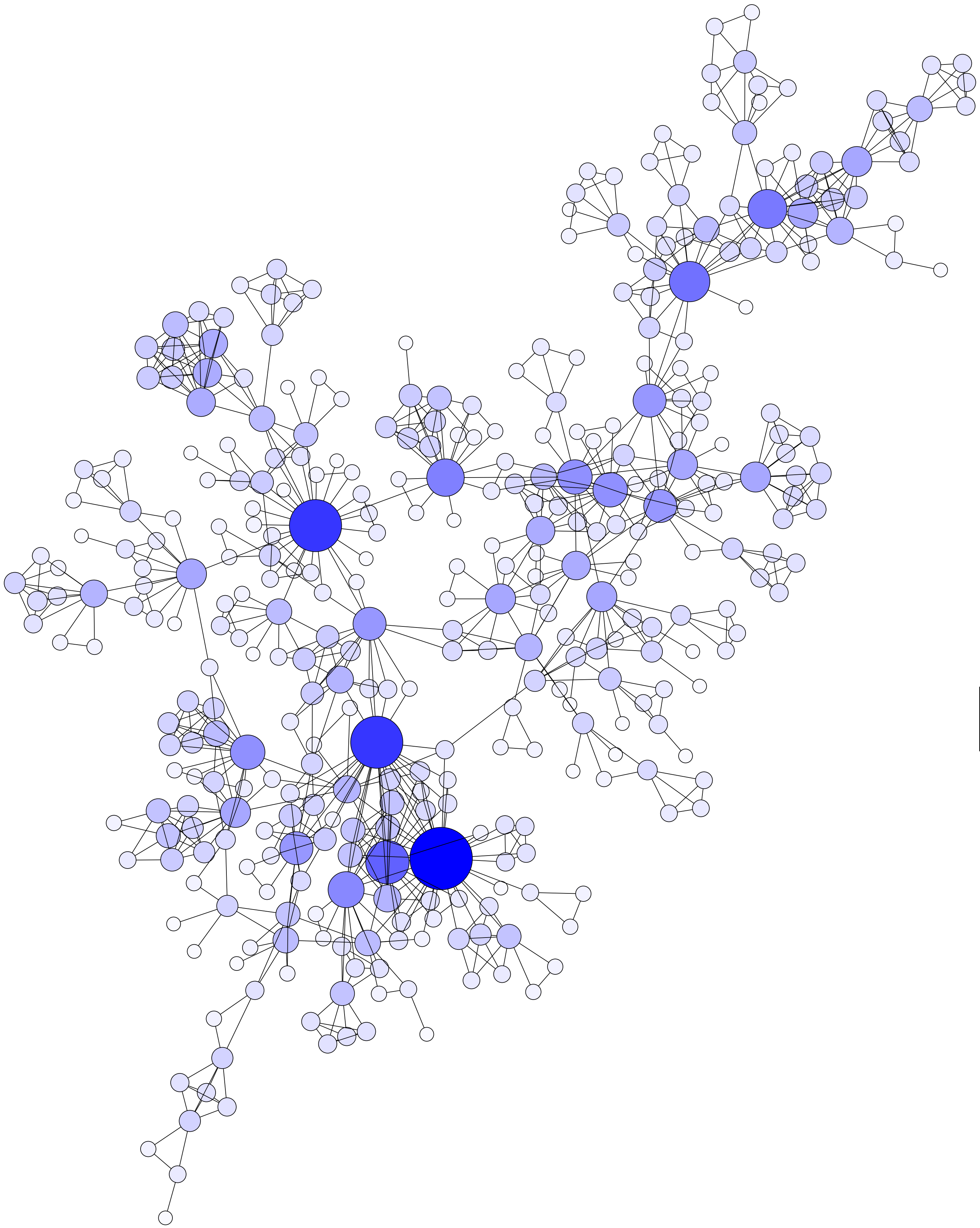
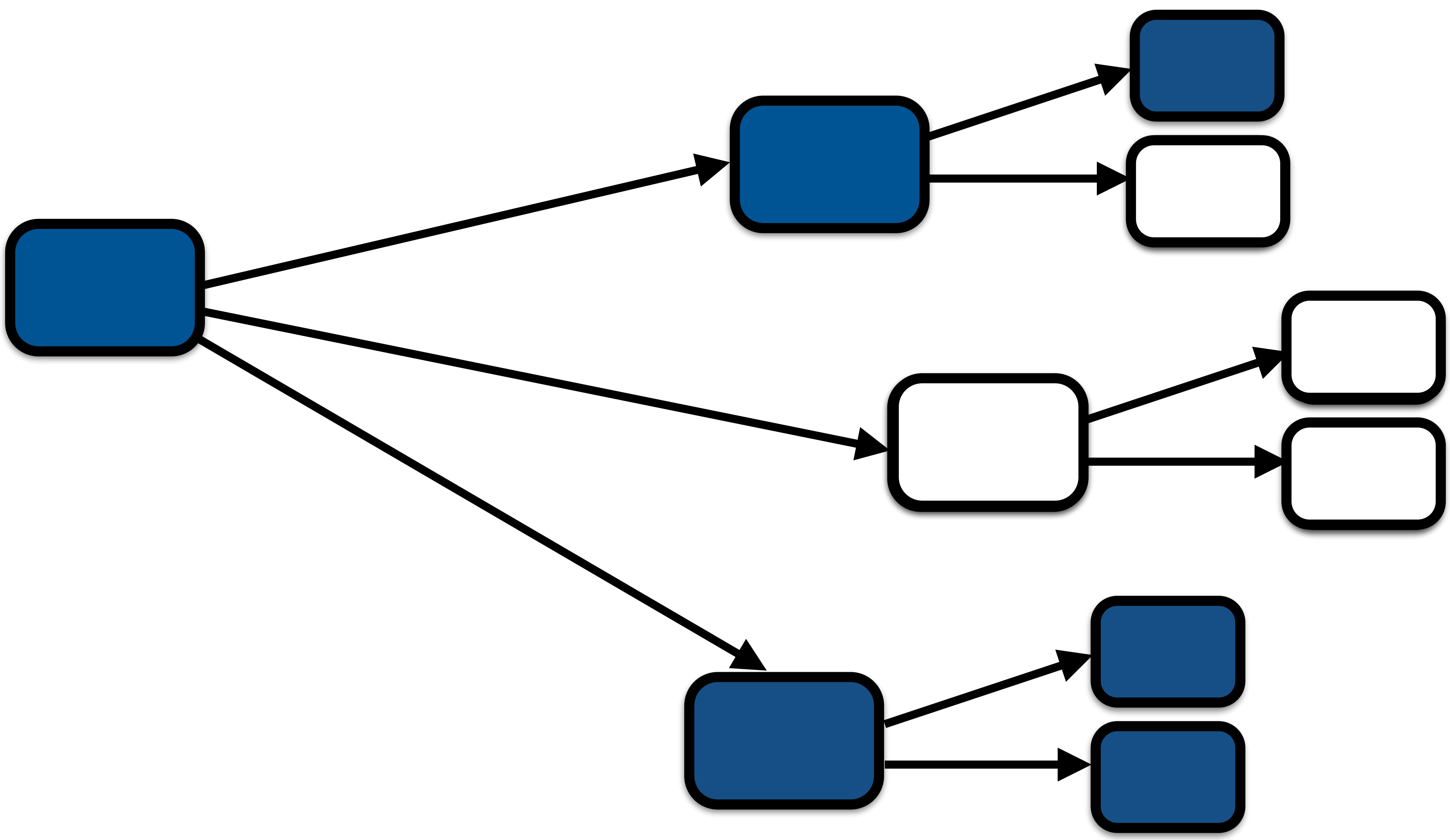


Diffusion on a Network (Part 1): Epidemics

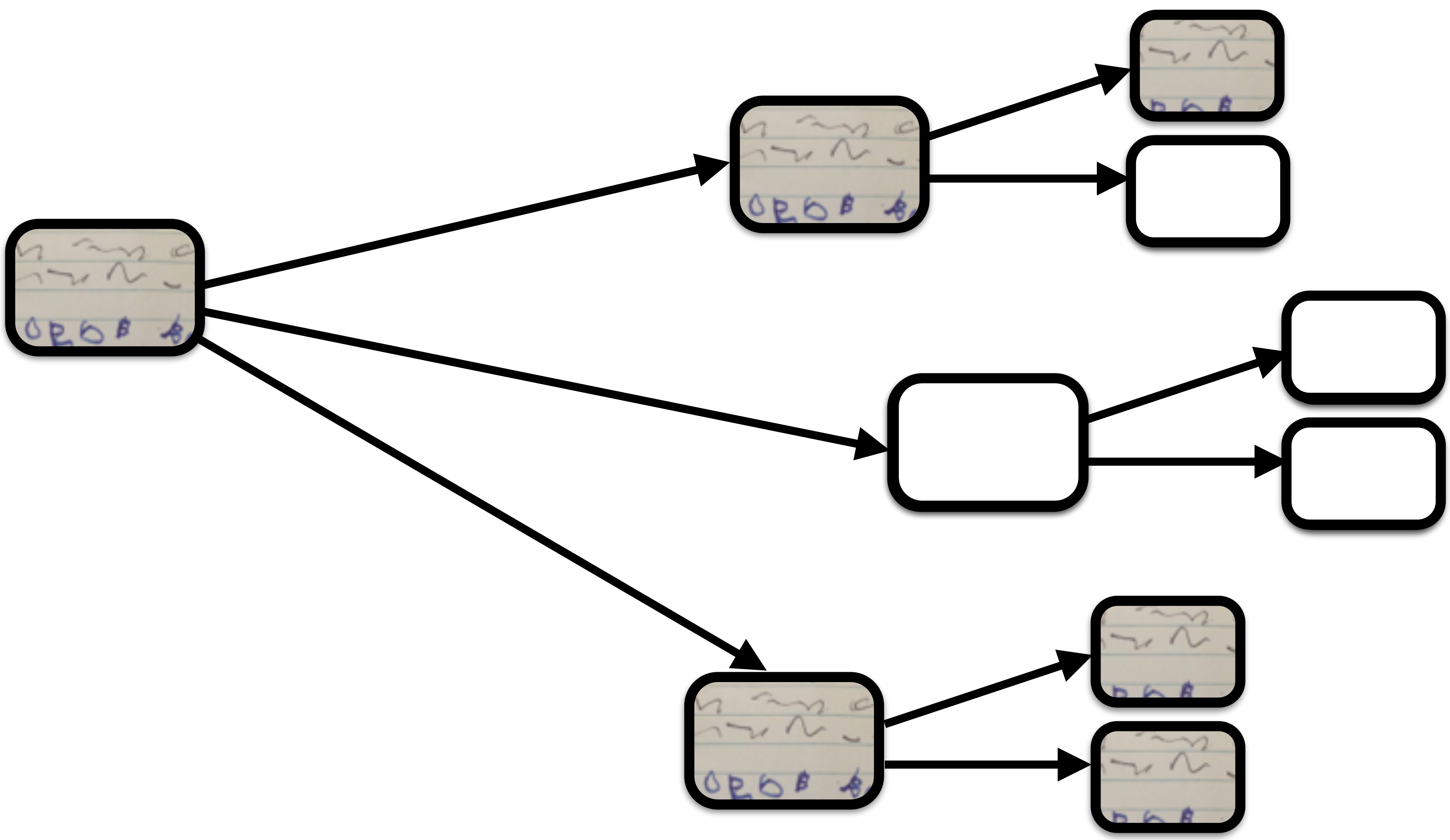
Diffusion

When we talk about things spreading through a population, we call it *diffusion*



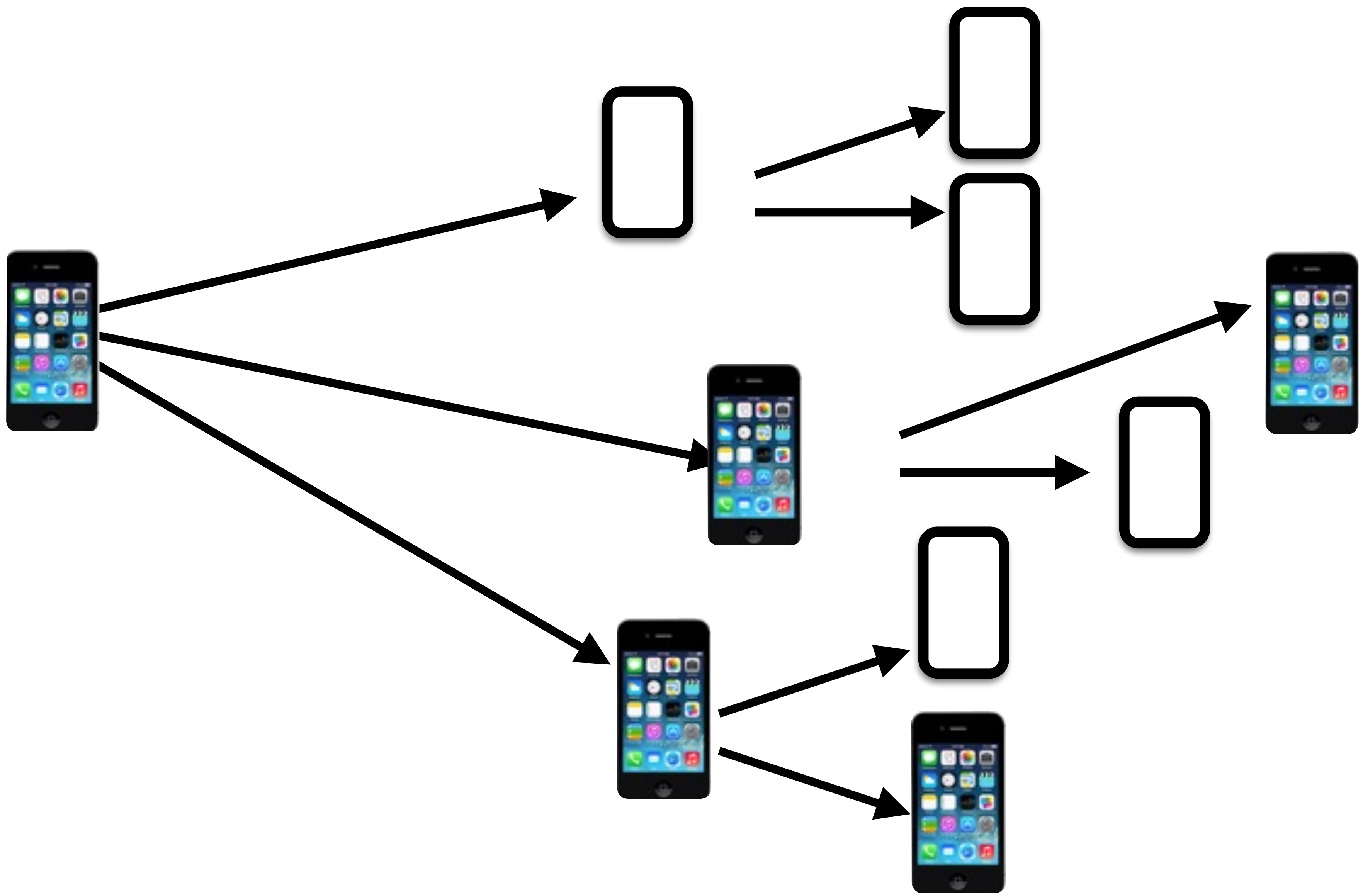
Diffusion

There are many different things that can diffuse over a social network…information



Diffusion

...technology...



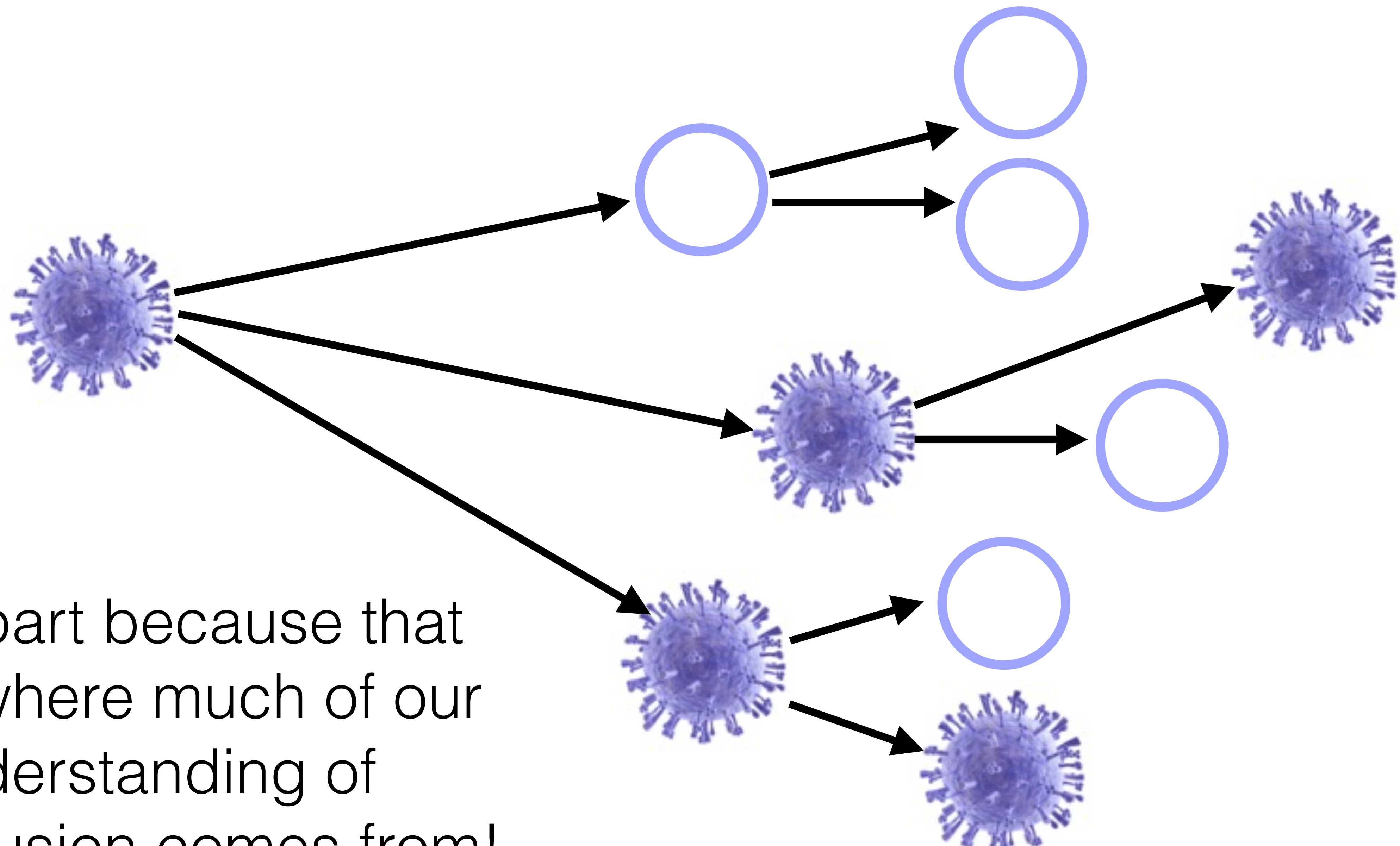
Diffusion

And lots of other things…

- Information
  - new tools and technology, job openings, stock tips, weather conditions
- Computer Viruses
  - Melissa virus, forwarding email spam
- Social Norms
  - emperor's new clothes, pedestrian crossings text message and email conventions
- Political Opinions
  - support for candidate, cause, or position
- New Technology
  - cars, telephones, micro-finance, building techniques and materials (water bottle lights), farming techniques, stoves
- Fads
  - hula hoops, beanie babies, parachute pants, big hair, "green" technology, local food, urban gardening
- Reputation
  - contractors, freelance workers online sellers

Diffusion

Today, we are going to talk about diseases…



In part because that is where much of our understanding of diffusion comes from!

Diseases: Epidemiology

Epidemiology: The study of the spread (and control) of disease

What epidemiologists want to know:

1. How is a given disease likely to spread in a given population, without any intervention?

2. What are the best methods to control the spread of the disease?

Diseases: Epidemiology

Oversimplifying (more than a little), how a disease spreads in a particular population depends on two things:

1. Characteristics of the disease (mutation rate, deadliness, transmission rate)

2. Characteristics of the network the disease spreads on (average degree, average distance, degree distribution)

Diseases: Epidemiology

What characteristics of the disease affect it's spread?

- How it is transmitted
- It's life cycle
- Deadliness Immunity
- Mutation Rate

airborn, sexually transmitted, contact with bodily fluids

incubation period, when contagious vs symptomatic, survival rate outside of host

what fraction of hosts die or gain immunity, how quickly do hosts die or recover

how long does immunity last?

Diseases: Epidemiology

Contrasting different viruses:

Influenza:
- Airborn
- Doesn’t live long on surfaces
- Host is contagious even when not symptomatic
- Not particularly deadly
- Hosts are mobile while infectious
- Mutates quickly

HIV:
- Blood-born
- Infectiousness depends on type of exposure
- Host is contagious even when not symptomatic
- Used to be very fairly deadly, but not any more
- Very slow to mutate

Ebola:
- Requires contact with bodily fluids
- Extremely infectious, given contact
- Can live for a long time on surfaces
- Extremely deadly
- Host can remain contagious after recovery
- Slow to mutate

Diseases: Epidemiology

Contrasting different viruses:

Influenza: airborne + host contagious without knowing it + not deadly + quick to mutate = frequent, widespread outbreaks, difficult to control

HIV: blood-born + host contagious without knowing it + not deadly + slow to mutate = size of outbreaks depends on external factors (eg: awareness), possible to control

Ebola: direct contact + not contagious unless symptomatic + very deadly + slow to mutate = infrequent outbreaks, tend to be easily controlled, but crops up again and again

Diseases: Epidemiology

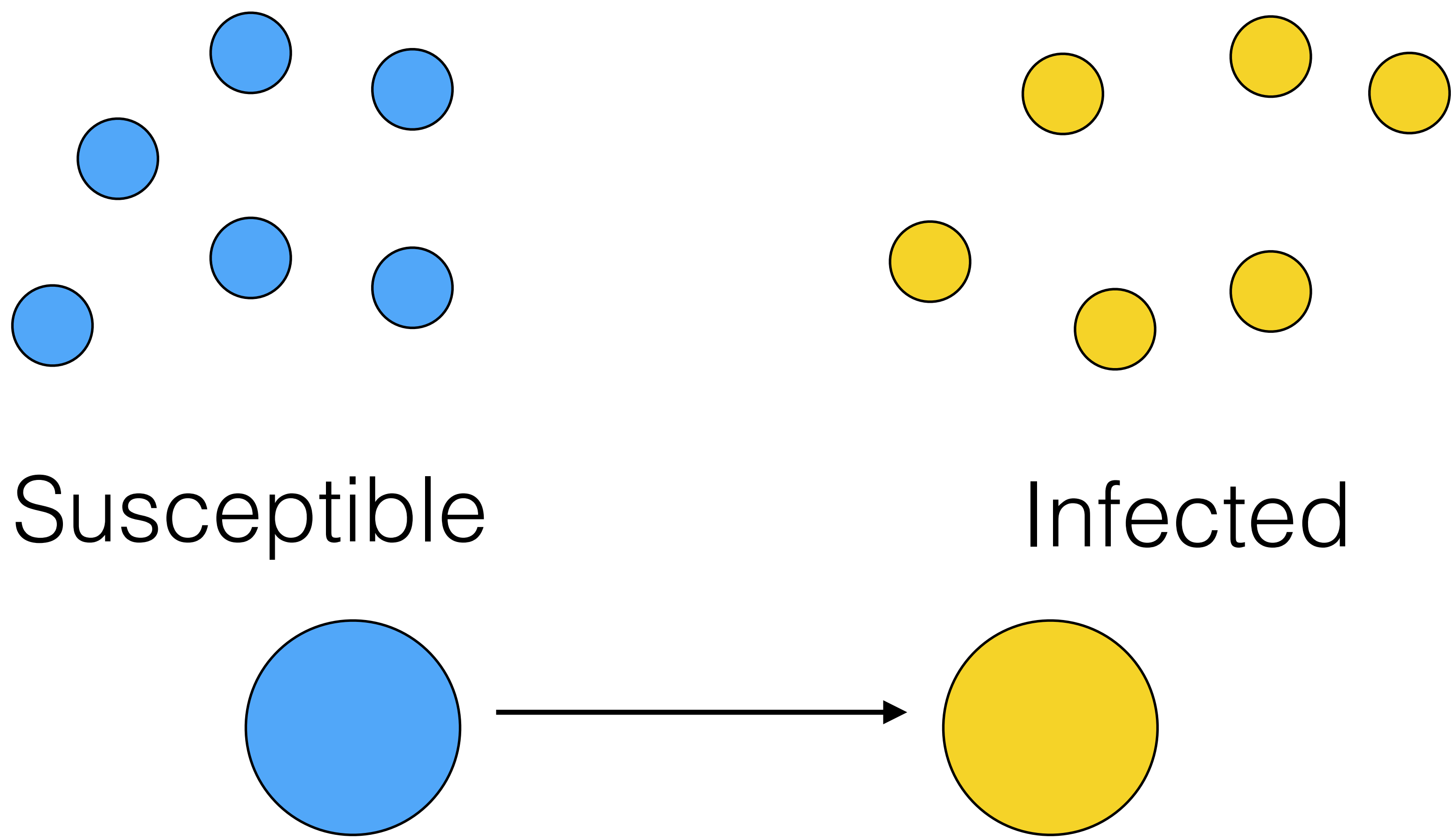
Suppose we are modeling the spread of flu on campus:

- To ignore the network, assume that every day you interact with a random group of people.
- Each day, you have the same probability of interacting with any other person in the university
- This is called “perfect mixing”

It is the basis of a classic model of disease spread: the *SI model*

Diseases: The SI Model

People fall into one of two groups:



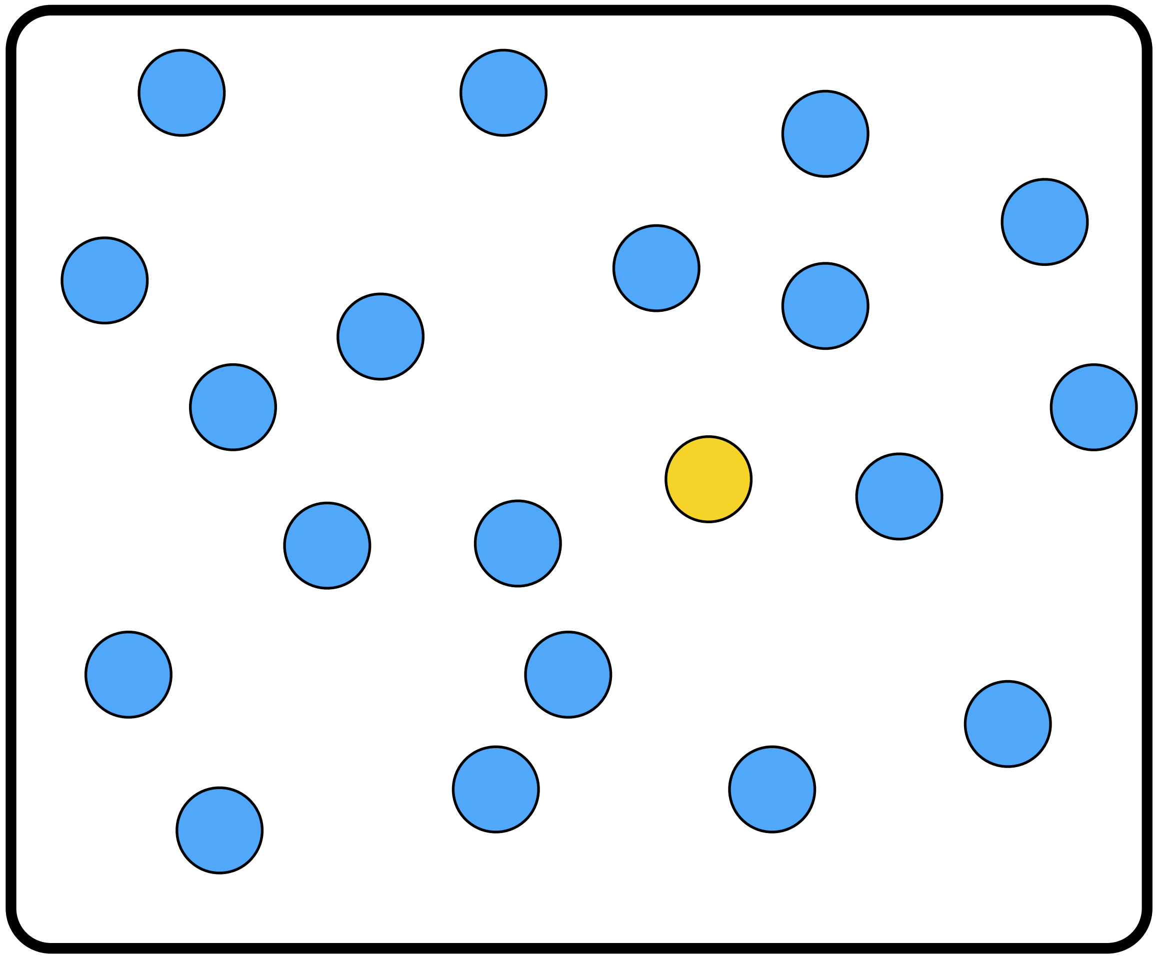
People can move from being susceptible to being infected

Diseases: The SI Model

Population of N people

Initially, I_0 people are infected

The rest are susceptible to infection: $S_0 = (N - I_0)$

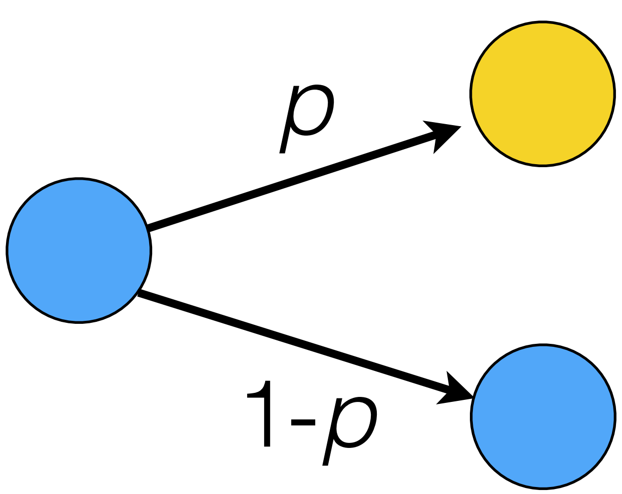


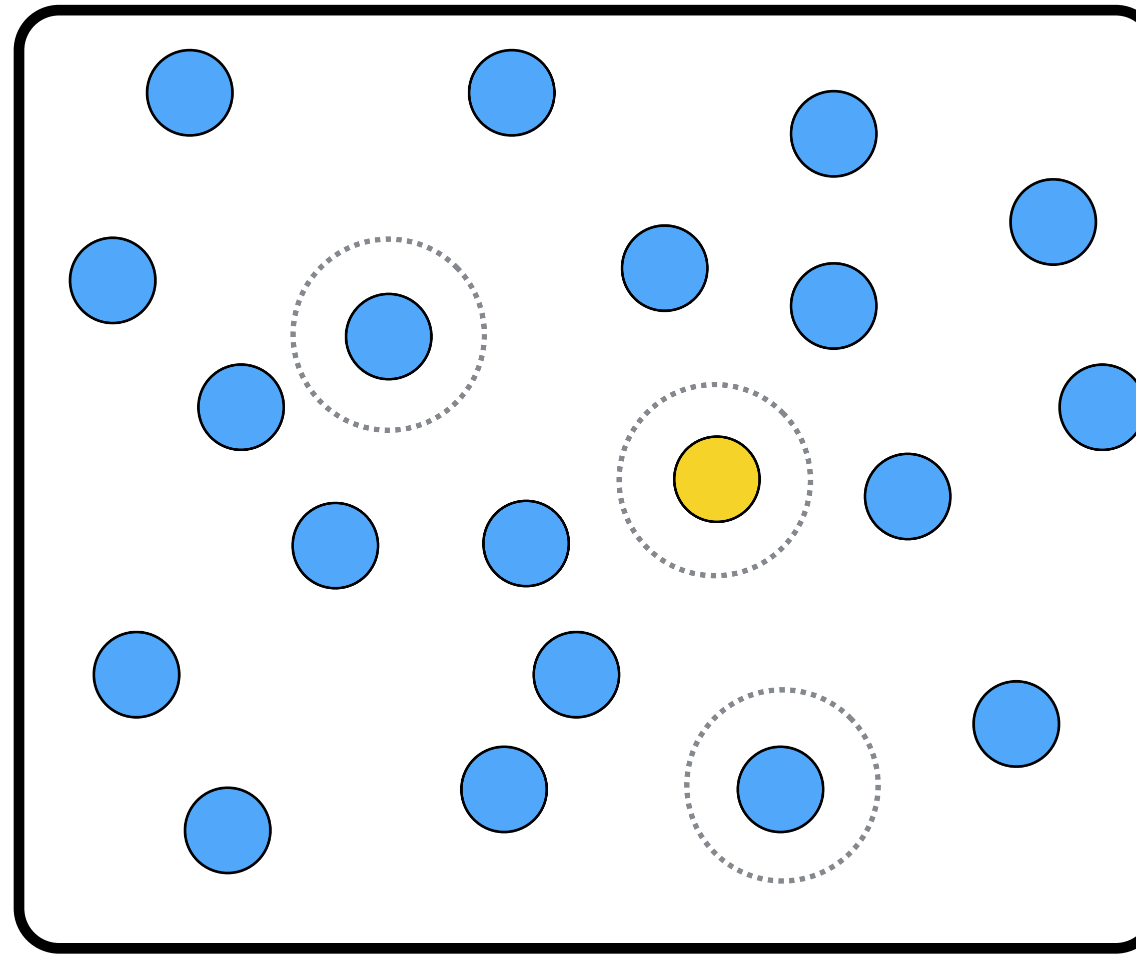
$N = 20$
$I_0 = 1$
$S_0 = 19$

Diseases: The SI Model

In each time period, you encounter *c* other people at random

For each infected person you encounter, you become infected with probability *p*





With probability (*1-p*), you aren't infected, and remain susceptible

Diseases: The SI Model

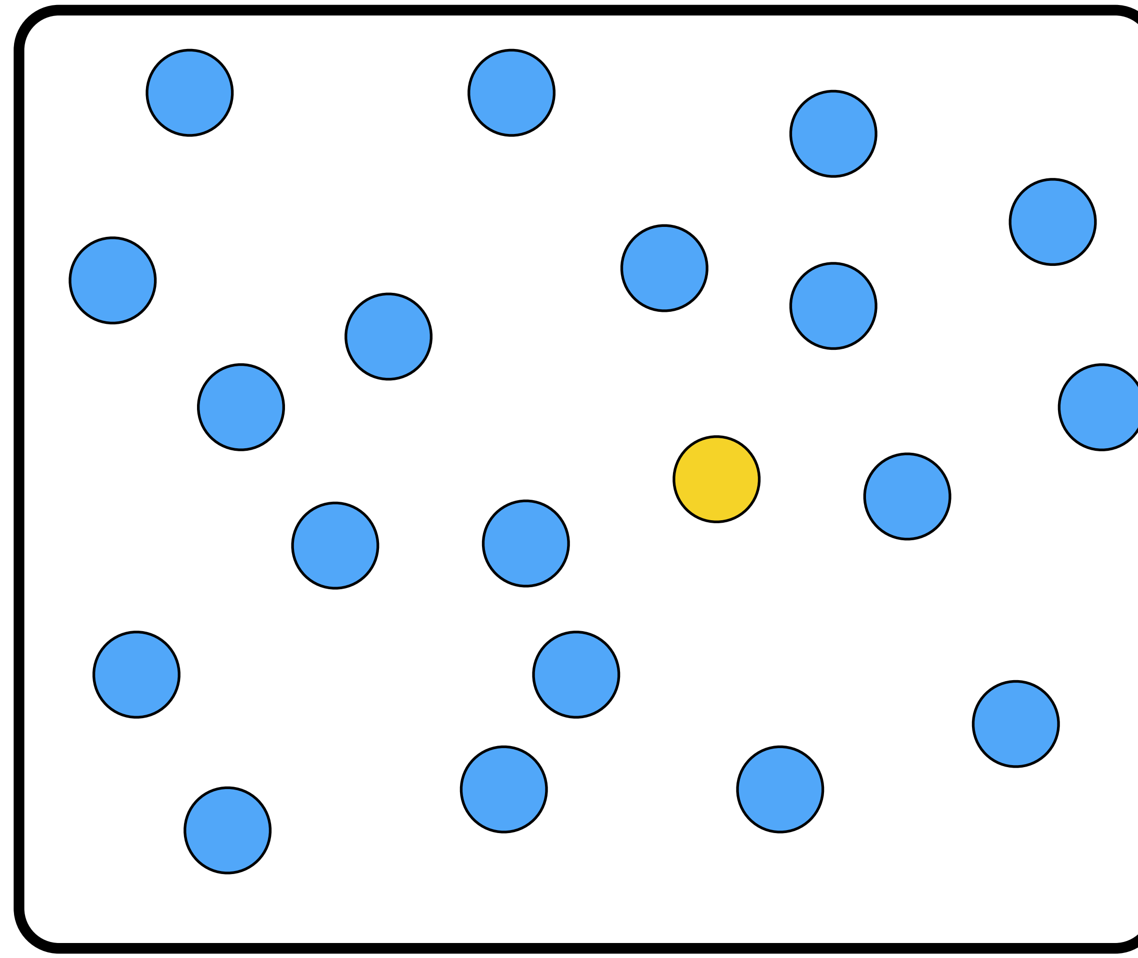
In the first period, the probability that you get infected = prob(meet an infected individual) * prob(they infect you)

$$c\left(\frac{I_0}{N}\right)p$$

number of people you meet

chance any one of them is infected

prob you catch the disease, given that you meet someone who is infected



Diseases: The SI Model

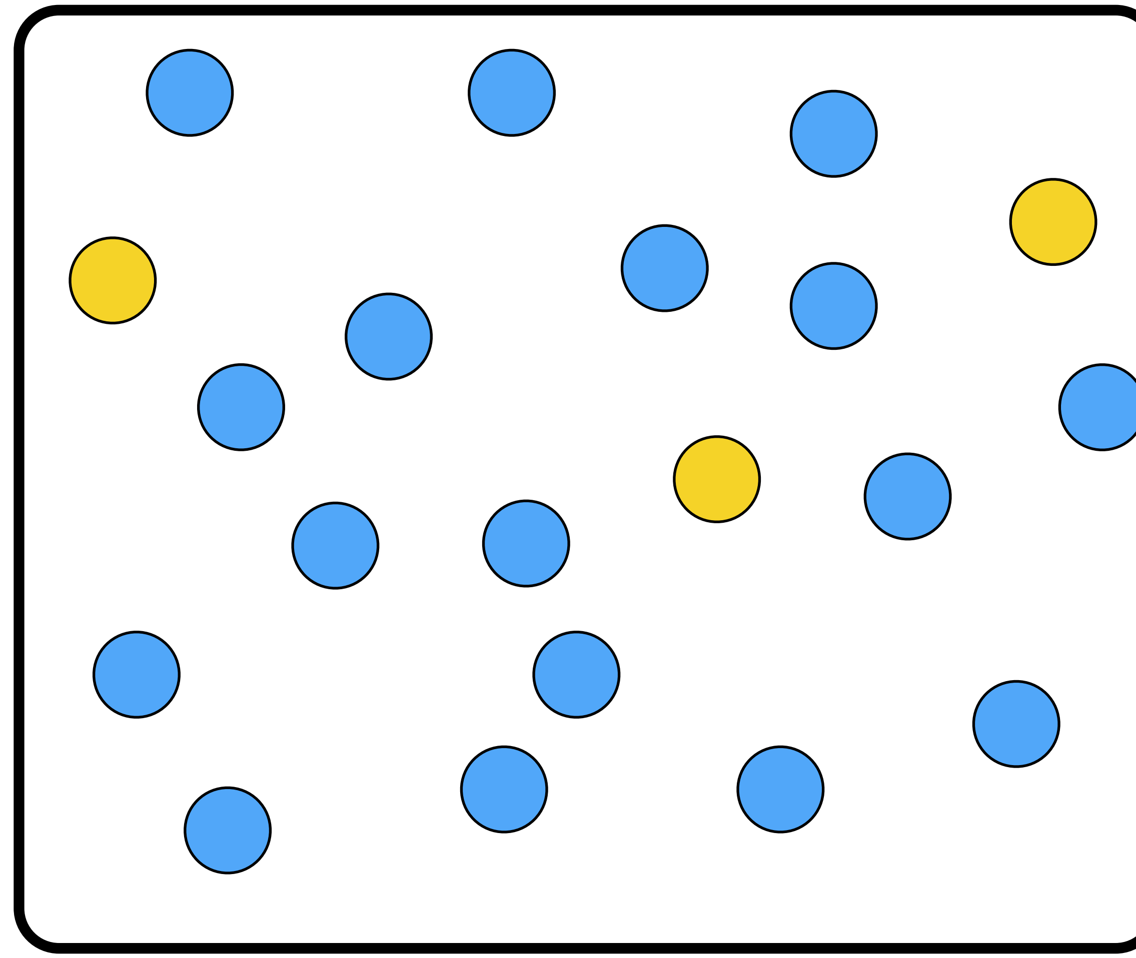
So the total number of people infected after the first period is:

$$I_1 = S_0 \; c \left(\frac{I_0}{N}\right) p + I_0$$

number of people who could be infected

chance any one of them is infected

number already infected



So if c = 3 and p = .66, then

$$I_1 = 19 \left(3 \left(\frac{1}{20}\right) \left(\frac{2}{3}\right)\right) + 1 = 2.9$$

Diseases: The SI Model

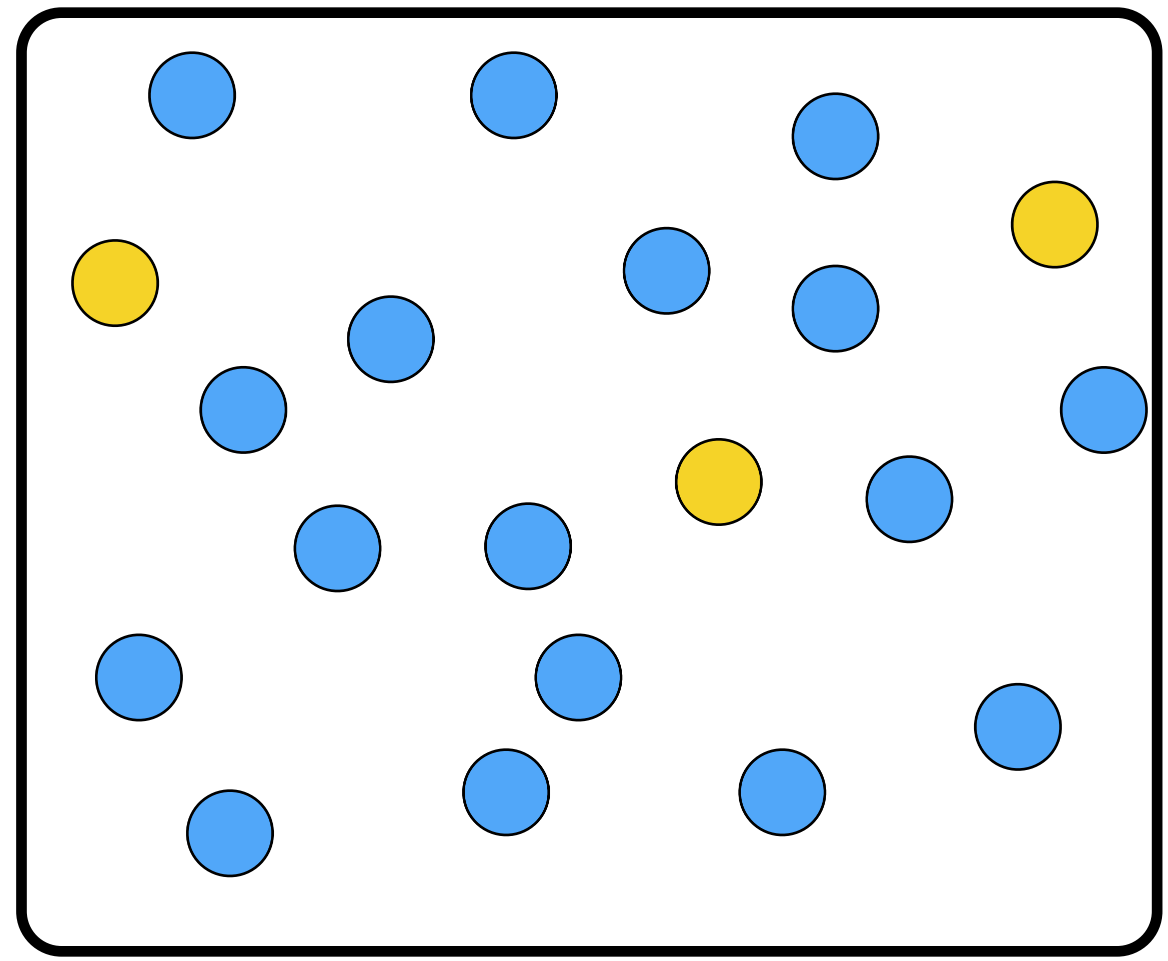
Now repeat…the chance you get infected in the second period is

$$c\left(\frac{I_1}{N}\right)p$$

c: number of people you meet

$\frac{I_1}{N}$: chance any one of them is infected

p: prob you catch the disease, given that you meet someone who is infected



Diseases: The SI Model

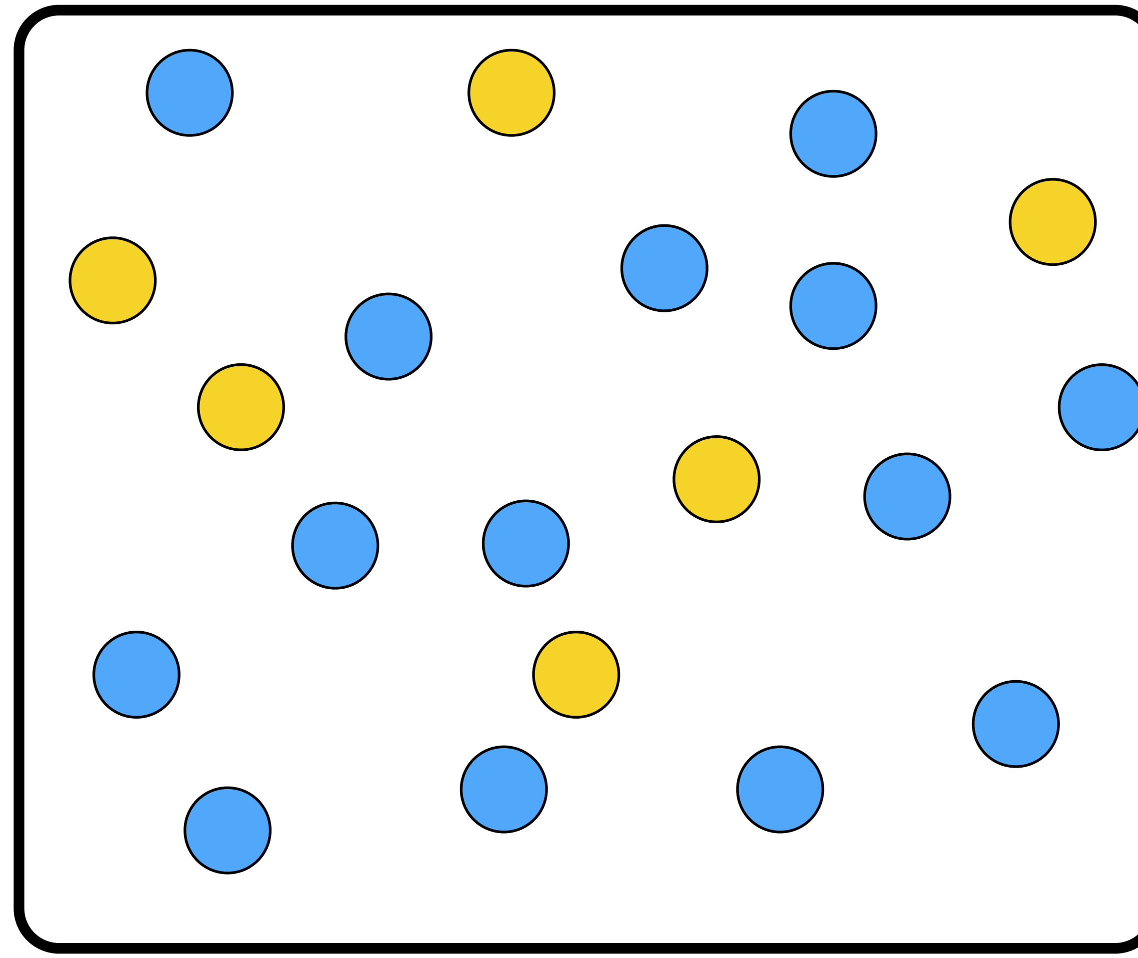
So the total number of people infected after the second period is:

$$I_2 = S_1 c \left(\frac{I_1}{N}\right) p + I_1$$

number of people who could be infected

chance any one of them is infected

number already infected



So if c = 3 and p = .66, then

$$I_2 = 17.1\left(3\left(\frac{2.9}{20}\right)\left(\frac{2}{3}\right)\right) + 2.9 = 5.959$$

Diseases: The SI Model

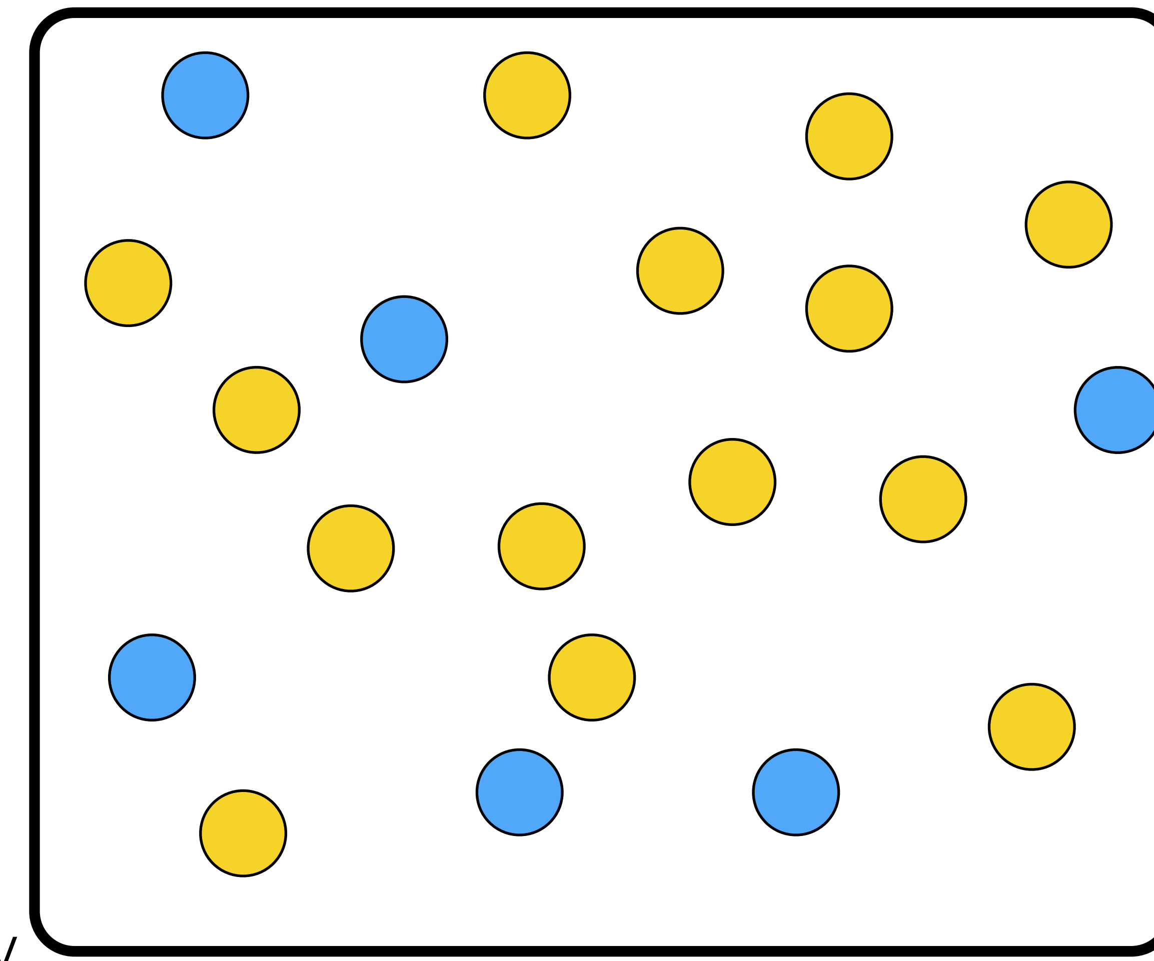
And by extending the pattern, the number infected after *t* periods is

$$I_t = S_{t-1} c \underbrace{\left(\frac{I_{t-1}}{N} \right) p}_{} + I_{t-1}$$

number of people susceptible after the last period

chance any one of them is infected

number already infected



So if c = 3 and p = .66, then

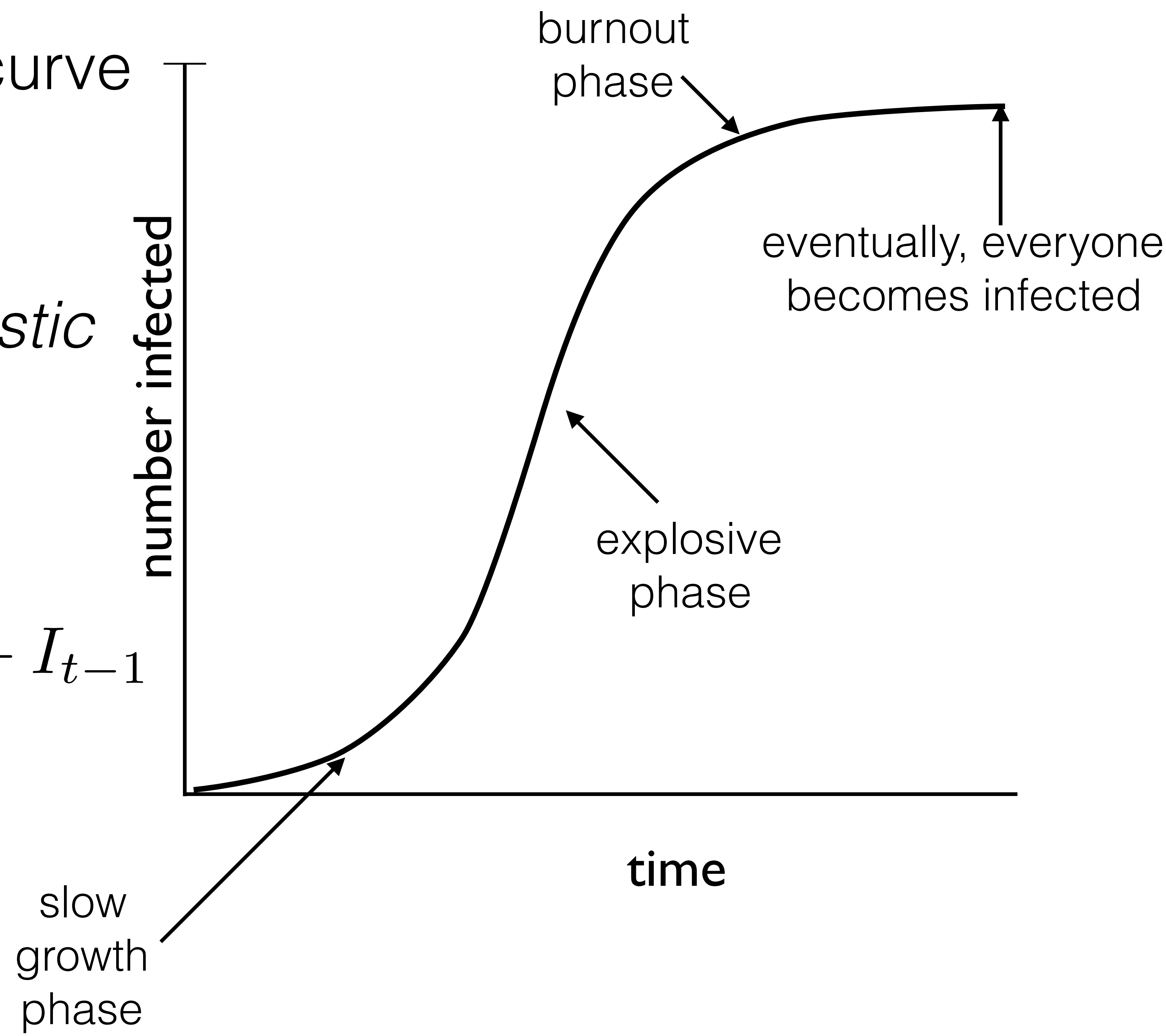
$$I_3 \approx 14$$

Diseases: The SI Model

Result is a growth curve that looks like this:

This is called a *logistic* growth curve

$$I_t = S_{t-1} c \left(\frac{I_{t-1}}{N} \right) p + I_{t-1}$$

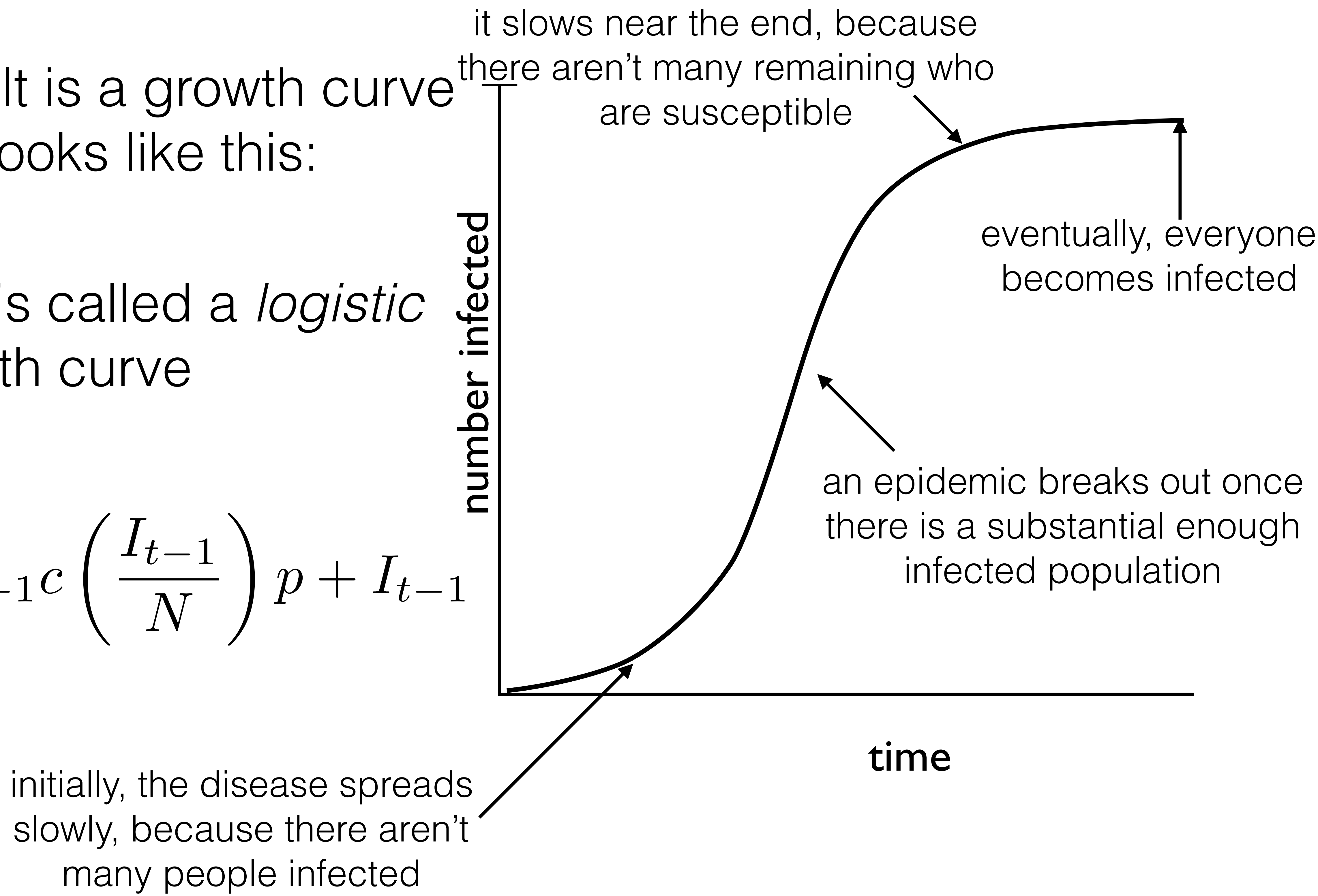


Diseases: The SI Model

Result is a growth curve that looks like this:

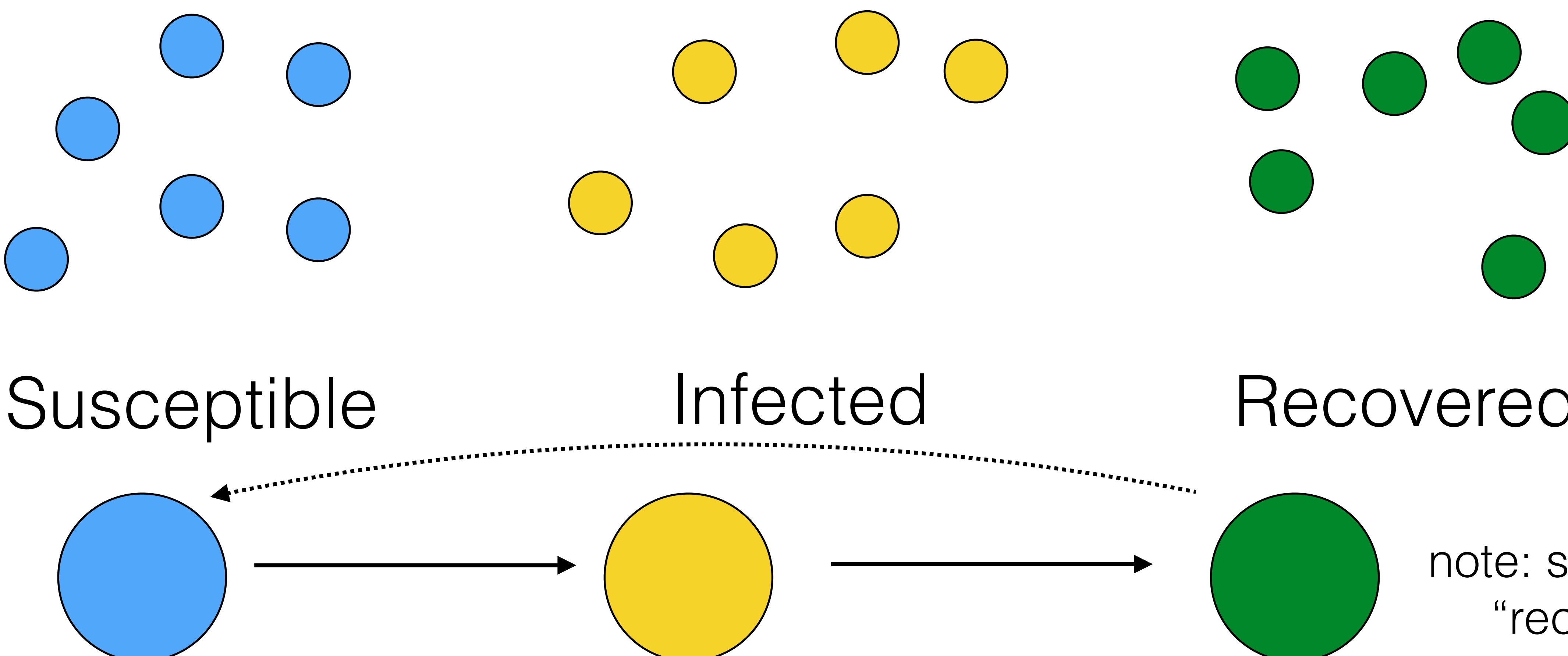
This is called a *logistic* growth curve

$$I_t = S_{t-1} c \left(\frac{I_{t-1}}{N} \right) p + I_{t-1}$$



Diseases: The SIR Model

The SIR model modifies the SI model: now there are three groups



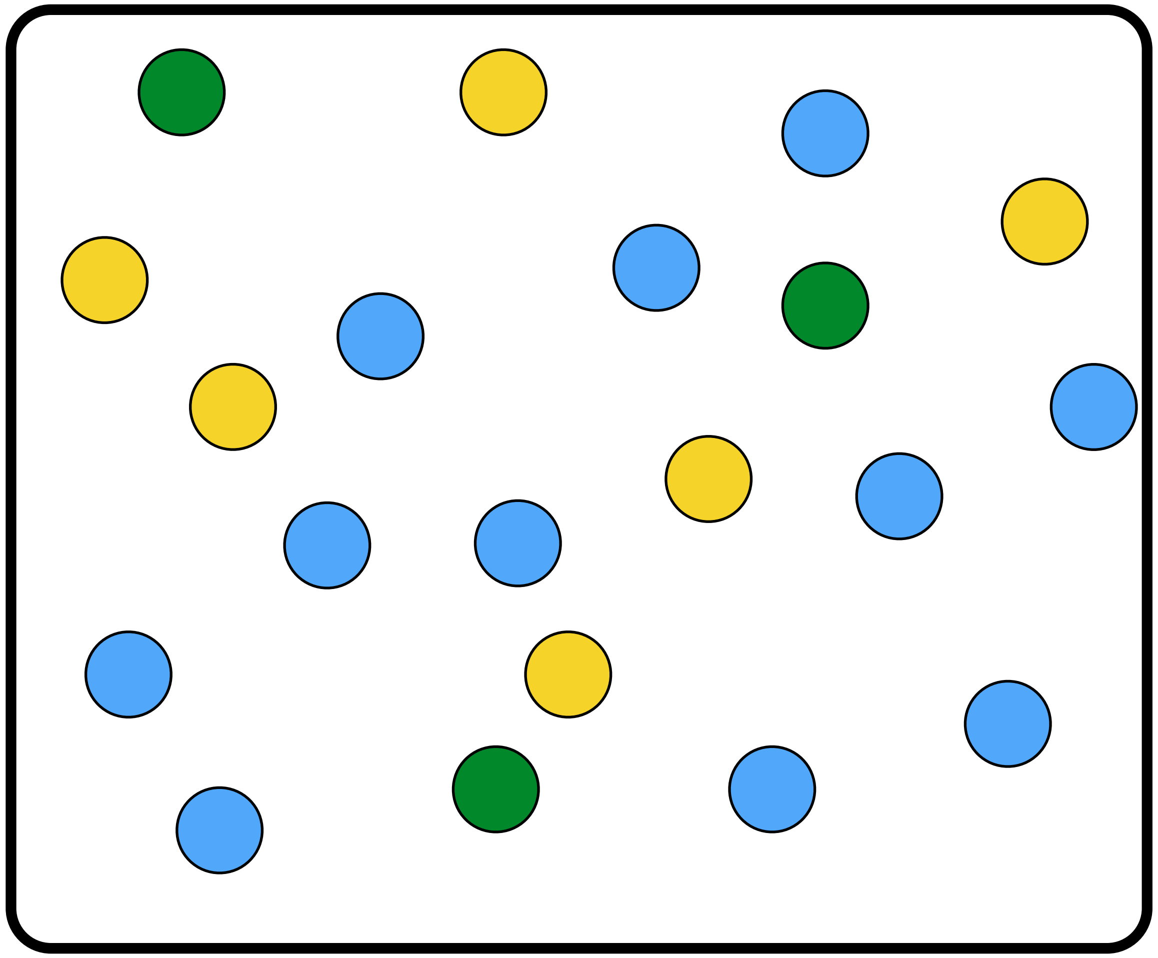
note: sometimes "recovery" actually means death

People move from susceptible to infected to recovered, and sometimes back to susceptible

Diseases: The SIR Model

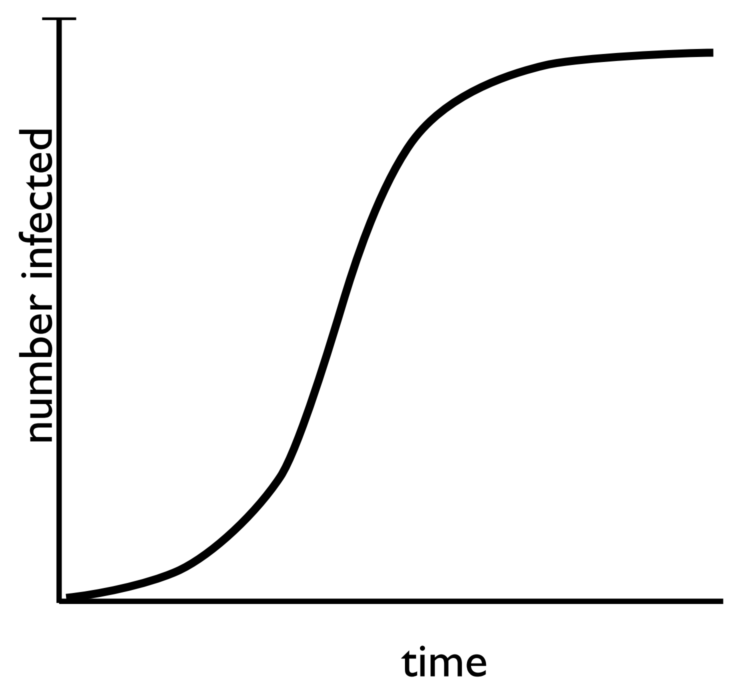
Now, people recover with probability *r*

People who recover are not susceptible to the disease, and they also can't infect anyone



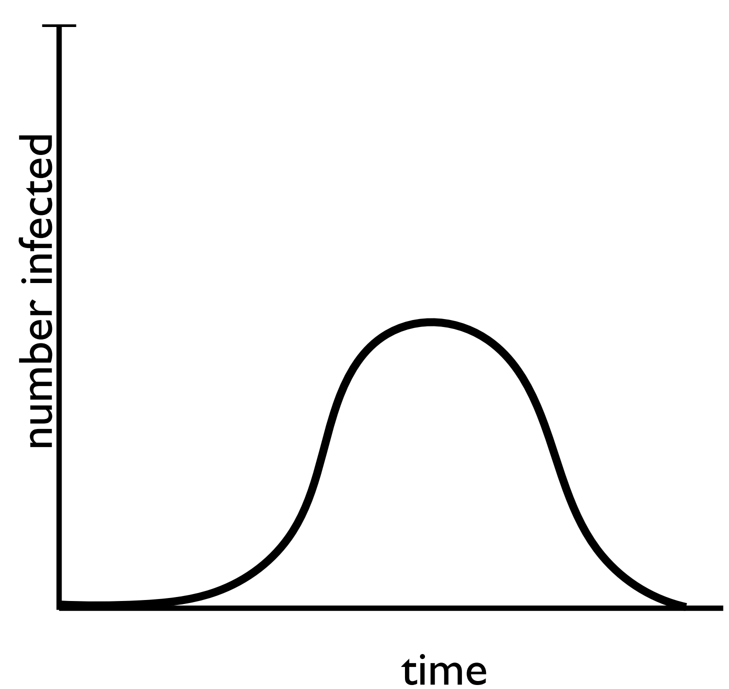
Diseases: The SIR Model

In the SI model, there was only one possible outcome (everyone is infected).



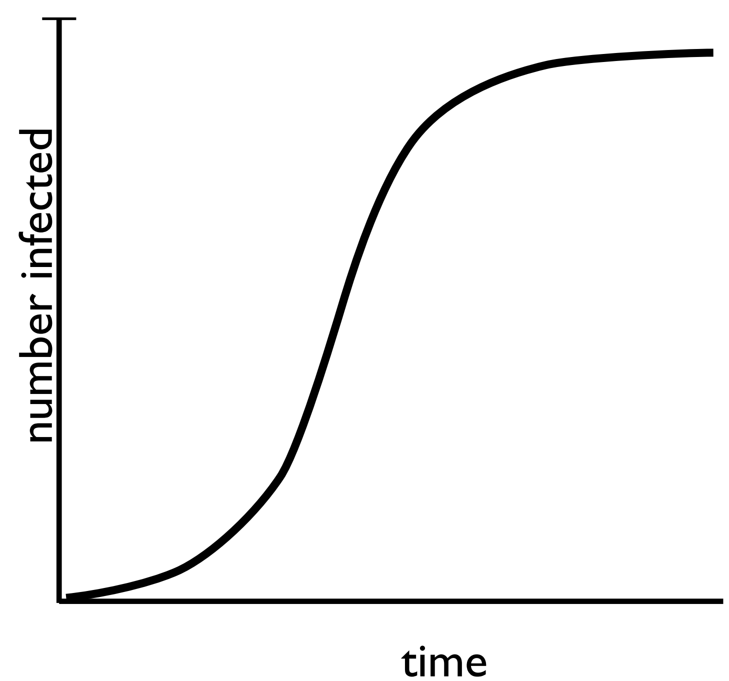
Now there are several possible outcomes:

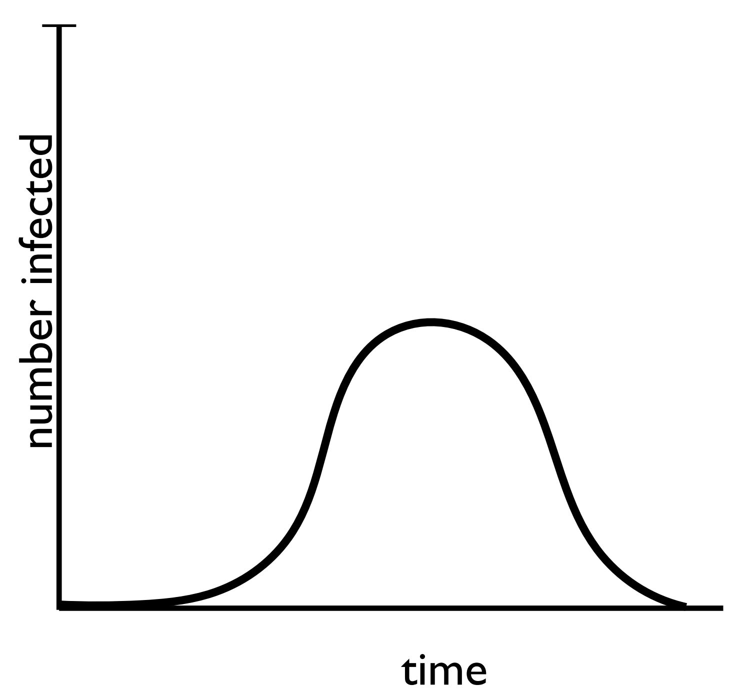
1) An epidemic (everyone becomes infected)
2) The disease dies out before infecting everyone



Diseases: The SIR Model

Whether the outbreak turns into an epidemic depends on the balance between the recovery rate (*r*) and the infection rate (*p*)

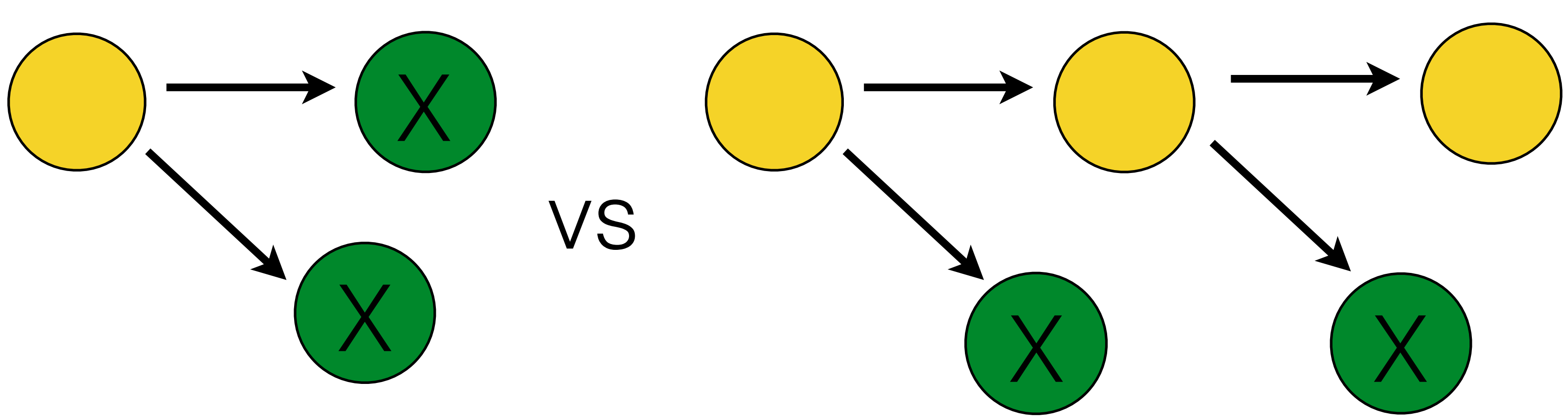




Diseases: Epidemiology

Characteristics of a disease that affect the probability of an epidemic:

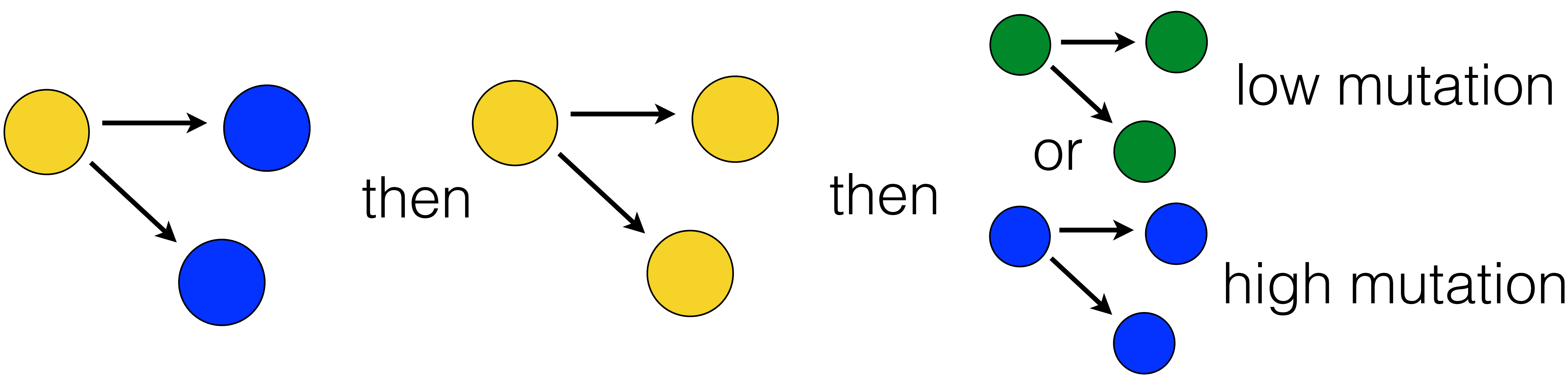
- deadlier diseases kill people before the disease can spread
- low mutation rates mean that recovered people remain immune
- low transmission rates mean that you're less likely to get the disease from incidental contact



Diseases: Epidemiology

Characteristics of a disease that affect the probability of an epidemic:

- deadlier diseases kill people before the disease can spread
- low mutation rates mean that recovered people remain immune
- low transmission rates mean that you're less likely to get the disease from incidental contact



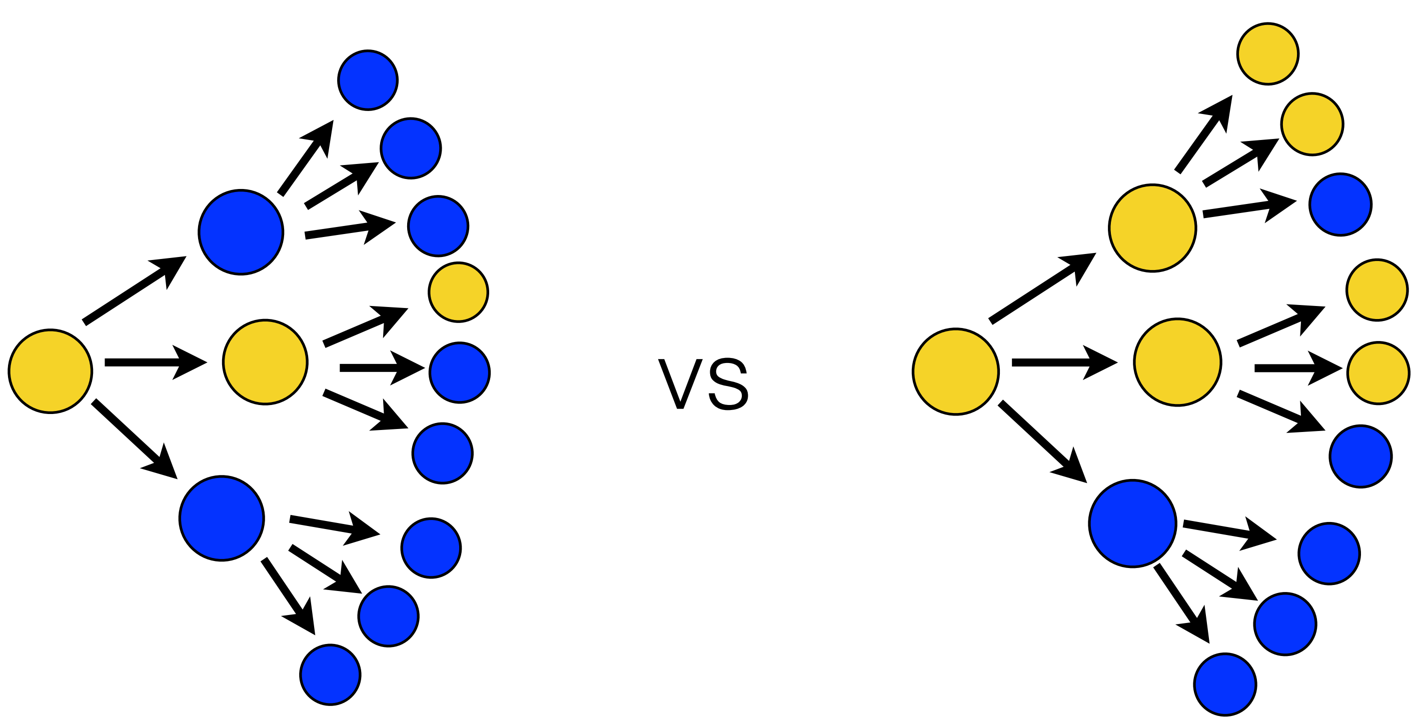
Diseases: Epidemiology

Characteristics of a disease that affect the probability of an epidemic:

deadlier diseases kill people before the disease can spread

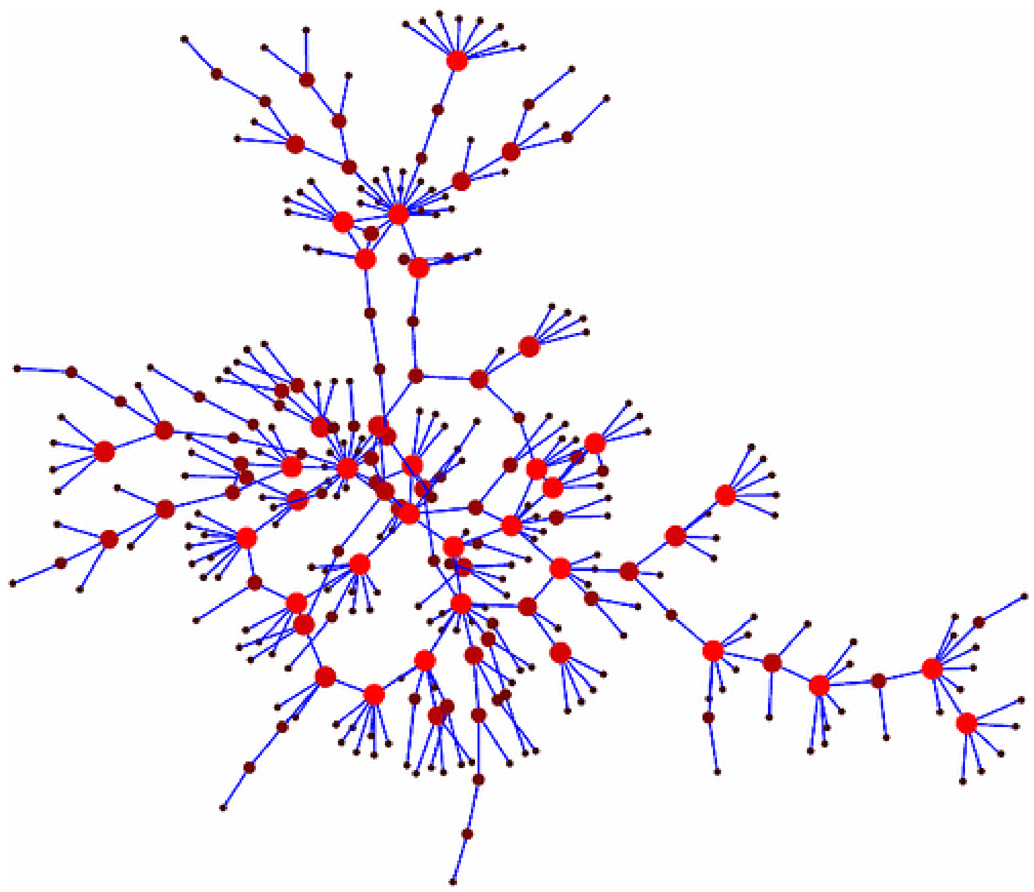
low mutation rates mean that recovered people remain immune

low transmission rates mean that you're less likely to get the disease from incidental contact



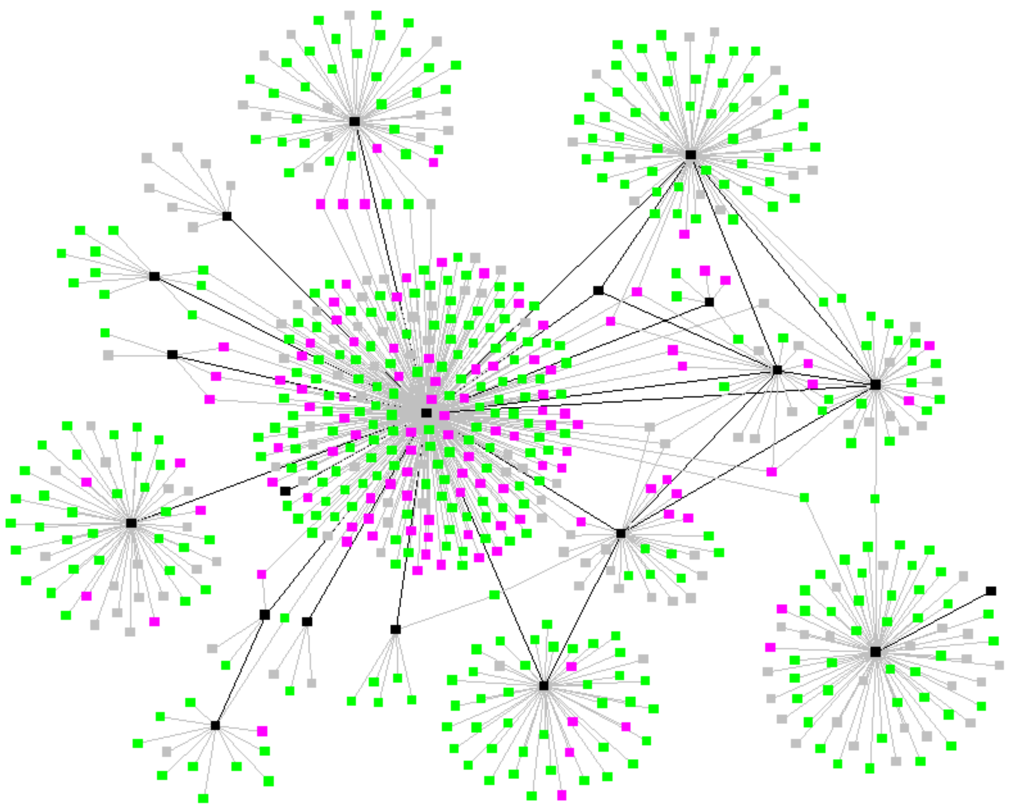
Diseases: The role of the network

So now, what if people don't interact at random?



Sexual Contacts

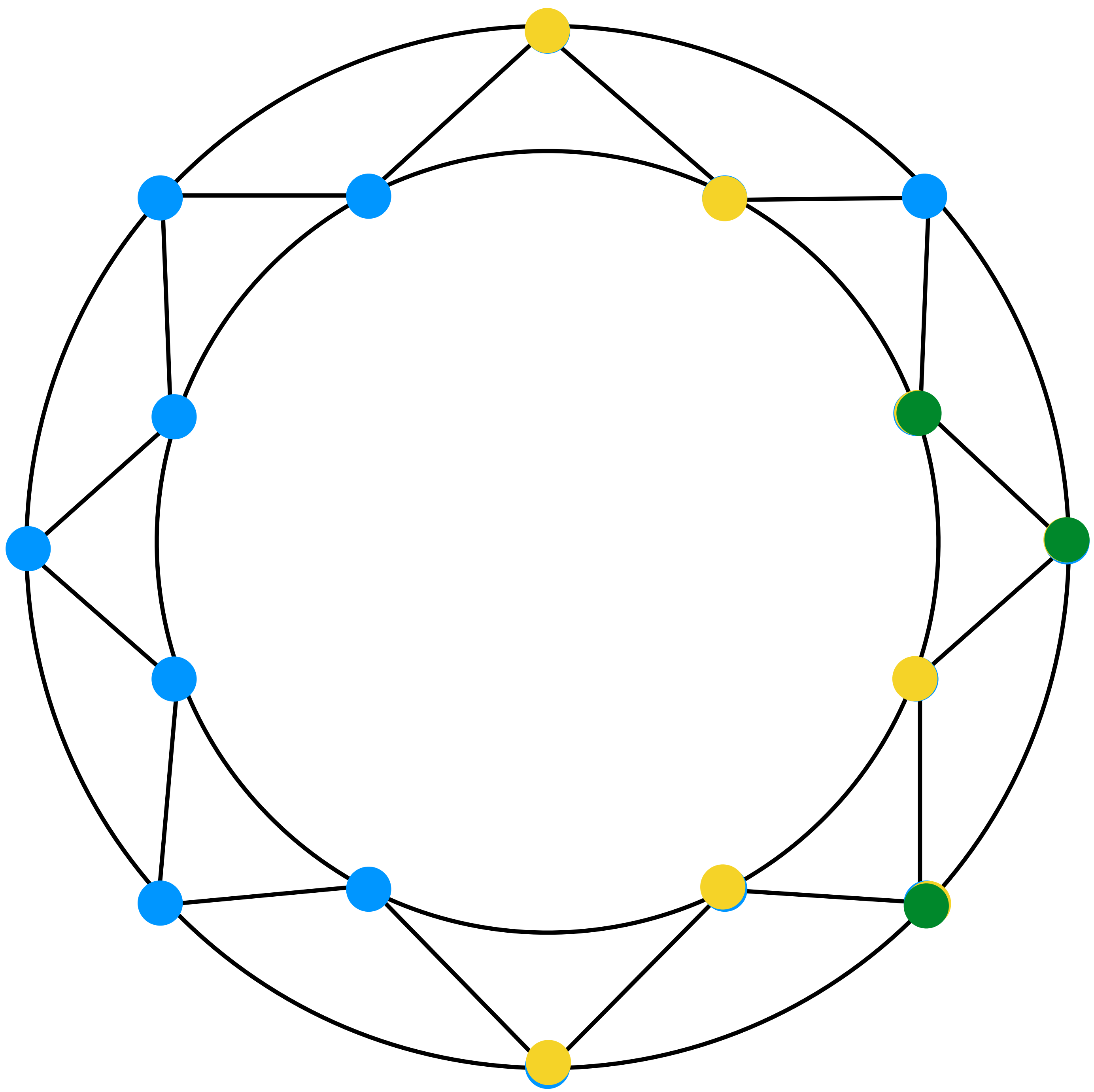
ref: Potterat et al (2002)



All Contacts

ref: orgnet.com

SIR Model on a Network



Now what if the infection spreads on a network?

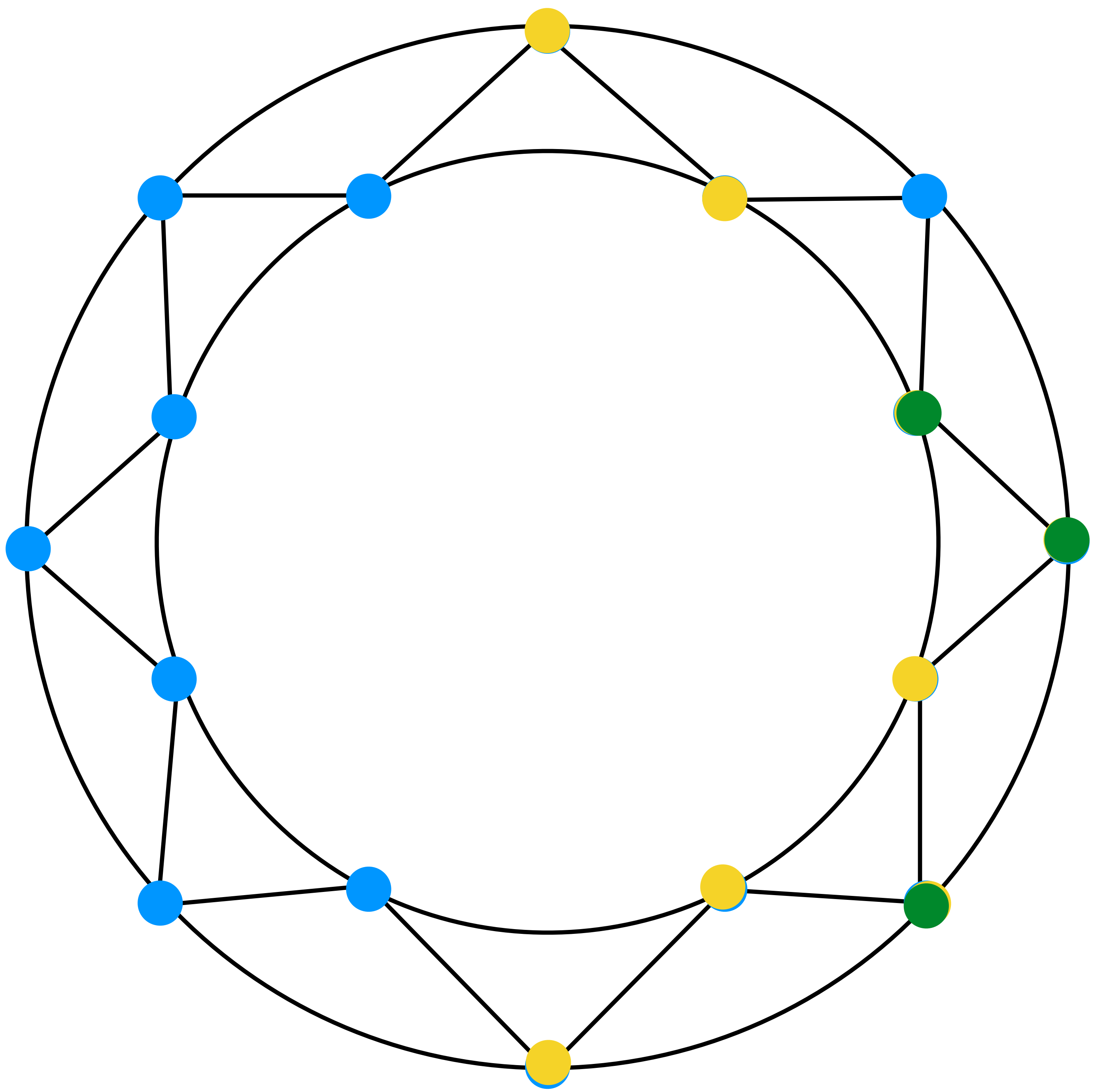
Still have S, I, R groups

Every time period, you are exposed to your direct neighbors

Prob(infected if friend is infected) = p

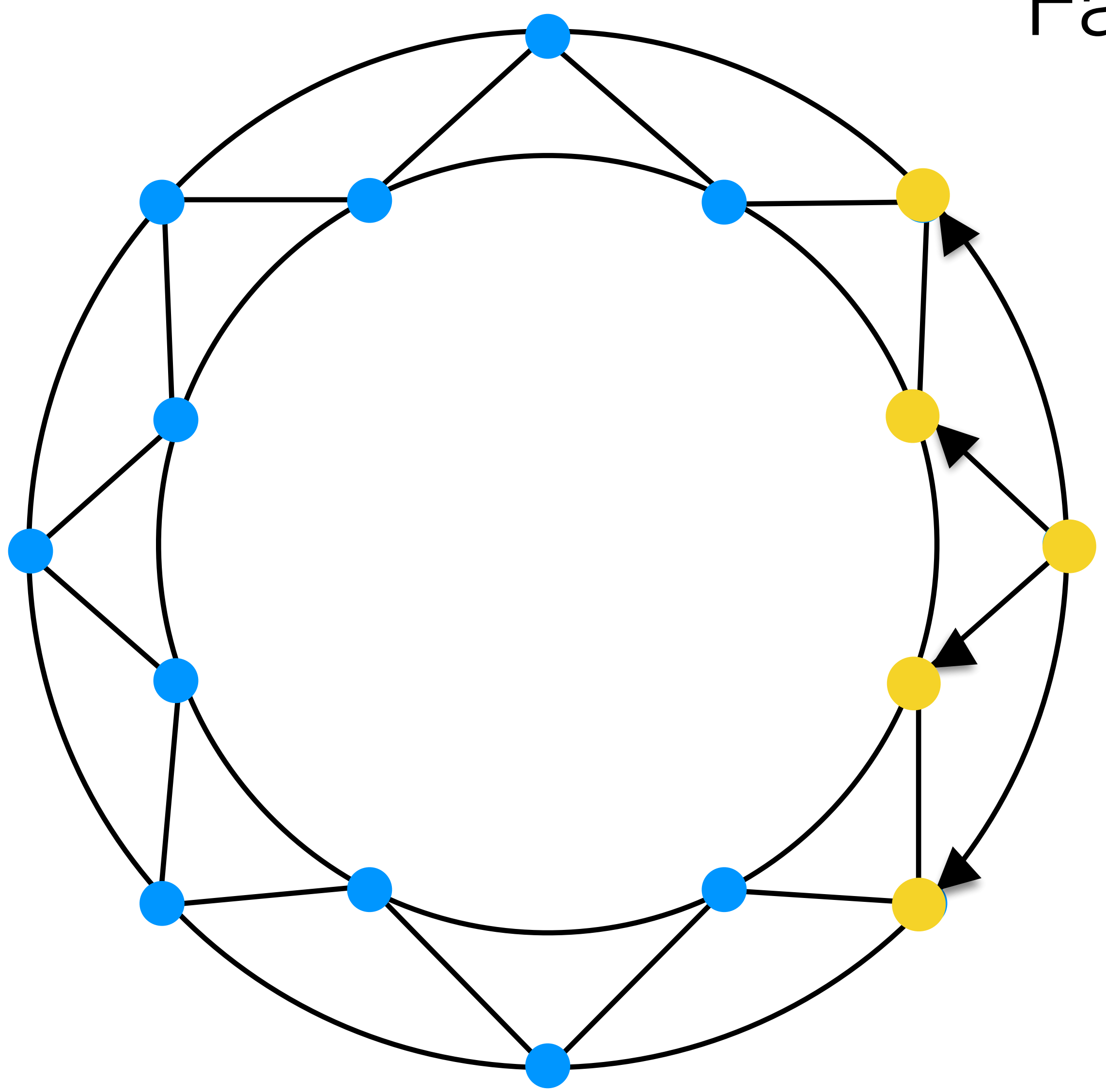
Prob(recovery) = r

SIR Model on a Network



The probability of an epidemic still depends on the infection rate (p) and the recovery rate (r), but it also depends on network structure

Network Structure and Diffusion

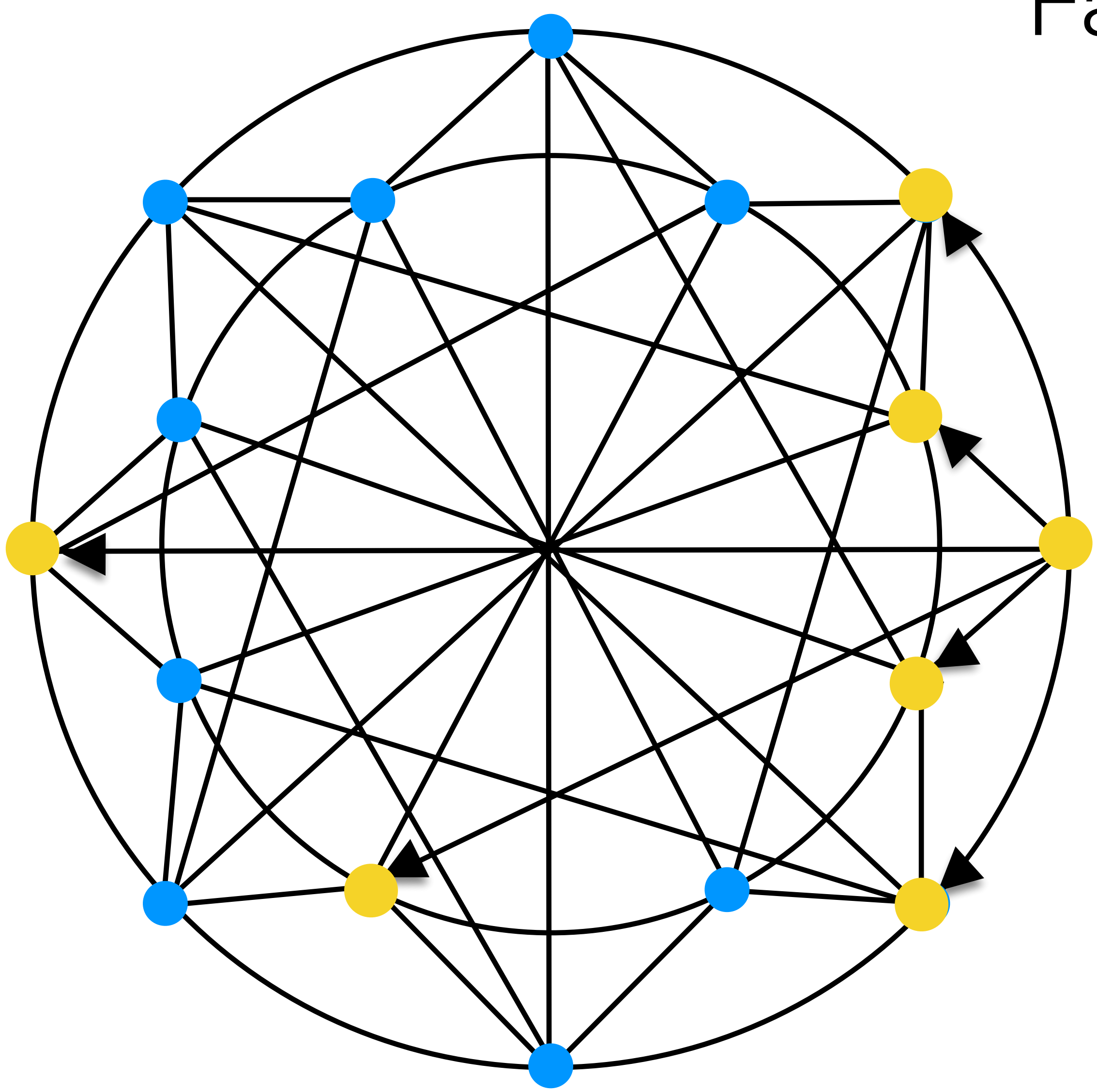


Factor 1: average degree

A node has the potential to infect everyone she is connected to

→ A higher average degree means more chances for an infected person to expose others

Network Structure and Diffusion



Factor 1: average degree

A node has the potential to infect everyone she is connected to

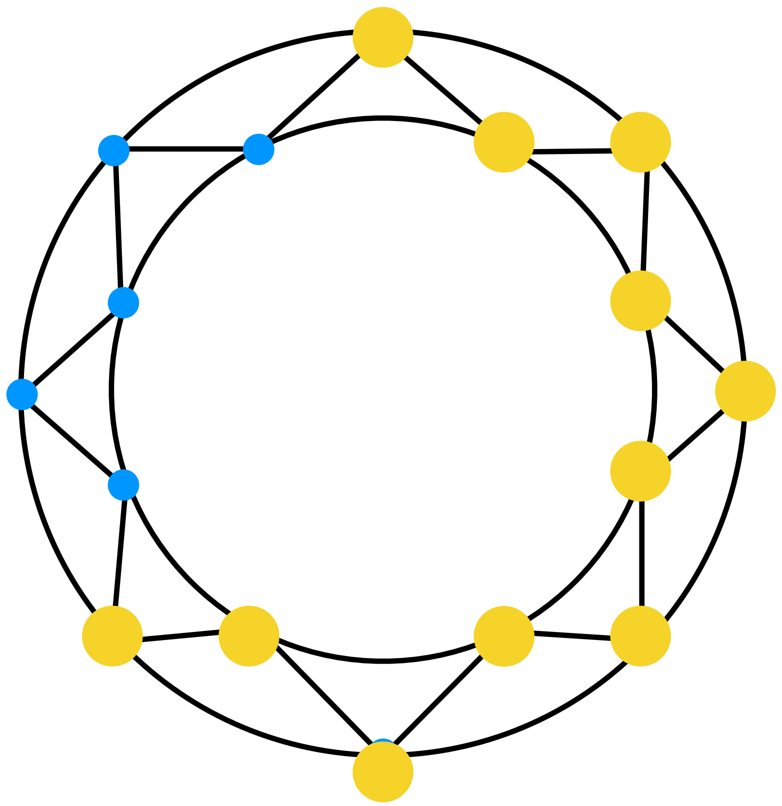
→ A higher average degree means more chances for an infected person to expose others

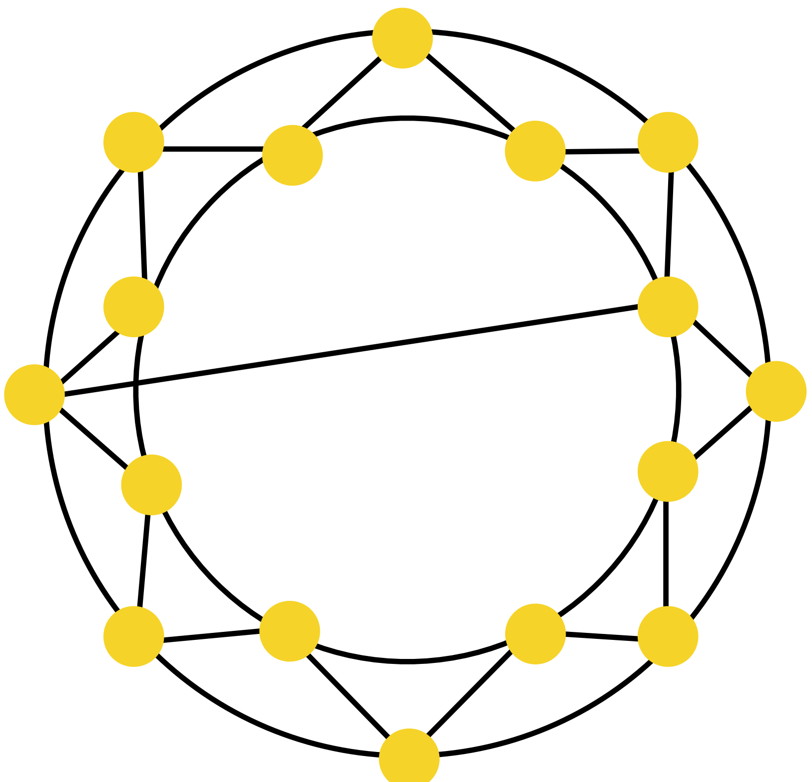
Network Structure and Diffusion

Factor 2: Average Distance

When people have contact with others in far-flung regions of the network, then disease can spread more widely.

→ Epidemics are more likely in a small world

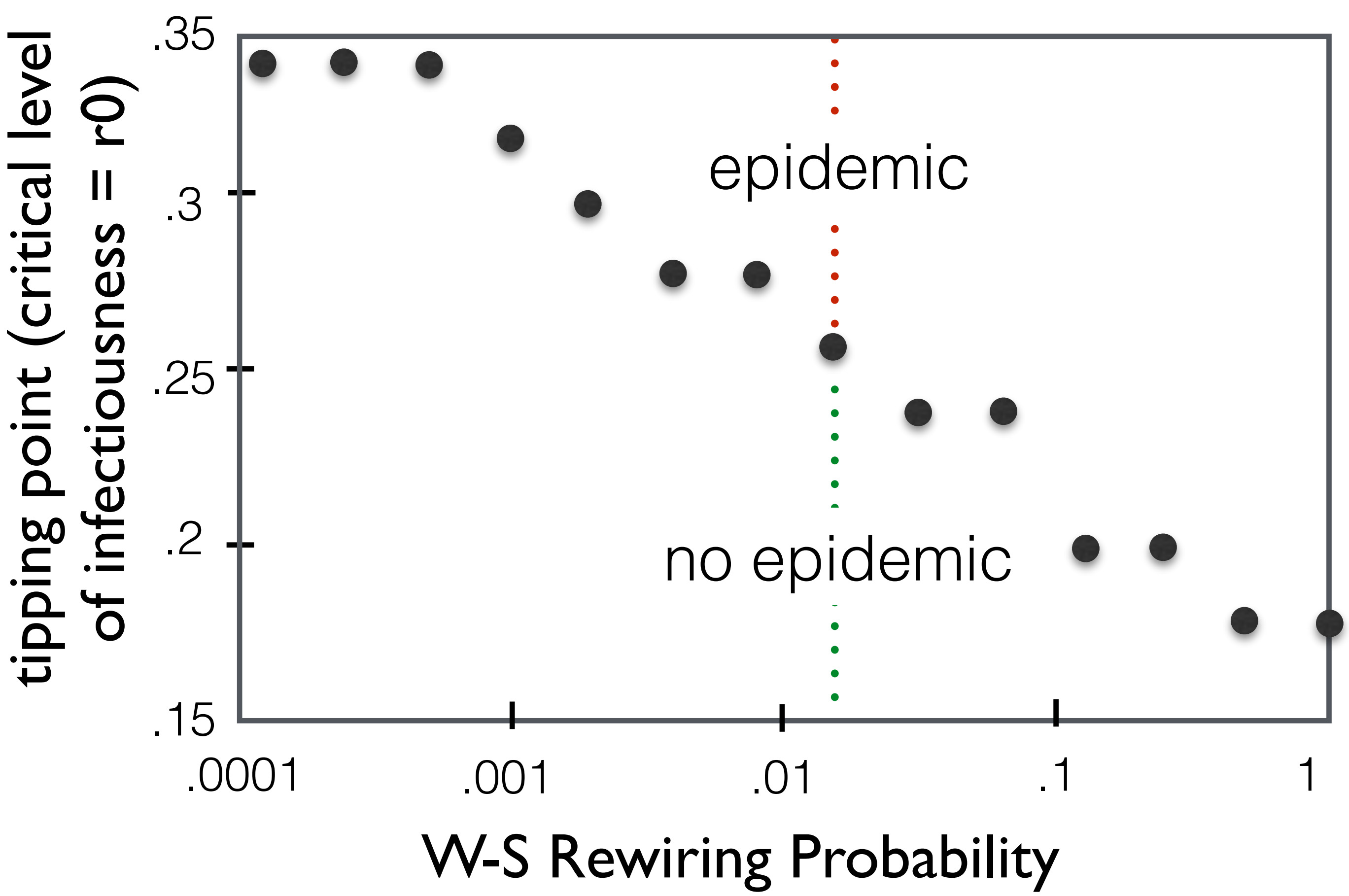




Network Structure and Diffusion

Factor 2: Average Distance

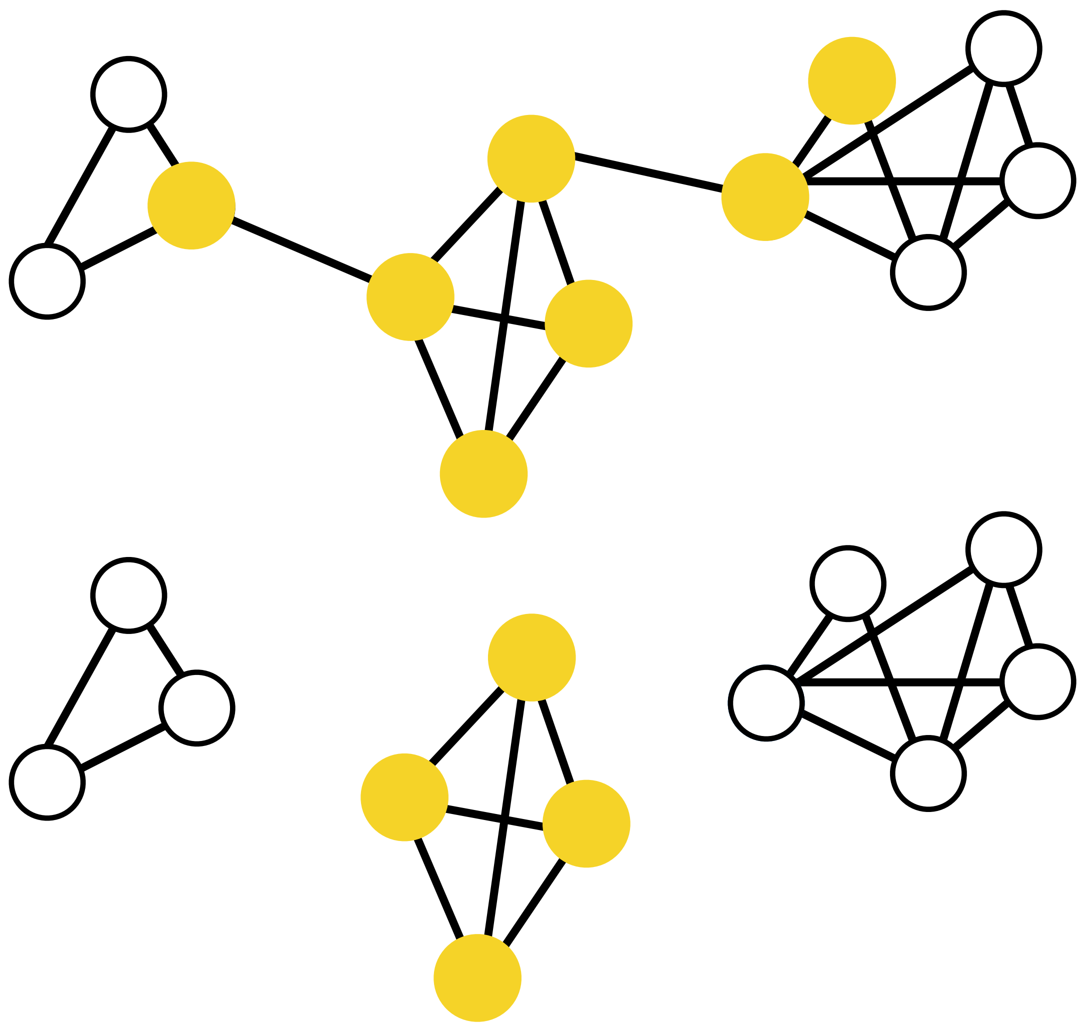
You can formalize this by looking at a Watts-Strogatz network



A few long-distance connections make it much easier for the disease to spread, which lowers the critical level of infectiousness (r0). This increases the likelihood of an epidemic.

Network Structure and Diffusion

Factor 3: The size of the largest connected component



The disease can only spread to people who are connected to the original point of infection

→ disease will not spread as far when the network is broken into chunks

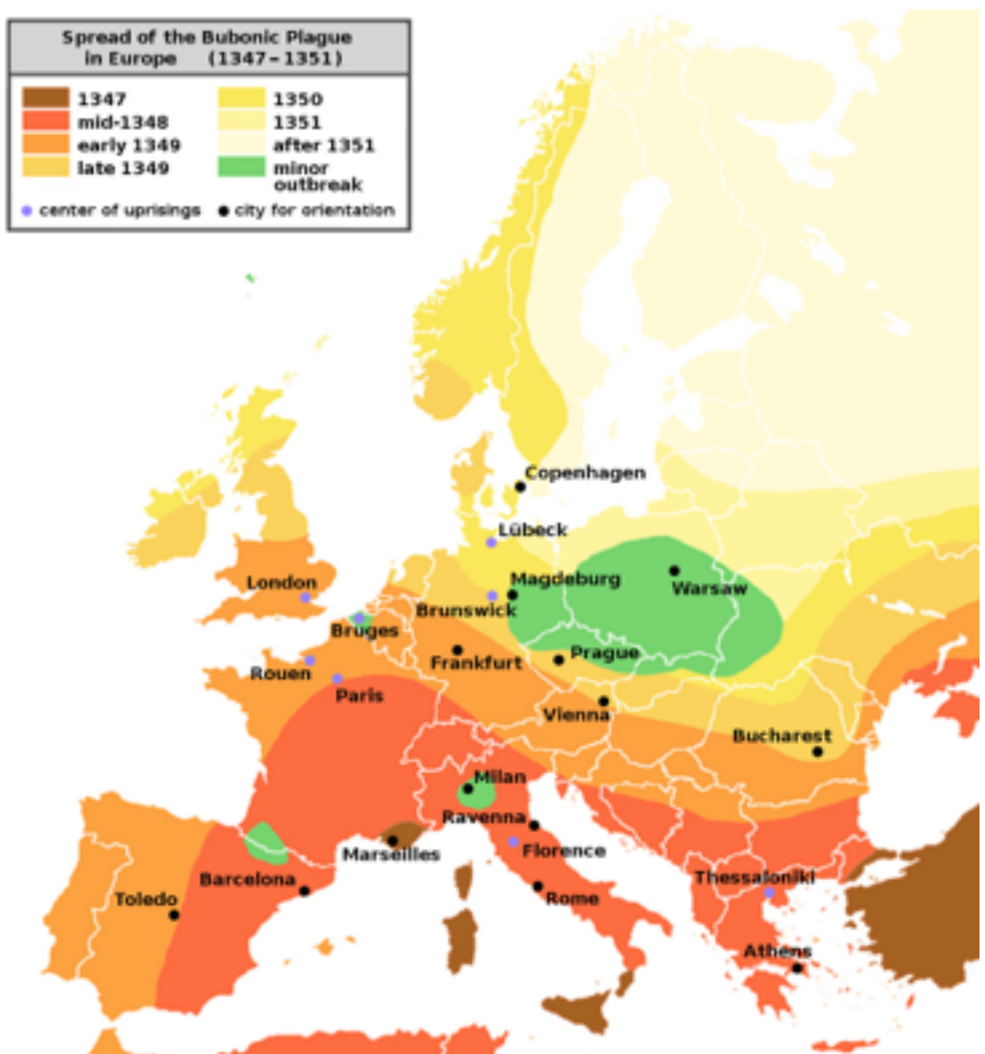
What affects structure?

The structure of social networks that affect the spread of disease depends on characteristics of the disease, and characteristics of the society the disease lives within:

- Social and cultural norms → burial practices, greetings, condom use, sexual practices, norms of personal space, medical beliefs, family structure, visiting practices
- Environmental factors → sanitation systems, transportation systems, housing
- Health care policies → quarantine periods, immunization programs, awareness, public event cancelation, response time
- Transmission characteristics → type of contact required, duration of contact required, infectiousness, lifespan outside of host

Network Structure and Diffusion

The effects of network structure can be seen in the differences between the spread of disease in the middle ages and the spread of disease today.



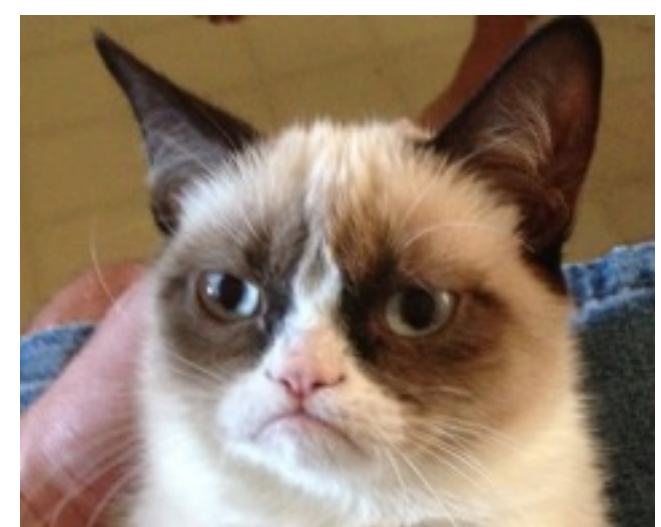
The spread of the plague was certainly terrifying in it's time, but it's nothing compared to the pandemics of today's more connected world…

Network Structure and Diffusion

Although we haven't yet talked about the diffusion of *ideas*, you can also see the effects of network structure in the spread of viral content

Fads have existed forever. But they spread more quickly and more widely than they used to.

→ "Epidemics" of viral content have become "pandemics"



Controlling Diffusion

The relationship between network structure and diffusion suggests ways that we could prevent epidemics

→ change who interacts with who

→ change who is susceptible

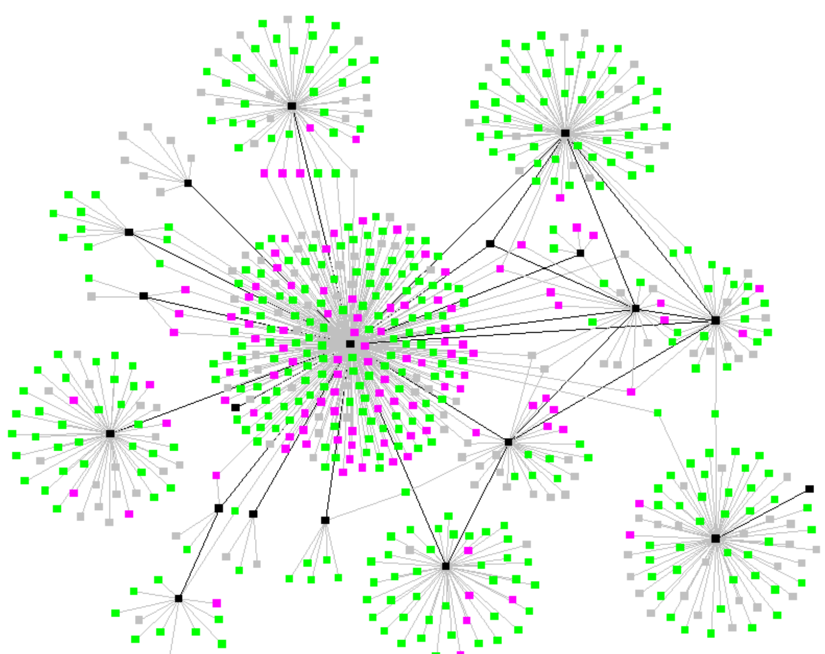
→ change who is infected

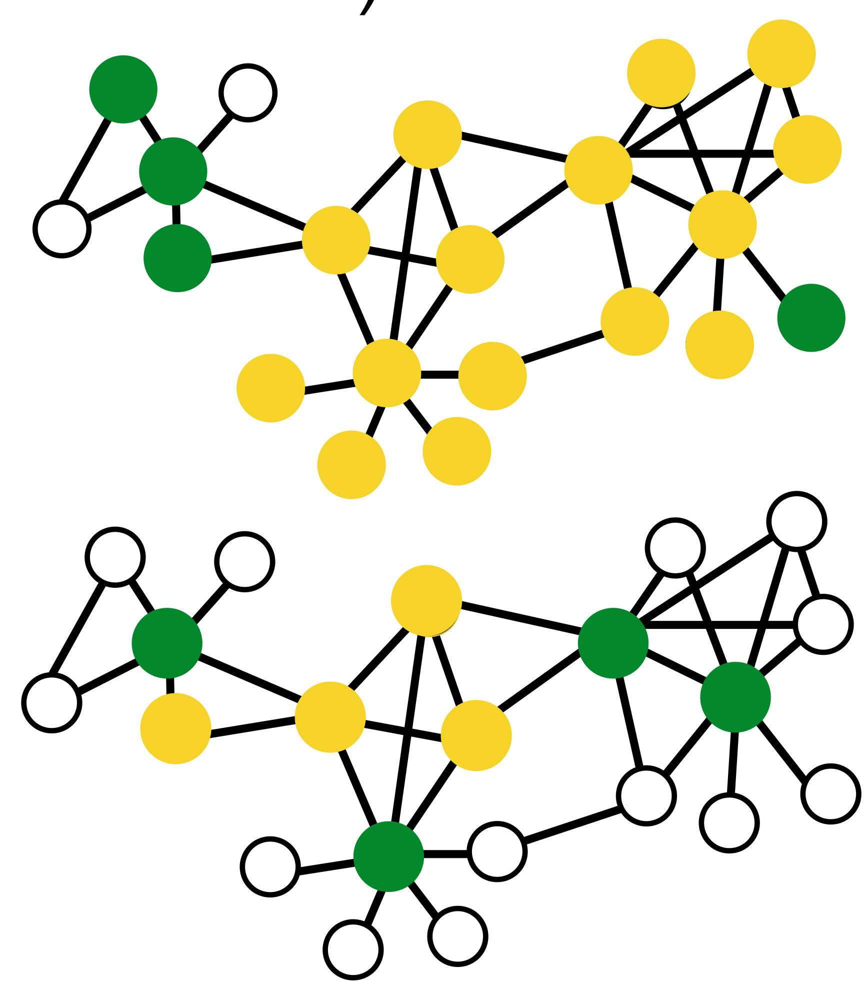
Controlling Diffusion: Disease

Vaccinate a small number of people

→ Vaccinating the right people can prevent the spread of the disease with very few resources

→ Good people to vaccinate are hubs (health care workers) and bridges (bus drivers)



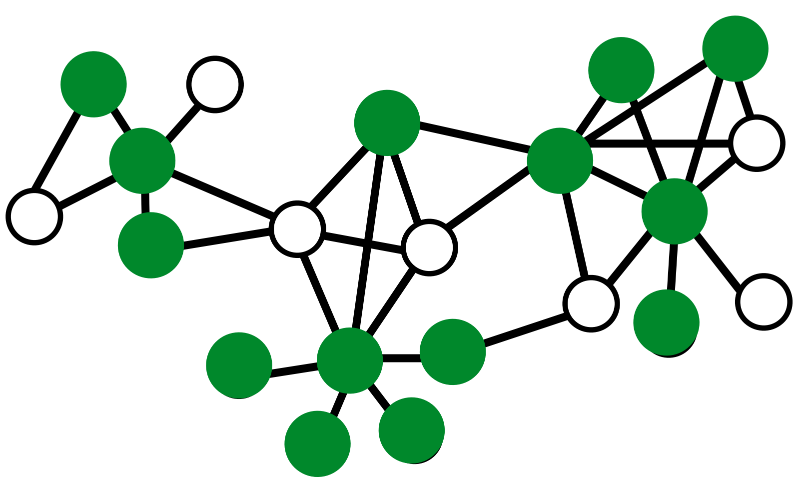


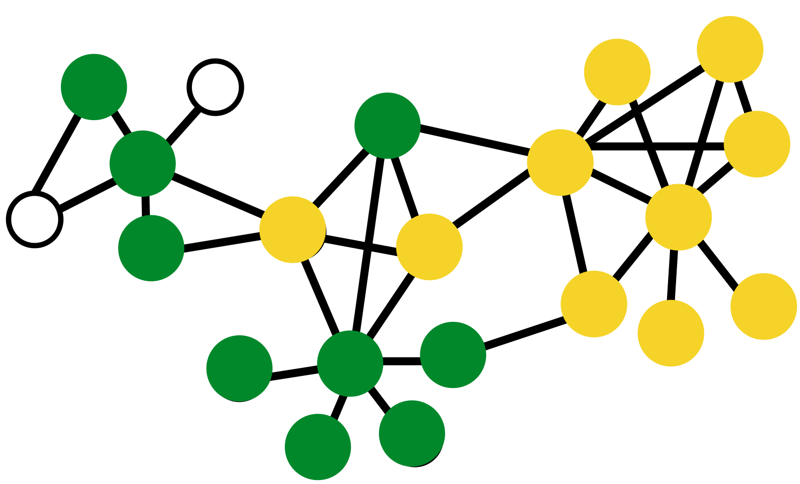
Controlling Diffusion: Disease

Vaccinate a large number of people

→ Vaccinating enough people effectively reduces the size of the largest component to 0

→ The fact that the network is broken up protects people who are unable to be vaccinated. This is called “herd immunity”.

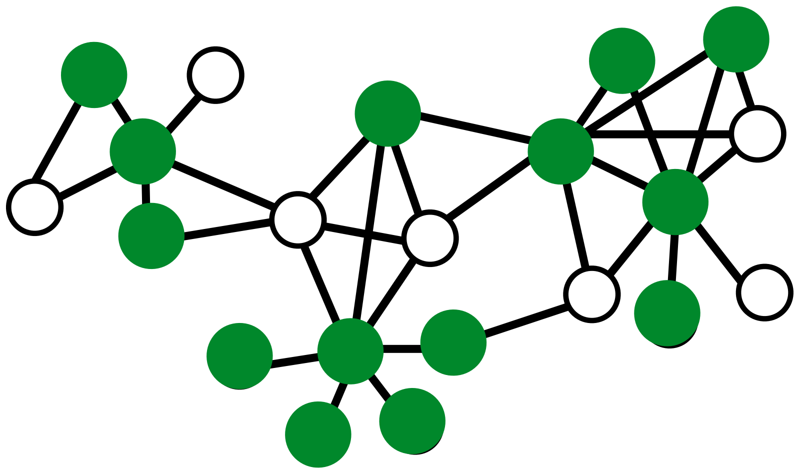


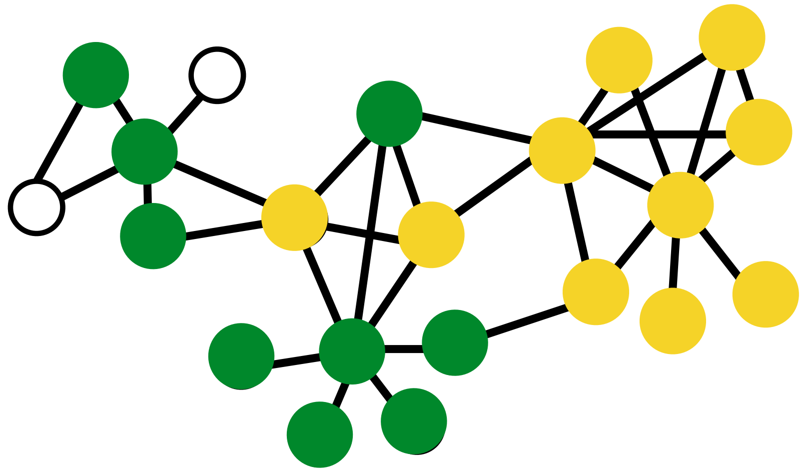


Controlling Diffusion: Disease

→ If the population is vaccinated for long enough, the disease won't have enough hosts and will die out (eg: small pox)

→ Subpopulations that fail to vaccinate may lose herd immunity and have outbreaks, which may infect people in other communities who can't be vaccinated



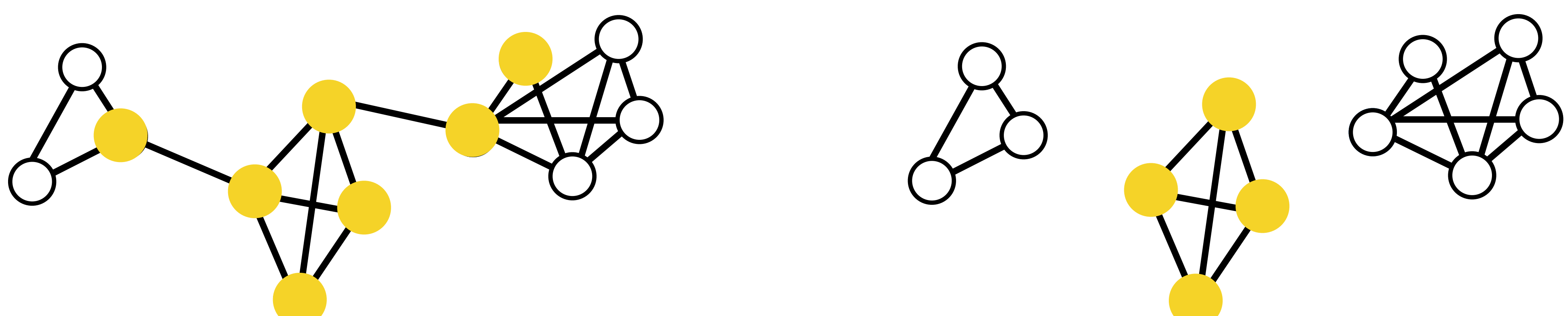


Controlling Diffusion: Disease

Quarantine

→ Quarantines break up the network into much smaller chunks (eg: households)

→ Limiting large gatherings (eg: baseball games) and travel restrictions can achieve the same thing.



Summary

Takeaways from today:

- Diffusion is how things like disease and information spread through a population
- A lot of our knowledge of diffusion comes from epidemiology—the SI model and the SIR model
- Diffusion on a network gives us some ideas of how to start to control the process

Next time:

- Information: fads, virality, network effects, strategy, and cascades