Contagion

4

> I have your disease in me now.
>
> —Dorothy Vallens, *Blue Velvet* (1986)

Usually, when you give something to someone, you don't have it anymore. But there are some things that you can keep and still give to others. Diseases. Knowledge. Beliefs. Behaviors. We call such things **contagions**. How contagious things spread one from person to another is an important area of research in many fields. Epidemiologists and ecologists want to understand how a disease will spread through a population, as well as what might impede its spread. Social scientists want to understand the spread of ideas or behaviors. Marketers want to know what influences the adoption of products and innovations.

In this chapter, we will consider the dynamics of how contagions spread. As with the previous chapters, we'll mess around with agent-based models. However, this chapter also represents a watershed in the course of our study of social behavior, because it is in this chapter that we begin to incorporate the use of purely mathematical models. For those who cringe at the thought of doing math, I cannot help you—math is wonderful, and you should feel bad for your discrimination against one of the great intellectual achievements of human civilization. For those who worry that your math skills are rusty or underdeveloped, I promise to go slowly and to restrict our analysis to mathematical techniques you should remember from high school—algebra and functions. For those of you with advanced math training who scoff at simple approaches, I urge you to try your hand at it and see how powerful such approaches can be. If you crave more of a mathematical challenge, there are many resources available that can help you to go deeper.[1]

The coronavirus pandemic of 2020–2023[2] made nearly everyone in the world aware of how important it is to understand contagion processes. And in this chapter, we *will* cover dynamics specific to infectious diseases. To start, however, I want to focus on the diffusion of information or innovations. I have two reasons for this. First, it's not entirely obvious that the

[1] For example, see Keeling and Rohani (2008) or Bjørnstad (2018). For a more lighthearted approach, see also Smith? (2014).

[2] One can only hope that some sort of end is in sight.

adoption of a new innovation actually *is* anything like a contagion, and modeling may help to convince you that it indeed is. Second, the innovation diffusion models are formally quite simple yet provide good scaffolds for the more complicated contagion models that are introduced later in the chapter. We'll start with models of spontaneous adoption and of adoption by social influence. After that, we'll tackle more disease-relevant issues related to recovery and immunity, as well as the importance of public health measures such as vaccination and social distancing to reduce the severity of an epidemic.

4.1 The Diffusion of Innovations

How do new innovations spread, or diffuse, through a population? In 1962, the first edition of Everett Rogers's now-classic book, *Diffusion of Innovations* was published.[3] Rogers studied how a variety of innovations spread, with examples ranging from the adoption of hybrid seed corn by Midwest farmers to the adoption of ham radio among tech enthusiasts to the adoption of new ideas among French intellectuals. Rogers found that in all cases, when the cumulative adoption rate was plotted over time, it typically yielded a sigmoidal or S-shaped curve. That is, adoption increased slowly at first, then increased rapidly, then slowed down again.

What explains this apparently universal pattern of adoption? Rogers proposed an explanation based on the idea that individuals differ in their proclivity for adoption. He posited that a population could be reliably partitioned into several types: initial innovators, early adopters, early majority, late majority, and laggards. To make this taxonomy work, he also had to posit fairly specific proportions of the population that fall into each category: approximately 2.5% innovators, 13.5% early adopters, 34% early majority, 34% late majority, and 16% laggards (Figure 4.1).

Rogers's explanation involves two major assumptions. First, all individuals are assumed to have access to the same information about the availability of a new innovation. Second, individuals are assumed to vary in a very particular way: in terms of how long that information takes to provoke adoption of the innovation. The first assumption seems plausible; after all, if information spreads rapidly through a population—much more rapidly than the time it takes to actually be influenced by that information and adopt the innovation—then the exposure to information might be effectively simultaneous. The second assumption strikes me as more problematic. It assumes an awful lot about why people adopt and why those differences might occur, and requires very specific distributions of personality types to produce the sigmoid adoption curve. Is it plausible that these same distributions occur across contexts and cultures? It's possible, but unlikely. For now, let's reject the second assumption and assume instead that individuals respond to information pretty much identically. We can then ask if the first assumption of identical exposure to information is sufficient to generate the sigmoid adoption curve.

4.2 Spontaneous Adoption

We're going to build a model in which a new innovation is introduced into a population. Everyone perceives the innovation in the same way, and everyone has the same proclivity to adopt. In this model, there is no social influence, so there is no need to model social

[3] The fifth edition was published in 2003, shortly before Rogers's death.

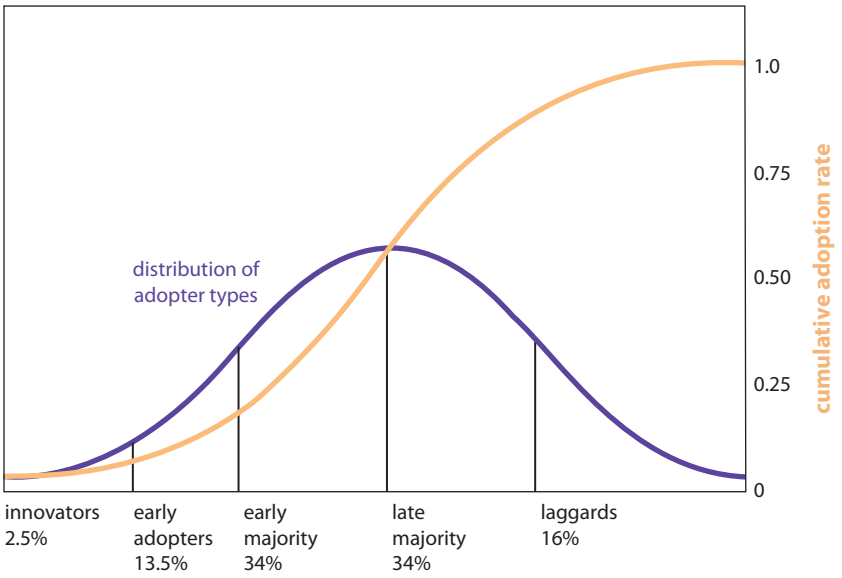

Figure 4.1 Rogers's diffusion model. The yellow curve shows the overall adoption rate over time. The blue curve shows the distribution of adopters according to the timing of their adoption.

structure. However, we might want to add social influence in the future so that we can allow agents to come into contact with one another and also vary the rate of that contact. Luckily, we have a simple model that can do the trick: it's our basic Particle World model from chapter 2. In that model, agents are embodied on a 2-D space with toroidal boundaries and move around using a correlated random walk. The speed and turning angle can be adjusted to affect the rate at which agents contact new social partners. For simplicity, we'll ignore collisions and flocking. In this way, we have a world in which agents move around in physical or social space and come into contact with others at varying rates.

In the Particle World model, agents differed from one another only in terms of their current position and directional heading. Now we'll add one more agent-level property: infection status. Agents in our new model are in one of two possible states: **susceptible** or **infected**. This terminology is taken from the idea of disease contagion, but it can be applied to the adoption of products or behaviors if you think of "infected" as simply indicating that the agent has adopted the product or behavior. We can initialize the model with a small number of infected agents; if we are thinking about innovation diffusion, these are our innovators.

The spontaneous adoption model works as follows. We consider a population of N agents, with some number $n_0 < N$ of agents as initial adopters (or innovators). At each time step, every agent who has not yet adopted will adopt with probability α. That's it. That's the entire model description, at least as far as assumptions that influence the dynamics of adoption. Our code also allows agents to move around, but because social influence is (for now) irrelevant to adoption, movement doesn't matter. Notice that there's no way for an agent to become *uninfected*, though we'll consider that later on as well. If agents cannot become uninfected, the population is always going to end up with everyone infected. However, exactly *how* that happens isn't necessarily obvious, so we'll focus on the temporal dynamics of infection. Will we see the telltale sigmoid adoption curve under spontaneous adoption?

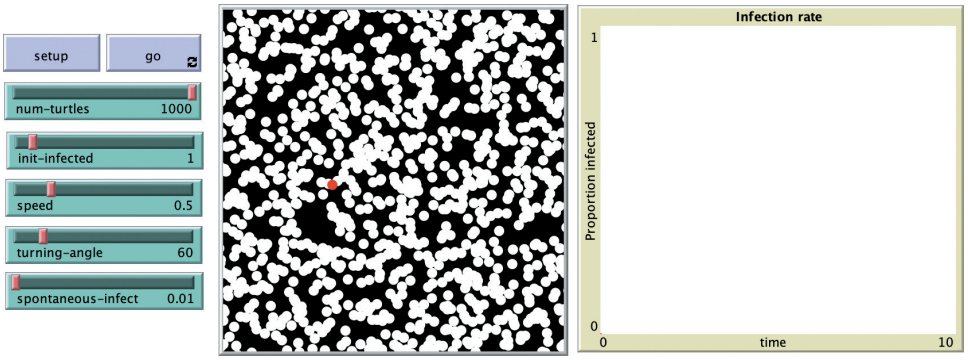

Figure 4.2 The Interface window for the spontaneous adoption model. Uninfected agents are white, and there is a lone red infected agent.

4.2.1 Coding the Model

The NetLogo code for this model is **contagion_spontaneous.nlogo**. Because much of the code will be carried over from the Particle World model, I'll focus on what is new here. Figure 4.2 depicts the model's Interface window.

As before, we will include global parameters for the population size and for the speed and turning angle of the random walk. In addition, we'll also add parameters for the number of initially infected agents (`init-infected`) and the rate of spontaneous infection (`spontaneous-infect`). Each agent will keep track of its own infection status, which we can code as an agent-level Boolean variable, as follows:

NetLogo code 4.1

```
turtles-own [
  infected?
]
```

At initialization, the agents are created and placed in random locations on the grid. We'll initialize each agent as uninfected, and then infect an initial few. We could infect initial agents probabilistically, but I have chosen to control the precise number of initially infected agents. What *is* probabilistic is exactly *which* agents become infected. This is controlled in the code by the procedure `setup-infected`, which is called from `setup`, and which chooses n_0 agents at random and infects them:

NetLogo code 4.2

```
to setup-infected
  ;;can't infect more agents than exist
  if init-infected > num-turtles
  [set init-infected 1]
  ask n-of init-infected turtles [
    set infected? true
  ]
end
```

For the purposes of visualization, I have made infected agents red and uninfected agents white. This is controlled by a procedure called `recolor`. The colors do not influence the

model dynamics, but they do help us keep track of which agents are and aren't infected. Later, we'll see how we can use color as redundant information about infection status to avoid undesired modeling artifacts regarding the timing of infection and recovery.

The model dynamics work as follows. The simulation stops if everyone is infected, since no further infections are possible. This is useful for two reasons: first, it reduces computational time for running batches of simulations, and second, it allows us to record the time it takes for the contagion to diffuse throughout the population. If at least some agents are susceptible, the procedure `infect-susceptibles` is called. This causes each susceptible agent to become infected with probability `spontaneous-infect` (technically, we don't need to differentiate between infected and uninfected agents, because agents who are already infected cannot change their state, but this won't be the case when agents can recover or disadopt). Again, we model a probabilistic event by comparing a random draw from a uniform distribution to some threshold.

NetLogo code 4.3

```
to infect-susceptibles
  ask turtles with [not infected?] [
    if random-float 1 <  spontaneous-infect
    [set infected? true]
  ]
end
```

The provided code then recolors agents based on their infection status and has them move using a random walk, but neither of these are essential to the model dynamics. With our model now coded, it will be useful to be able to plot the proportion of the population that is infected as a function of time. In NetLogo, this is done by inserting the following command into a plot in the Interface window. We can then take a look at the temporal dynamics of adoption under spontaneous adoption.

NetLogo code 4.4

```
plot (count turtles with [infected?]) / num-turtles
```

4.2.2 Analyzing the Model

When every agent has the same information and the same proclivity for adoption, it readily becomes apparent that the time course of adoption is *not* described by a sigmoid curve. Instead, we observe an r-shaped curve, characteristic of asymptotic growth (Figure 4.3). The rate of adoption is maximal at first and then slows as there are fewer and fewer individuals who haven't yet adopted. This makes sense. When all susceptible agents become infected at the same rate, the number of new infections will always be proportional to the number of uninfected agents, which is maximal at initialization. As this number declines, so too does the rate of new infection.

Because local interactions don't actually matter in this model, its dynamics can be computed much more simply than the way we have done it. Instead of modeling the individual agents, we can use straightforward numerical simulation of a **recursive function**, or more simply, a **recursion**. Similar to a recursive procedure in programming, a recursive function takes its own prior value as input to compute the next value. Consider a population of size

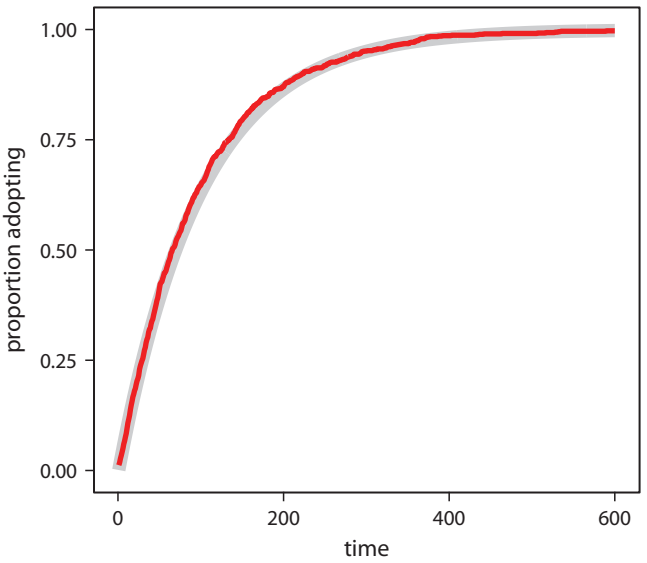

Figure 4.3 Adoption dynamics under spontaneous infection, with a population of $N = 1000$ and a spontaneous adoption rate of $\alpha = .01$. The red line shows the results of a single run of the agent-based model. The grey line shows the results of numerically simulating the mathematical model of spontaneous adoption.

N, where the number of infected agents at time t is given by I_t, and where the rate of spontaneous adoption is α. The number of infected agents at time $t + 1$ is given by the following equation:

$$I_{t+1} = I_t + \alpha(N - I_t) \tag{4.1}$$

It is sometimes useful to represent these dynamics using a **difference equation** instead of a recursion. Difference equations place the focus on how the function changes (differs) between consecutive time steps. We can rewrite the spontaneous adoption model as a difference equation by simply subtracting the value of I_t from I_{t+1}:

$$\Delta I = I_{t+1} - I_t = \alpha(N - I_t). \tag{4.2}$$

The capital Greek letter Δ (Delta) is very commonly used to represent change, and so this equation dictates how the value of I changes from one time to the next. Note that this model, like all the models in this book, uses **discrete-time dynamics**. This means that we define some discrete increment of time, and then specify how the system changes between two moments in time separated by this increment. We can make the time between events as small as we like, but the increments are nevertheless discrete. Models of dynamical systems can also be studied with continuous-time dynamics using differential equations rather than difference equations. Such an approach has several advantages, notably that one can sometimes obtain closed-form solutions for differential equations. Discrete-time modeling is nevertheless quite powerful, and also maps well onto our agent-based models, which also use discrete time. The interested reader can learn more about modeling with differential equations in Box 4.1: Coupled Differential Equations.

Numerically computing the temporal dynamics of the model from our difference equation can be easily done in any programming language and requires only that we provide

values for N, α, and I_0. While NetLogo isn't particularly advantaged for this type of analysis, it can still be done with relative ease by using lists. These allow us to store an ordered list of values, which we can use to represent the number of infected agents at each point in time. The following code does the trick:

NetLogo code 4.5

```
to analytical-solve
  let num-infected []
  set num-infected fput 0 num-infected ;;set first entry to 0
  let i 1
  while [i < 1000] [
    let x (item (i - 1) num-infected)
    set num-infected lput (x +  spontaneous-infect *
        (num-turtles - x)) num-infected
    set i (i + 1)
  ]
  show num-infected
end
```

We first create an empty list called `num-infected` and initialize its first value as zero. For 999 more time steps, we add a proportion α of the currently susceptible agents to the previous number of infected agents. Figure 4.3 shows that the curve resulting from the numerical simulation very closely matches the curve resulting from the agent-based simulation. The ABM introduces more stochasticity, and so the resulting curve is a little less smooth (and thus probably more realistic) than the one from the numerical simulation, but it is also gratifying to see that the numerical simulation accurately predicts the results of a more stochastic process. In general, when mathematical approaches are possible for the models we examine, I will usually try to include some discussion of those approaches. It is also worth noting that, while detailed descriptions are often necessary for communicating agent-based models, mathematical models are usually much easier to describe because the models are fully captured by the variables and their governing equations.

The r-shaped curve resulting from the spontaneous adoption model is distinctly different from the sigmoid adoption curve that is observed empirically for the diffusion of innovations. So we can confidently say that the sorts of innovations studied by Rogers don't diffuse this way, with everyone having the same proclivity for adoption and the same information. Must we therefore posit strongly prescribed individual differences in adoption proclivity, as Rogers did? As we will see, there is another way.

4.3 Social Influence: The SI Model

For our next model, let's stick with the assumption that everyone has the same proclivity to adopt. Wait a minute, you say, we already showed that doesn't work. Not so fast. A key assumption in the spontaneous adoption model was that all individuals had access to the same information *at the same time*. What if *that* assumption no longer held?

A key assumption of the spontaneous adoption model is that simply knowing about an innovation's existence or utility is what drives adoption. Instead, let us assume that people are influenced primarily by direct interpersonal interaction or exchange—such as directly witnessing the innovation in use by another person. Even though we will continue to ignore

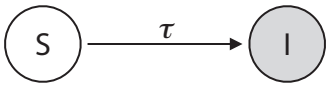

Figure 4.4 The SI model, in which τ is the transmissibility of the contagion.

differences in individuals' proclivity to adopt, this new assumption permits heterogeneity among individuals in their social interactions. If I meet lots of people who have already adopted and you do not, then I will be more likely than you to adopt the innovation. This sort of heterogeneity is not intrinsic to the individual agents in the sense of psychological characteristics, but rather emerges through variation in their social encounters. Models of this sort, in which susceptible agents become permanently infected through social influence, are sometimes called SI models (the letters stand for susceptible-infected, not social influence; see Figure 4.4).

The key idea here is that innovations might spread sort of like diseases. Assume that every exposure to an infected individual carries some risk that a susceptible individual will be convinced and adopt the innovation. To build this model, we need to consider just *how* exposure translates into adoption. In our agent-based model, agents move through space using random walks, and we can use proximity in this space to represent opportunities for exposure. Unlike in the spontaneous adoption model, the rate of interpersonal interactions matters, because agents will be influenced by encounters with their spatial neighbors. Location also matters, because some areas of space may have more infected individuals than others. The strength of social influence can be dictated by the **transmissibility** of the innovation, similar to how contagious a disease is. Presumably, not all interactions lead to transmission. Transmission depends on a number of factors. If we are talking about an innovation, these factors might include the intrinsic attractiveness and affordability of the innovation. If we are talking about a disease, factors influencing transmissibility include the health of susceptible individuals, the physiology of the pathogen, and the duration of contact between susceptible and infected individuals.

4.3.1 Transmissibility

Let τ be the transmissibility of the contagion, the probability of being influenced to adopt by someone who has already adopted. Consider a susceptible agent surrounded by several infected agents at the same time. What is the probability that this susceptible agent will become infected? If the agent has only one infected neighbor, then we're set: the probability of adopting is τ. However, things get more complicated if there are multiple neighbors who have adopted. The joint probability[4] of being infected by each of two infected neighbors is τ^2. Because τ is a probability, it's bounded in $[0, 1]$, which implies that $\tau^2 \leq \tau$. This makes sense, because being infected by each of two individuals is less likely than being infected by only one of them. But it doesn't matter, for present purposes, which neighbor successfully infects an individual or whether the infection spreads from one or more neighbors. The joint probability τ^2 is therefore not what we need in order to model how an individual with multiple infected neighbors becomes infected.

What we really want is the probability that *at least one* infected neighbor transmits the contagion. To do this, we first need to calculate the probability that the contagion *doesn't*

[4]Joint probability refers to the probability that each of two events both occur. For more on probability theory, see chapter 8.

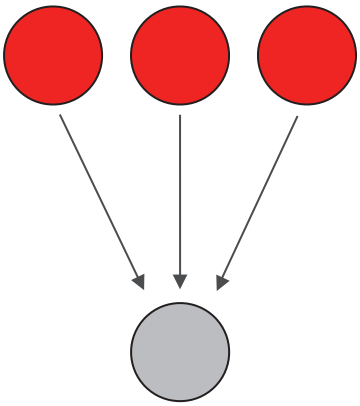

Figure 4.5 A susceptible agent (grey) encounters three infected agents (red).

spread. With multiple infected neighbors, this is equivalent to the joint probability that all of them fail to transmit the contagion. The probability that one infected neighbor fails to transmit the contagion is $1 - \tau$. If three neighbors have adopted (Figure 4.5), the probability of the contagion simultaneously failing to spread to a susceptible agent from any of its neighbors is $(1 - \tau)(1 - \tau)(1 - \tau) = (1 - \tau)^3$. More generally, with n infected neighbors, the probability of the contagion failing to spread is $(1 - \tau)^n$. So, the probability of this *not* happening—that is, of at least one neighbor transmitting the infection—is simply the inverse:

$$\Pr(\text{infection}) = 1 - (1 - \tau)^n \tag{4.3}$$

With this in mind, we can update our code, building on the code for the spontaneous adoption model. In our new model, agents will move around as before, using a random walk. When agents are sufficiently close, within distance r from one another, infected agents can transmit the contagion to susceptible agents. It's worth noting the importance of physical proximity in this model—other assumptions would be required to accurately model behavioral contagions that can spread through long-distance communication or diseases that spread via vectors such as biting flies.

4.3.2 Coding the Model

The NetLogo code for the SI model is **contagion_SI.nlogo**. To our previous model parameters we need only add the `transmissibility` of the contagion. The initialization of this model is exactly the same as the spontaneous adoption model: agents are placed in random locations on the 2-D space, and a select few of them are infected. The dynamics of the model work somewhat differently, however, with space and movement becoming critical.

At each time step, the simulation once again ends if all agents have become infected. If at least one susceptible agent remains, each susceptible agent counts the number of infected agents nearby. Of course we have to define more precisely just what is meant by "nearby." I have defined nearby agents as those whose locations are less than one spatial unit away from the focal agent, which captures the physical closeness needed for some contagions (though not others). Once we know how many infected agents are nearby, the agent becomes infected with a probability dictated by Equation 4.3. The new `infect-susceptibles` procedure is written as follows:

NetLogo code 4.6

```
to infect-susceptibles
  ask turtles with [not infected?] [
    let infected-neighbors (count other turtles with
     [color = red] in-radius 1)
    if random-float 1 <  1 - (((1 - transmissibility) ^
     [infected-neighbors) * (1 - spontaneous-infect))
    [set infected? true]
  ]
end
```

Each agent is then recolored according to its infection status and takes a step using the `move` procedure we've seen before, which involves turning a random angle dictated by `turning-angle` and then taking a step of size `speed`. The model dynamics continue until all agents are infected.

4.3.3 Analyzing the Agent-Based SI Model

Let's now examine what happens when adoption spreads via direct social influence rather than by a fixed response to global information. In the beginning of a simulation run, when few agents have adopted, most interactions are between pairs of susceptible agents and there are few new adoptions. The rate of growth will therefore be slow. As more agents adopt, the adoption rate will increase until a majority of agents have adopted. After this, adoption will slow down again since most interactions will be between agents who have already adopted. In other words, the model generates just the sort of S-shaped, or sigmoid, adoption curve observed by Rogers and others (Figure 4.6).

This simple model points to a potential explanation for why so many different innovations all have sigmoid adoption curves, and suggests that social influence likely does drive

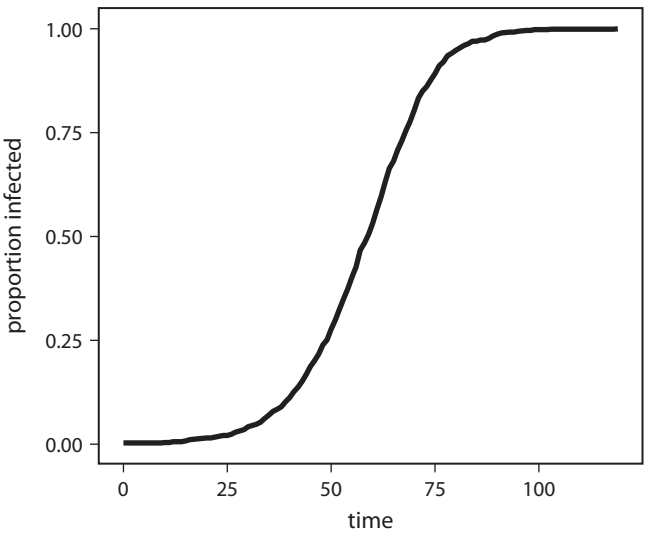

Figure 4.6 Adoption dynamics under social influence. For this run, I used 1000 agents on a 33 × 33 continuous square grid with toroidal boundaries, `speed` = 0.5, `turning-angle` = 60, and `transmissibility` = 0.05.

adoption. Indeed, SI models are regularly used to model the diffusion of innovations and behaviors. What else can we learn from this model? As I've said, it's often valuable to simply play with a model to get a feel for the sort of dynamics that occur when we change the parameters. I encourage you to do so with this model. Even experiments where the results are obvious are valuable as a check to make sure your model is running as it should. For example, if you vary the transmissibility of the contagion, you should observe that the contagion spreads much faster when `transmissibility` is 0.1 than when it is 0.01. Obviously, innovations that are more intrinsically attractive or otherwise contagious will diffuse more rapidly through a population.

Other outcomes are perhaps less obvious. The fewer agents there are—the less dense the population—the fewer interactions will occur per unit time, and the longer it will take for the innovation to diffuse. This might be counterintuitive if it seems like innovations should take longer to spread in larger populations with more people. But if adoption events are independent and rely on the density of social networks, then denser populations will adopt more rapidly. This is certainly the case with the spread of infectious diseases. Relatedly, the speed of diffusion depends on how much mixing goes on—the rate at which adopters will interact with new people who haven't yet adopted. We can use the `speed` and `turning-angle` parameters to control the contact rate and ultimately the speed of diffusion. Play around and see what kinds of relationships you can uncover.

It may seem like we have built a fairly complicated model to generate fairly simple insights. Of course, one person's simple is another person's complicated. However, the fact that the models generate smooth curves and make only simple assumptions about agents' states and dynamics is a good indication that our system may be a good candidate for analytical mathematical modeling.

4.4 The Analytical SI Model

Agents in our agent-based SI model interact at random, moving around using a random walk. This sort of random mixing is often modeled mathematically as a **well-mixed** population, in which interactions are completely random and there is no consideration of geography or social relationships. If there are multiple types of individuals in the population (such as infected and uninfected individuals), then the probability that a randomly encountered individual is of type A is simply the frequency of type A individuals in the population. You might picture a bustling city in which people are constantly moving all over the place and encountering one another willy-nilly. In later chapters we will consider group- and network-structured interactions. For now, we will derive a mathematical SI model in a well-mixed population.

Assume a large population of N individuals, for which some number, I, are infected with the contagion. The remaining $S = (N - I)$ individuals are uninfected but susceptible to infection. If two susceptible individuals meet, nothing happens, since neither is infected. If two infected individuals meet, nothing happens, since both are already infected. The interesting scenario is when a susceptible individual meets an infected individual. We therefore need to ask how often a randomly selected interaction will be one of these interesting cases that involves the possibility of a new infection. We can consider a focal individual and ask about their probability of infecting a new individual. By considering infections only from this perspective, and not the perspective of the focal individual becoming newly infected, we avoid double counting when we sum over the entire population. Because of our assumption that individuals meet at random, the probability of an encounter that could lead to a

new infection in this way is the joint probability that the focal individual is infected and the individual they encounter is susceptible:

$$\Pr(S, I) = \frac{I}{N}\left(\frac{N-I}{N}\right) \tag{4.4}$$

Note that this is the probability that the interaction *could* lead to a new infection; it says nothing about whether it will.

As before, let τ be the transmissibility of the contagion. The probability that a random interaction leads to a new transmission is the joint probability that the interaction is between an infected individual and a susceptible one and also involves a transmission event:

$$\Pr(\text{new infection}) = \tau\frac{I}{N}\left(\frac{N-I}{N}\right) \tag{4.5}$$

This equation characterizes a new infection resulting from a single interaction. However, we are interested in changes to the population at large—the sum total of all interactions. We therefore need to describe how the number of infected individuals changes over time.

There are N individuals who could be the focal agent in an interaction. For each of those N individuals, we take the joint probability that the focal agent is infected, the individual they encounter is susceptible, *and* such an encounter leads to a new infection. The number of new infections at each time step is therefore given by

$$\#\text{ new infections} = \Delta I = N\tau\frac{I}{N}\left(\frac{N-I}{N}\right) \tag{4.6}$$

To model the dynamics of infections over time, we can turn this equation into a recursion for the number of infected individuals at time $t+1$:

$$I_{t+1} = I_t + N\tau\frac{I_t}{N}\left(\frac{N-I_t}{N}\right) \tag{4.7}$$

This equation is still a bit busy, so let's do some algebra to simplify it. There are a lot of N's in the equation above. Using the magic of mathematics, we can cross some of them out, leaving us with:

$$I_{t+1} = I_t + \tau I_t\left(1 - \frac{I_t}{N}\right) \tag{4.8}$$

Inspecting Equation 4.8 tells us a few things. First, the infection will spread faster if τ increases. This makes sense, and it also fits with what we observe in our simulations. If you have a more contagious infection, the number of infected individuals will increase more quickly. How does the overall time trajectory of the contagion play out? When I_t is very small, few interactions will lead to new transmissions, because most interactions will involve two susceptible individuals, neither of whom can spread the contagion. When I_t is large and close to N, the term $(1 - I_t/N)$ will be close to zero, and the rate of new transmissions will similarly be small because most individuals will already be infected. This indicates that the rate of change should be largest when an intermediate fraction of the population is infected, and indeed this is exactly what we find if we run numerical simulations of Equation 4.8. Figure 4.7 shows that the infection curve takes a nice sigmoid shape, where the overall time to universal infection is determined entirely by τ.

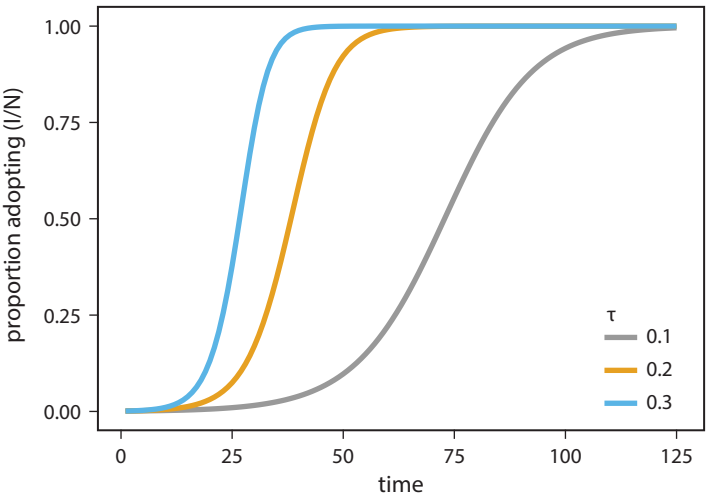

Figure 4.7 Dynamics of the SI model for varying τ.

Notice how the factors related to agent movement that we discussed in terms of the agent-based model—density, speed, and turning angle—appear to be absent from our analytical model. Indeed, all the factors contributing to variation in transmission are captured by the single variable, τ. In some models, this term is explicitly decomposed into two components: τ is used to represent the direct transmissibility that occurs during a contact, and another term, c, is used to represent the per capita contact rate, so that the **effective transmissibility** of the contagion, β, can be characterized as $\beta = c\tau$. This is the joint probability that an individual has a new encounter *and* that the encounter leads to a new infection (conditional on it being between susceptible and infected individuals). Framing it this way highlights that one could decrease the rate of new infections either by decreasing the probability of transmission during an interaction or by decreasing the rate at which interactions occur. This formulation also has the advantage of neatly capturing key sources of variation in transmission as well as mapping more directly onto our agent-based model. That said, the agent-based model still affords us a more nuanced view of how interpersonal contact can vary. Movement speed might represent the duration of individual interactions, while turning angles might indicate the degree to which social networks are cliquish and assortative. In general, analytical and agent-based models are often complementary, and insights are often most abundant when the two approaches are combined.

4.5 Getting Better: The SIS Model

A key assumption of the SI model is that once people are infected—or once they have adopted whatever product or behavior we're talking about—that's it. They remain infected for as long as they live (or at least for as long as the model is relevant). This is a pretty strong assumption. In terms of infection, sometimes people recover and are no longer infected. In terms of products or behaviors, sometimes people decide that they don't want or need them anymore and so stop using them. They might give up the product or behavior forever, or they might simply take a hiatus. Let's first focus on the latter case, so that people can recover (or disadopt), and thereafter return to a renewed state of susceptibility. For the

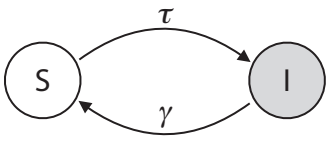

Figure 4.8 The SIS model.

remainder of this chapter, I will default to talking about contagions in terms of infectious diseases. However, many of the lessons will continue to apply to the spread of innovations or behaviors.

In technical terms, we now move from the SI model to the SIS (susceptible-infected-susceptible) model (Figure 4.8). How does recovery work? The simplest assumption is that infected individuals recover with a fixed probability for every unit of time they are infected. We therefore need to add just one additional parameter to our model: γ, the *recovery rate*. Every time step, a proportion γ of infected individuals recover and become susceptible again, so that the number of susceptible individuals is increased by γI and the number of infected individuals is reduced by that same amount. This means that the expected duration[5] of an infection is $1/\gamma$. We can update our recursion for the number of infected individuals at time $t+1$ accordingly:

$$I_{t+1} = I_t + \tau I_t \left(1 - \frac{I_t}{N}\right) - \gamma I_t \tag{4.9}$$

At each time step, susceptible individuals are becoming newly infected, but infected individuals are also recovering and becoming newly susceptible. So what happens? If the contagion effectively spreads, the population will settle to an equilibrium where the rate of new infections equals the rate of recovery. Because of this, the entire population never becomes fully saturated with the contagion, as it did in the SI model, because infected individuals are recovering at the same time as new individuals are getting infected. This model achieves what is known as a **dynamic equilibrium**, in which a fixed proportion of the population remains infected even though the individuals in the population continue to change their infection status.

We can use our mathematical model to calculate the precise proportion of the population that will remain infected at equilibrium. This occurs when there is no longer any change in the total number of infected individuals; that is, when $I_{t+1} = I_t$. We begin by rewriting Equation 4.9 and letting $I = I_{t+1} = I_t$:

$$I = I + \tau I \left(1 - \frac{I}{N}\right) - \gamma I \tag{4.10}$$

I've omitted the time subscripts here because we are considering the case where there is no change between times t and $t+1$, and so time, in a sense, becomes irrelevant. We can cancel out some of the I terms, which leaves us with:

$$\tau I \left(1 - \frac{I}{N}\right) - \gamma I = 0 \tag{4.11}$$

[5]Using a fixed probability of recovery implies that infection durations will be exponentially distributed, which in turn implies that the most common infection duration will be a single time step. This is an unrealistic but mathematically useful assumption. See Yan (2008) for a discussion of models incorporating other distributions.

Now let's rearrange the equation by moving γI over to the right side of the equals sign, and have some fun with algebra:

$$\tau I\left(1-\frac{I}{N}\right)=\gamma I \tag{4.12}$$
$$\tau\left(1-\frac{I}{N}\right)=\gamma$$
$$1-\frac{I}{N}=\frac{\gamma}{\tau}$$

One more rearrangement gets us to

$$\boxed{\frac{I}{N}=1-\frac{\gamma}{\tau}} \tag{4.13}$$

This is the equilibrium infection rate, the proportion of the population that always remains infected. We can use this equation to check our intuitions about the model. The equilibrium infection rate goes down when transmissibility is lower or when recovery rate is higher, which makes sense.

It's worth taking a bit of time to get some additional intuition about why this equilibrium occurs. One approach is to plot the difference equation as a function of the infection rate. Dividing all sides by N to get rates rather than quantities of infection, we get the difference equation

$$\Delta I=\tau I(1-I)-\gamma I \tag{4.14}$$

If we plot ΔI as a function of I, we end up with the plot in Figure 4.9 (left side). This sort of graph is very useful. Note that when I is below 0.75 (the equilibrium infection rate for the values of τ and γ used in this example), the rate of change is positive. This means that the number of infections will increase. However, when I is above 0.75, the rate of change becomes negative, meaning that infections will decrease. In this way, a stable equilibrium is

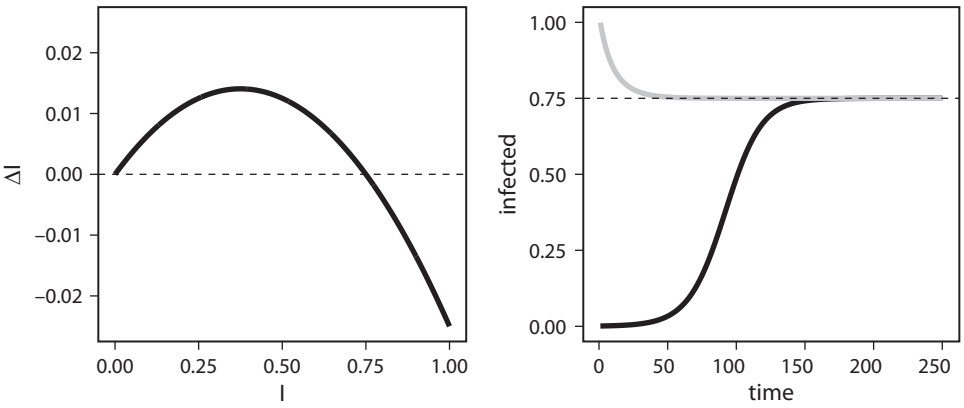

Figure 4.9 Equilibrium dynamics of the SIS model. Left: the difference equation for ΔI plotted against I, where I is the rate of infection. Right: infection dynamics starting at either $I(0)=0.001$ or $I(0)=0.999$. In all cases, $\tau=0.1$ and $\gamma=0.025$.

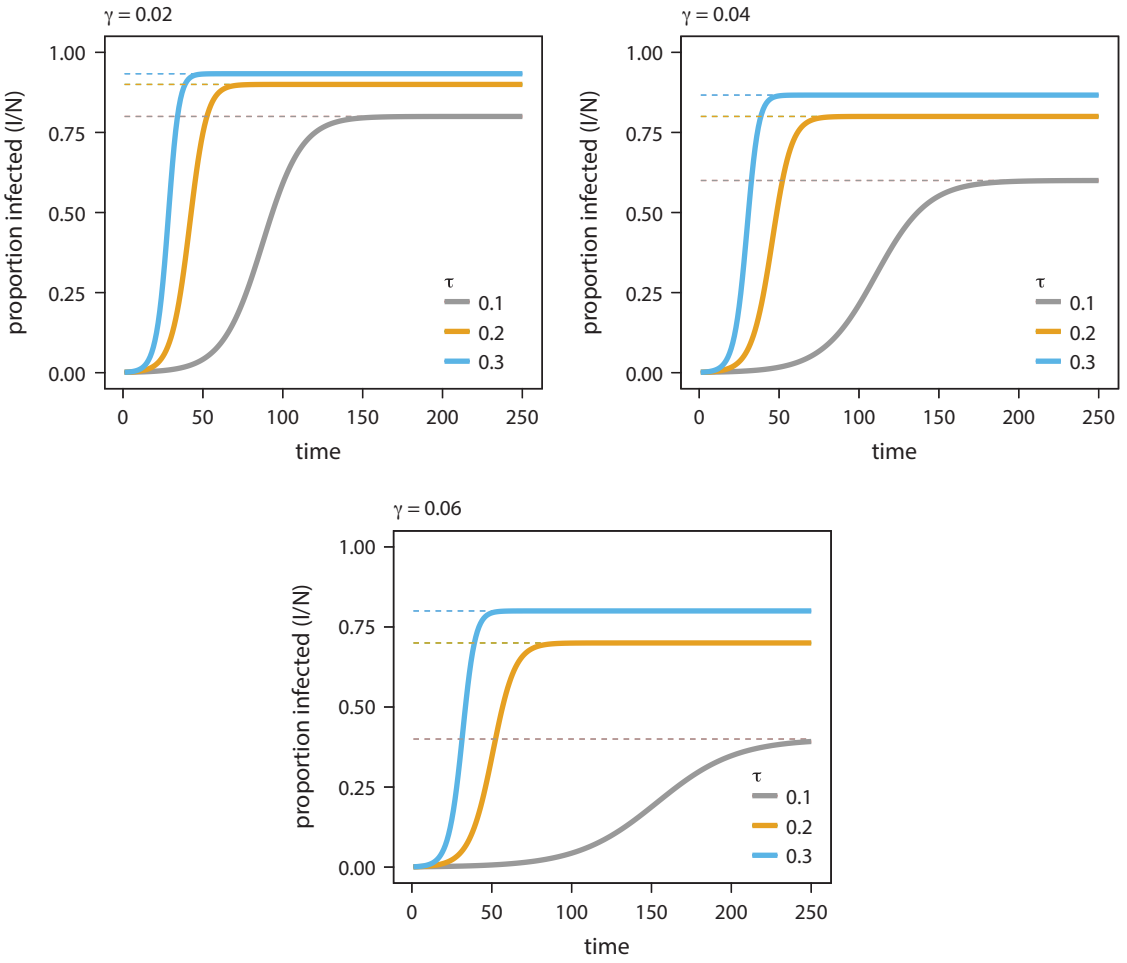

Figure 4.10 SIS model dynamics for several values of effective transmission rate ($c\tau$) and recovery rate (γ). The dashed lines indicate the equilibrium frequencies of infection.

reached at $I = 0.75$, where the rate of change is zero. Note also that there is another value of I for which the rate of change is zero: $I = 0$. This makes sense, because there can be no new infections when no one is infected. However, this is an **unstable equilibrium**, because a slight perturbation (that is, the introduction of new infections) will yield a positive rate of change and push the system to its stable equilibrium. This situation is further illustrated by the right side of Figure 4.9, which plots numerical simulations of the infection rate starting when the infection rate is either close to zero or close to one. In both cases, the infection rate returns to its stable equilibrium value.

While the ratio γ/τ determines the equilibrium proportion of the population infected, each parameter contributes to the dynamics with which the system reaches that equilibrium. Figure 4.10 shows the dynamics for the same parameters used above in our analysis of the SI model (Figure 4.7), for different values of γ. Compare this figure with the SI model in Figure 4.7 (the equivalent of the SIS model in which $\gamma = 0$). Notice that, in addition to the equilibrium infection rate dropping, the speed at which the infection reaches that

equilibrium is slower for higher recovery rates, and the overall course of the contagion through the population is slower. This makes sense if you consider that individuals are recovering as well as becoming infected at each time step.

The analytical model we have just derived indicates that if we know the effective transmissibility rate and the recovery rate for the contagion, we can determine the equilibrium frequency of infection (or adoption) with confidence, and that infection frequency should remain stable. Let's investigate this claim further and see what happens when we model individuals explicitly as agents moving in space who become infected due to physical proximity with an infected agent. This will introduce a very small amount of the heterogeneity that real-world systems have in spades.

4.5.1 The Agent-Based SIS Model

The NetLogo code for this model is **contagion_SIS.nlogo**, and it builds upon the code for the SI agent-based model. We must add one new global parameter, the recovery rate (`recovery-rate`). The initialization of this model is the same as the SI model, and the dynamics differ primarily in the inclusion of recovery, though I have also added the provision that the simulation will stop if zero agents are infected, indicating a failed outbreak.

The principal dynamics in the SI model involved three stages at each time step. First, susceptible agents could become infected from contact with infected agents. Second, agents were recolored based on infection status. And third, agents moved using a correlated random walk. To this we will add a new stage before the recoloring: recovery. During this stage, infected agents may recover and become susceptible again. Importantly, we don't want a new infection to be immediately negated by having a newly infected agent recover before the agent has the opportunity to infect its neighbors (or, more importantly, before it has actually been infected for any amount of time). Here is where careful consideration of one's code is crucial. If we use only the value of an agent's `infected?` parameter, an agent could flip from susceptible to infected and back to susceptible on a single time step. One way to avoid this is to use two different markers of agent infection. Luckily, we already have this redundancy built into our code, since an agent's infection status is represented both by the Boolean variable `infected?` and by the agent's color. If newly infected agents remain white until the end of the time step, we can restrict recovery to agents who are both infected and red. The code for the procedure `recover-infecteds` is as follows:

NetLogo code 4.7

```
to recover-infecteds
  ask turtles with [infected? and color = red]
  [
    if random-float 1 < recovery-rate [
      set infected? false
    ]
  ]
end
```

When simulations of this model are run, we should notice that the dynamics of infection do resemble those of the analytical model, in the sense that the proportion of infected individuals in the population climbs to its equilibrium value and then hovers around this

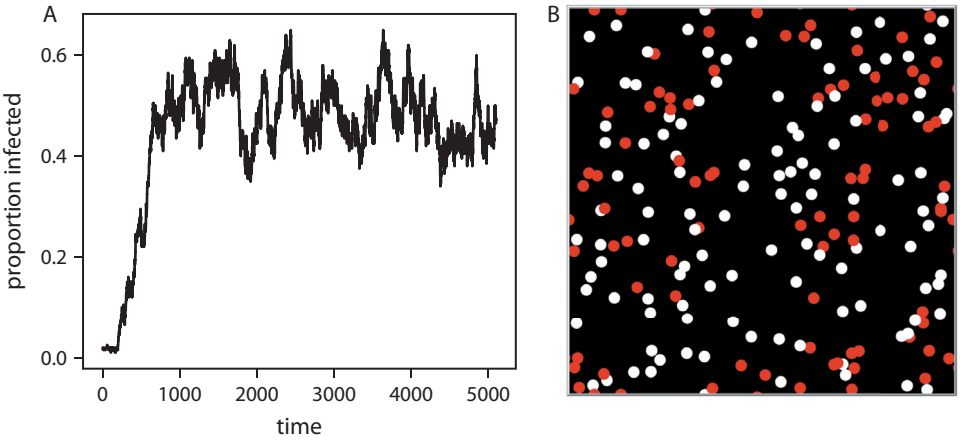

Figure 4.11 Agent-based SIS model dynamics. (A) The proportion of infected agents plotted over time. (B) Spatial positions of infected (red) and uninfected (white) agents at $t = 3400$. For this simulation, `num-turtles` = 200, `speed` = 0.5, `turning-angle` = 360, `transmissibility` = 0.1, and `recovery-rate` = 0.02.

value indefinitely. An important *difference* between the agent-based and analytical models is that the infection rate in the former does not completely stabilize as neatly as predicted by the latter. Rather, the infection rate oscillates around the equilibrium value due to stochastic fluctuations in the spatial distribution of infected agents. Decreasing the mobility of agents by decreasing the population density and increasing their maximum turning angle exacerbates this effect and can create dynamic spatial clusters of infection and non-infection (Figure 4.11).

This disparity again highlights the complementary nature of the analytical and agent-based approaches. The analytical model shows us how to calculate the equilibrium infection rate. The ABM shows us that this calculation is merely an estimate of the average, and that fluctuations around that average are likely and will probably cluster in nonrandom ways on social networks. In general, it is almost always useful to be able to look at the same problem from multiple angles. Agent-based models of contagion are especially useful when one wants to consider highly structured interactions and/or heterogeneity in individuals or environments. Analytical models are useful because one can calculate certain parameters exactly and explore the consequences of other parameters instantly without needing to run computationally demanding simulations. In the next subsections, we will consider two useful values that can be derived analytically: the basic reproduction number, R_0, and the critical vaccination threshold needed for herd immunity.

4.5.2 Will It Spread? Calculating R_0

The preceding analyses implicitly assumed that, following an outbreak, a contagion will spread until it reaches its equilibrium infection level. However, another outcome is possible: the initial outbreak may fizzle out before it spreads very far. This can happen if the first individuals to become infected recover before they have a chance to transmit the disease further. What factors might determine whether this happens? After some thought, your intuition should hopefully be that whether or not an initial outbreak spreads will depend on the contact rate, the transmissibility of the contagion, and the recovery rate. If contact rate

is low (so that people have few chances to transmit the contagion), transmissibility is low (so that contact events rarely lead to new transmissions), and the recovery rate is high (so that infected individuals recover quickly before they can spread the infection), an epidemic might be avoided. These intuitions are valuable, but they can be improved by quantifying them.

Recall that the dynamics of the SIS model are captured by Equation 4.9:

$$I_{t+1} = I_t + \tau I_t \left(1 - \frac{I_t}{N}\right) - \gamma I_t$$

The term in parentheses, $(1 - \frac{I_t}{N})$, will be very close to one at the start of an outbreak in a large population, because $\frac{I_t}{N}$ will be close to zero (that is, almost everyone is uninfected). Thus, we can rewrite the recursion equation as:

$$I_{t+1} \approx I_t + I_t(\tau - \gamma) \tag{4.15}$$

We can rewrite this as a difference equation, which allows us to drop the subscripts:

$$\Delta I \approx I(\tau - \gamma) \tag{4.16}$$

This equation should make it clear that the number of infected individuals will increase only if $\tau - \gamma > 0$. Rearranging this, the infection will spread if

$$R_0 = \frac{\tau}{\gamma} > 1, \tag{4.17}$$

where R_0 (usually pronounced "R-naught") is known as the **basic reproduction number**. This is the expected number of new infections an infected person will generate in a fully susceptible population before they recover. Calculating this confirms our intuition that an infection is more likely to spread if the effective transmission rate is high and the recovery rate is low. Specifically, it will spread if the rate of new infections is greater than the rate of recovery. Note that R_0 is emphatically *not* purely a property of the contagion. Rather, it is affected by *all* the factors that influence transmission and *all* the factors that influence recovery. It is an emergent property of the interaction between a contagion and the environment, including the social environment. Thus, social and behavioral measures can change R_0. If social behaviors are relatively predictable, at least in the aggregate, then R_0 can be—and often is—calculated for specific contagions, which can help researchers to prioritize certain outbreaks over others. Calculating R_0 is additionally useful because it gives us access to another tool. It allows us to study a phenomenon called **herd immunity**.

4.5.3 Vaccines and Herd Immunity

We have just learned that we can decrease the chance that an infection will spread throughout a population by decreasing its effective transmissibility. One way to reduce transmission is through vaccination. Of course, not all infections have available vaccinations, and not all vaccinations are perfect. Vaccines are a politically loaded topic in some circles, and there are important ethical and medical concerns surrounding vaccines. Vaccines have saved millions of lives, and as such their benefits usually outweigh their societal costs. However, we are not

here to debate the benefits of vaccines but instead to do what modelers do best, which is to consider the consequences of particular assumptions.

We will assume that (1) vaccines are administered at random to a proportion V of the population before the initial outbreak of infection, and (2) vaccines are perfectly effective, so that a vaccinated individual can neither become infected nor transmit the infection to others. These assumptions can later be changed as needed, but making them simplifies the math and allows us to establish a baseline. We can now write a new difference equation for the change in the number of infected individuals that incorporates vaccinations:

$$\Delta I = \underbrace{\tau\left(1-\frac{I}{N}\right)}_{A}\underbrace{(1-V)}_{B}I - \gamma I \tag{4.18}$$

In this equation, A is the probability of encountering a susceptible individual and transmitting the infection, and B is the probability that the susceptible individual is not vaccinated and therefore *becomes* infected. At the onset of an outbreak (when I is small), this equation can be approximated as

$$\Delta I \approx I\left(\tau(1-V)-\gamma\right) \tag{4.19}$$

Following the same logic as in the previous section, the infection will spread when rare if and only if

$$\frac{\tau}{\gamma}(1-V) > 1 \tag{4.20}$$

The left side of this equation can be rewritten as $r_0 = (1-V)R_0$, where r_0 is known as the *effective* basic reproductive number.

What good is this? Well, if we have information about the basic reproduction number for an infection and we have an effective vaccine, we can calculate the proportion of the population we need to vaccinate to stop the spread of disease, V^*. Importantly, we don't need to vaccinate everyone, just enough to drop the rate of transmission below the rate of recovery. This phenomenon is known as *herd immunity*. The unvaccinated proportion of the population is kept safe by vaccinating enough individuals to stop the disease from spreading. Note that herd immunity does not mean that unvaccinated individuals will never become infected, just that only a small number of infections will occur before the infection dies out.

We can calculate the threshold vaccination rate for herd immunity by letting V^* denote the smallest value for which $r_0 \leq 1$. We can derive this as follows:

$$\begin{aligned}(1-V)R_0 &\leq 1 \\ R_0 - R_0V &\leq 1 \\ R_0 - 1 &\leq R_0V \\ 1-\frac{1}{R_0} &\leq V\end{aligned} \tag{4.21}$$

The threshold vaccination rate for herd immunity is therefore:

$$V^* = 1-\frac{1}{R_0} \tag{4.22}$$

It is clear that when the basic reproduction rate, R_0, is larger—that is, when the disease will tend to spread more rapidly—larger proportion of the population will need to be

Figure 4.12 The SIR model, in which τ is the transmissibility of the contagion and γ is the recovery rate.

vaccinated to avoid an outbreak. If we combine this information with what is known about the transmission rates of certain diseases in high density areas (e.g., where R_0 will tend to be highest), we can calculate the minimum vaccination rate needed for herd immunity for different diseases. If vaccine rates are high enough, the disease will fizzle out via recovery before it can infect a large number of individuals. If the proportion of vaccinated individuals drops below V^*, however, an outbreak can occur in which most or all of the unvaccinated individuals will contract the disease. This is why vaccination cannot be viewed merely as a choice for individuals to make. Vaccination only provides herd immunity—which is especially important for at-risk people who either cannot be vaccinated or for whom vaccines are ineffective—if sufficiently high numbers of people vaccinate.

Calculations of quantities like R_0 and V^* should be treated with caution, however, because they stem from simple models that make a lot of strong assumptions. For example, the models we have examined assume that individuals don't change their behaviors when they are infected, that infected individuals are instantly contagious, that vaccines are perfectly effective, and, most importantly, that populations are perfectly well mixed. Incorporation of these and other factors into a model's assumptions may change the conclusions one can draw from that model. We will return to this discussion of model complexity at the end of the chapter. For now, we will continue extending our simple model for one last hurrah.

4.6 Staying Better: The SIR Model

The SIS model assumes that a recovered individual can immediately become infected once again. This isn't the case for many contagions. Antibodies may provide immunity. More pessimistically, infection may lead to death or removal from the population by other means. In the case of innovations or behaviors, an individual may abandon them for good.

In the SIR model, individuals can be in one of three states: susceptible, infected, or removed.[6] Susceptible agents can become infected at a rate proportional to the transmissibility τ, and infected agents become removed at the recovery rate γ, just as in the SIS model. The main difference is that, in the SIR model, removed individuals remain removed and cannot again become infected (Figure 4.12).

The SI, SIS, and SIR models belong to a family of models known as **compartment models**, because they can be viewed as keeping track of flows between different "compartments," which represent the different states in which individuals can be found. With only two compartments, as in the SI and SIS models, systems can be modeled using a single difference equation, because individuals who are not in one compartment are necessarily in the other one. With three compartments, we need at least two equations, with the third compartment implied by the proportion of the population not in the other two compartments. For clarity,

[6]The "R" sometimes stands for "recovered," with the understanding that recovered agents are immune. However, I think this too easily leads to confusion between the SIS and SIR models, so I have chosen to use "removed," which is also consistent with most common usage in mathematical epidemiology. In its most general sense, removal from the population of susceptible or infected individuals can occur via recovery with immunity as well as death.

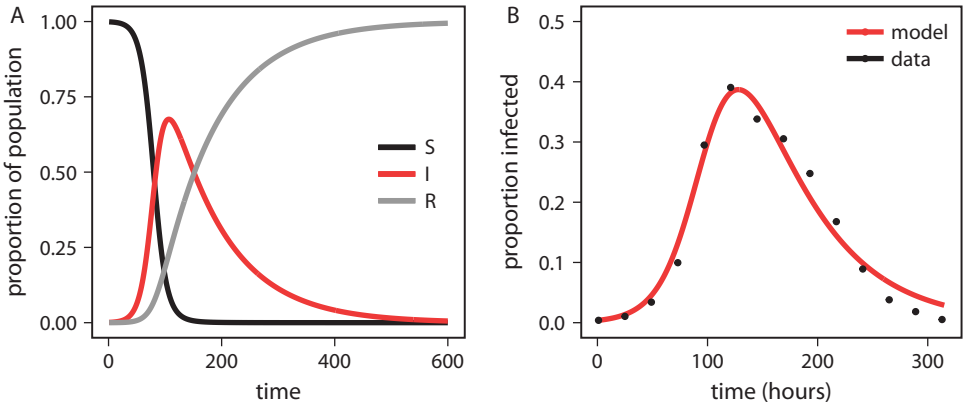

Figure 4.13 (A) Dynamics of the discrete-time SIR model from numerical simulation, using $\tau = 0.1$, $\gamma = 0.01$. (B) Fit of an SIR model to data from an influenza outbreak in a British boarding school in 1978.

I will use three equations. I have chosen to show difference equations rather than recursions, in order to place the focus on how the size of each compartment changes over time:

$$\Delta S = -\tau S \frac{I}{N} \tag{4.23}$$

$$\Delta I = \tau S \frac{I}{N} - \gamma I$$

$$\Delta R = \gamma I$$

The number of susceptible individuals decreases by the expected number who encounter an infected individual, times the probability of transmission. The number of infected individuals increases by this same number and decreases by the expected number of already-infected individuals who recover or are removed. And the number of removed individuals increases by the same number.

The only equilibria for this model are for everyone to either remain susceptible or become removed—infection is an unstable state of being. As such, the number of infected individuals will first rise and then fall. I've numerically simulated Equations 4.24 with some arbitrary values for τ and γ; Figure 4.13A shows the path of the resulting epidemic.

The SIR model is foundational for mathematical epidemiology, because it is the simplest model that can capture the full time course of an epidemic. As mentioned, it and the other contagion models we've considered make the simplifying assumption of a well-mixed, closed population. This assumption will often fail to hold in a strong sense, though extensions can be made to the model to accommodate many common objections. However, the model does provide a good baseline. There's a well-known case in epidemiology of an influenza outbreak in a British boys' boarding school in 1978 (Murray, 1989). A boy came back sick from a holiday, and infections were tracked during which time boys neither entered nor left the school. As seen in Figure 4.13B, an SIR model can be made to very closely match the true path of the outbreak. There are, however, differences between the data and the model. One source of difference stems from the fact that the mathematical model allows for fractional individuals to be infected, which thereby increases the infection rate and prolongs the time course of the epidemic. This type of assumption is reasonable in a

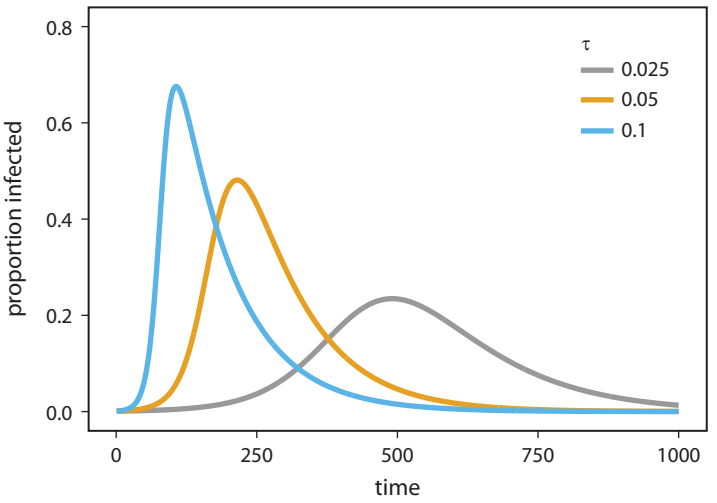

Figure 4.14 Temporal dynamics of infection in the analytical SIR model for different values of transmissibility, τ. In all cases $\gamma = 0.01$.

very large population but leads to issues in a relatively small population like a school (which had a total enrollment of 763). Agent-based models avoid this problem, but at the cost of lowered mathematical tractability.

4.6.1 Flattening the Curve

Starting in 2020, the COVID-19 pandemic brought considerable attention to the importance of "flattening the curve," referring to a reduction in transmissibility in order to decrease the maximum number of individuals infected at any given time. This allows more time for public health researchers to develop treatments, and also decreases the chance that health care facilities will become overwhelmed with too many patients. The SIR model helps us to understand exactly what those advocates were talking about. Consider Figure 4.14, which plots infection trajectories for three different values of τ, calculated via numerical simulation of the SIR recursions.

It should be clear that measures that reduce effective transmissibility are a good thing, all things being equal. It *should* be clear. However, it is not always easy to imagine how various factors really do influence transmissibility. Individuals are not a monolith. Rather, each of us operates in our own private world, aware primarily of our own existence and that of our closest social relations. It can be difficult to see how our own behavior contributes to the whole. In the case of adopting measures to reduce transmissibility, the assumption of homogeneity has another glaring mismatch with reality, which is that public health measures are almost never adopted uniformly or instantaneously. It is difficult to capture these heterogeneities in an analytical model. For a complementary approach, we will once again turn to our agent-based model.

4.6.2 The Agent-Based SIR Model

The NetLogo code for this model is **contagion_SIR.nlogo**. We need only to make very minimal changes to our code for the SIS model. In fact, it will be useful to be able to easily

toggle back and forth between the two models, since their only difference is whether or not recovered individuals become immune to further infection. This is generally a good strategy when working with a family of related models—it is advisable to create code where different model assumptions can be toggled by varying the value of certain parameters. In this case, we'll introduce a global Boolean variable called `remove-recovered?`. When this is true, we'll be working with the SIR model, in which recovered agents are "removed" and so become immune to further infection. When it is false, we'll revert to the SIS model in which recovered agents return to a susceptible state. The SIR model does not require the addition of any new global parameters.

Previously, we added an agent-level Boolean variable that tracked whether or not the agent was infected. To this, we'll add another agent-level Boolean variable that tracks whether an uninfected agent can become infected—I've called it `immune?`. When this is true, the agent is immune and cannot become infected. At initialization, `immune?` will be false for all agents. In NetLogo, the declaration of the agent-level variables looks like this:

NetLogo code 4.8

```
turtles-own [
  infected?
  immune?
]
```

Other than the addition of the `immune?` variable, the initialization of the SIR model is identical to the SIS model. The stages of the model dynamics are also the same: first, susceptible agents can be infected; second, infected agents can recover; third, agents are recolored based on their infection status; and fourth, the agents move. The devil, however, is in the details, and we will need to update the first, second, and third stages to account for removed agents. Let's start with recoloring. Since color is used not only for visualization but also as information to guide agent behavior, we need a third color to signify removed status. I have chosen to make removed agents grey. Here is our simple procedure to recolor agents, now only slightly less simple:

NetLogo code 4.9

```
to recolor
  ask turtles [
    ifelse infected?
    [set color red]
    [ifelse immune?
      [set color grey]
      [set color white]
    ]
  ]
end
```

The procedure to infect susceptibles is exactly the same as before, but we now have to exclude removed agents from the set of potential new infections. Here is the code to accomplish this:

NetLogo code 4.10

```
to infect-susceptibles
  ask turtles with [not infected? and not immune?][
    let infected-neighbors (count other turtles with
     [color = red] in-radius 1)
    if (random-float 1 <  1 - (((1 - transmissibility) ^
      infected-neighbors) * (1 - spontaneous-infect)))
    [set infected? true]
  ]
end
```

Similarly, we have to update the recovery procedure so that agents who recover become immune, but only if we are using an SIR model. Recall that we introduced a Boolean variable, `remove-recovered?`, in order to toggle between the SIR and SIS models. When this Boolean variable is true, we need to set recovered agents to become not only uninfected, but also immune.

NetLogo code 4.11

```
to recover-infecteds
  ask turtles with [infected? and color = red]
  [
    if random-float 1 < recovery-rate [
      set infected? false
      if remove-recovered?
      [set immune? true]
    ]
  ]
end
```

That's it. You should confirm the model works as expected, and play around with it. Notice that there are sometimes a few agents who never become infected. The number of such agents should increase as the effective transmissibility of the infection decreases and as the recovery rate increases. With this in mind, I want to return to our earlier discussion about flattening the curve.

Recall that in the analytical model, τ is really the *effective* transmissibility, which combines the probability that an interaction occurs with the probability that it leads to a transmission. In the ABM, these factors are explicitly separated. The probability that contact leads to a transmission event is captured by our `transmissibility` parameter. The probability of that contact occurring in the first place emerges endogenously from agents' movement strategies (controlled by the speed and turning angle parameters) and the population density (controlled by the population size and the size of the space). This helps bring into focus the fact that many things contribute to the spread of a contagion by virtue of their contributions to its effective transmissibility. As an example, let's focus on the parameter `speed`. I ran several simulations of the SIR ABM, holding all parameters fixed except `transmissibility` and `speed`. Figure 4.15A shows that, as expected, reducing direct transmissibility can flatten the curve. The probability of interactions stays the same, but the probability of an interaction leading to a new infection is diminished. You can think of this as akin to wearing face masks that block the spread of airborne droplets. Figure 4.15B shows that, even without reducing direct transmissibility, reducing agents' speed—which in turn

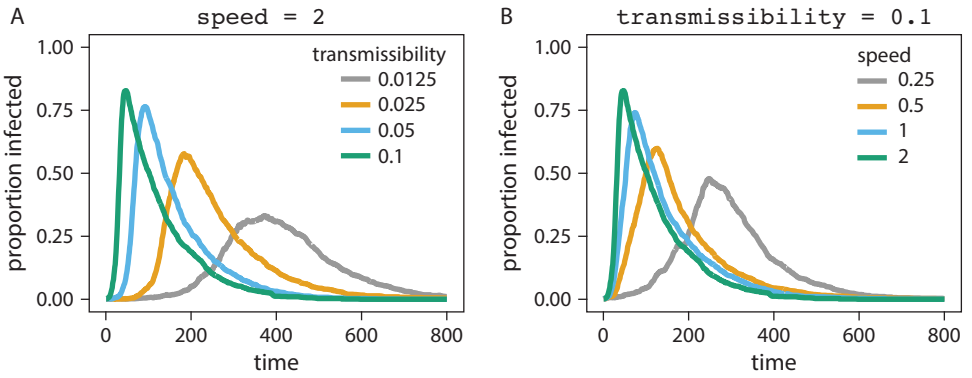

Figure 4.15 Temporal dynamics of infection in the agent-based SIR model. (A) Reducing transmissibility flattens the curve, holding speed constant. (B) Reducing speed flattens the curve, holding transmissibility constant. In all cases, `num-turtles` = 1000, `init-infected` = 1, `turning-angle` = 360, `recovery-rate` = 0.01.

reduces the number of new individuals an agent interacts with in a given time period—also flattens the curve. Think of this as akin to social distancing and avoiding crowds. It is left as an exercise to the reader to show that when both transmissibility and speed are reduced, the effect can be additive, but also that when adoption of these measures is sufficiently low, the effects can be muted.

BOX 4.1: Coupled Differential Equations

We have modeled the change in the relative sizes of the S, I, and R compartments under the assumption of discrete time units. This assumption is intuitive and allows us to analyze complex social dynamics using simple algebraic manipulation. However, dynamical systems are often modeled over *continuous* time, using **differential equations**. If you have ever studied physics or had a course in differential calculus, you will be familiar with the idea that a differential equation can specify a function for the instantaneous rate of change in some variable for any arbitrary time t. What makes differential equations powerful is that one can define the rate of change in a variable as a function of its own current value. When the differential equations defining how multiple quantities change over time include one another as variables, the equations are said to be *coupled*. For example, our SIR model can be rewritten as a set of three coupled differential equations like this:

$$\begin{aligned} \frac{dS}{dt} &= -\tau S \frac{I}{N} \\ \frac{dI}{dt} &= \tau S \frac{I}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{4.24}$$

Note that the right side of each equation is exactly the same as in the discrete-time difference equations, but now each defines the *instantaneous* rate of change at any time t.

What do we get out of this? Well, other than being pretty, differential equations lend themselves to a number of useful analysis techniques, and can sometimes even yield closed-form analytical solutions by which we can obtain the state of the system for any arbitrary time t without numerical simulation. Consider our simple model of spontaneous adoption. The differential equation defining this system is:

$$\frac{dI}{dt} = \alpha\,(N - I(t)) \tag{4.25}$$

Knowledge of differential equations allows us to obtain a closed-form solution for the infection rate:

$$I(t) = N\left(1 - e^{-\tau t}\right) \tag{4.26}$$

This means that if we know that a system is governed by Equation 4.26 and we know the initial conditions, we can precisely predict the infection rate at any other time. Coupled differential equations can also generate cyclical or even chaotic dynamics. Their use is worth mastering.

Despite their value, I have avoided the use of differential equations throughout this book. I do this for several reasons. First, one of the goals of this book is to make the skills and ideas presented here available to those with minimal mathematical training. Second, social behaviors are often best represented as discrete events rather than as things unfolding in continuous time. And third, discrete time is the default in most agent-based models, so connecting analytical models to agent-based simulations is more intuitive using discrete-time models. For an excellent introduction to the mathematical and numerical analysis of dynamical systems, with a heavy emphasis on differential equations, see Strogatz (2015). For innovative social science approaches, see Turchin (2003).

4.7 Reflections

This chapter has provided a brief introduction to a very large domain of modeling. The field of mathematical epidemiology is built on sophisticated extensions to the sort of compartment models we examined here, as well as on the use of agent-based models that incorporate things like geography, network topology, demography, and life history. The use of contagion models to study behavior change and the diffusion of innovations is also widely used in sociology, marketing, and other areas of computational social science.

The first draft of this chapter was begun in 2018, when contagion seemed like an important topic with which students of human social behavior should engage. I then revised this chapter throughout 2020, 2021, and 2022, while the world was reeling from the COVID-19 pandemic. Understanding contagion now seems absolutely paramount. The treatment it has been given here is very introductory. There are many books written on the mathematics of disease contagion alone,[7] and there is a growing literature on its relation to "epidemics" of information and misinformation.[8] Nevertheless, I hope that I have provided you with a firm

[7] For example, see Bjørnstad (2018).

[8] For example, see O'Connor and Weatherall (2019).

foundation from which to learn more elsewhere. Indeed, this is a primary goal for many of this book's chapters.

Contagion is a complex phenomenon, and the models presented here are, as usual, radical simplifications. We can and should probe more deeply at multiple levels. Here are a few noteworthy limitations that can and should be addressed by more advanced models. First, individuals vary in their ability to transmit and adopt contagions, as well as their ability to recover or become reinfected. Second, individuals do not meet up at random, but live and move in structured populations. Indeed, people move both within and between communities in nonrandom ways. Third, individuals are affected by cultural norms, socioeconomic circumstances, and identity. This affects who they interact with and who they learn from. All of these factors (and others) affect how contagions spread.

In the models I have presented I have treated the spread of behaviors and innovations as perfectly analogous to the spread of disease, but they are not really the same. Ideas are not discrete packets that are transmitted whole cloth between minds. Information is multidimensional, multimodal, and influenced by a host of cognitive and social forces. Engaging with some of that complexity is the topic of the next chapter, where we will encounter the spread of continuous opinions.

4.8 Going Deeper

Contagion is a rich topic, and our simple models barely scratch the surface. Here are some brief descriptions of further directions to be explored when considering contagion.

The models we considered in this chapter assumed that interactions are either well mixed or, at minimum, random, with contagions spreading perfectly and instantaneously. In reality, interactions are often constrained by physical space, institutions, and social relationships. How interactions are structured can play an important role in the spread of a contagion. One of the most important developments in the study of social interactions has been the rise of network theory beginning in the late 1990s. This has given rise to both the subfield of network epidemiology (Danon et al., 2011) as well as the study of behavioral contagions on networks (Young, 2006).

Many contagion models, even those considering structured populations, assume that the ability or opportunity to transmit or contract a contagion is unaffected by one's infection status or by variation in informational states. In reality, these all interact. Disease state may influence behavior, as when sick individuals stay home and refrain from mixing with others. Intelligent organisms respond to new information, further influencing disease transmission. Contagion of a disease might interact with contagion of the *knowledge* of that disease. We may avoid overtly sick individuals. Awareness that an epidemic is spreading may further alter behaviors in ways that can either impede or hasten the spread of a disease (Funk et al., 2010). Social influences on behaviors like vaccine refusal (Mehta and Rosenberg, 2020) or mask wearing (Smaldino and Jones, 2021) might also spread like contagions. **Coupled contagion** models, which incorporate both behaviors and diseases, can be useful to study the simultaneous spread of multiple phenomena when each influences the spread of the others (Bedson et al., 2021).

Another aspect to consider, particularly in terms of informational or behavioral contagions, is that not all individuals are influenced in the same way by the same people. You may like a product, especially when your friends also adopt it (producing synergies like reduced costs and coordination benefits), but you may also dislike adopting the product if it becomes associated with a group that you do not identify with. In marketing experiments, subjects

disadopted or rated poorly products that were associated with outgroup individuals, even if those outgroups were not particularly disliked (Berger and Heath, 2007, 2008). If the probability of adoption is positively influenced by ingroup adoption but negatively influenced by outgroup adoption, a number of outcomes are possible. These include delayed adoption by one group, suppressed adoption (in which one group initially begins to adopt and then fully disadopts), or polarization, in which products are associated with different groups in different regions (Smaldino et al., 2017; Smaldino and Jones, 2021).

The models we considered in this chapter assume that contagions spread through direct exposure. Some pathogens are transmitted through other means, such as by sexual contact or by vectors like mosquitos. For beliefs and behaviors, adoption may require multiple exposures to the contagion from multiple "carriers." There is a large and growing literature on the variety of social learning strategies used across contexts for the adoption of adaptive behaviors (Laland, 2004; Kendal et al., 2018), and these are sometimes incorporated into contagion models under the banner of **complex contagion** (Centola and Macy, 2007; Eckles et al., 2019). The dynamics of such models can be quite different from those in the "simple" contagion models seen here. We will revisit contagion dynamics under different social learning strategies when we tackle networks in chapter 9. Other researchers have criticized the notion that the spread of cultural artifacts like beliefs and behaviors can be usefully modeled as contagions at all, and argue that incorporating more sophisticated cognitive mechanisms is needed to model social transmission; this is an exciting area of research at the boundaries of the cognitive and social sciences (e.g., Goldberg and Stein, 2018; Rabb et al., 2022; Falandays and Smaldino, 2022).

Finally, the models we considered in this chapter assumed that, for a behavioral contagion, exposure to one adopter was equivalent to exposure to any other. But there are good reasons why this won't always be the case. In an uncertain environment, conformity to the majority behavior can be beneficial (Boyd and Richerson, 1985). If the value of behaviors is opaque, copying successful or prestigious individuals can be a good strategy (Henrich and Gil-White, 2001). An extensive literature has considered the dynamics and evolutionary implications of various transmission biases, also called **social learning strategies** (Laland, 2004; Kendal et al., 2018), which may be important to include in models of behavioral contagion. Models have also indicated how empirical patterns in the rate of adoption of a particular behavior might be used to test for the psychological biases being used for adoption decisions, such as a bias for conformity (Smaldino et al., 2018a).

4.9 Exploration

1. **It's not my thing.** Consider the contagious adoption of consumer products. What factors influence the adoption of products or behaviors? How well do the SI or SIS models capture these contributing factors? If you have concerns, how might you extend or alter one of these models to address your concerns?

2. **Spontaneous adoption.** In most of the models considered in this chapter, we assumed that the only way to pick up an infection was to catch it from direct contact. This is true for most diseases, but for products or behaviors, an individual might also adopt spontaneously even in the absence of social influence. This indicates we might want to combine the spontaneous adoption and SI models.

   (a) Assume that once individuals are infected, they stay infected. However, susceptible individuals can become infected either spontaneously or by contacting

an infected individual. Write a recursion equation that describes the change in adoption rate over time, such that susceptible agents spontaneously adopt with a probability α and become infected upon contact with an infected individual with probability τ.

(b) Using either numerical simulation or an agent-based model, compute the time it takes the infection to reach saturation in a population as a function of effective transmissibility (τ, in the range [0.001, 1]) for several values of α: $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. Plot the relationship between τ and time to saturation on a log-log plot. What pattern do you observe?

3. **Don't move.** Modify the SIS ABM so that infected individuals change their behavior and move differently from susceptible individuals. They either stop, or slow down and increase their turning angle (so that they minimize their movement). Explore variations on this assumption. What happens to the contact rate when infected agents restrict their movement? How is the trajectory of the contagion affected by these changes? Explain why this violates the assumptions of the analytical SIS model.

4. **3-2-1 Contact.** Starting with the SIS ABM, write new code and calculate the rate of contact in the SIS agent-based model. That is, determine the probability that a given agent interacts with another agent, and thereby has an opportunity to acquire the contagion, per unit time. The actual number of contacts will vary over time, so you will probably want to get an average.

5. **Zombie invasion.** Modify the spatial SI ABM to study how to handle the *ZOMBIE APOCALYPSE*. In this model, infected agents become zombies. A healthy agent bitten by a zombie becomes a zombie. Zombies move more slowly than uninfected agents, and eventually die if they go too long without feasting on human flesh. You will need to add at least two new parameters: one controlling the movement speed of infected agents, the other controlling the death rate of infected agents. Holding the movement behavior of healthy agents constant, investigate the extent of the zombie apocalypse (how many agents become zombies) as a function of the zombie movement speed and death rate.

   (a) Write up a formal description of the model, including your new parameters.
   (b) Write code for your new model, including any necessary outcome measures.
   (c) Perform batch runs to answer the research questions, and plot the results. How do you interpret your results?

6. **Flatten the curve.** Use the SIR ABM to demonstrate how reducing disease transmissibility and factors influencing mobility both serve to "flatten the curve" of infection over time.

   (a) Write code to produce a reporter called `prop-infected` that reports the proportion of agents that are infected when called.
   (b) Use BehaviorSpace to run simulations varying `transmissibility`, `speed`, and `turning-angle`. Plot the resulting dynamics of infection for at least two values of each variable. What do you observe?

7. **Thank you for your compliance.** Use the SIR ABM to study how compliance with social distancing measures influences disease transmission. Start with a population of $N = 500$ and one initially infected individual. Let `turning-angle` $= 360$ and `recovery-rate` $= 0.01$. Define two classes of behavior. *Safe* agents wear face

masks and practice social distancing, which we will model so that they move with `speed` $= 0.1$ and, if infected, transmit the disease with probability $\tau = 0.01$. *Unsafe* agents do none of these things. They move with `speed` $= 1$ and, if infected, transmit the disease with probability $\tau = 0.1$.

(a) Write code for a model that implements these changes. Include a slider for a variable called `prop-safe` that determines the proportion of agents who are safe (the rest are unsafe).
(b) Write code to track the maximum infection rate over the course of an epidemic. Plot the trajectory of the population's infection rate over time for several (at least 3) values of `prop-safe`, and report the maximum infection rate for each run. How much does widespread compliance seem to matter? Note that here, the infection rate refers to the proportion of agents who are infected, not to how that proportion changes over time.
(c) Use BehaviorSpace to run simulations to record the maximum infection rate, varying `prop-safe` between 0 and 1 in increments no greater than 0.1. Do this for at least 3 population sizes: $N = \{50, 200, 500\}$. Run at least 5 simulations per condition. Plot your results. What do you conclude about the importance of compliance and its relationship to population density?

8. **Vaccinating agents.** Modify the SIS ABM so that a fixed proportion of the population is vaccinated. Start with a single infection and consider whether the infection spreads or dies out (so that no agents are infected). Run a small batch of runs to consider the proportion of runs for each parameter condition in which the contagion failed to spread, varying the transmissibility and recovery rate for at least two values each. Do your simulation results match the analytical predictions?

9. **Herd immunity.** Using the analytical SIR model, consider an outbreak of a new disease. Given that some of the infected individuals may recover before they can spread it to others, we want to calculate the conditions for when the outbreak will or will not spread.

(a) Assume that when an outbreak is new, $1 - I/N \approx 1$. Show that the infection will spread only if $\frac{\tau}{\gamma} > 1$. This quantity on the left side of the inequality is sometimes called the "basic reproduction number," R_0.
(b) Assume we have a vaccine that works perfectly, so that anyone who is vaccinated is protected against the outbreak. However, vaccinating everyone is infeasible for a variety of reasons. Luckily, we don't need to vaccinate everyone to prevent a disease from spreading, even if $R_0 > 1$. If enough people are vaccinated, the disease will be unable to spread effectively before it dies out, having only infected a few people. This is known as *herd immunity*. How many do we need to vaccinate? It depends on R_0. Let V be the proportion of individuals vaccinated, so the change in the number of infected individuals at any time is given by

$$\Delta I = \tau \left(1 - \frac{I}{N}\right)(1 - V)I - \gamma I.$$

Show that the minimum proportion of the population that must be vaccinated for herd immunity is

$$V^* = 1 - \frac{1}{R_0}.$$

10. **Vaccine efficacy.** Consider the analytic SIS model with vaccines. Assume a proportion V of the population has been vaccinated, and that both interactions and vaccinations are uniformly distributed at random in the population. In the chapter, we showed that under the assumption of perfectly effective vaccines, the vaccine threshold for herd immunity is $V^* = 1 - \frac{1}{R_0}$. Now, let's assume that the vaccine has an efficacy e, meaning that the vaccine prevents infection with probability e. The probability that an interaction between an infected individual and a vaccinated susceptible individual leads to a new infection is therefore $\tau(1-e)$. In this case, the difference equation governing the rate of new infections is:

$$\Delta I = \tau I \left(1 - \frac{I}{N}\right)((1-V) + V(1-e)) - \gamma I.$$

    (a) Prove that, when the outbreak is new, the threshold vaccination rate for herd immunity is $V^* = \frac{1}{e}\left(1 - \frac{1}{R_0}\right)$.
    (b) Plot the equation for the threshold vaccination rate as a function of the efficacy, e. Use $\tau = 0.1$ and $\gamma = 0.05$. Plot both e and V^* in the range $[0.5, 1]$. What does the resulting curve imply for the relationship between herd immunity and vaccine efficacy?

11. **Complex contagion.** Adoption of some behaviors or technologies may require exposure to multiple "infected" individuals. Some sociologists have referred to this phenomenon as a "complex contagion." This differs from "simple" contagions in which contact with a single infected individual is sufficient for spread.

    (a) Create a new NetLogo model on a 121 × 121 grid, using patches only. Patches can be either infected or susceptible. Initialize the model so that all patches are uninfected, except for a 3 × 3 square of patches in the center of the grid. Assume each patch is in contact with its nearest eight neighbors (Moore neighborhood). Establish a slider for a parameter called `threshold` that varies from 1 to 20. At each time step, each uninfected patch that has at least `threshold` infected neighbors becomes infected (when `threshold` > 1, the contagion is considered "complex"). Is there a maximum threshold for the spread of infection? If so, why?
    (b) Plot the proportion of patches that are infected as a function of time for several values of `threshold`. How long does it take for the infection to completely diffuse through the population in each case, if indeed it *does* diffuse? (Use BehaviorSpace to run multiple simulations as needed.)
    (c) Create a switch for a Boolean variable called `asynchronous?`. If `asynchronous?` = true, patches can become infected and immediately infect other agents on the same time step. If false, all patches first determine if they will become infected on a given time step, then all of those who will change status (from uninfected to infected) do so at the same time. Does this switch change the spatial dynamics of the contagion?
    (d) The maximum threshold for the spread of infection may depend in part on the network connectivity—that is, how many other agents each is connected to. Let's introduce a more strongly connected network. Create a new reporter called `big-neighbors`, which when called by a patch returns the set of patches corresponding to the focal patch's nearest 24 neighbors (its Moore neighborhood

with $r = 2$; you may want to use the NetLogo primitive `patch-set`). Create a switch for a Boolean variable called `bigger-neighborhood?`. If this is true, agents will use these bigger neighborhoods instead of just their closest 8 neighbors. Initialize the simulation so that the initially infected agents form a 5×5 square. How does the network connectivity (8 vs. 24 neighbors) influence the maximum value of `threshold` regarding whether or not the contagion will diffuse?