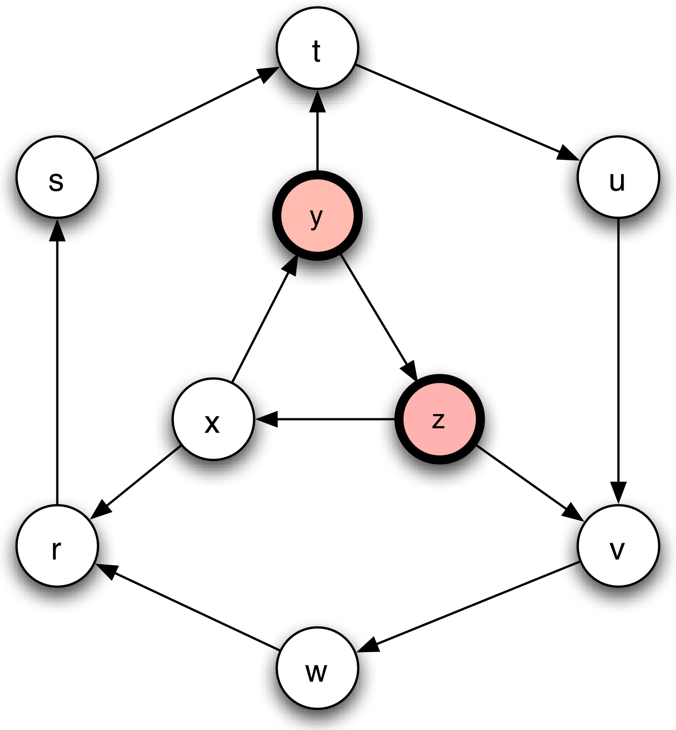
Initially some nodes in **I** state, rest in **S** state.

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**

SIR Epidemics on Networks

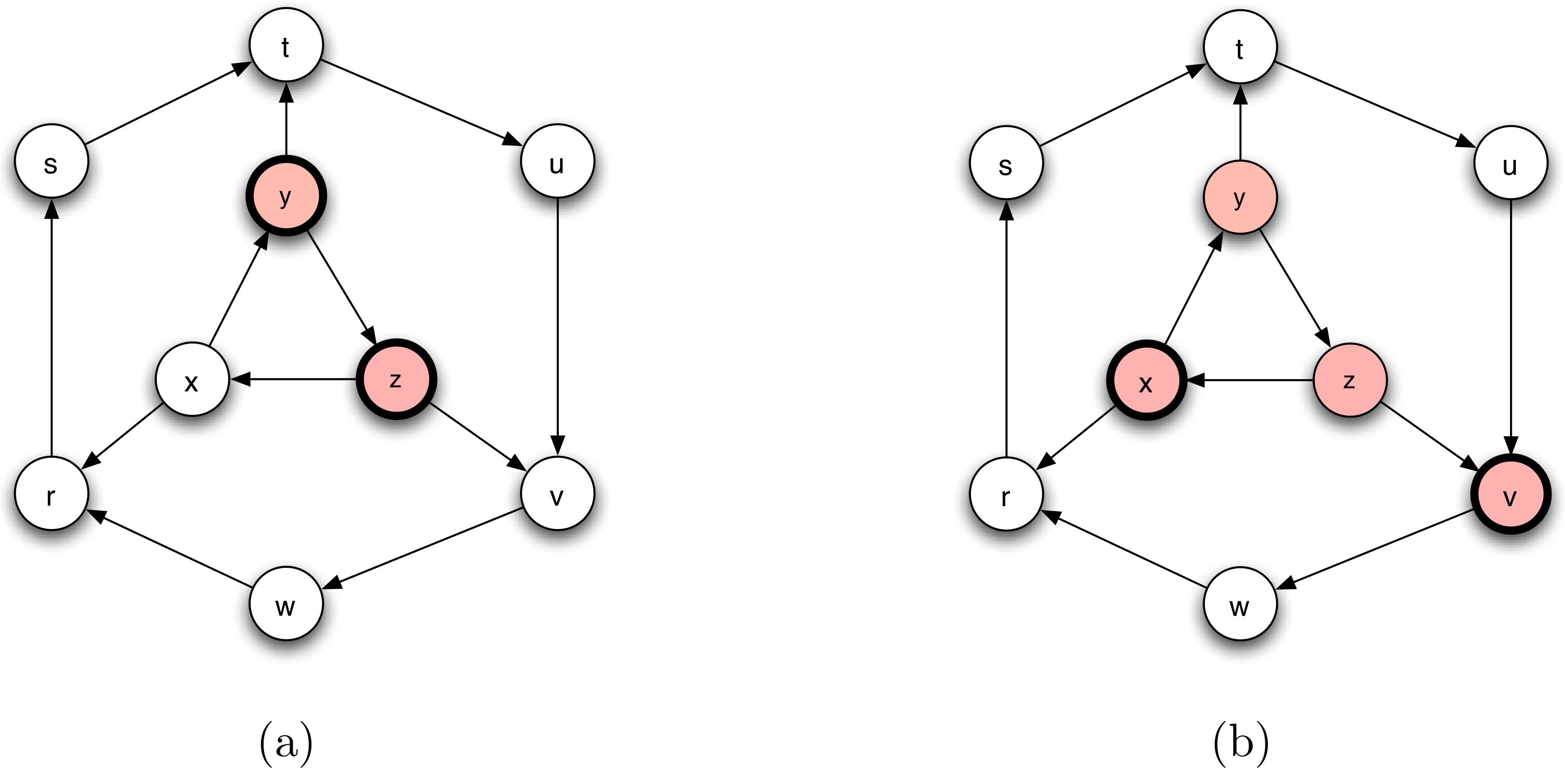


(a)

p=1/2

$\mathbf{t_I}$ =1

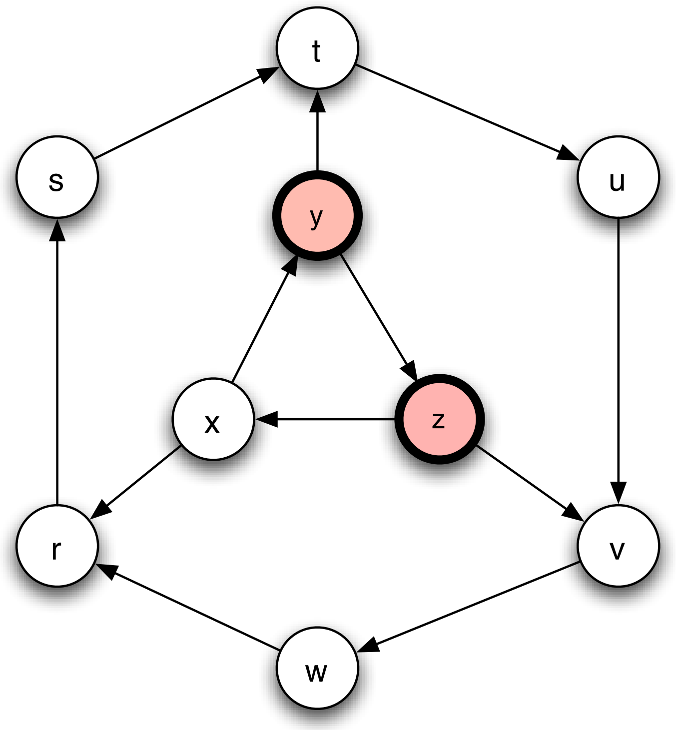
SIR Epidemics on Networks



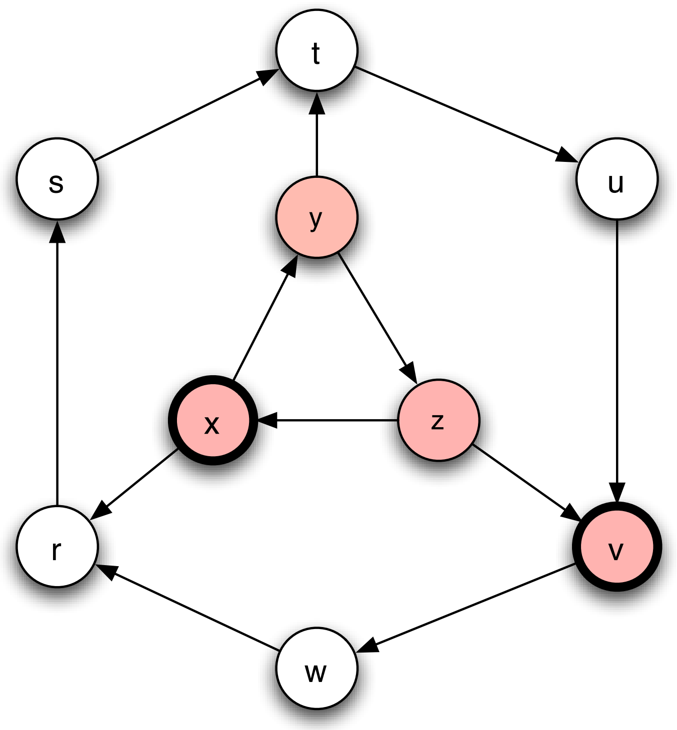
p=1/2

t_I =1

SIR Epidemics on Networks



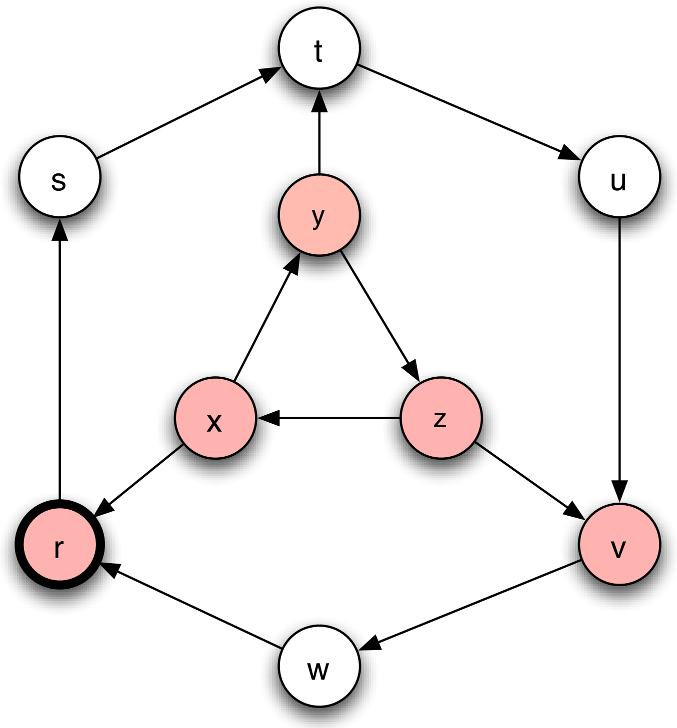
(a)



(b)

p=1/2

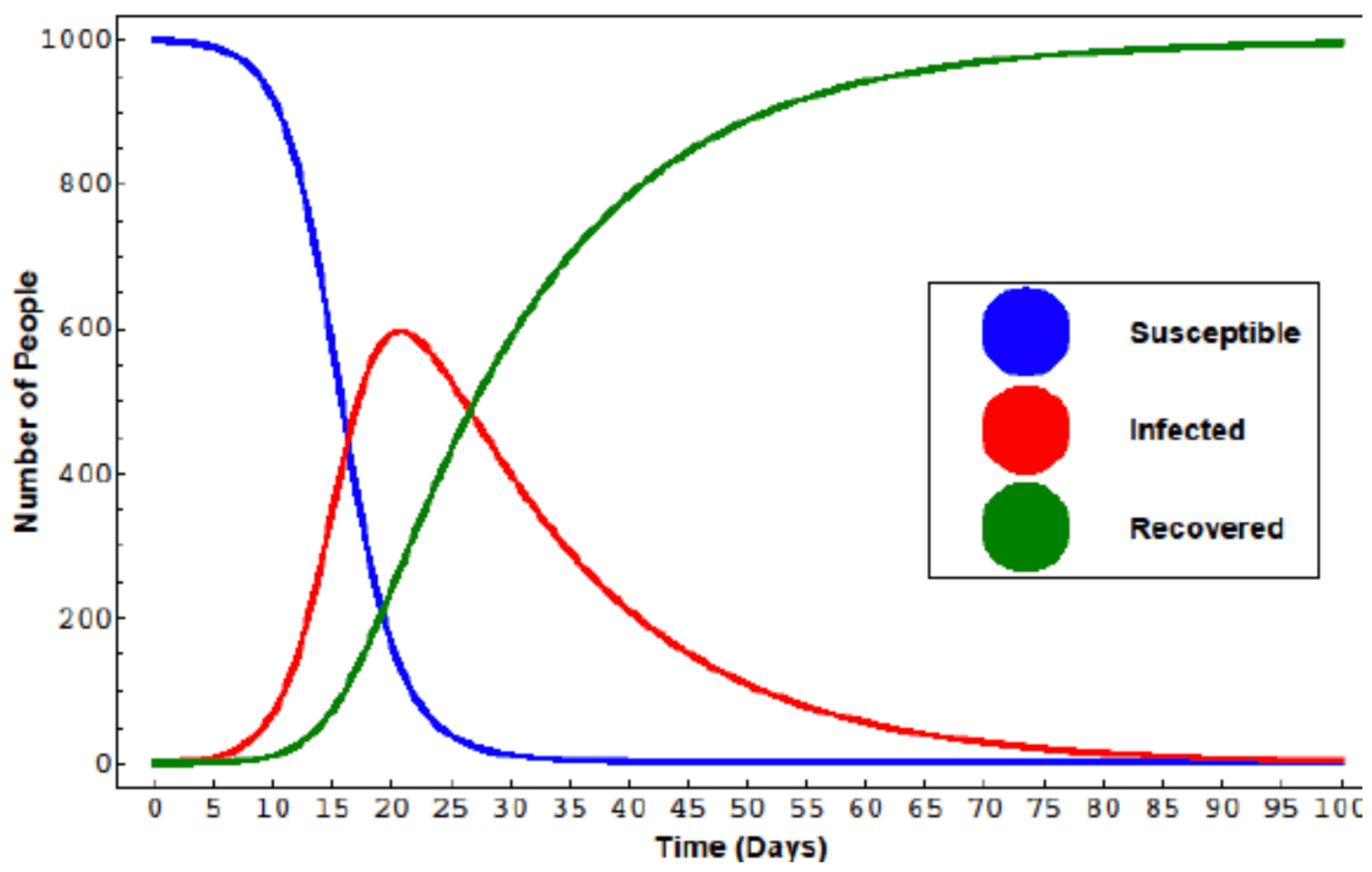
$\mathbf{t_I}$ =1



(c)

SIR Epidemics on Networks

Typical run of SIR on a graph representing a contact network



Big questions in epidemiology: how many will an epidemic infect?
How will the spread change with changes in parameters?
Based on that, what are best defences?

SIR Epidemic Extensions

Many extensions to accommodate different parameters

Some contacts more likely than others:

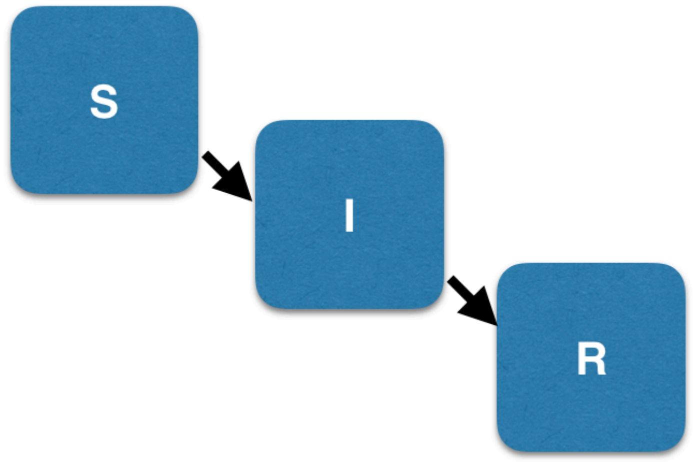
→ probability $\mathbf{p_{uv}}$ that is pair-dependent

Disease goes through different stages (infectious incubation, then less infectious symptomatic transmission):

→ **SEIR or S"III"R:** either **E**xposed state or several different infectious states (with different **p**'s or **t**'s)

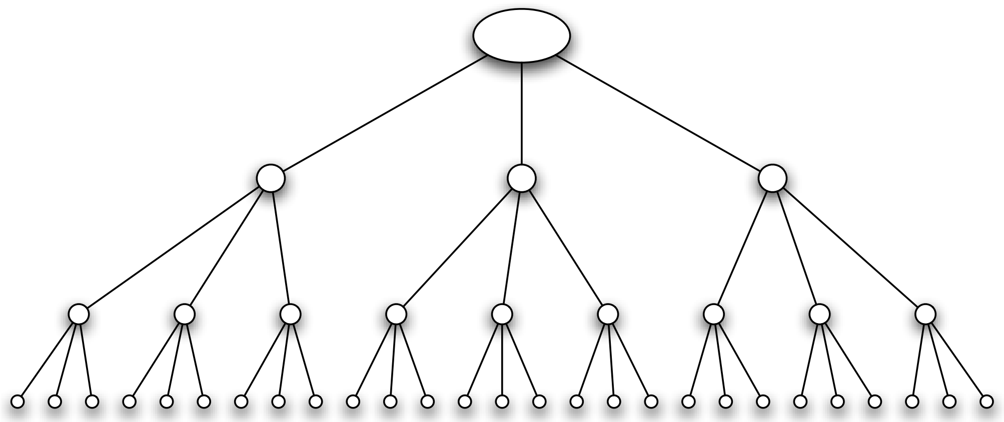
SIS: later in the lecture

Mutations (infectiousness, breaking immunity, etc)



From trees to networks

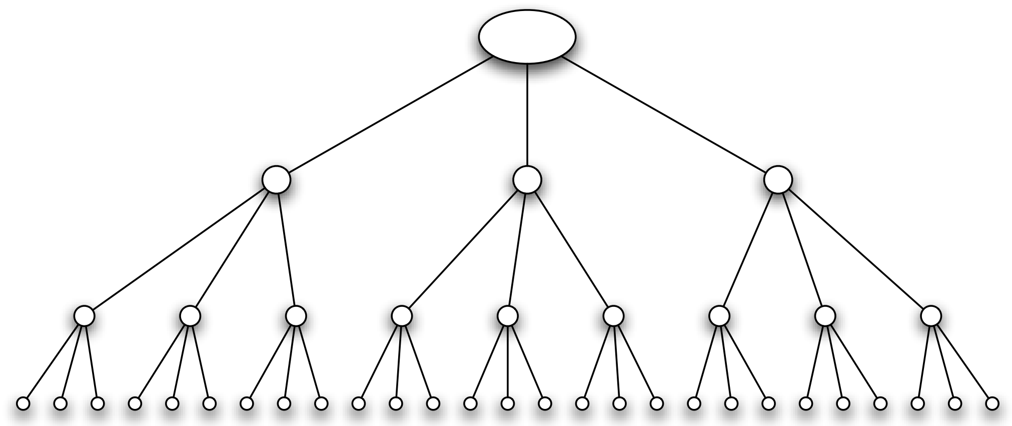
Recall that analysis of R_0 was for trees:



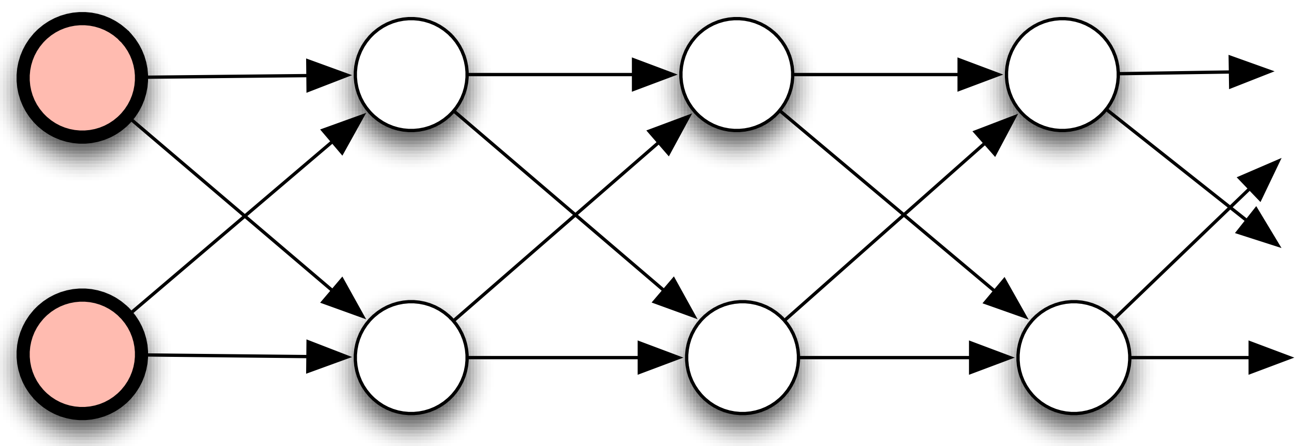
Do we have the same knife-edge R0 ~ 1 result in general graphs?

From trees to networks

Recall that analysis of R_0 was for trees:

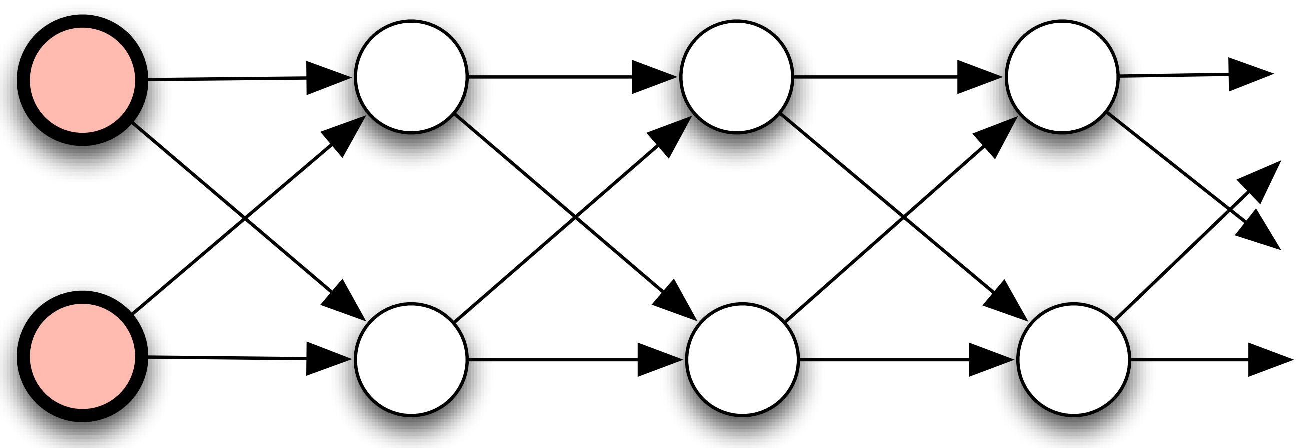


What happens on other networks? Consider **p=2/3, k=2.**



From trees to networks

What happens on other networks? Consider **p=2/3, k=2.**



Calculate R_0 as number of expected new cases per node

R0 = (2/3)*2 = 4/3 > 1

But this will almost certainly die out: $(1/3)^4 = 1/81$ chance that all four edges fail even if both nodes are infected

Prob that this happens after finite number of steps converges to 1

Now: SIS Epidemic Model

S = Susceptible

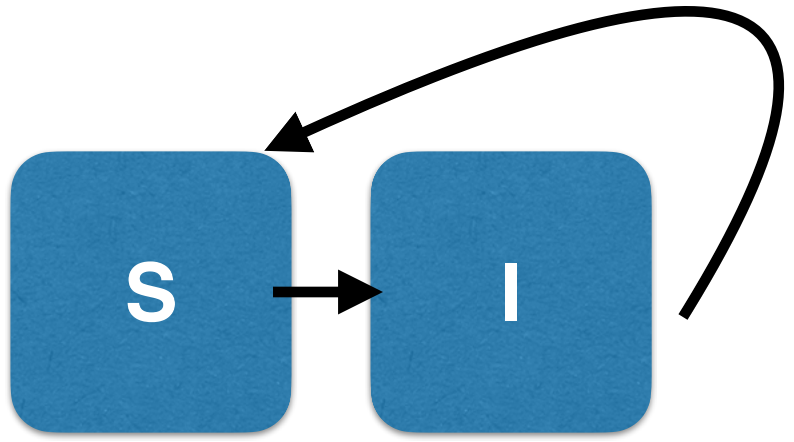
I = Infectious: node is infected and infects with prob **p**

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

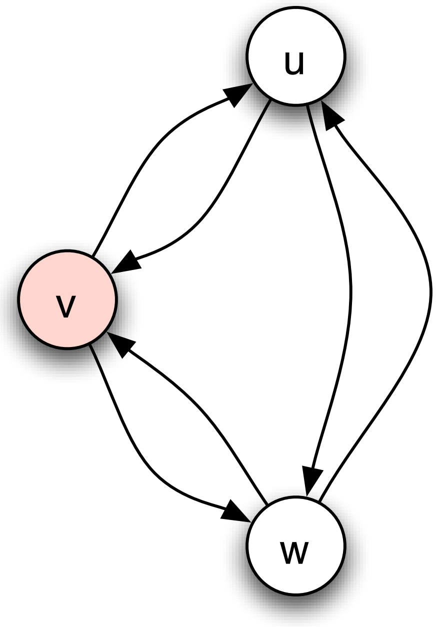
After $\mathbf{t_I}$ time steps, node **returns to S**



SIS Epidemic Example

p= 1/2

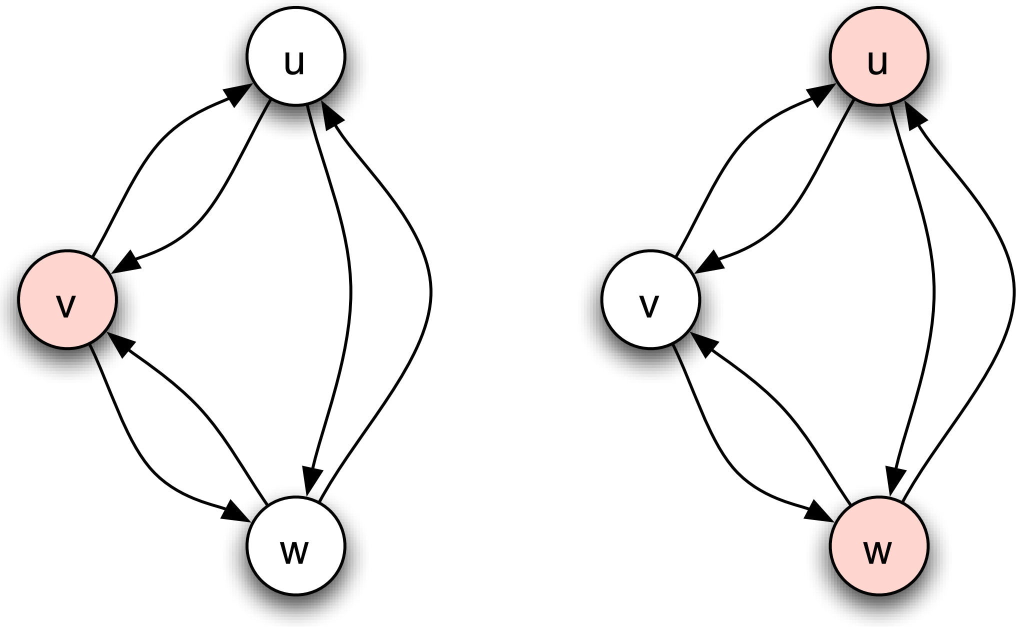
$\mathbf{t_I}$ = 1



SIS Epidemic Example

p=1/2

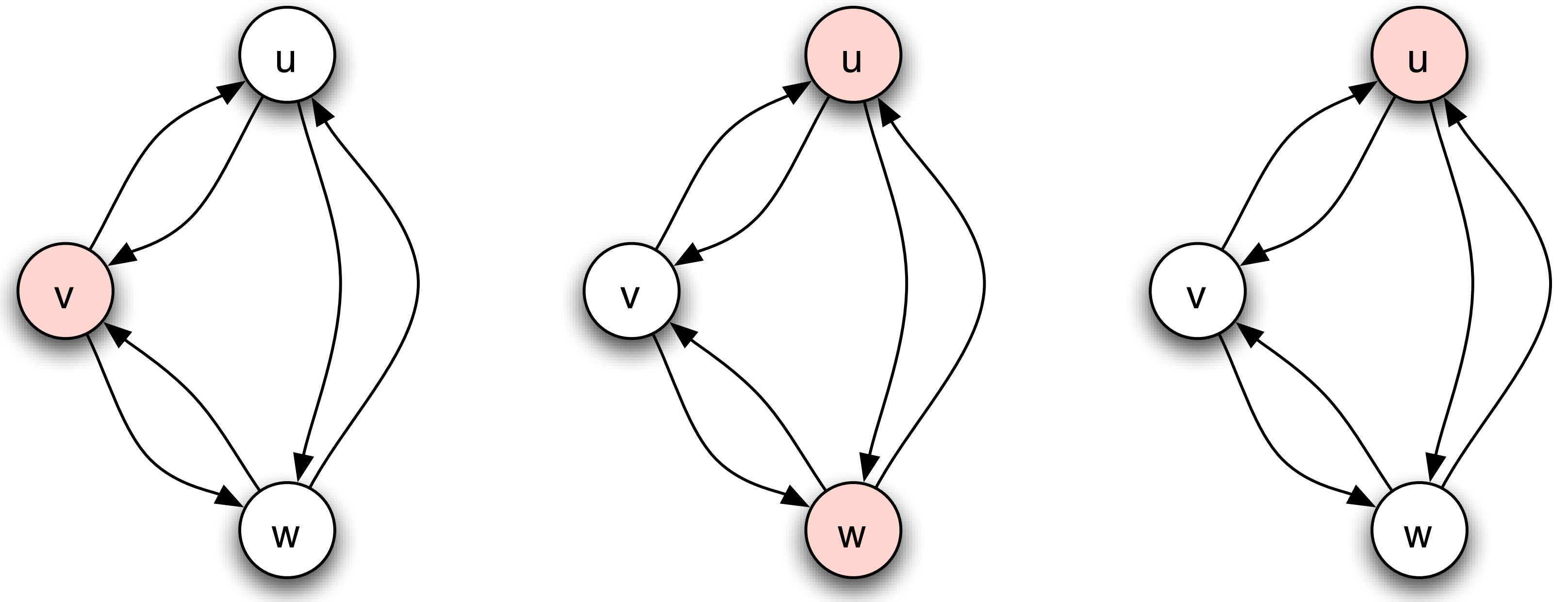
$\mathbf{t_I}$ =1



SIS Epidemic Example

p=1/2

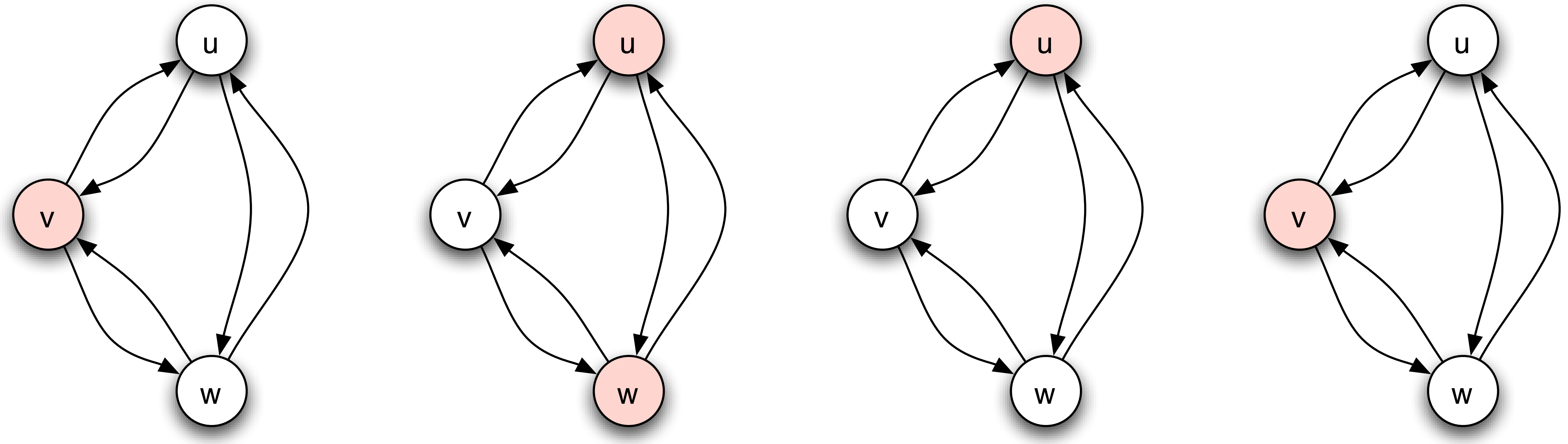
$\mathbf{t_I}$ =1



SIS Epidemic Example

p= 1/2

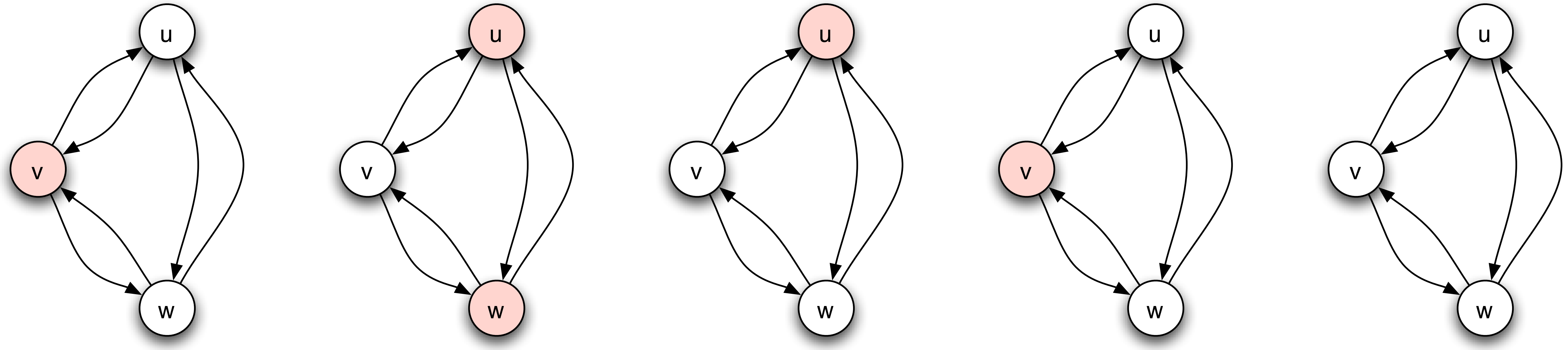
$\mathbf{t_I}$ = 1



SIS Epidemic Example

p=1/2

t$_I$ =1

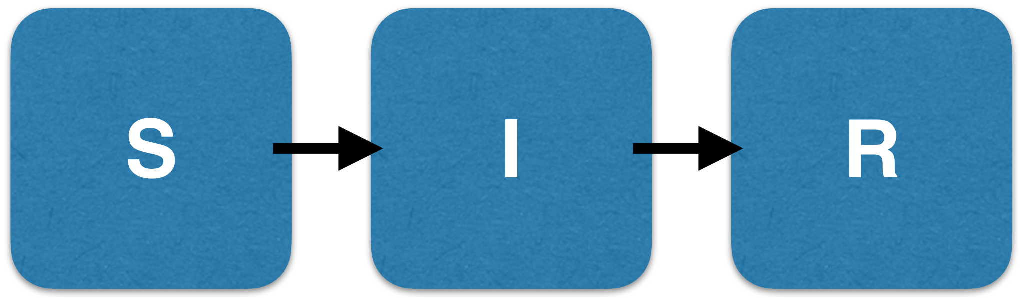


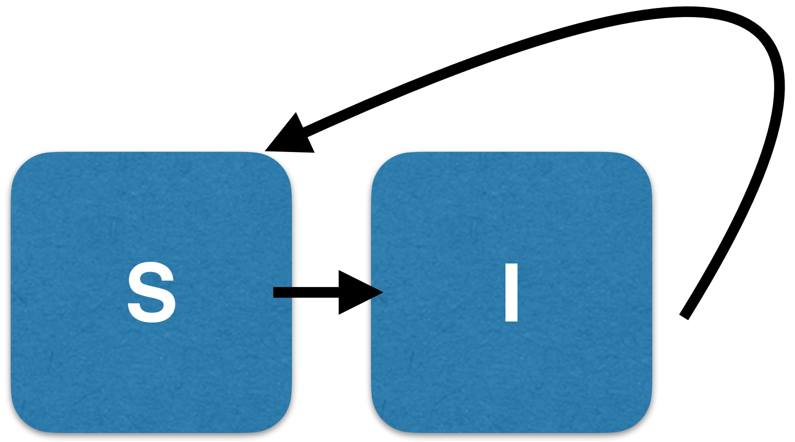
SIR vs. SIS

SIR: "burning through" a finite supply of susceptible

SIS: can run for a very long time, cycling through targets

SIS, like SIR, has a **critical threshold** ("knife-edge"/"tipping point"); trickier mathematical analysis. On non-trees both depend on more than just R_0.

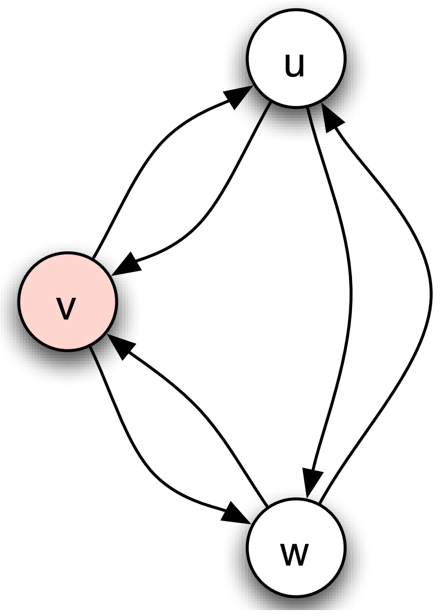




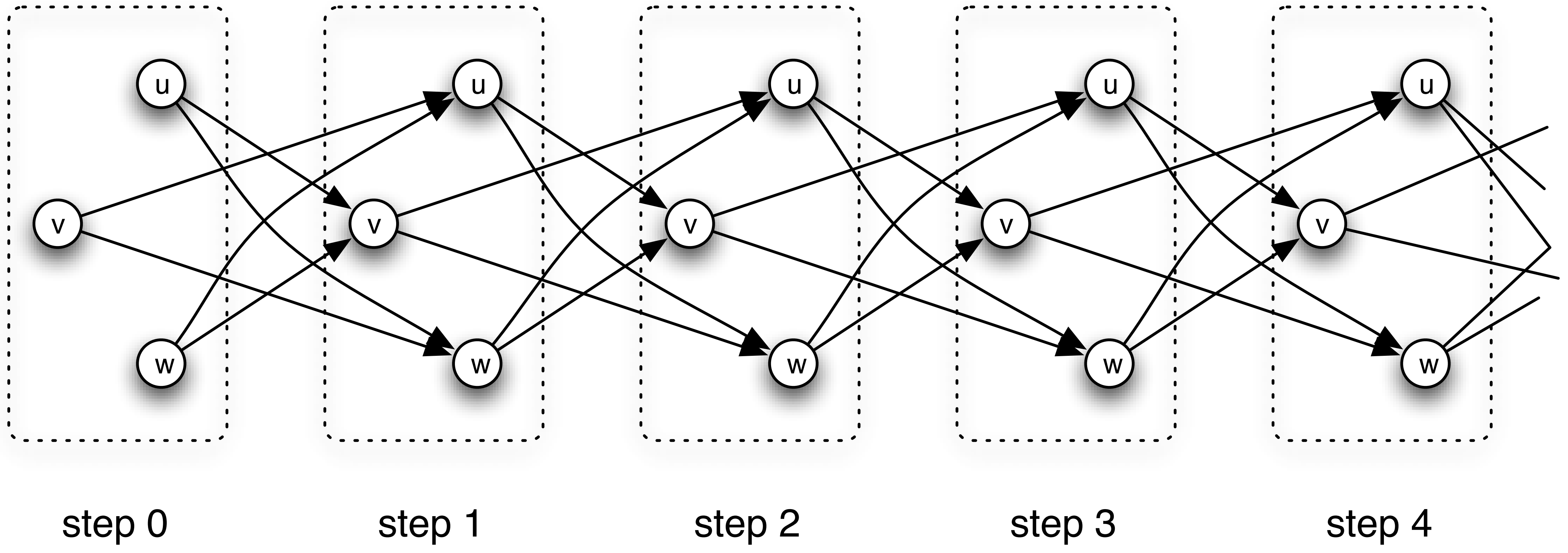
SIS as SIR on a bigger network

Consider **time-expanded** network: if u connects to v in network, have u_t connect to v_{t+1}

SIS is SIR on a time-expanded network.



(t_I = 1 here)

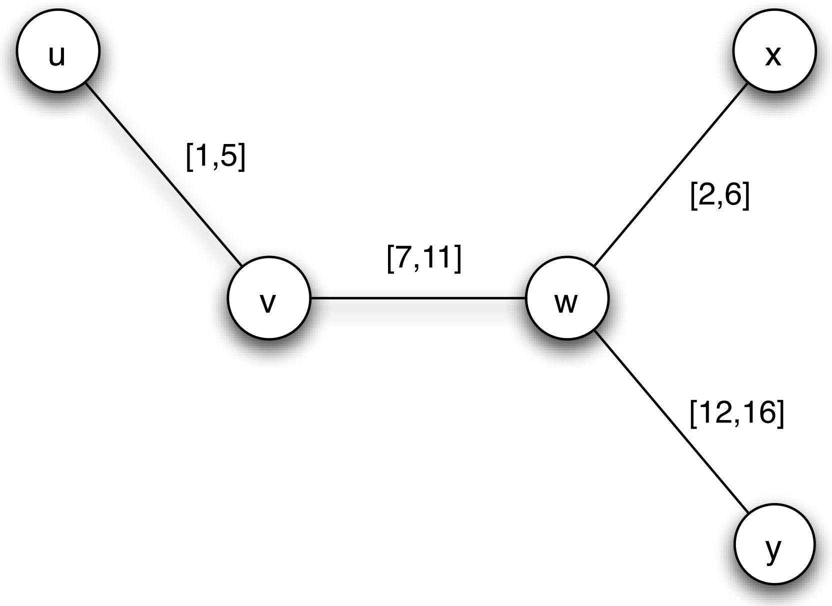


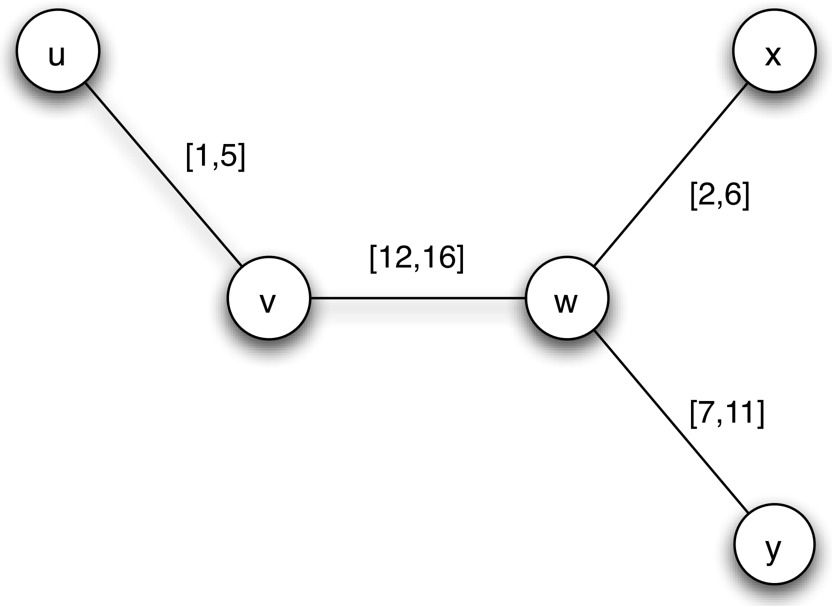
Transient Contacts & Concurrency

- So far, we've been analyzing **static** networks
- This is reasonable when the rate of transmission is **typically much faster** than edge creation/ deletion
- But some epidemic diseases last for years (HIV)
- **When edges are active** becomes very important

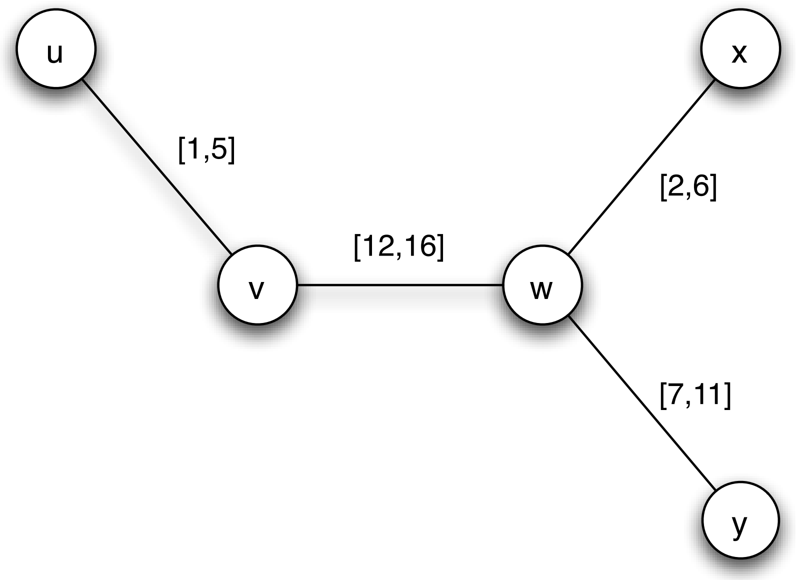
Transient Contacts & Concurrency

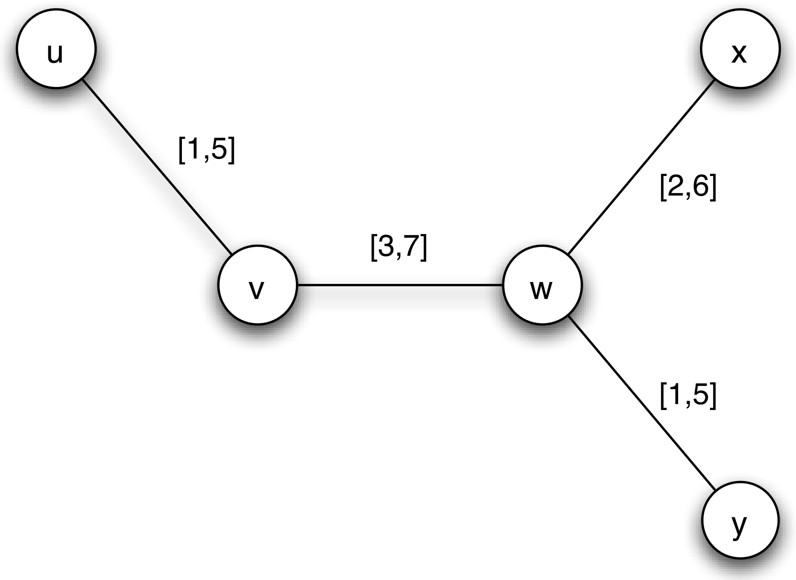
A less random model: it matters in what order contact is made in the contact network.





Concurrency: having two or more contacts at once.



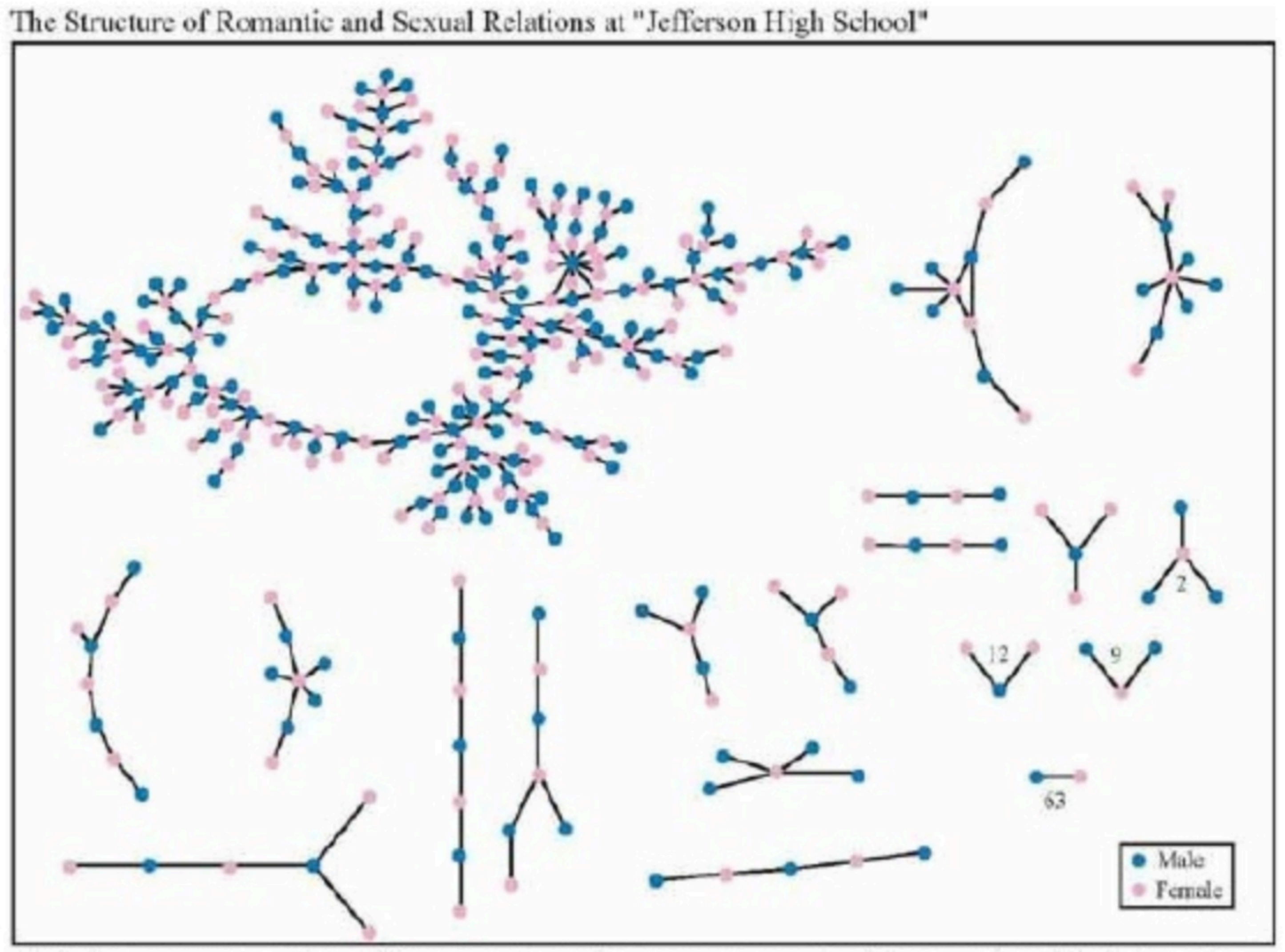


Transient Contacts & Concurrency

Small changes in times can produce **large differences in global epidemic spread**

There are **rich classes of network models** incorporating transience and concurrency

It's not enough to just know the structure



Oscillations

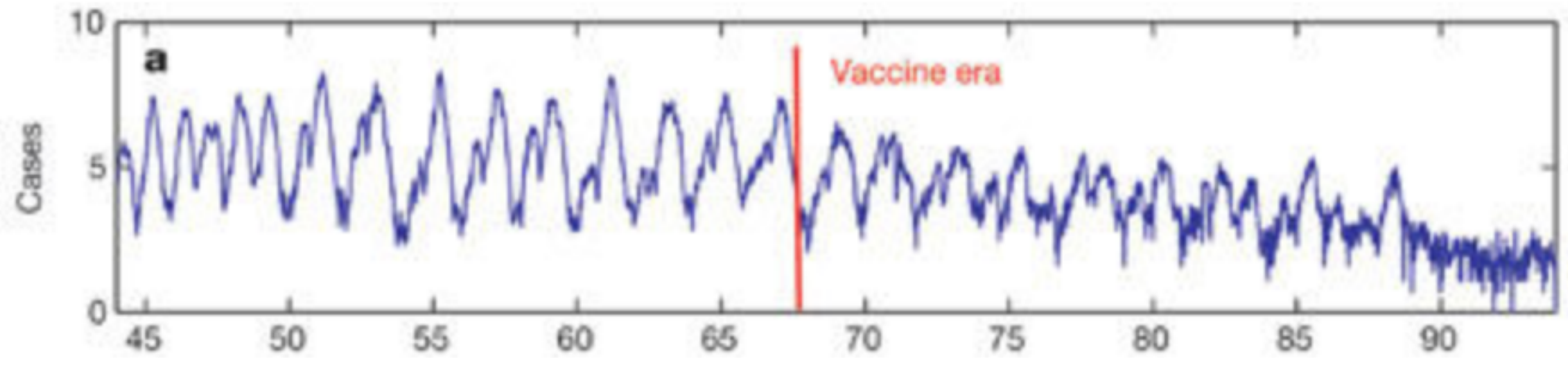
Diseases can be **cyclical** / have **oscillations** (like measles and syphilis)

To model this, vary the model so nodes have **temporary immunity**

SIRS: Susceptible, Infected for I steps, Recovered for R steps, then Susceptible again

This can **produce oscillations in very localized** parts of the network

But for large fluctuations at the global network level, need **small-world structure** (random long-range contacts)

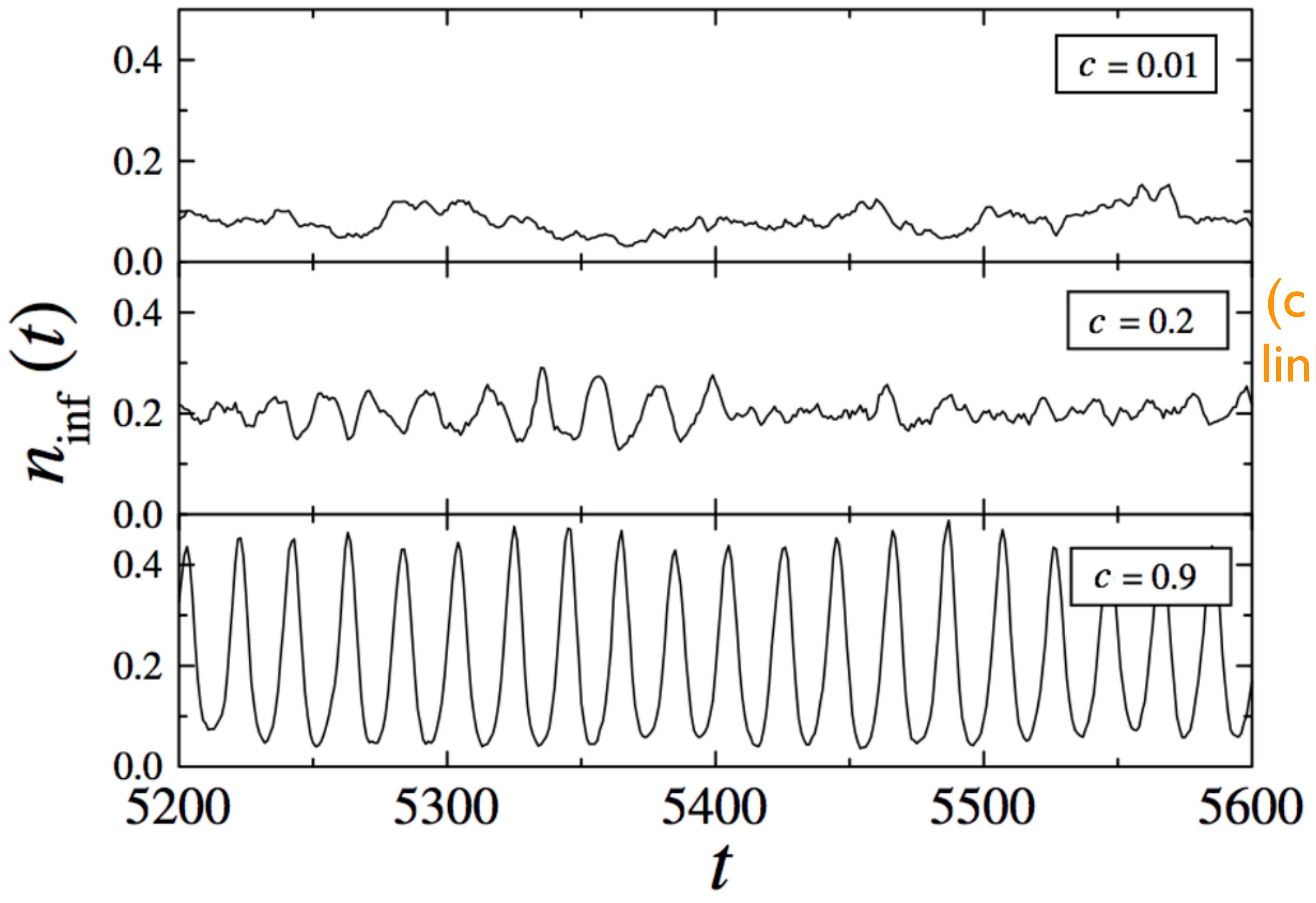


Measles

Oscillations

Diseases can be **cyclical** / have **oscillations** (like the flu)

But for large fluctuations at the global network level, need **small-world structure** (random long-range contacts)



(c = long-range link probability)

Epidemics vs. Behaviour

In epidemic models, nodes get infected from one particular other node

To model information spread, people often use epidemic models ("viral diffusion")

But many social phenomena (behaviours, beliefs, practices, etc.) are *complex*: costly, risky, uncertain, etc.

Epidemics vs. Behaviour

When a behaviour is risky, costly, or uncertain, you may not do it just because one of your friends is (but this is what epidemic diffusion looks like)

Social movements, health technologies, political activism, etc.

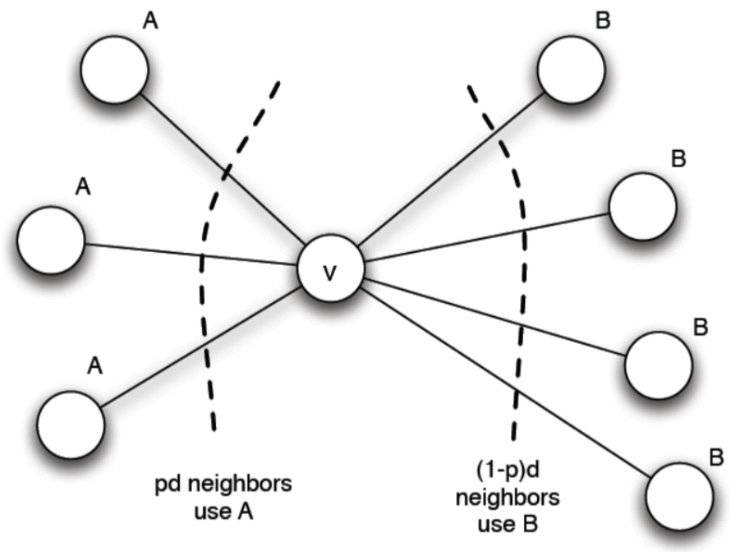
E.g.: PrEP medication is the best latest in HIV prevention → one pill a day gives 90% prevention.

But in two trials in sub-Saharan Africa, it didn't work ... because no one was taking it! (fears of discrimination, etc.)

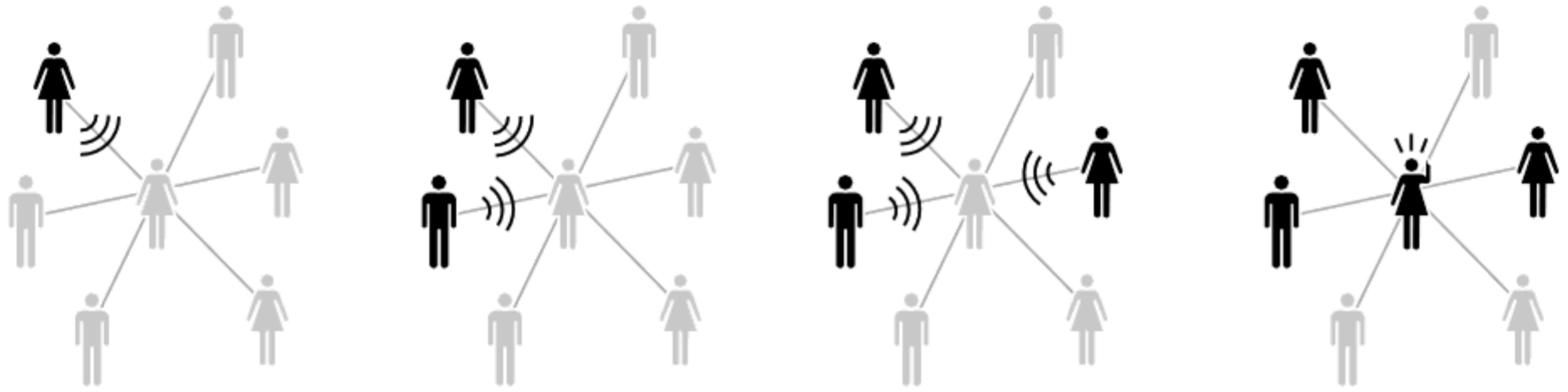
How do you get behaviour to diffuse?

Epidemics vs. Behaviour

Previously we saw a model of behaviour diffusion based on utility



This is an example of *complex* diffusion: in general, need more than one neighbour to adopt before you adopt a behaviour.

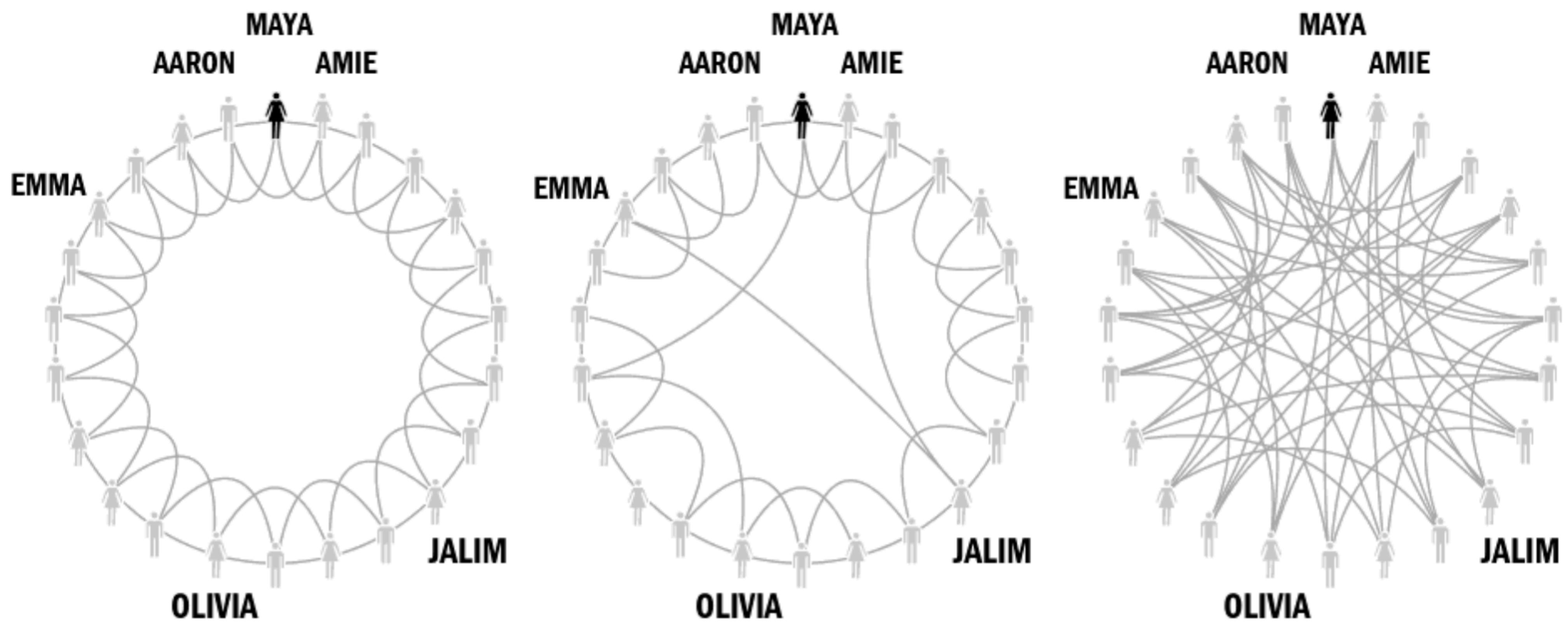


Epidemics vs. Behaviour

Simple vs. complex diffusion
Epidemics vs. behaviour

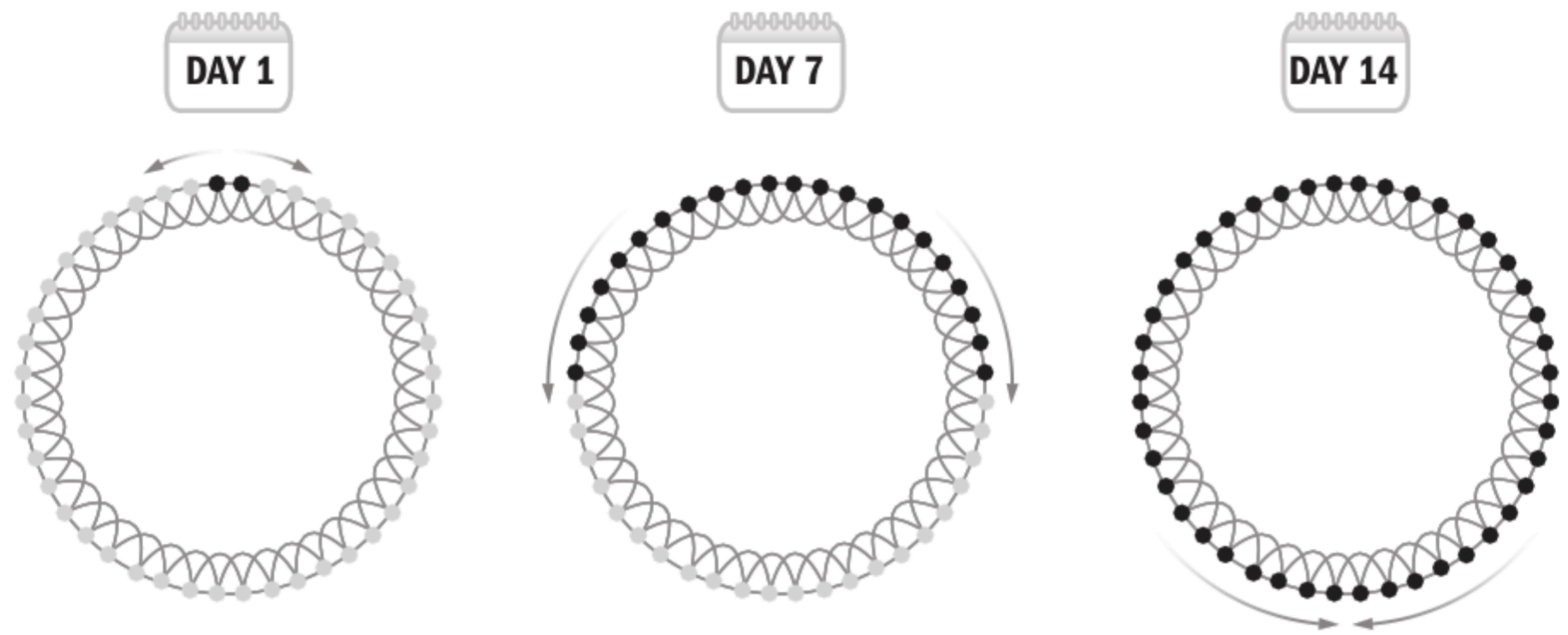
What's the difference?

Recall the small-world model

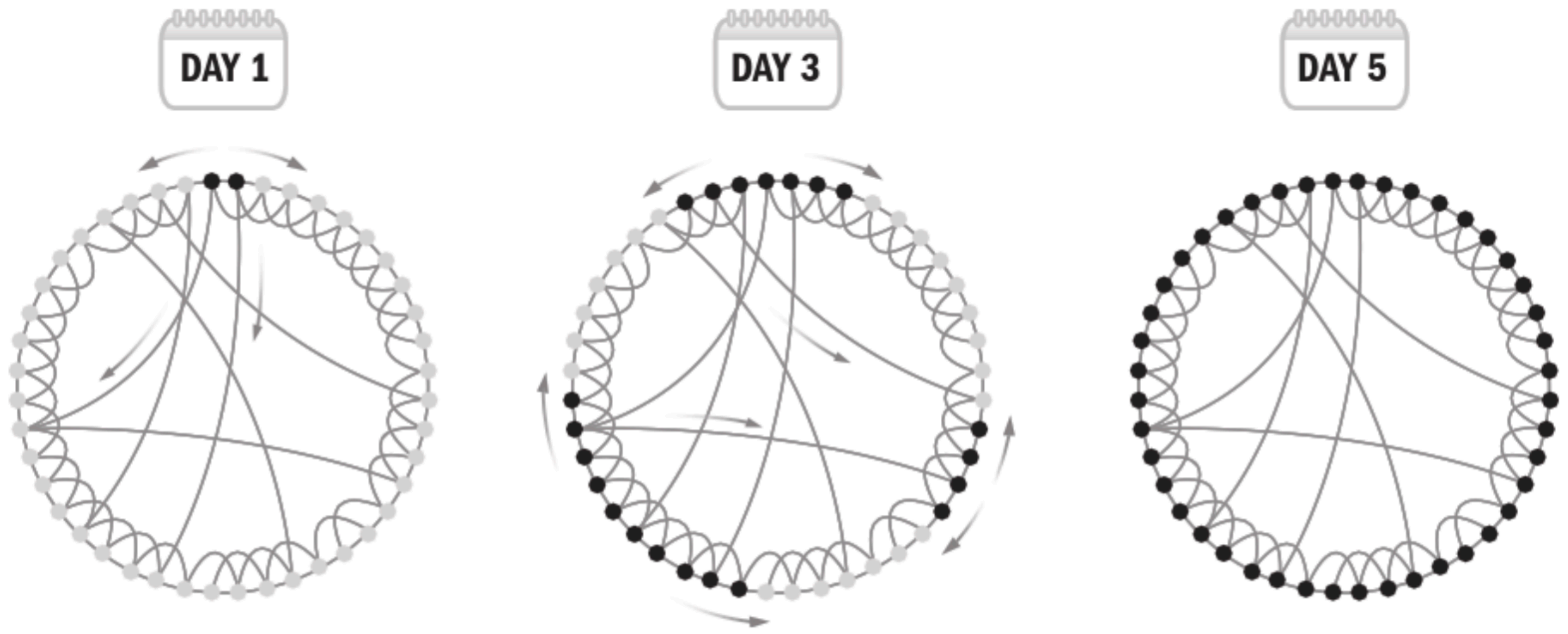


Simple Diffusion

Large world:

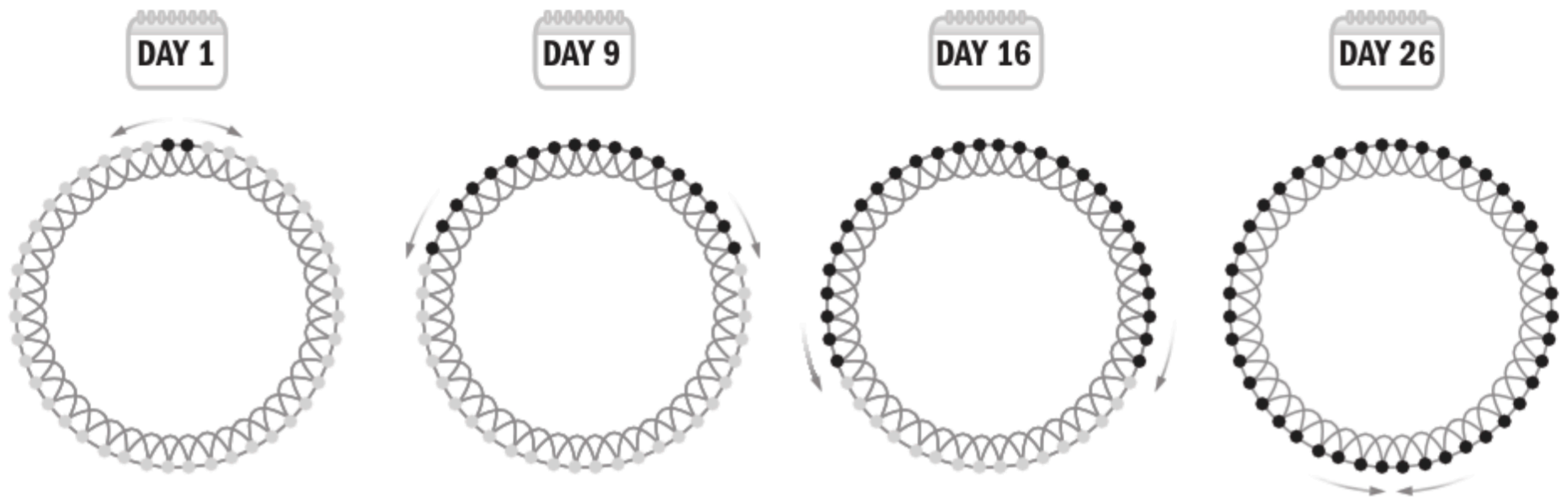


Small world:

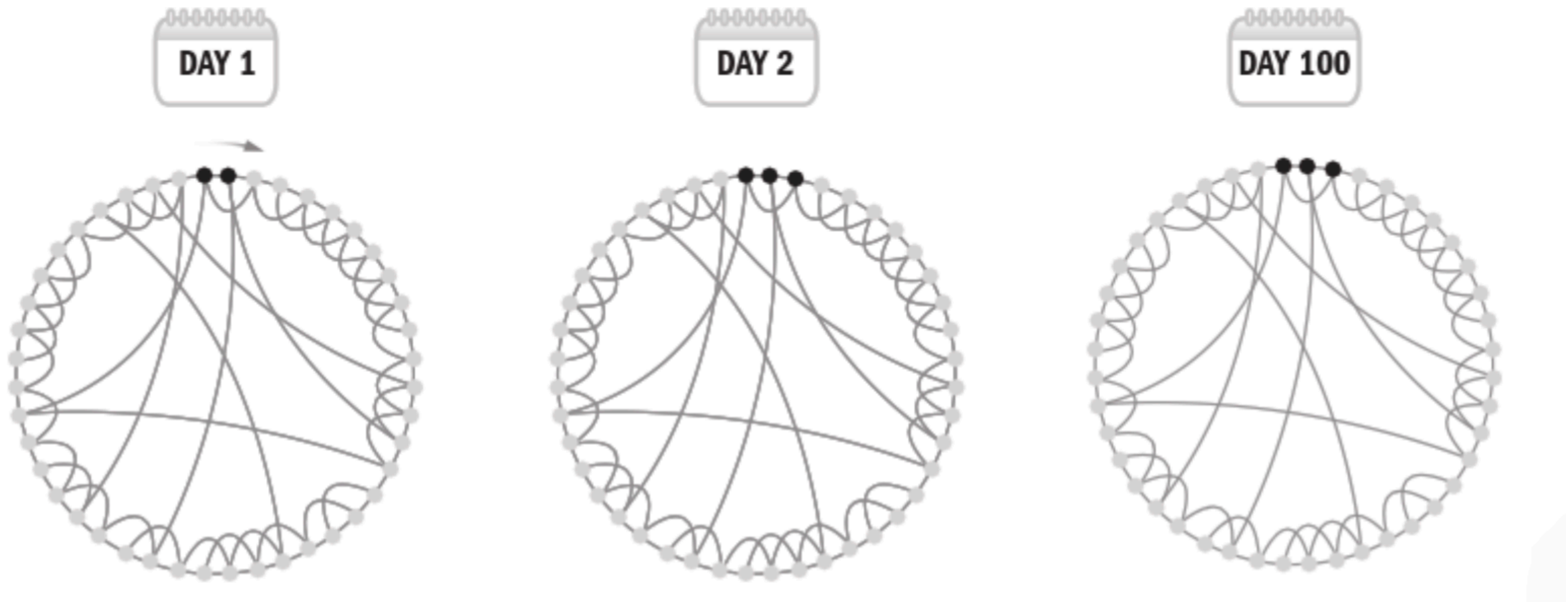


Complex Diffusion

Large world:

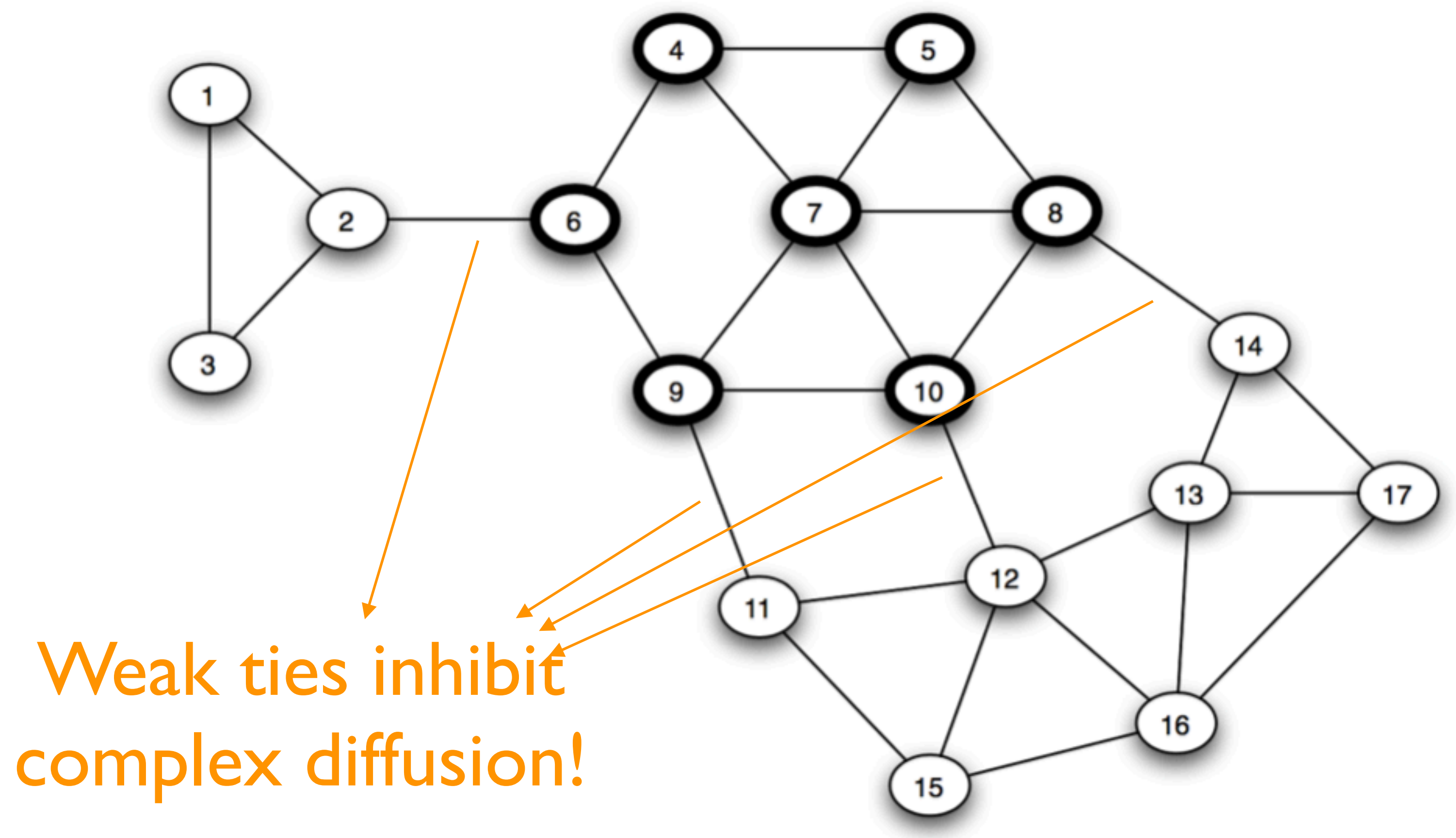


Small world:



Simple vs. Complex Diffusion

Weak ties are extremely useful for simple diffusion and contagion, but they inhibit complex diffusion!



Voting

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…

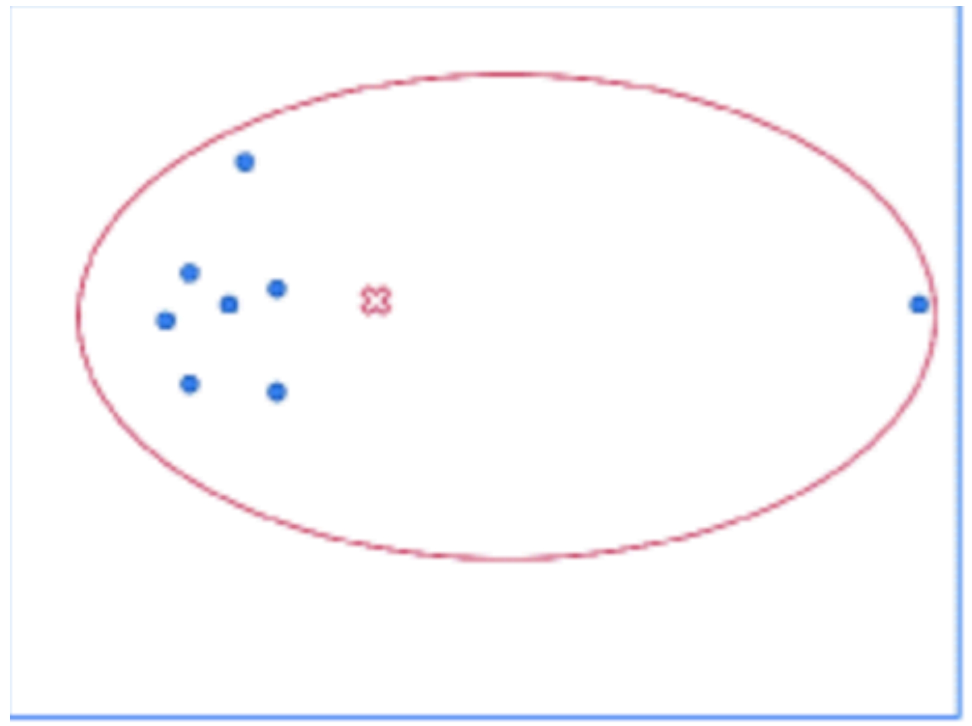
Why voting is hard

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

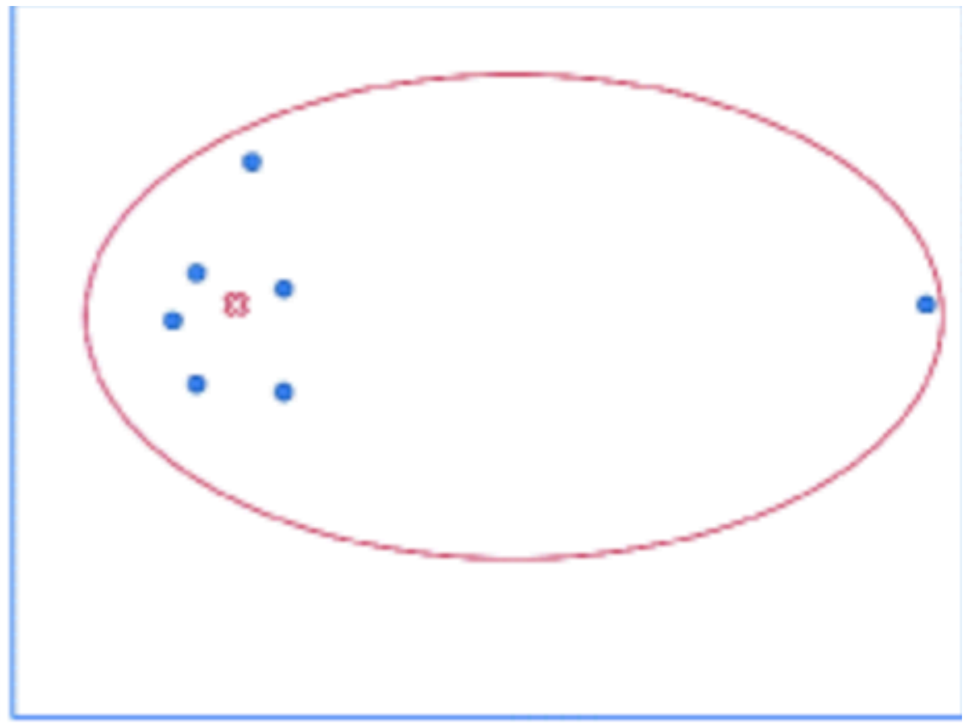
- President
- Award-winner
- Policy decision

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

 >  **OR**  > 

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

 >  **AND**  >  **=>**  > 

Individual preferences

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

Nodes: alternatives

Directed edges: $Y \longrightarrow X$ 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!

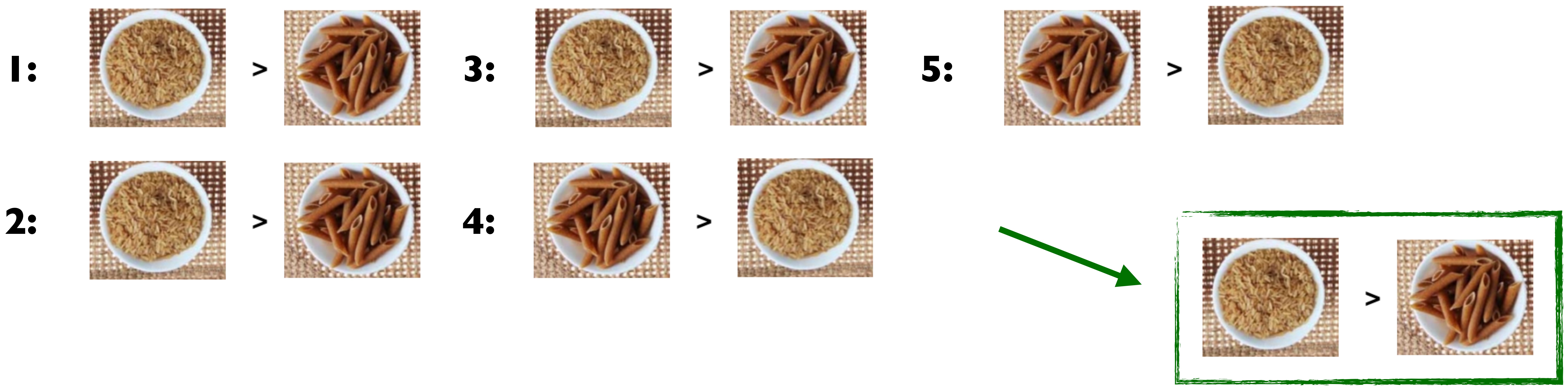
Voting Systems

Voting system: 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?

1:  $>_1$  $>_1$ 

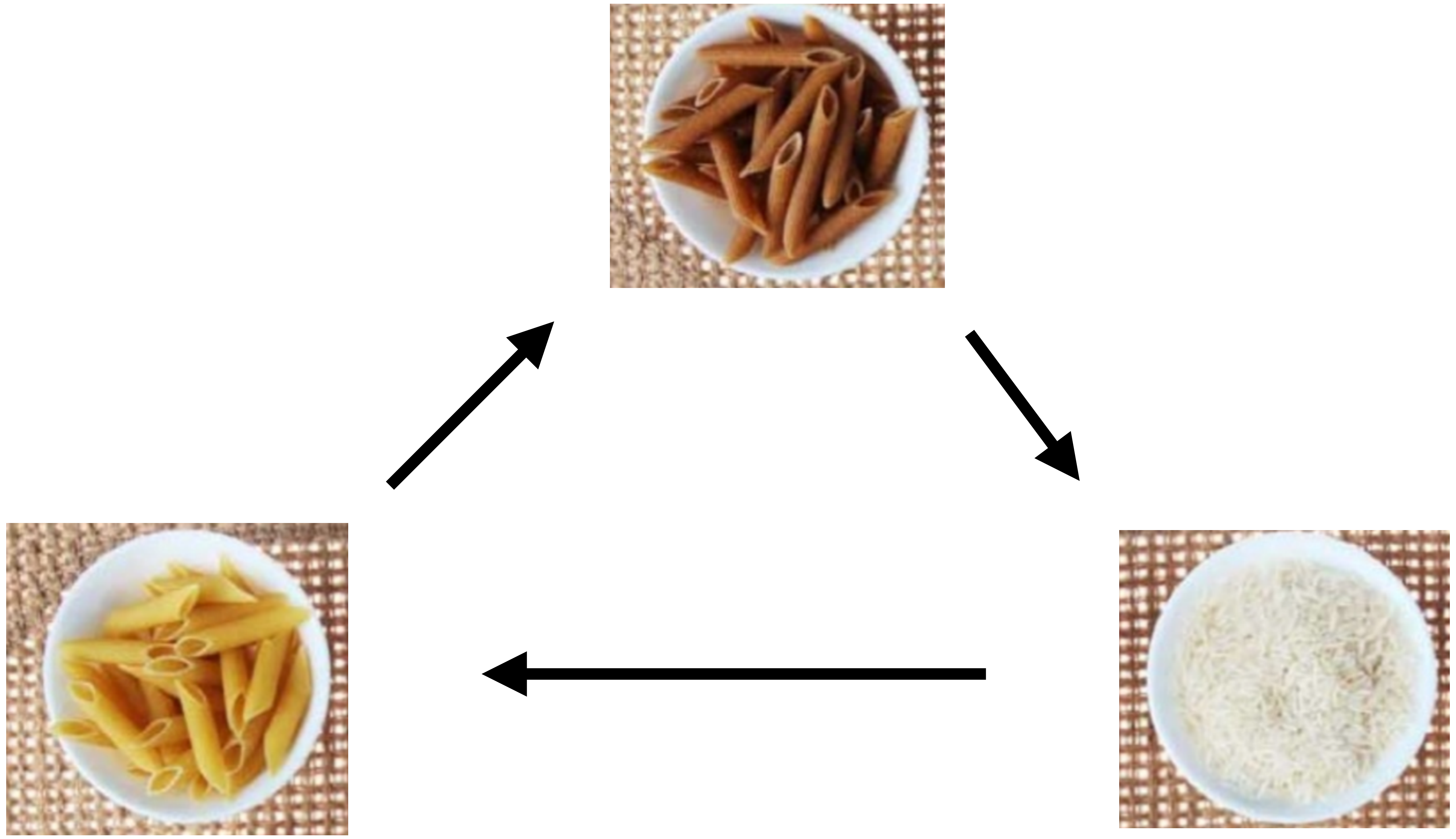
2:  $>_2$  $>_2$ 

3:  $>_3$  $>_3$ 

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!

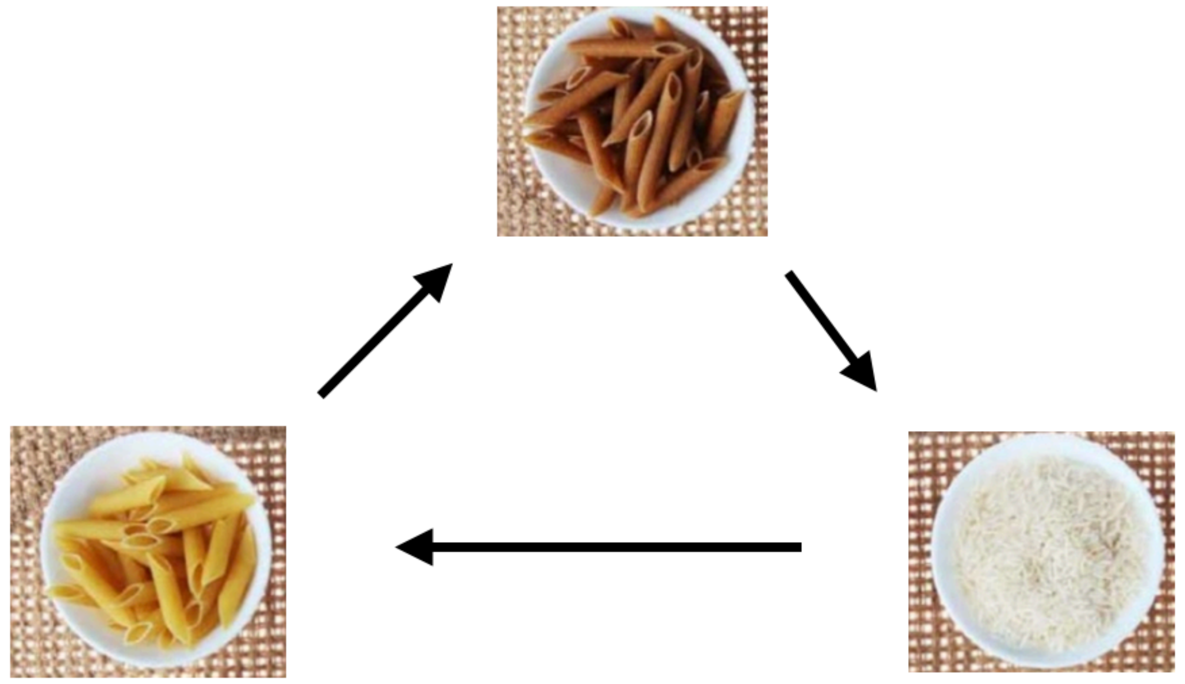
Condorcet Paradox

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

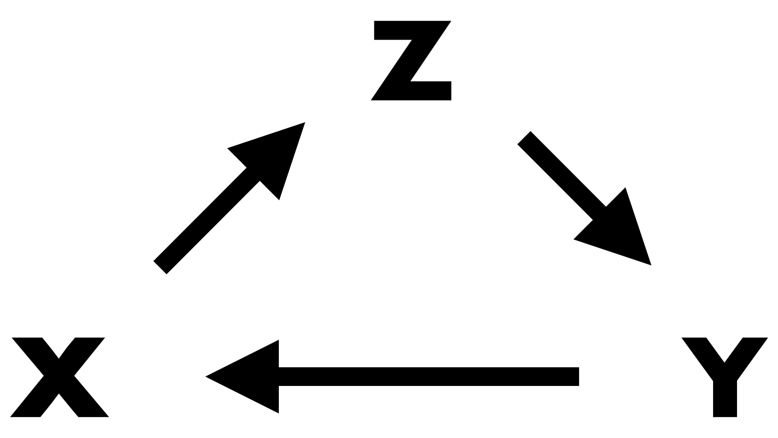
Condorcet Paradox can even happen within a single individual person

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.

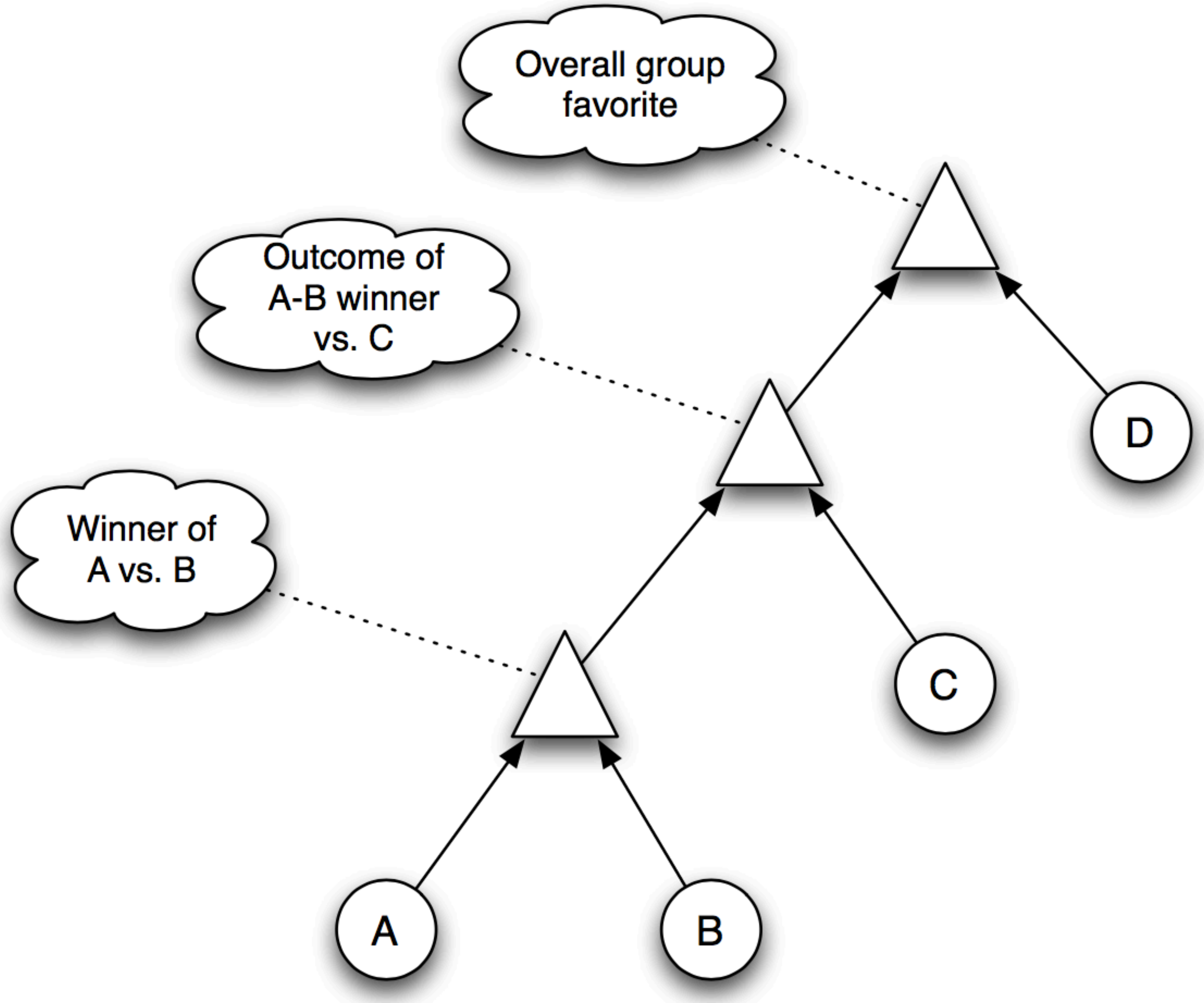
Winner of the final comparison is the group favourite

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

→ Then recurse!

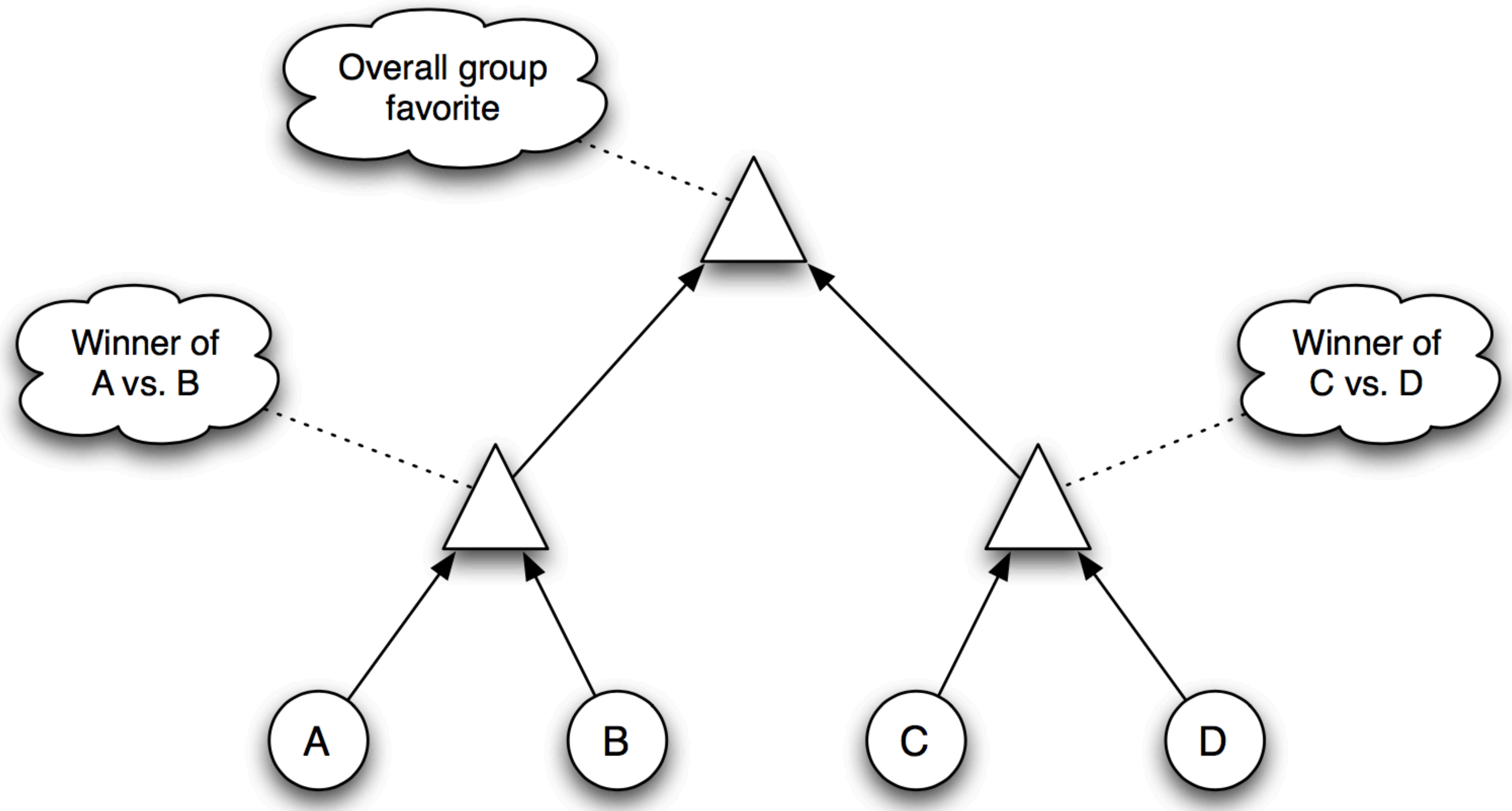
Majority Rule: Other Ideas

Graphically:



Majority Rule: Other Ideas

Other kind of elimination tournament:

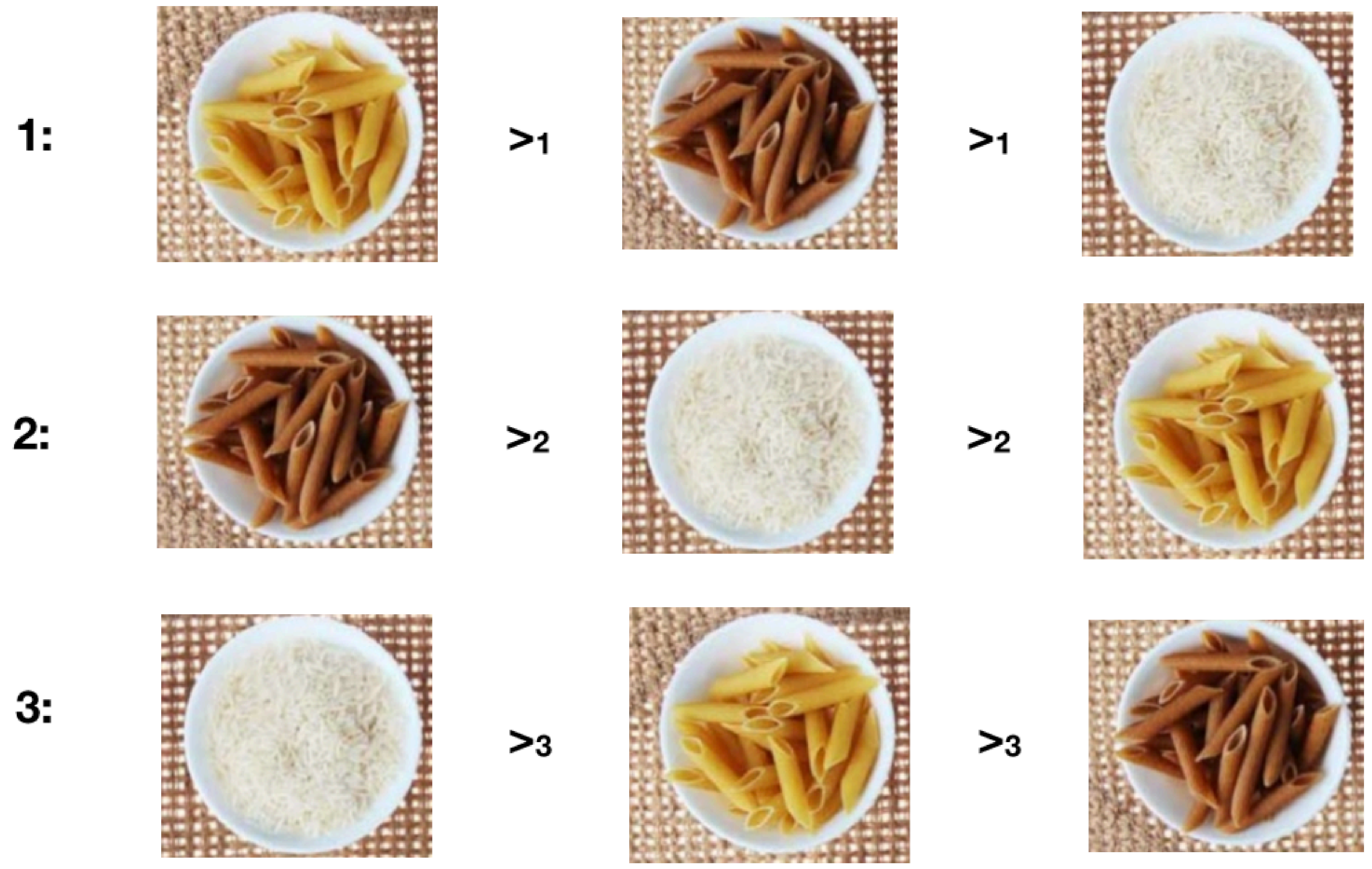


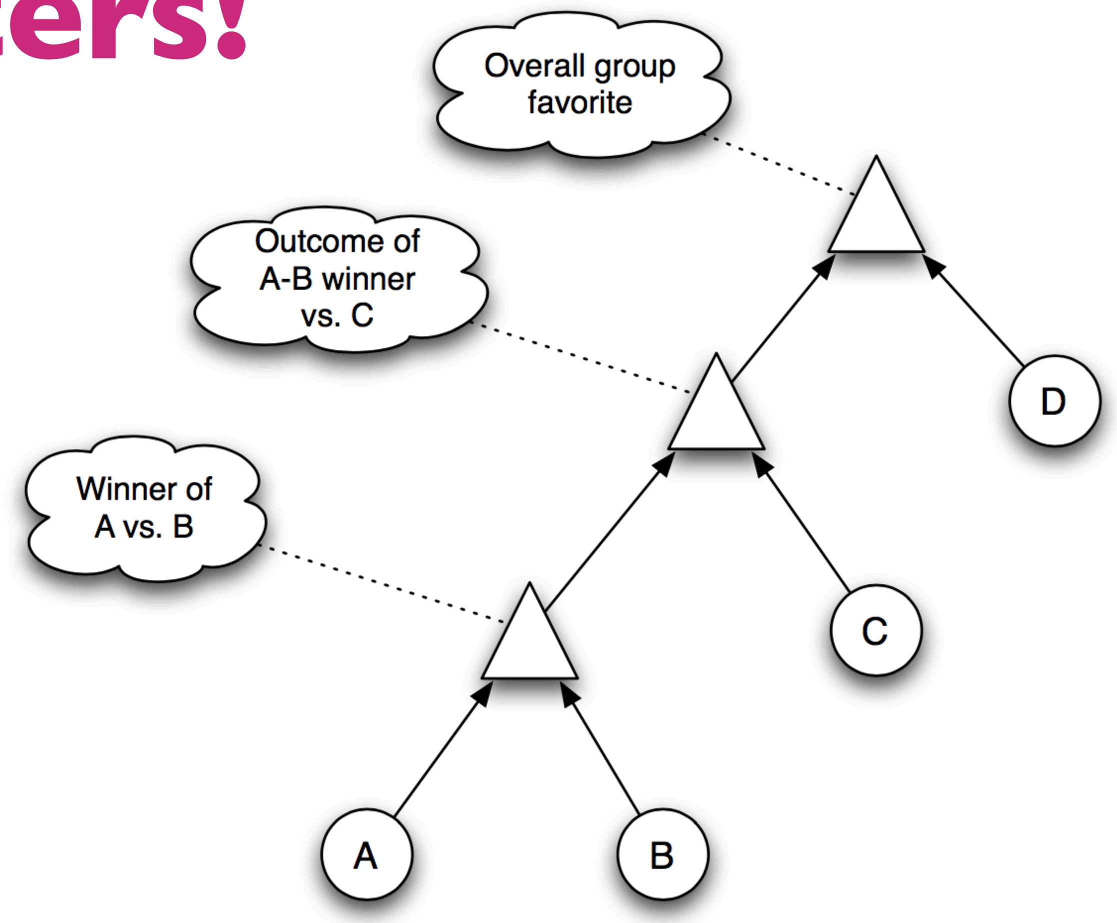
Majority Rule: Other Ideas

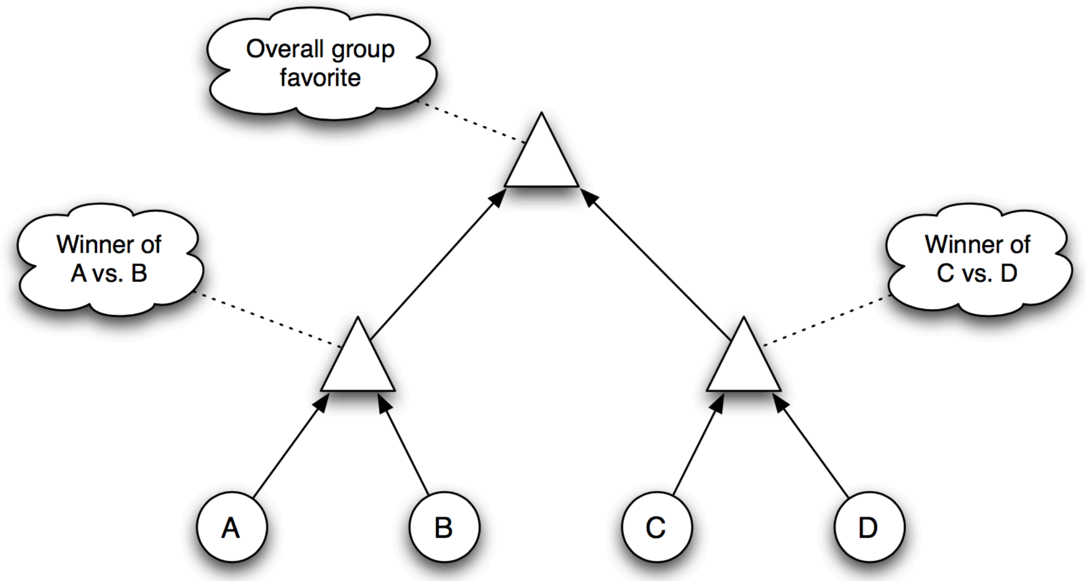
What's wrong with this?

Strategic agenda setting: order matters!

Consider example from before:



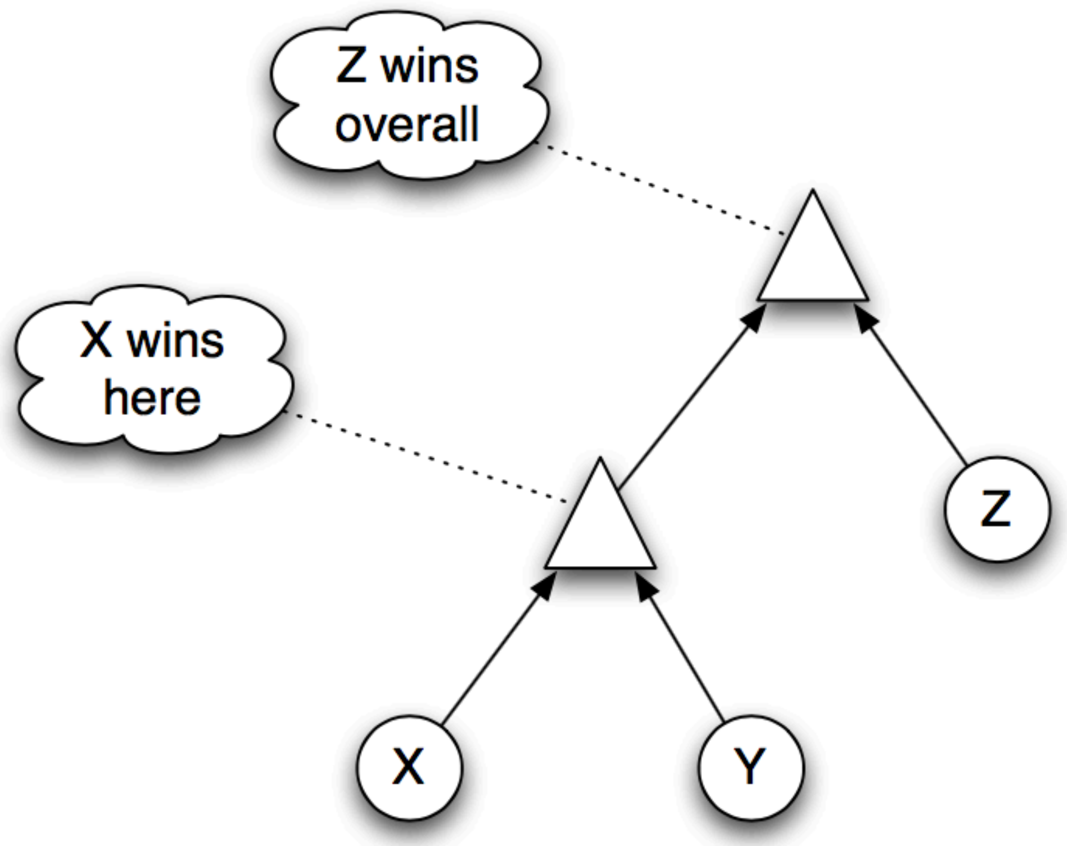


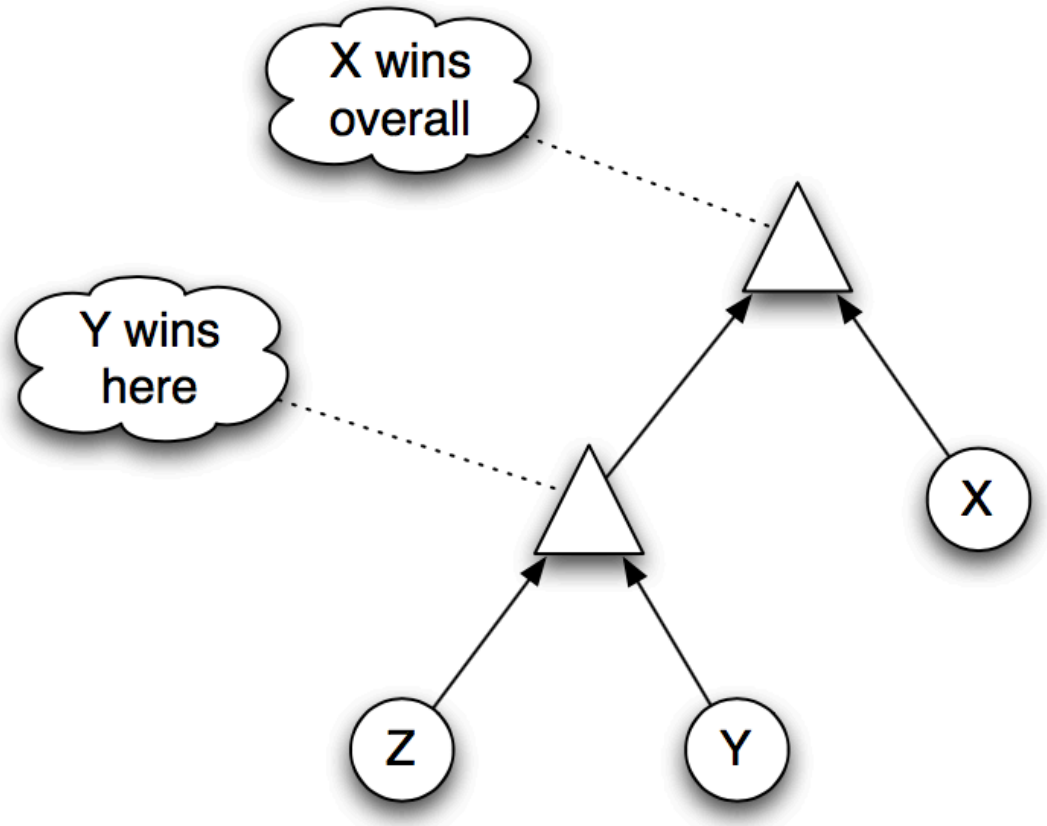


In what order do we evaluate the alternatives?

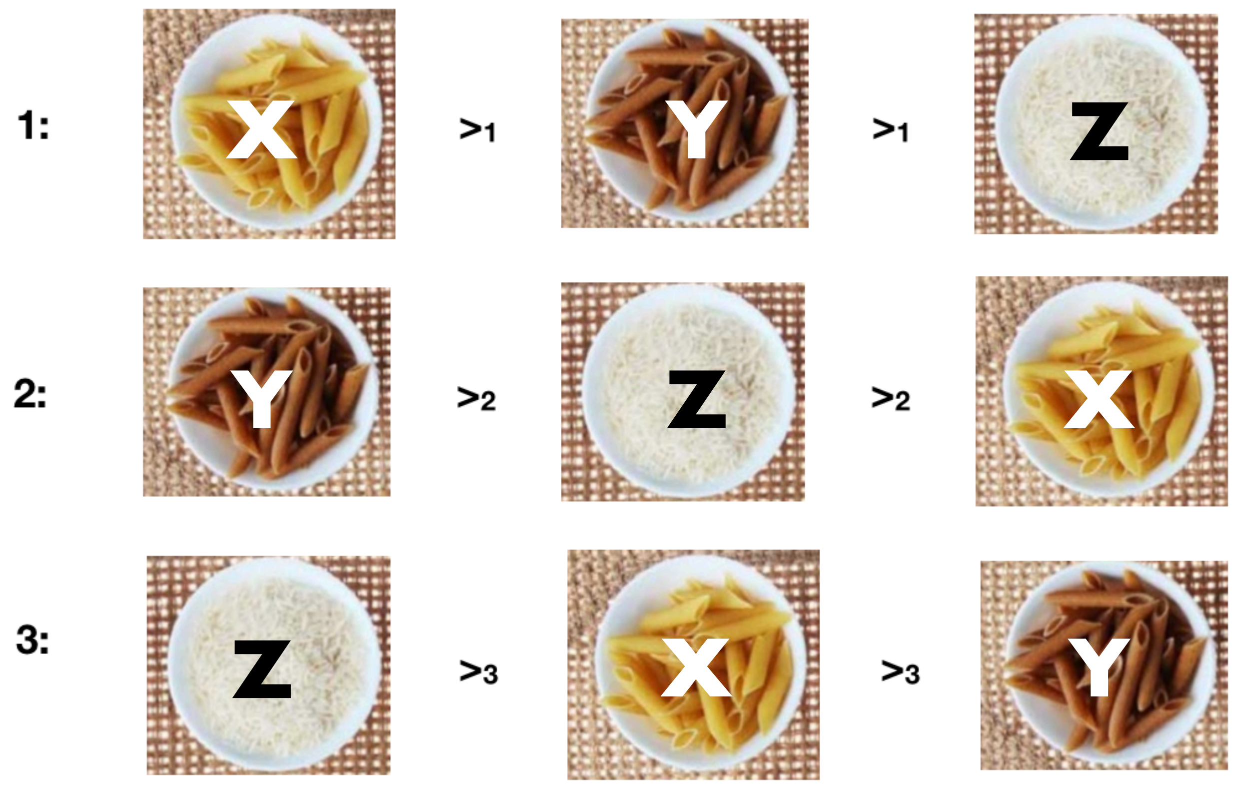
Majority Rule: Other Ideas

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?

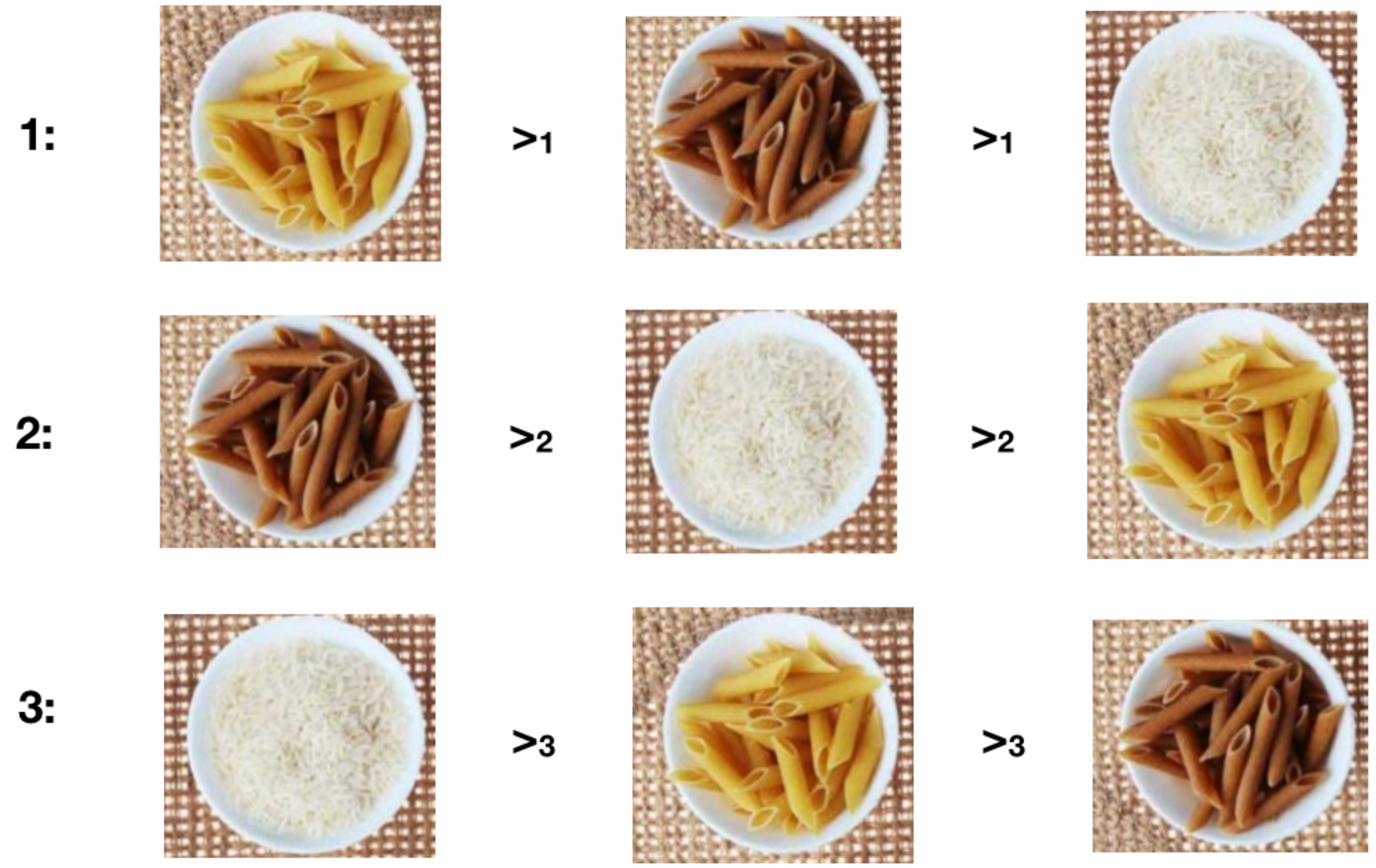
If you're faced with an impossibility result, you don't just give up

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**

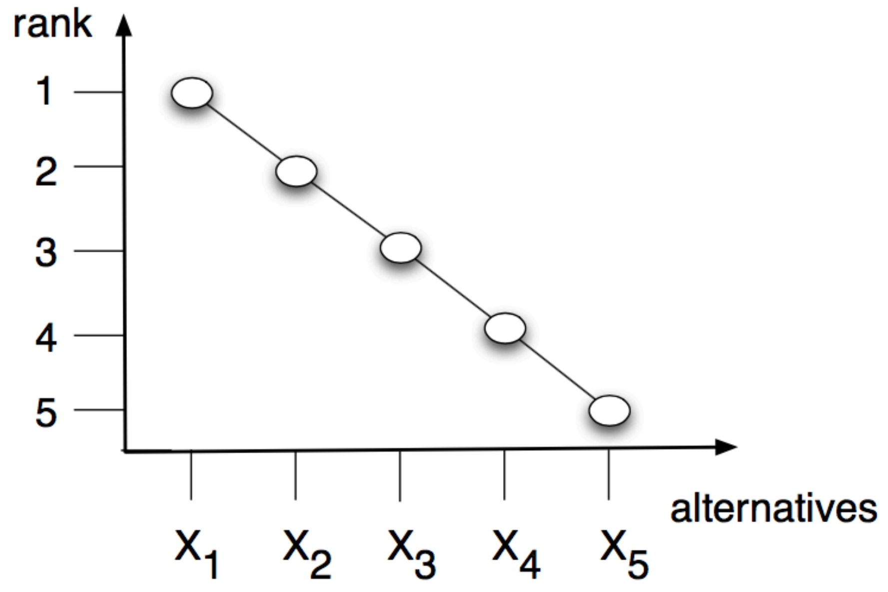
Not impossible, but they're more unusual

Ideal Points

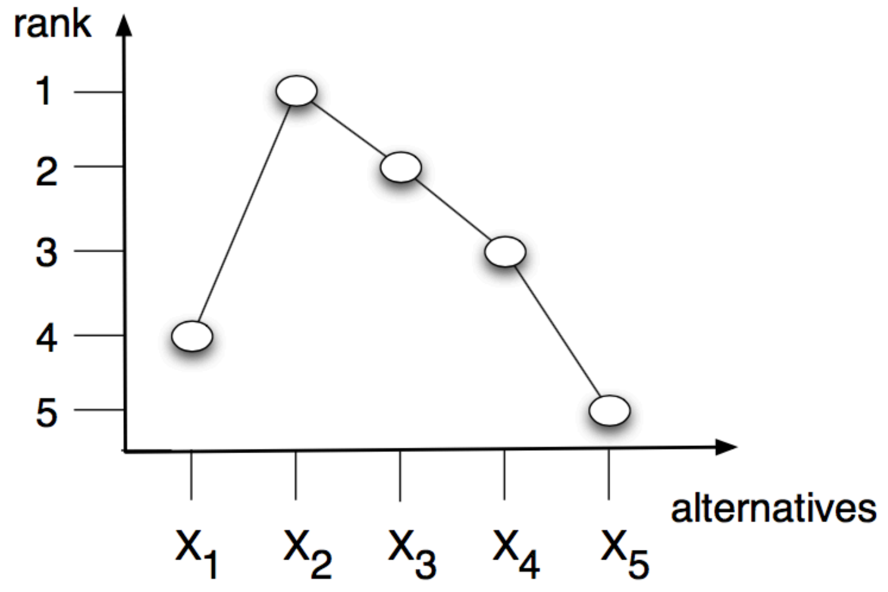
Assume the preferences lie on a one-dimensional spectrum, and each voter has an "ideal point" on the spectrum

They evaluate alternatives by proximity to this ideal point

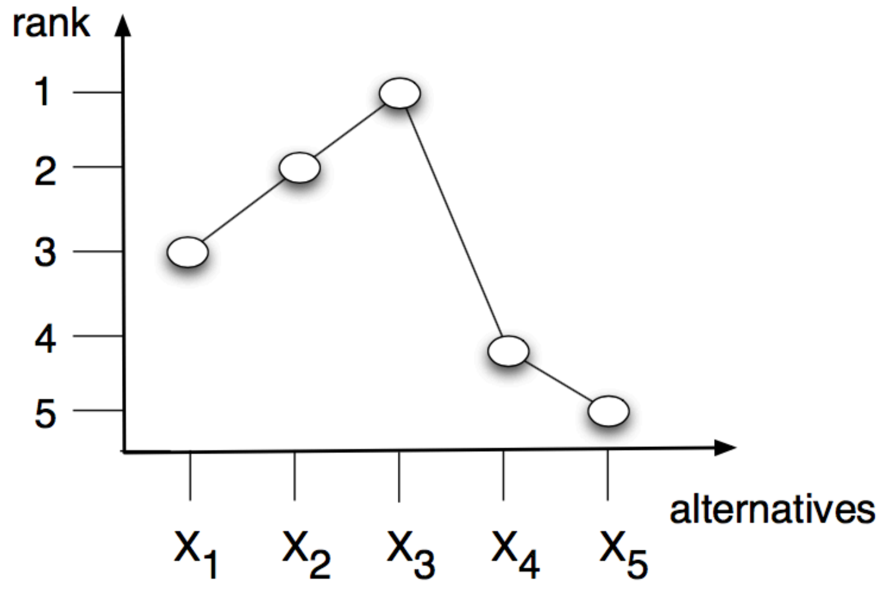
Actually we can assume something weaker: each voter's preferences "fall away" consistently on both sides of their favourite alternative



(a) *Voter 1's ranking.*



(b) *Voter 2's ranking.*



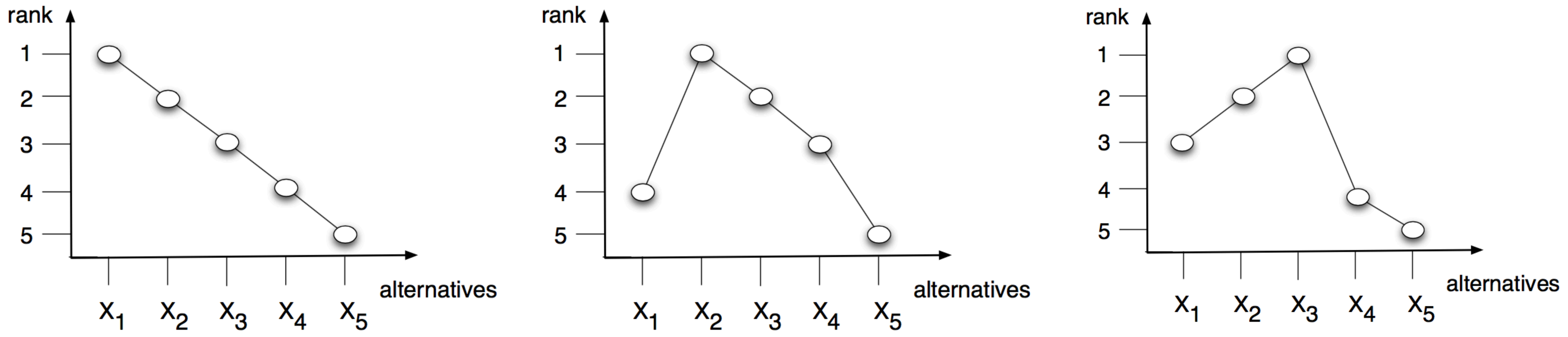
(c) *Voter 3's ranking.*

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!

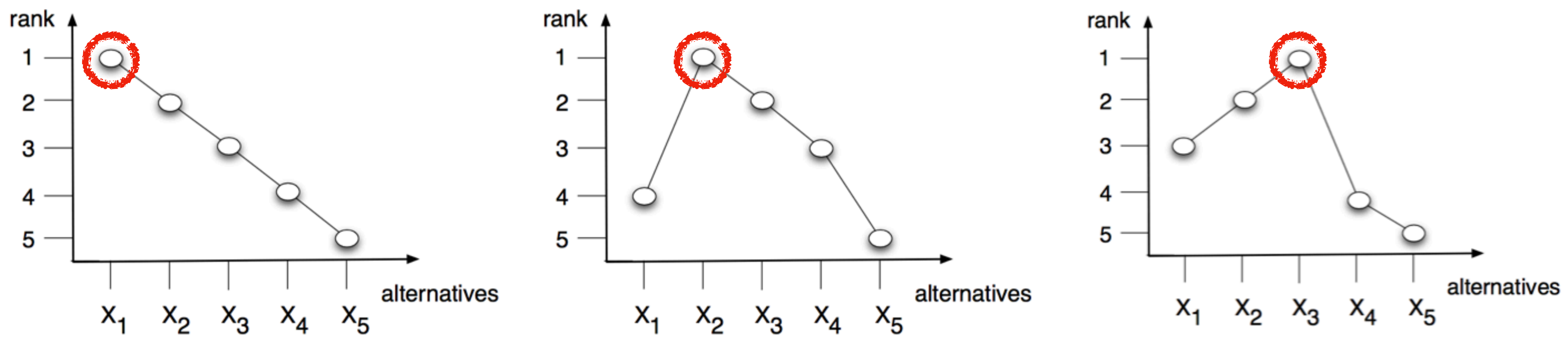
Proof

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

Proof

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.

Let X_m be the median individual favourite, and X_t some random

Assume $t > m$, and sort the voters in order of their favourites

Proof

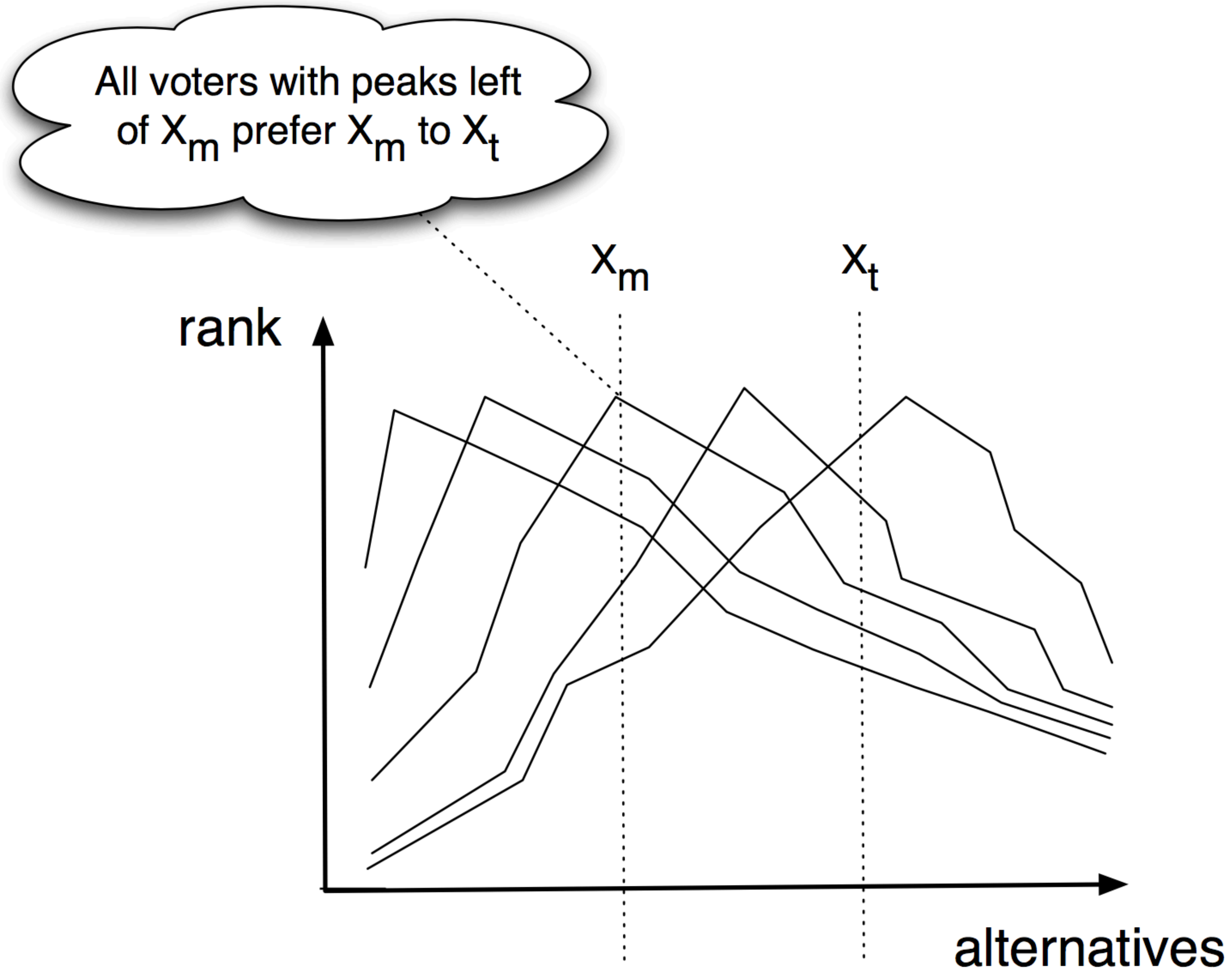
[remember k is odd]

By definition X_m is in position (k+1)/2 of sorted favourites

In the first (k+1)/2 positions, everyone either prefers X_m or their favourite is to the left of X_m

For everyone in the latter group, both X_m and X_t on the right-hand "down-slope" of preferences, but X_m is higher than X_t

So (k+1)/2 people prefer X_m to X_t — which is a strict majority!



Proof

So X_m can gather a majority of support against any other alternative X_t (case to the left of Xm is completely symmetric)

Build up group favourites one at a time: start by finding global favourite, then remove from individual rankings and recurse

Notice that removing it from the rankings still maintains the single-peaked property of everyone's preferences!

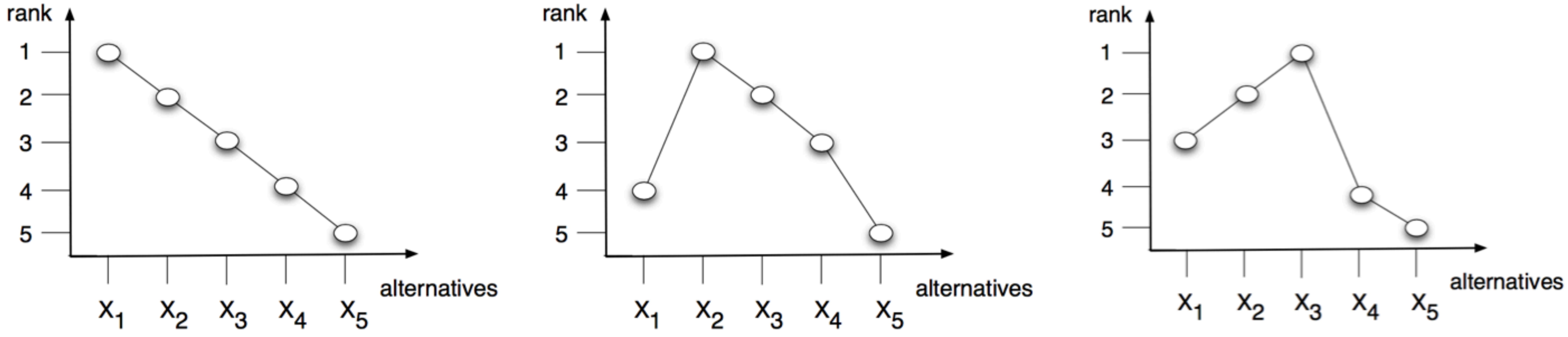
It just chops the peak off

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?

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

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

- Since $q>1/2$, $(1\text{-}q)^{k-2}$ 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

Arrow's Theorem

- "Voting profile" = set of k individual rankings
- 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**
  - **Another way to say it:** voting profile restricted to X and Y deletes everything except X and Y. Voting system should order X and Y same way on profile as on profile restricted to X and Y
- Criterion 3 **"Non-Dictatorship"**: the group ranking should not just always be what one particular voter thinks