CHAPTER 2

How to Construct a Model

Chapter Goals:

- To describe the steps involved in developing a model
- To derive equations that describe the dynamics of a biological phenomenon

Chapter Concepts:

- Discrete-time model
- Continuous-time model
- Recursion equations
- Differential equations
- Life-cycle diagrams
- Flow diagrams
- Mass action

2.1 Introduction

If you have seen mathematical models but never constructed one, it may seem like an overwhelming task. Where do you start? What is the goal? How do you know whether the model makes sense? This chapter outlines the typical process of modeling and gives helpful hints and suggestions to break down the overwhelming task into manageable bits. The most important piece of advice is to *start*. Start thinking about problems that puzzle you. Grab a piece of paper and start drawing a flow diagram illustrating the various processes at work. The biggest hurdle preventing most biologists from modeling is the paralysis one feels in the face of mathematics; most of the technical problems that pop up along the way can be surmounted or sidestepped (at the very least by simulation). You will certainly make mistakes (we all do), but there are telltale signs of mistakes, and they can be corrected. Over time, you will learn more tools and techniques that will allow you to avoid pitfalls and to get further with the problems that interest you. Your intuition will develop to help you "see" when something is wrong with your model and to help you interpret your results.

Models can describe any biological phenomenon. In the core of this book, we focus on *dynamical* models, which describe how a system changes over time. Dynamical models are very common in biology as they provide insight into how various forces act to change a cell, an organism, a population, or an assemblage of species. Within dynamical models, two broad classes are distinguished: deterministic and stochastic. "Deterministic" is shorthand for the assumption that the future is entirely predicted (determined) by the model. "Stochastic" is shorthand for the assumption that random (stochastic) events affect the biological system, in which case a model can only predict the probability of various outcomes in the future. In the remainder of this chapter, as well as in Chapters 3–12, we focus on deterministic models. The steps for constructing stochastic models are similar, but we postpone further consideration of stochastic models until Chapters 13–15.

Box 2.1 describes, in seven steps, how to construct a dynamical model. This is like describing how to ride a bike in a series of steps; obviously we can only give an idea about how the process works. Mastering the steps requires practice, and the remainder of this chapter contains a series of seven sections, each corresponding to one of the seven steps in Box 2.1.

Box 2.1: Seven Steps to Modeling a Biological Problem

Step 1: Formulate the question
What do you want to know?
Describe the model in the form of a question.
Boil the question down!
Start with the simplest, biologically reasonable description of the problem.

Step 2: Determine the basic ingredients
Define the variables in the model.
Describe any constraints on the variables.
Describe any interactions between variables.
Decide whether you will treat time as discrete or continuous.
Choose a time scale (i.e., decide what a time step equals in discrete time and specify whether rates will be measured per second, minute, day, year, generation, etc.).
Define the parameters in the model.
Describe any constraints on the parameters.

Step 3: Qualitatively describe the biological system
Draw a life-cycle diagram (see Figure 2.2) for discrete-time models involving multiple events per time unit.
Draw a flow diagram to describe changes to the variables over time.
For models with many possible events, construct a table listing the outcome of every event.

Step 4: Quantitatively describe the biological system
Using the diagrams and tables as a guide, write down the equations.
Perform checks. Are the constraints on the variables still met as time passes? Make sure that the units of the right-hand side equal those on the left-hand side.
Think about whether results from the model can address the question.

Step 5: Analyze the equations
Start by using the equations to simulate and graph the changes to the system over time.
Choose and perform appropriate analyses.
Make sure that the analyses can address the problem.

Step 6: Checks and balances
Check the results against data or any known special cases.
Determine how general the results are.
Consider alternatives to the simplest model.
Extend or simplify the model, as appropriate, and repeat steps 2–5.

Step 7: Relate the results back to the question
Do the results answer the biological question?
Are the results counterintuitive? Why?
Interpret the results verbally, and describe conceptually any new insights into the biological process.
Describe potential experiments.

2.2 Formulate the Question

The first step, coming up with a question, can be more difficult than it sounds. In most biology classes, students are told what the questions are and what answers have been found. Rarely are students asked to formulate scientific questions for themselves. This is very unfortunate because, in any scientific enterprise (modeling or otherwise), the process begins with a question. One hint is to keep an eye out for things that do not make sense or that seem to conflict—there very well might be an interesting and nonintuitive resolution. For now, start simple and don't worry about how profound your question is. Look around you, find a living object, and think up one question about how it might change over time. We did this and came up with the following three questions, which we will use in this chapter to illustrate model construction. (i) How does the number of branches of a tree change over time? (ii) How does a cat change the number of mice in a yard? (iii) How does the number of people with the flu change over the flu season?

The above three questions are "toy" examples that will make it easier to show the steps of modeling. Nevertheless, these simple examples also embody many of the key elements that come together in various combinations when constructing more complicated and realistic models. As we will see, the tree branching model is a special case of a model describing population growth. The mouse model incorporates an important component of immigration that is commonly used in ecology. For example, Blower et al. (2000) used a similar model of immigration to describe individuals moving into the gay male community of San Francisco. Finally, the flu model highlights some important concepts related to interactions among variables. For example, the way that we will model flu transmission is fundamentally similar to the way that Phillips (1996) modeled the infection of cells by HIV. Thus, these toy models provide an excellent background for tackling more complex models.

2.3 Determine the Basic Ingredients

Once you have a question in mind, proceed to Step 2 in Box 2.1. First, think about what entities might change over time; these entities are the *variables* in your model. The number of variables will depend on the question of interest. In our toy examples, we might choose to follow (i) the number of branches on a tree, (ii) the number of mice in a yard, and (iii) the number of people with the flu and the number without the flu. In choosing variables to track, we must always simplify reality. For example, in keeping track of the number of branches, we lose information about their size and age. As a general principle, start simple, adding more variables only when the model fails to address the question.

Next, we assign a letter to represent each variable—it is easier to write "x" than "the number of branches on a tree." The letters n, p, x, and y are commonly used to represent variables, but the choice is arbitrary. A good idea is to choose letters that help you remember what the variable represents, e.g., "n" for number or "p" for proportion. If a model contains multiple variables

that are similar in nature, placing subscripts on the variables can help to emphasize their similarity, e.g., n_1 and n_2 for the numbers of two different species. For our models, we will use (i) $n(t)$ for the number of branches on a tree, (ii) $n(t)$ for the number of mice in a yard, and (iii) $n(t)$ for the number of people with the flu and $s(t)$ for the number of susceptible people.

To remind ourselves that a variable, say n, varies over time, we can write it as $n(t)$ where t represents time and there is no space between the n and the (t). The parentheses tell us that our variable is a function of something else (time), and we read $n(t)$ as "n at time t." This notation helps to avoid math errors. For example, without this notation, we might forget that n takes on different values at different times and mistakenly treat it as a constant. Be aware, however, that not all authors use the same notation; they might write n_t instead or might simply state that n is a variable and not write it explicitly as a function of time. The important thing is to be consistent and to remember that, if we write a variable as $n(t)$, we mean "n at time t" not "n times t."

Another way to avoid math errors is to keep a list (at least a mental list) of any constraints that must remain true about the variables. For example, the number of branches on a tree should never become negative. The number of people with the flu and the number without the flu should never be negative and should sum up to the total population size. If a variable describes a frequency, a probability, or a fraction of a whole (e.g., the fraction of the total population with the flu), it should always lie between zero and one ($0 \leq p(t) \leq 1$). Ensuring that your equations and results obey the list of constraints is a good way to check that no errors have crept in.

Once you have a preliminary list of variables, the next step is to choose a type of dynamical model to describe changes in these variables. There are two main types of dynamical models, *discrete time* and *continuous time*, depending on whether time is represented in discrete steps or along a continuous axis. Discrete-time models describe how the variables change from one time unit (e.g., day, year, or generation) to the next. Continuous-time models track the variables over any period of time. Both discrete-time and continuous-time models are idealizations of reality, and they make somewhat different assumptions.

A *discrete-time model* tracks changes to variables in discrete time steps.

A *continuous-time model* allows variables to change at any point in time (i.e., time is treated as continuous).

Discrete-time models assume that changes cannot compound within a time unit. For example, in a discrete-time model for the number of branches on a tree, branches that arise during a time unit cannot give rise to new branches within the same time unit. As long as the time unit is short enough (e.g., a day), this assumption is often reasonable. If the time unit were long (e.g., a year), however, then some new branches might very well branch again within the year. These branching events would not be counted in a discrete-time model if the new branches were not present at the beginning of the year.

Continuous-time models assume that variables can change at any point in time, with increments or decrements occurring even within tiny intervals of time. As a consequence, it is possible for a change to occur in one small interval of time followed by the same type of change in the next small interval of time. But this may not be biologically realistic. For example, a continuous-time model might allow a newly formed branch to immediately produce its own new branch. In reality, the new branch must undergo enough cell divisions to

produce a new bud, which takes time. If the rate of branching is small, then this won't be much of a problem because the average time between branching events will be large. But if the rate of branching is high, then a continuous-time model will generate incorrect predictions unless it takes into account the time lag between the formation of a branch and the formation of buds on this new branch.

Because discrete- and continuous-time models treat the timing of events in different ways, they display different temporal dynamics. In discrete-time models the variables "jump" from one value to another from one time unit to the next, and the size of these jumps can be small or large depending upon the parameters of the model. In continuous-time models, on the other hand, the variables change smoothly over time. This means that, as a variable goes from one value to another, it passes through all intervening values along the way (Figure 2.1).

In either case, we must also choose a time scale over which changes to the variables are measured. We use a "day" as the basic unit of time for the toy models considered in this chapter. Specifically, we assume that each time step in discrete-time models reflects the passage of 24 hours and that all processes in continuous-time models occur at a rate measured per day.

Just as time can be modeled discretely or continuously, so too can the variables themselves. For example, the number of branches on a tree, the number of mice in a yard, and the number of people with the flu are all discrete, integer-valued quantities (i.e., they are integers such as 0,1,2, . . . , etc.). On the other hand, an organism's metabolic rate or an organism's weight can take on any of a continuum of possible values. Regardless of the true nature of the variables, the majority of models in ecology and evolution treat variables as being continuous, an approach that we follow throughout most of the book (except in Chapters 13–15, which incorporate random events and explicitly track the numbers of each type). There are three main justifications for treating variables as continuous. First, for many questions, the variables of interest take on large enough values that treating them as continuous will introduce very little error in the results (e.g., the number of HIV particles in the blood). Second, a reinterpretation of the variable (e.g., as the total biomass of mice rather than the number of mice) can sometimes justify the use of a continuous variable. Third,

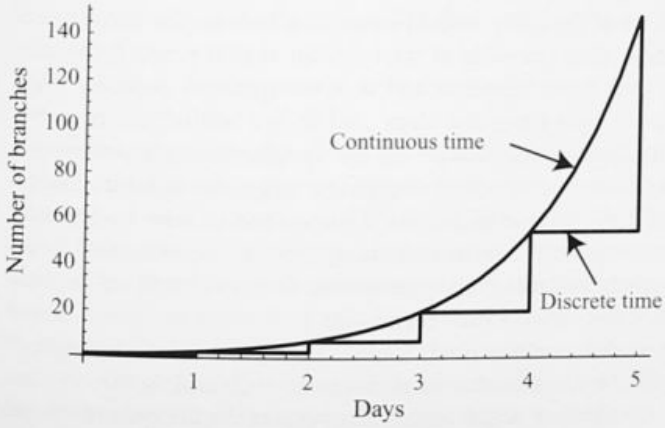

Figure 2.1: Tree branching. A plot of the number of branches on a tree over time using a discrete-time model and a continuous-time model.

it is typically easier mathematically to treat variables as being continuous rather than discrete. Remember, all models are abstractions of biological reality, and treating variables as continuous is often a reasonable abstraction.

A *recursion equation* describes the value of a variable in the next time step.

In discrete-time models, we track changes to a variable using a *recursion* equation, which describes the value of a variable (say, n) in the next time unit as a function of the variable in the current time unit:

$$n(t + 1) = \text{"some function of } n(t)\text{."} \tag{2.1a}$$

Such equations are called recursions, because one can apply them recursively to find out how the variable changes across a number of time units (from t to $t + 1$, then from $t + 1$ to $t + 2$, etc.). An equivalent way to track a variable is to use a *difference* equation. A difference equation specifies how much a variable changes from one time unit to the next, and it is just the difference between the recursion equation for $n(t + 1)$ and the current value of the variable $n(t)$:

$$\Delta n = n(t + 1) - n(t) = \text{"some function of } n(t)\text{,"} \tag{2.1b}$$

where the capital Greek letter Δ ("Delta," see Table 2.1) denotes "change," and we read Δn as "the change in the variable n." Recursion equations are more commonly used to describe the value of a variable in discrete-time models, but we will occasionally use difference equations when we want to understand how much a variable changes across a time step.

In continuous-time models, equations specify the rate of change of the variables over time:

$$\frac{\mathrm{d}(n(t))}{\mathrm{d}t} = \text{"some function of } n(t)\text{."} \tag{2.1c}$$

A *differential equation* describes the rate at which a variable changes over time.

Such equations are called *differential* equations. Differential equations are distinct from the more familiar derivatives taught in introductory calculus courses (see Box 2.2). You can think of a differential equation as a description of the ebb and flow in a variable over time. To get a better feel for a differential equation, imagine plotting the value of the variable $n(t)$ as a function of time (see Chapter 4). The slope of the curve would be $\mathrm{d}(n(t))/\mathrm{d}t$ because the derivative of a function at a point gives the slope of the function at that point. If the variable is increasing over time, the slope and thus $\mathrm{d}(n(t))/\mathrm{d}t$ are positive. If the variable is decreasing over time, the slope and thus $\mathrm{d}(n(t))/\mathrm{d}t$ are negative. When the magnitude of $\mathrm{d}(n(t))/\mathrm{d}t$ is small, the variable changes slowly over time, whereas when the magnitude of $\mathrm{d}(n(t))/\mathrm{d}t$ is large, the variable changes rapidly. As we will see, this mental picture is the reverse of how we typically construct models. We usually start by describing how various biological forces change the value of the variable (i.e., contribute to $\mathrm{d}(n(t))/\mathrm{d}t$), and we then try to infer the value of the variable itself (i.e., $n(t)$).

Which type of model should be used? Sometimes, there is a natural choice. If you want to model the number of annual plants on an island, a discrete-time model using a year as the time unit is appropriate because the life cycle of annual

TABLE 2.1
Greek letters. Here, we list the Greek letters commonly encountered in biological models (with alternative characters in parentheses).

Lower case	Upper case	Name
α	A	alpha
β	B	beta
χ	X	chi
δ	Δ	delta
ε	E	epsilon
$\phi(\varphi)$	Φ	phi
γ	Γ	gamma
η	H	eta
ι	I	iota
κ	K	kappa
λ	Λ	lambda
μ	M	mu
ν	N	nu
o	O	omicron
$\pi(\varpi)$	Π	pi
$\theta(\vartheta)$	Θ	theta
ρ	P	rho
$\sigma(\varsigma)$	Σ	sigma
τ	T	tau
υ	Y	upsilon
ω	Ω	omega
ξ	Ξ	xi
ψ	Ψ	psi
ζ	Z	zeta

plants is itself discrete; that is, the seeds produced during one year will not germinate until the following year. By contrast, if you want to model your blood sugar levels after a meal, a continuous-time model would be more natural because there are no clear demarcations in time. Conceptually, it is sometimes easier to think in terms of discrete-time models where changes describe what happens over an interval of time rather than continuous-time models where changes are described by instantaneous rates. Mathematically, however, continuous-time models can be easier to analyze because one can utilize the various rules of calculus summarized in Appendix 2 (see Chapter 6). As we discuss later (Box 2.6

Box 2.2: Derivatives and Differential Equations

Calculus is the mathematical study of rates of change. The most important concepts and rules of calculus are summarized in Appendix 2, including formulas for differentiating and integrating a variety of functions. For example, the derivative of the polynomial $y = ax^2 + bx + c$ with respect to x is $dy/dx = 2ax + b$. Here, the rate of change of the *dependent* variable y is a function only of the *independent* variable x. In many biological problems, however, the rate of change of the dependent variable is a function of the dependent variable itself, e.g., $dy/dx = \alpha y + \beta$. Notice that the variable on the right-hand side is y not x. An equation relating the derivative of a variable to a function of the variable itself is called a *differential equation*. Equations (2.8)–(2.10) are differential equations. For example, in equation (2.8), the derivative of the dependent variable describing the number of tree branches, $n(t)$, with respect to the independent variable (time t) is a function of $n(t)$, not t. Differential equations naturally arise in continuous-time biological models because we often expect the rate of change of a variable to be a function of its current value. For example, large trees can have more new branches, a cat can eat more mice if there are more mice available, and more people can catch the flu if there are more susceptible people within the population.

A derivative or differential equation describes how a variable changes. But what we usually want to know is the *value* of the dependent variable (e.g., $n(t)$) as a function of the independent variable (e.g., t). In a typical calculus course, we are taught how to solve for y by taking the antiderivative or integral of both sides. In other words, we could solve the equation $dy/dx = 2ax + b$ for $y(x)$ by integrating both sides with respect to x to obtain its *solution*, $y = ax^2 + bx + c$ (see Appendix 2), which gives us the value of y for any value of x. A common error that students make when they first encounter differential equations is to integrate the left-hand side of an equation like $dn(t)/dt = bn(t)$ with respect to t but the right-hand side with respect to $n(t)$. This would give $n(t) = bn(t)^2/2$. To see that this is incorrect, take the derivative of both sides with respect to t (see Appendix 2). This would give $dn(t)/dt = bn(t)dn(t)/dt$, which incorrectly has $dn(t)/dt$ on the right-hand side. The error in this procedure crept in when we took the antiderivative of the left-hand side with respect to t, but the antiderivative of the right-hand side with respect to a different variable, $n(t)$. To solve for $n(t)$ we would have to take the antiderivative of *both sides* with respect to t, i.e.,

$$\int \frac{dn(t)}{dt}\, dt = \int bn(t)\, dt. \tag{2.2.1}$$

The left-hand integral is $n(t)$, as before, but we cannot evaluate the right-hand integral because doing so requires $n(t)$, which is what we are trying to find. In Chapter 6, we will see how to obtain solutions to certain types of differential equations, like the ones presented in this chapter. For now, it is enough to recognize the distinction between derivatives and differential equations and to remember that care must be taken when integrating differential equations.

Before leaving the subject, it is worth mentioning that the term "differential equation" encompasses several types of equations, all of which arise in biology. Differential equations can be written as functions of more than one dependent variable. For example, in our flu model, the

(continued)

Box 2.2 *(continued)*

differential equation (2.10a) for the number of people with the flu, $dn(t)/dt$, will depend on both the number of people with the flu, $n(t)$, and the number of susceptible individuals in the population, $s(t)$. Differential equations can also be written as functions of *both* the dependent variable $n(t)$ and the independent variable t. Such differential equations arise whenever we expect a variable to change as a function both of its current value and of time. For example, in a seasonal environment, the budding rate of a tree should depend on the time of year as well as on the number of branches on a tree. We can model this by treating b as some function of time, $b(t)$, rather than a constant. In addition, differential equations might depend on the past state of a variable as well as (or instead of) its current state. For example, in the tree branching example, the production of new branches at time t might depend on the total number of branches τ days ago, or $n(t - \tau)$, as these branches are now large enough to branch again. Revising equation (2.8) gives $dn(t)/dt = bn(t - \tau)$. Such equations, known as "delay differential equations," arise naturally when describing biological processes involving time lags.

All of the above examples have only one independent variable (time). These fall into the category known as "ordinary differential equations" (ODE). Many biological problems involve more than one independent variable (e.g., space as well as time), and such differential equations are known as "partial differential equations" (PDE).

and Chapter 4), discrete-time and continuous-time models can sometimes exhibit similar behavior over time, and it is possible to predict when they should behave similarly. Thus, in many cases, one is free to choose between the two.

The next step is to describe the *parameters* of the model; these are the various quantities that influence the dynamics of the model, but that remain fixed over time as the variables change. As with variables, each parameter is given its own symbol, which you are free to choose. Commonly used symbols for parameters are italicized roman letters (e.g., a, b, c, d, m, and r) and lower-case greek letters (e.g., α, β, Table 2.1).

A chief difference between discrete-time and continuous-time models is that parameters representing events per unit time are described as the *number of events* (or fraction of the population undergoing the event) per time step in discrete-time models but as the instantaneous *rate of events* per unit time in continuous time. In contrast, parameters that do not represent events per unit time (e.g., the probability that an event is one type or another) retain the same definition in the two types of models. We will discuss the difference in parameter units between discrete- and continuous-time models at greater length in Box 2.6, once we have described how their dynamical equations are derived.

Potential parameters for our discrete-time models include (i) the number of new branches that bud off each old branch per day, b; (ii) the fraction of mice in the yard eaten by the cat per day, d, and the number of mice born per mouse per day, b; (iii) the fraction of healthy people that are exposed to a flu carrier per day, c, and the probability of transmission of the flu between a healthy person and a flu carrier upon exposure, a. The analogous parameters in a continuous-time

model would be (i) the rate of budding for each old branch, b; (ii) the rate of consumption of mice, d, and the rate of births per mouse, b; and (iii) the rate of contact between a flu carrier and a susceptible person, c, and the probability of transmission of the flu between a carrier and a healthy person per contact, a. These parameters represent events per unit time and so have slightly different definitions for the discrete-time and continuous-time models except a, which always represents the probability of contracting the flu per contact.

As with variables, one should also keep track of any constraints imposed on each parameter. For example, can a parameter be negative? Does a parameter represent a fraction, proportion, or probability, in which case it must fall between zero and one? These constraints might well depend on the type of model. For example, the parameter d in the cat-mouse model is restricted to lie between zero and one in discrete-time models (because it represents the *fraction* of mice eaten by the cat), whereas the analogous parameter d in the continuous-time model can have any positive value (because it represents the *rate* of consumption of mice per unit time). This is another common difference in the parameters between discrete- and continuous-time models (described more fully in Box 2.6).

In addition to the absolute constraints on each parameter, it is worth keeping track of the range of parameter values that are biologically reasonable. For example, it is reasonable to assume that the number of new branches that bud off each old branch per day is small for most trees ($b << 1$). Similarly, the number of mice born per mouse per day (b) will be much less than one ($b << 1$). We write $b << 1$ to imply that b is much smaller than one. How much smaller depends on the context, but typically this statement implies that b is 0.1 or less. Having a list of constraints and reasonable ranges for parameters can help in two important ways. First, reasonable parameter values must be chosen to carry out realistic simulations and to plot relevant graphs. Second, results from a model often depend on the values of the parameters, e.g., whether a parameter is positive or negative, large or small, so that making accurate predictions from a model depends on choosing appropriate parameter values.

Before proceeding to the next step, it is a good idea to construct a table of all the variables and parameters in your model, as well as any constraints on these terms. You can later revisit this table to ensure that it includes the variables and parameters needed to capture the essence of the biological process and to address the question of interest. It is very common that the first version of a model includes too many variables and parameters, causing the model to be unnecessarily complex, or too few variables and parameters, causing a model to behave in unintended ways (e.g., populations grow to infinite size, or nobody ever recovers from the flu). If a model displays unintended behavior, then think about whether the biological system being modeled includes other processes that should also be incorporated into the model (e.g., competition, recovery).

2.4 Qualitatively Describe the Biological System

Before writing equations down, it is a very good idea to organize your model conceptually with the aid of a diagram or table. Diagrams and tables make it easier to see whether the necessary variables and parameters are included and

make it easier to write down dynamical equations (recursion equations or differential equations). We describe three organizational techniques: a life-cycle diagram, a flow diagram, and a table of events.

2.4.1 Life-Cycle Diagrams

A graphical technique, which we call a *life-cycle diagram*, keeps track of the various events occurring during a single time step, along with their order of occurrence. Such diagrams are useful only for discrete-time models, where there is a discrete time period during which various events can occur. As a simple example, consider the tree branching model. Each time step represents a single day, and only one type of event can happen during any given day: the growth of more branches. As result, the life-cycle diagram is extremely simple (Figure 2.2a).

> A *life-cycle diagram* illustrates the order of events that occur within each time step (for discrete-time models).

The tree branching model is so simple that a life-cycle diagram is not really required to organize things. Life-cycle diagrams become indispensable when multiple events occur during a single time step. Consider the model of mice being eaten by a cat. Now there are three events that occur each day: mice give birth, mice move in from neighboring areas, and the cat eats mice. In a discrete-time model, one must choose an order for these events, as well as a point in time when the population is censused (e.g., when we count the number of mice, $n(t)$). For example, Figure 2.2b illustrates the case where events occur in the following order: a census, followed by predation by the cat, mouse births, mouse migration, and finally the next census. These events cause changes to the number of mice, which we describe as $n(t)$ at the census point, $n'(t)$ after predation, $n''(t)$ after births, and $n'''(t)$ after migration. Because migration is assumed to occur last in the daily life cycle, the number of mice at the next census, $n(t+1)$, will equal $n'''(t)$. Alternatively, we might instead assume that births happen first, then migration, and then predation, yielding the life cycle in Figure 2.3. As we shall see, the order of events in a life cycle can affect the results of a model, sometimes substantially.

Finally, consider constructing a life-cycle diagram for the model of flu transmission. The time step is again one day, and as with the tree branching model, there is only a single event that can happen during each day: transmission of the flu. There is an additional wrinkle with this model, however, in that there are now two variables that we are tracking (healthy individuals and people with the flu). As a result, we could construct a life-cycle diagram for each of the variables (Figure 2.2c). But because there is only one event per cycle, these life-cycle diagrams are again not very useful (as was the case with the tree branching model).

2.4.2 Flow Diagrams

A second method for organizing a model, which is often more useful for models containing multiple variables, is a *flow diagram*. A flow diagram illustrates the interconnections among the variables and provides a schematic picture of how each variable affects its own dynamics as well as the dynamics of the other variables. In a typical flow diagram, each circle represents one variable within the model. Returning arrows that exit and come back to the same circle represent a

> A *flow diagram* illustrates how each variable affects its own dynamics and those of other variables.

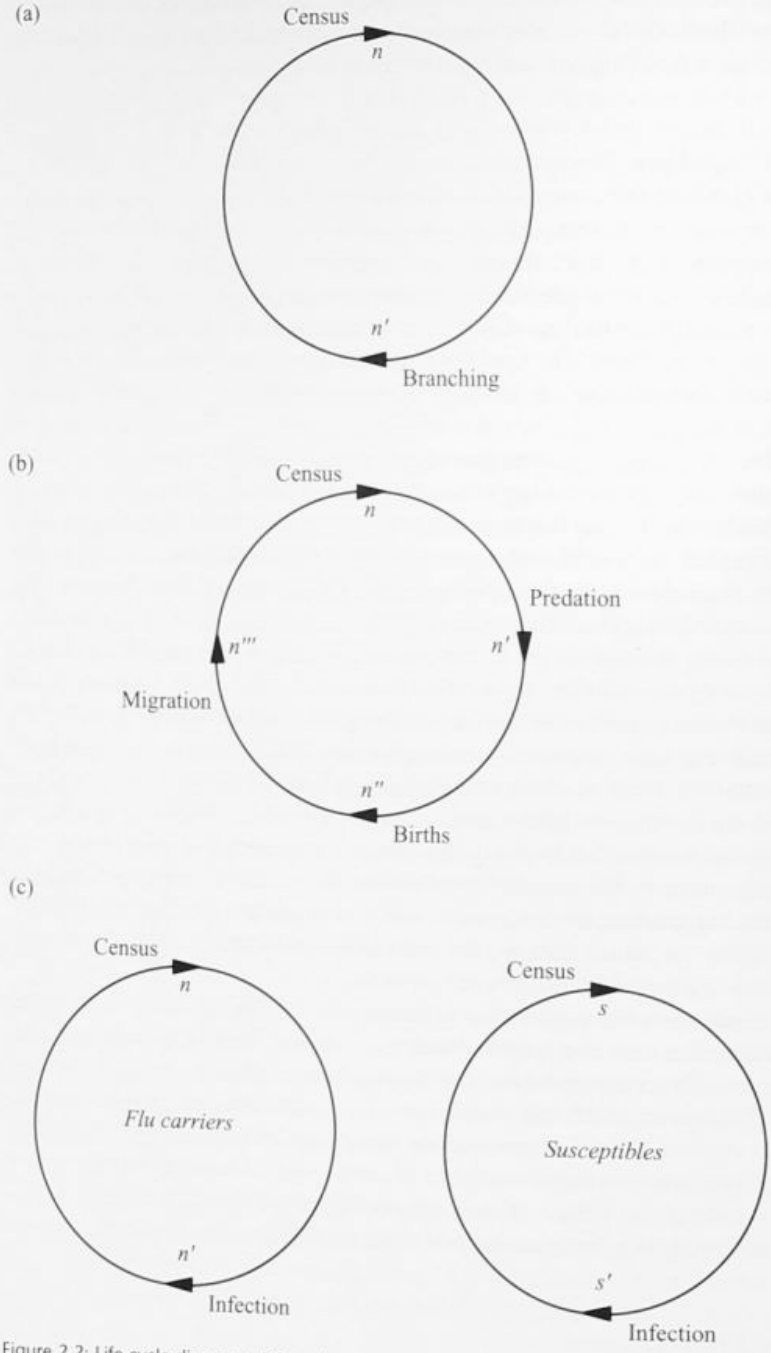

Figure 2.2: Life cycle diagrams. Life cycle diagrams for the three toy models explored in this chapter: (a) the number of tree branches, (b) the number of mice, (c) the number of people with and without the flu.

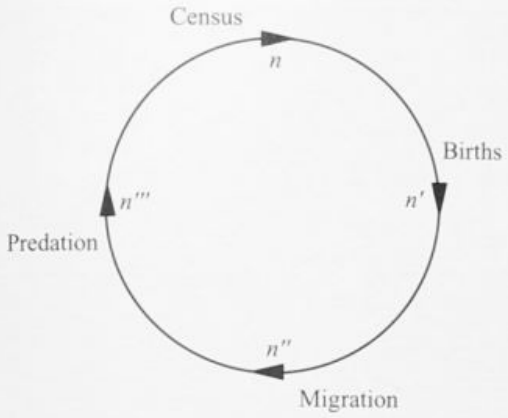

Figure 2.3: Alternate life cycle diagram. A life cycle diagram for the number of mice with a different ordering of events than occurs in Figure 2.2b.

variable that can generate more of itself. As a very simple example, a flow diagram for the tree branching model has a single returning arrow, representing the budding of a tree branch into new tree branches (Figure 2.4a).

Arrows leading into a circle represent the different ways in which a variable can go up over time, while arrows exiting a circle represent the different ways in which the variable can go down over time. A flow diagram for the mouse population has one of each of these, representing immigration and deaths, respectively, along with a returning arrow representing births (Figure 2.4b). Unfortunately, flow diagrams for discrete-time models are cumbersome when multiple events can occur within a single time step (as in this mouse model) because it is difficult to depict the ordering of the events. To be consistent with the ordering of events in Figure 2.2b, we need to consider flow across the death arrow first, then the birth arrow, and finally the migration arrow, updating the variable after each event (e.g., from $n(t)$ to $n'(t)$ after the first event).

Flow diagrams become really useful when there are multiple variables. In our flu model, we are tracking the number of susceptible and infected individuals. There must be an interaction (contact) between an infected and a susceptible person for transmission to occur, and this can be represented on a flow diagram in a variety of ways. In Box 2.3, we describe a convention for building flow diagrams, which is designed to facilitate the process of converting flow diagrams into mathematical equations. According to this convention, an interaction between two variables is represented by the merging of arrows emanating from two circles (Figure 2.4c). Different people use different conventions, but sticking to the same convention is important to avoid mistakes along the way.

Flow diagrams are constructed in the same way for continuous-time and discrete-time models (Box 2.3). For continuous-time models, however, the arrows represent events occurring continuously over time at certain rates, and we do not have to worry about the order in which events take place (e.g., we do not have to update the variables from $n(t)$ to $n'(t)$ after the first event).

On a flow diagram, it is very useful to specify (mathematically) the flow represented by each arrow directly on the diagram, including how this flow depends on the variable(s) themselves (step 8 of Box 2.3). This convention allows us to distinguish between a constant *number* exiting a circle (e.g., D, if

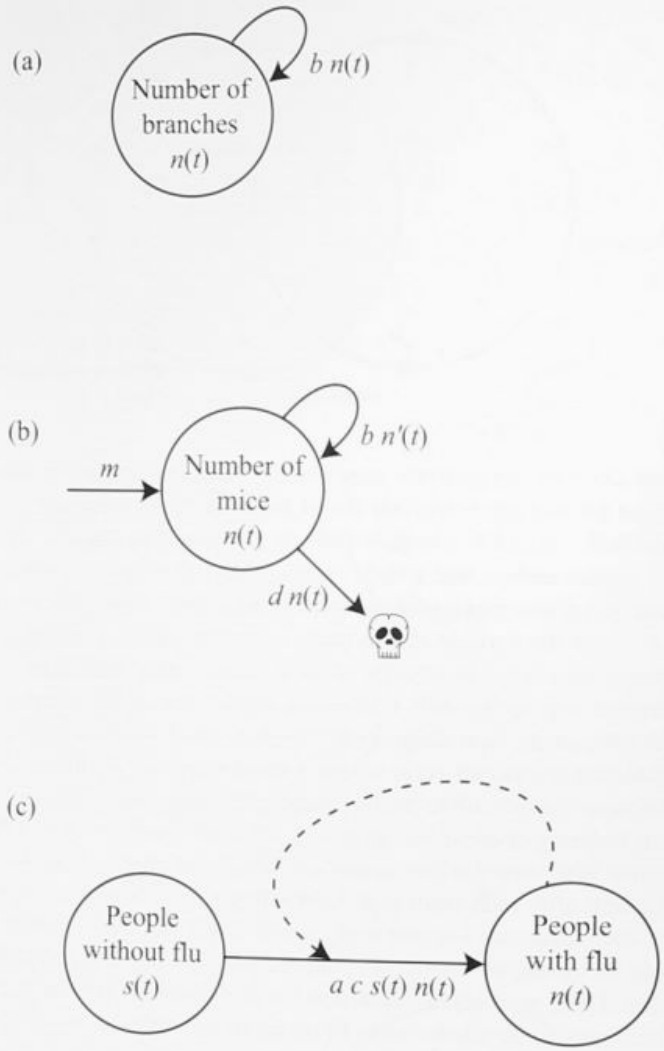

Figure 2.4: Flow diagrams. Flow diagrams for the examples explored in this chapter: (a) the number of tree branches, (b) the number of mice, (c) the number of people with and without the flu.

your cat eats a constant number D of mice per day) and a constant *fraction* exiting a circle (e.g., $d\,n(t)$, if your cat eats a constant fraction d of available mice per day). Specifying the flows is the hardest step in constructing a flow diagram, because it forces us to be very specific about the biological processes that we are modeling.

A *mass-action* interaction assumes that the rate of interaction between two variables is proportional to the values of each.

Specifying the flow for each arrow also forces us to describe if and how the variables interact. Interactions come in many different forms depending on the type and complexity of the interaction. For the flu model in Figure 2.4c, we have used the simplest form of interaction, known as a *mass-action* interaction. Mass action refers to the assumption that two types interact at a rate that is proportional to the number (or density) of the first type times the number (or density) of the second type, just as if the two types were moving about and bumping into each other at random. In the flu model, we assume that individuals with

Box 2.3: Drawing Flow Diagrams

(1) Draw a separate circle to represent each variable in your model.

(2) Use a solid arrow to indicate when a process removes an amount of the variable (arrow exits circle) or contributes an amount to the variable (arrow enters circle).

(3) Use an arrow that comes from nowhere but that enters a circle to indicate when there is an external source for one of the variables (e.g., mice from another field).

(4) Use an arrow that comes from a circle but goes to nowhere (or to a skull) to indicate when a variable exits the system (e.g., by death or emigration).

(5) Use an arrow that starts at one circle and goes to another circle to indicate when one type can become converted into another type (e.g., a susceptible individual catches the flu).

(6) Use a dashed arrow to indicate when a variable influences the flow into another circle but does not represent a decline in the variable from which the arrow begins (e.g., in Figure 2.4c, a carrier of the flu does not lose the flu by passing it on).

(7) Include an arrow that exits and returns to the same circle ("a returning arrow") whenever a variable can generate more of itself (e.g., by new births). A returning arrow can represent changes due to births only, or can describe the net change following both births and deaths.

(8) Write down the total flow along each arrow, specifying how this flow depends on the variable from which the arrow comes and on any interacting variables. If the flow across an arrow represents a conversion from one type to another (e.g., from number of prey to number of predators), there may be a conversion factor (e.g., one prey might represent only $\varepsilon = 1/100$ of the resources needed to produce one predator). Write this factor as "times ε" at the end of the arrow.

(9) For discrete-time models, decide on an ordering for the various events that occur during each time step and put a prime after the variable to indicate its state "after the first event," a double prime after the variable to indicate its state "after the second event," etc.

(10) Check to make certain that your variables are linked together in the way that you want.

(11) Check to make sure that each arrow has a flow rate written by it.

(12) Check to see if there are any variables that are completely unconnected to the rest of the diagram.

(13) Check to see if there are any parameters in your model that do not appear on the flow diagram.

(14) For a discrete-time model, check that there is never more than 100% of a variable leaving a circle.

the flu interact with susceptible individuals at a rate proportional to $n(t)\ s(t)$. Specifically, an infected individual has a probability c of contacting any given one of the $s(t)$ susceptible individuals per day, giving a total number of contacts per day of $c\ s(t)$ per infected individual. With $n(t)$ infected individuals in the population, we expect a total of $c\ n(t)\ s(t)$ contacts per day across the whole population. The probability that any one contact results in the transmission of the influenza virus is a, and therefore we expect a total of $a\ c\ n(t)\ s(t)$ new cases of flu per day. This is written below the arrow in Figure 2.4c.

The use of mass action in the flu model makes qualitative sense; if there are no people with the flu ($n(t) = 0$) or if there are no individuals susceptible to the flu ($s(t) = 0$), then there will be no new cases of the flu, because $a\ c\ n(t)\ s(t)$ is zero. There may, however, be times when you are not yet ready to specify the details of such interactions. If so, you can write the flow in general terms as $g(s(t),n(t))$, indicating that the flow rate is some function $g(\ )$ that depends on the variables $s(t)$ and $n(t)$. This lets you put off the decision of how the interactions depend on the variables until later.

Once a flow diagram has been labeled, steps (11)–(15) of Box 2.3 describe various checks to ensure that the flow diagram accurately reflects your model. If any problems arise, return to section 2.3 and revise your list of variables, parameters, and constraints, adding and subtracting as necessary. It is critical to repeat this process until you are happy that the flow diagram captures the essence of the biological process that you wish to model.

2.4.3 Tables of Events

For discrete-time models involving multiple events within a time step and multiple variables, neither life-cycle diagrams nor flow diagrams easily encapsulate all of the relevant information. In such cases, a *table of events* can be a useful organizational tool. We illustrate the construction of such a table using the flu model (Table 2.2). This model has only one event per time step (infection), so that a table is not really necessary. In more complex models, however, such tables of events are invaluable. For example, Table 8.1 organizes the events for an evolutionary model involving two genes, where we must consider

TABLE 2.2
Interaction table for the flu model. The first column lists every possible pair of individuals that could come into contact. The second column lists the number of each type of contact. The remaining columns list the change in number of infected and susceptible individuals resulting from such a contact. (Alternatively, we could have listed the number of each type after the contact, but here it is easier to list the changes.)

Interaction	Number of contacts	Result of contact	
		Infected	Susceptible
Infected × infected	$c\ n(t)\ n(t)$	No change	No change
Infected × susceptible	$c\ n(t)\ s(t)$	$+a$	$-a$
Susceptible × susceptible	$c\ s(t)\ s(t)$	No change	No change

several events (fertilization, selection, meiosis, and recombination) acting on every possible type within the population.

2.4.4 Rules of Thumb for Qualitatively Describing a Model

The main purpose of these qualitative descriptions is to clarify and organize the biological processes that you want to include in a model. How you decide to do this (using a life-cycle diagram, a flow diagram, a table, or some other approach) is partly a matter of taste, but we suggest the following rules of thumb. Life-cycle diagrams are useful for discrete-time models in which more than one event can occur during a single time step. Flow diagrams are most useful when there are multiple variables in either discrete or continuous time, although care must be taken to specify the order in which arrows should be considered in a discrete-time model. Alternatively, for discrete-time models with multiple events and multiple variables, a table of events is often the clearest way to describe a model.

2.5 Quantitatively Describe the Biological System

At this point, we are ready to derive dynamical equations for the model. Conceptually, dynamical equations track all of the factors that cause a variable to increase or decrease over time and have the form

$$n(t+1) = n(t) + \text{increase} - \text{decrease} \qquad \text{(recursion equations)}, \qquad (2.2a)$$

$$\Delta n = \text{increase} - \text{decrease} \qquad \text{(difference equations)}, \qquad (2.2b)$$

$$\frac{\mathrm{d}(n(t))}{\mathrm{d}t} = \text{rate of increase} - \text{rate of decrease} \qquad \text{(differential equations)}. \qquad (2.2c)$$

For discrete-time models, we describe the value of the variable in the next time step using (2.2a) or the change in the variable across the time step using (2.2b). If the model involves multiple events per time step, we must specify an order to these events (e.g., with a life-cycle diagram) and apply (2.2a) or (2.2b) after each event, updating the value of the variable before the next event. In practice, it is often easiest to first derive the recursion equation (i.e., 2.2a) and then, from this, construct the difference equation (i.e., 2.2b) if desired. For continuous-time models, the procedure is simpler. We sum all of the factors causing the variable to increase or decrease, regardless of how many events occur in the model. In a continuous-time model, we do not have to worry about the order of events within a time step, because the time step is so small (infinitesimally small) that no two events occur at exactly the same point in time.

To make this process more concrete, let us derive dynamical equations for our toy models (Table 2.3). These equations can be derived directly from an understanding of the models or with the aid of the life-cycle diagrams, flow diagrams, or tables of events. Because the type of qualitative description of the model usually depends on whether it is a discrete- or continuous-time model, we will consider these two cases separately.

TABLE 2.3
Dynamic equations derived in this chapter. Type refers to (1) recursion equation in discrete time, (2) difference equation in discrete time, and (3) differential equation in continuous time.

Model	Type	Equation	
Tree-branch model	1	$n(t+1) = n(t) + b\,n(t)$	(2.3a)
	2	$\Delta n = b\,n(t)$	(2.3b)
	3	$\dfrac{\mathrm{d}(n(t))}{\mathrm{d}t} = b\,n(t).$	(2.8)
Mouse population	1	$n(t+1) = (1+b)(1-d)\,n(t) + m$	(2.4)
	2	$\Delta n = -d\,n(t) + b\,(1-d)\,n(t) + m$	(2.5)
	3	$\dfrac{\mathrm{d}(n(t))}{\mathrm{d}t} = b\,n(t) - d\,n(t) + m$	(2.9)
Flu dynamics	1	$n(t+1) = n(t) + a\,c\,n(t)\,s(t)$	(2.7)
		$s(t+1) = s(t) - a\,c\,n(t)\,s(t)$	
	2	$\Delta n = a\,c\,n(t)\,s(t)$	(from 2.7)
		$\Delta s = -a\,c\,n(t)\,s(t)$	
	3	$\dfrac{\mathrm{d}(n(t))}{\mathrm{d}t} = a\,c\,n(t)\,s(t)$	(2.10)
		$\dfrac{\mathrm{d}(s(t))}{\mathrm{d}t} = -a\,c\,n(t)\,s(t)$	

2.5.1 Discrete-Time Models

Let us start with the simplest of our discrete-time models, the branching model. This model has a single variable (the number of branches) and only a single event can happen during a time step. Given the life-cycle diagram in Figure 2.2a, we can derive the recursion equation by specifying the number of branches existing after the first event (the branching event) occurs.

Recalling that b is the number of new branches that bud off each old branch per day, the total number of branches after the first (and only) event is the number of old branches plus the number of new branches, or $n'(t) = n(t) + n(t)\,b$. The next event on the life-cycle diagram is the census at the time $t+1$, so that $n(t+1)$ equals $n'(t)$. This gives us the recursion equation

$$n(t+1) = n(t) + b\,n(t). \tag{2.3a}$$

From equation (2.3a) we can readily construct the difference equation by subtracting off the current number of branches, $n(t)$:

$$\begin{aligned}\Delta n &= n(t+1) - n(t)\\ &= b\,n(t).\end{aligned} \tag{2.3b}$$

Equation (2.3a) tells us the total number of branches at time $t + 1$, while equation (2.3b) describes how many more branches there are at time $t + 1$ than at time t.

Let us now consider the mouse model (Figure 2.2b). To derive the recursion equation for this model we work our way around the life cycle, updating the value of the variable after each event. Following the logic used in the tree branching example, we have

$$n'(t) = n(t) - d\,n(t) \quad \text{after predation by the cat,}$$

$$n''(t) = n'(t) + b\,n'(t) \quad \text{after births,}$$

$$n'''(t) = n''(t) + m \quad \text{after migration.}$$

After migration, the next event on the life-cycle diagram is the census at the time $t + 1$, so that $n(t + 1)$ equals $n'''(t)$. Plugging the first equation for $n'(t)$ into the second equation, we get $n''(t) = (n(t) - d\,n(t)) + b\,(n(t) - d\,n(t))$, which factors to give $n''(t) = (1 + b)\,(1 - d)\,n(t)$. Plugging this result into the third equation gives the complete recursion

$$\begin{aligned} n(t + 1) &= n'''(t) \\ &= n''(t) + m \\ &= (1 + b)\,(1 - d)\,n(t) + m, \end{aligned} \tag{2.4}$$

which describes the number of surviving mice in the yard on the next day.

Equation (2.4) can be given a relatively simple explanation. A fraction d of mice are eaten by the cat and the remainder $1 - d$ survive, leaving $(1 - d)n(t)$ mice. Next, each surviving mouse gives rise to themselves plus, on average, b babies, resulting in $(1 + b)\,(1 - d)n(t)$ mice. Finally, m new mice arrive, giving equation (2.4).

This general process is summarized in Recipe 2.1:

Recipe 2.1

Writing Recursion Equations from Life-Cycle Diagrams (Discrete-Time Models)

Step 1: Use $n'(t)$, $n''(t)$, $n'''(t)$, etc. to denote the value of the variable after the first, second, third, etc., event in the life cycle and obtain recursions for these according to equation (2.2a).

Step 2: Set $n(t + 1)$ to the value of n after the final event in the life cycle.

Step 3: Substitute the recursion for $n'(t)$ into the recursion for $n''(t)$ and simplify. Then substitute the recursion for $n''(t)$ into the recursion for $n'''(t)$ and simplify, etc., until the resulting expression gives a recursion for $n(t + 1)$ solely in terms of $n(t)$.

If you wish to know the amount of change over the time step, the difference equation can be derived using Recipe 2.2:

Recipe 2.2
Deriving a Difference Equation from a Recursion Equation
Step 1: Calculate $n(t + 1)$ using Recipe 2.1.
Step 2: Subtract $n(t)$ from $n(t + 1)$ and simplify to get the difference equation, $\Delta n = n(t + 1) - n(t)$, describing the change in the variable per time step.

For the cat and mouse model, we get

$$\begin{aligned} \Delta n &= n(t + 1) - n(t) \\ &= (1 + b)(1 - d)\, n(t) + m - n(t) \\ &= -d\, n(t) + b\,(1 - d)\, n(t) + m. \end{aligned} \tag{2.5}$$

Taken together, these terms describe all of the changes in the mouse population per day.

The order of events can have a large impact on the predictions of a model. Consider a rather extreme case in the mouse model, where the cat catches 100% of the mice ($d = 1$), 10% of surviving mice give birth each day ($b = 0.1$), and mice arrive in droves ($m = 100$). If we start with one mouse ($n(0) = 1$) and plug these numbers into equation (2.4), we predict 100 mice after one day. If, however, predation is the last event rather than the first event (Figure 2.3), we have

$$\begin{aligned} n'(t) &= n(t) + b\, n(t) && \text{after births,} \\ n''(t) &= n'(t) + m && \text{after migration,} \\ n'''(t) &= n''(t) - d\, n''(t) && \text{after predation by the cat,} \end{aligned}$$

and the recursion equation will be

$$n(t + 1) = (1 - d)\,(n(t) + b\, n(t) + m). \tag{2.6}$$

Plugging in the same parameters, we now predict 0 mice rather than 100 mice after a day.

Which is the right answer? It depends on when we count the mice, when we let out the cat, and when mice tend to move about and give birth. Consider counting the mice at noon. Mice tend to be nocturnal, and it might be reasonable to assume that those that migrate do not immediately give birth that same night. If the cat is out only in the afternoon, equation (2.4) is a reasonable approximation to the system (afternoon: cat eats; night: mice give birth and then move in; noon: mice get counted). If the cat is out only in the morning, however, equation (2.6) is more appropriate (night: mice give birth and then move in; morning: cat eats; noon: mice get counted). Indeed, the difference in

the predicted number of mice makes sense even without a model—if you don't want mice around at your luncheon, then you'd better let the cat out in the morning, not after lunch.

The above example is extreme, but it emphasizes that ordering matters in discrete-time models, and it cannot be ignored. Lest you become too anxious about getting the order of events in a model perfectly right, however, the order typically does not have a large effect as long as little happens during any given time unit (specifically, when each term in the difference equation, Δn, is small relative to $n(t)$). In this case, the results depend less on what just happened within a time unit (which will be relatively little) and more on the value of the variable at the beginning of the time step, $n(t)$. Indeed, many discrete-time models are built by assuming that every change to the variables depends only on their values at the last census. To see this point, try comparing equations (2.4) and (2.6) with more moderate values of the parameters: $d = 0.1$, $b = 0.1$, $m = 1$, and $n(0) = 10$.

Our flu model has two variables. In Figure 2.2c, we drew a life-cycle diagram for each variable. You should try using Recipe 2.1 and Figure 2.2c to construct a recursion equation for each of the variables. Here, we will follow a different approach and use the flow diagram in Figure 2.4c.

Recipe 2.3

Writing Recursion Equations from Flow Diagrams (Discrete-Time Models)

Step 1: Considering each solid arrow in turn, update the value of each variable by taking its previous value

- plus the flow if the arrow enters the circle
- plus the flow if the arrow leaves and returns to the circle
- minus the flow if the arrows leaves the circle.

Step 2: Set $n(t + 1)$ to the value of n after the final arrow has been considered.

There is only one solid arrow in Figure 2.4c. Thus, we need only consider how it affects the number of people with the flu (plus $a\,c\,n(t)\,s(t)$ because the arrow enters the circle representing the number of flu carriers)

$$n(t + 1) = n(t) + a\,c\,n(t)\,s(t) \tag{2.7a}$$

and the number of susceptible individuals (minus $a\,c\,n(t)\,s(t)$ because the arrow leaves the circle representing the number of healthy individuals)

$$s(t + 1) = s(t) - a\,c\,n(t)\,s(t). \tag{2.7b}$$

Alternatively, these equations can be derived using a table of events (Table 2.2), by multiplying the number of contacts by the change caused to the number of infected and susceptible individuals.

2.5.2 Continuous-Time Models

For continuous-time models, differential equations are derived by summing the rates of all changes that occur to a variable, as described by a flow diagram. In fact, we can use the same flow diagrams (Figure 2.4) as before, remembering that the flows across the arrows are now described as rates and that we don't have to worry about the order of events (because continuous-time models consider infinitesimally small time intervals, during which two events are unlikely to occur simultaneously: Box 2.6).

Recipe 2.4:
Writing Differential Equations from Flow Diagrams (Continuous-Time Models)

$$\frac{\mathrm{d}(n(t))}{\mathrm{d}t} = \text{the flow rates along arrows entering the circle}$$
$$+ \text{ the flow rates along arrows leaving and returning to the circle}$$
$$- \text{ the flow rates along arrows exiting the circle.}$$

For the branching model (Figure 2.4a), there is only one way that the number of branches changes (by the budding off of new branches), and the differential equation is

$$\frac{\mathrm{d}(n(t))}{\mathrm{d}t} = b\,n(t). \tag{2.8}$$

The right-hand side is the same as the difference equation (2.3b) for Δn in the discrete-time model. This makes sense because both difference equations and differential equations describe changes to the variables. In contrast, the recursion equation (2.3a) also has $n(t)$ on the right-hand side because it describes the value of the variable rather than how it changes.

As mentioned above, the order of events within a time interval is irrelevant in continuous-time models because the change per time interval considered is infinitesimally small. Thus, in our mouse example (Figure 2.4b), we do not have to update the variable after each event, and we can drop the prime notation ($n'(t)$, etc.). Applying Recipe 2.4 to Figure 2.4b, the differential equation describing the number of mice is then

$$\frac{\mathrm{d}(n(t))}{\mathrm{d}t} = b\,n(t) - d\,n(t) + m, \tag{2.9}$$

whose terms take into account changes due to births, predation, and immigration, respectively. Now, however, the right-hand side does not look the same as the difference equation Δn. As you may have surmised, the reason is that the difference equation (2.5) allows only one bout of deaths, followed by births,

followed by migration, whereas the differential equation (2.9) allows these events to occur continuously throughout the day.

Finally, for the flu model (Figure 2.4c), we can apply Recipe 2.4 to translate the flow diagram into a pair of differential equations modeling the number of people with the flu and those that are susceptible:

$$\frac{\mathrm{d}(n(t))}{\mathrm{d}t} = a\,c\,n(t)\,s(t), \tag{2.10a}$$

$$\frac{\mathrm{d}(s(t))}{\mathrm{d}t} = -a\,c\,n(t)\,s(t). \tag{2.10b}$$

The above three toy examples illustrate how flow diagrams can be used to derive the equations of simple models. But the organizational techniques that we have described really become indispensable when constructing more complex models. To illustrate this, Boxes 2.4 and 2.5 derive the differential equations used in the HIV models introduced in Chapter 1. Box 2.4 develops Phillips' (1996) model for the dynamics of HIV within an individual. Phillips used these equations to predict how the numbers of virus particles within the bloodstream might change following infection by HIV. Box 2.5 develops the model of Blower et al. (2000) for the dynamics of HIV spread among individuals within the San Francisco gay male community. Blower *et al.* used these equations to predict the effects of antiretroviral therapies on the spread of HIV and on the total rate of death from AIDS. While these models are more complex and address more important biological questions, the steps involved in deriving the models are identical (Box 2.1).

Although we have derived the above differential equations with the aid of flow diagrams, they can also be derived directly from the discrete-time models by letting the time step shrink, as shown in Box 2.6. Box 2.6 also sheds light on several key differences between discrete- and continuous-time models. In particular, Box 2.6 clarifies the meaning of rate parameters and why constraints on these parameters differ in the two types of models. Box 2.6 also provides insight into when discrete- and continuous-time models will exhibit similar behavior and why they need not.

2.6 Analyze the Equations

At this point, we say that our model has been *fully specified*. We know the variables, the type of model, the parameters, and the equations describing changes in the variables (Table 2.3). The next step is to analyze the model. There are many different ways of analyzing equations, several of which we will discuss in this book. These include (in order of increasing difficulty)

- Graphical analyses (Chapter 4)
- Simulations (Chapter 4)
- Equilibrium and stability analyses (Chapters 5, 7, and 8)
- Deriving general solutions (Chapters 6 and 9)
- Determining long-term or asymptotic behavior (Chapter 10)
- Analyzing the model for periodic behavior (Chapter 11)

Box 2.4: Deriving the Equations in Phillips (1996)

We illustrate Phillips' (1996) model in the form of a *flow diagram* (Figure 2.4.1). The circles represent the number of susceptible CD4+ cells, $R(t)$, the number of latently infected cells, $L(t)$, the number of actively infected cells, $E(t)$, and the number of virions in the blood stream, $V(t)$. The arrows connecting these circles represent the rate per day at which one category leads to another, where the total flow rate is written beside each arrow. When two arrows meet, this represents an interaction that must occur between two categories to give rise to another category (e.g., an uninfected cell must encounter a virus to become infected).

Let us walk through this flow diagram from left to right. By doing so, we are essentially describing all of the assumptions made by Phillips (1996). At a rate of Γ per day, the immune system produces new uninfected CD4+ cells, of which a fraction τ become susceptible to attack by HIV. Even without HIV infection, CD4+ cells die or are eliminated from the body at a rate μ per susceptible cell per day, leading to a total flow out of the circle of $\mu\, R(t)$ per day. In addition, susceptible CD4+ cells become infected if they encounter a virus. New infections are assumed to occur at a rate $\beta\, V(t)$ per susceptible cell per day, leading to a total flow of $\beta\, V(t)\, R(t)$ per day. This is the simplest equation that captures the fact that cells should become infected at a faster rate if there are more cells to be infected ($R(t)$) or more viruses to do the infecting ($V(t)$). β is a constant that determines whether infections occur slowly (low β) or rapidly (high β); it is analogous to the product of the contact rate (c) and the probability of infection (a) in the flu model. The rate of new infections, $\beta\, V(t)\, R(t)$, employs the "mass-action" assumption.

Once infected, a CD4+ cell may harbor HIV in a latent, nonreplicating state or in its actively replicating state; Phillips lets p describe the probability that HIV becomes latent within a newly infected cell so that $1 - p$ is the probability that HIV becomes actively replicating. Because HIV

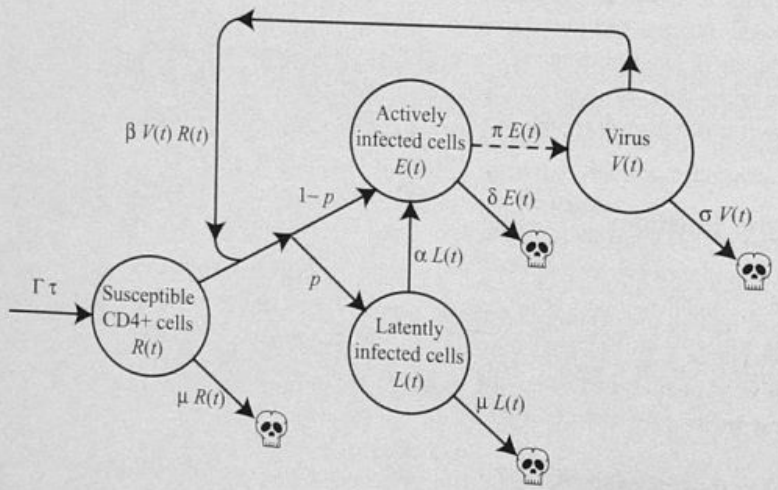

Figure 2.4.1: Flow diagram for viral load. The model describes the number of viruses in the blood stream after HIV infection (Phillips 1996).

(continued)

Box 2.4 *(continued)*

is hidden within the genome of latently infected CD4+ cells, it is assumed that these cells die at the same rate as uninfected cells (μ per cell per day). Latently infected cells may also be activated, however, which occurs at rate α per cell per day. Actively infected CD4+ cells are thus produced by two means: by the immediate conversion of an uninfected cell at rate $(1 - p)\,\beta\, V(t)\, R(t)$ or by the conversion of a latently infected cell at a rate $\alpha\, L(t)$. Actively infected cells die at a much faster rate δ per cell per day, due to the continual budding of virus particles at a rate π per infected cell per day. (We use a dashed arrow between actively infected cells and viruses because viral production by budding does not directly eliminate an infected cell.) Finally, virus particles degrade or are eliminated from the body at a rate σ per virion per day.

From the flow diagram illustrated in Figure 2.4.1, we can write down differential equations, describing the rate of change of each variable over time (e.g., $\mathrm{d}V(t)/\mathrm{d}t$ for the rate of change of virus particles). Each variable represented by a circle in Figure 2.4.1 changes at a rate equal to the sum of all of the arrows entering the circle minus all of the arrows exiting the circle:

$$\begin{aligned}
\frac{\mathrm{d}R(t)}{\mathrm{d}t} &= \Gamma\,\tau - \mu\, R(t) - \beta\, V(t)\, R(t),\\
\frac{\mathrm{d}L(t)}{\mathrm{d}t} &= p\,\beta\, V(t)\, R(t) - \mu\, L(t) - \alpha\, L(t),\\
\frac{\mathrm{d}E(t)}{\mathrm{d}t} &= (1 - p)\,\beta\, V(t)\, R(t) + \alpha\, L(t) - \delta\, E(t)\\
\frac{\mathrm{d}V(t)}{\mathrm{d}t} &= \pi\, E(t) - \sigma\, V(t).
\end{aligned} \tag{2.4.1}$$

(Technically, the rate at which virus particles infect susceptible cells should also be subtracted off from $\mathrm{d}V(t)/\mathrm{d}t$, but this rate is assumed small relative to the large number of virus particles in the bloodstream.) These equations were used by Phillips' (1996) to predict how the number of viral particles varied over time after initial infection with HIV (see Chapter 1 and Figure 1.5).

Box 2.5: Deriving the Equations in Blower et al. (2000)

Blower et al. (2000) developed a model to predict changes in HIV incidence in the San Francisco community of gay males. The authors were particularly concerned that effective antiretroviral therapies (ART) might cause people to be less cautious when engaging in behavior posing a risk for HIV transmission. Here we present a slightly simplified version of their model that ignores the evolution of HIV resistance. Their model assumes that ART has an influence on survival rates, sexual behavior, and the spread of HIV among gay men that are sexually active within San Francisco. In particular, it assumes that the average number of sexual partners with whom an HIV− individual has unprotected sex per year increases from c before ART to $c(1 + i)$.

(continued)

Box 2.5 *(continued)*

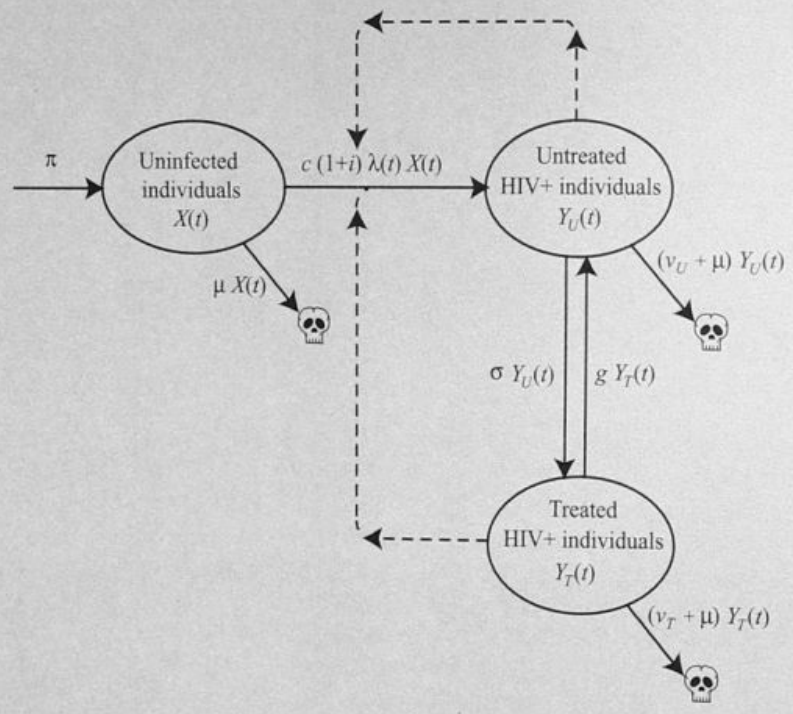

Figure 2.5.1: Flow diagram for HIV and AIDS cases. The model describes the number of cases of HIV and AIDS in the gay male community of San Francisco (Blower et al. 2000). See Table 2.5.1 for further description of the parameters.

A flow diagram for the model of Blower et al. (2000) is illustrated in Figure 2.5.1. The circles represent the number of uninfected individuals, $X(t)$, the number of infected individuals taking drug therapy, $Y_T(t)$, and the number of infected individuals not taking drug therapy, $Y_U(t)$. The arrows represent the rate per year at which one category leads to another category, where the total flow rate is written beside each arrow. When two arrows meet, this represents an interaction that must occur between two categories to give rise to another category (e.g., between infected and uninfected individuals). The parameters describing the flow rates in Figure 2.5.1 are defined in Table 2.5.1.

From the flow diagram, we can determine the rate of change of each variable using Recipe 2.4:

$$\begin{aligned}
\frac{dX(t)}{dt} &= \pi - c(1+i)\,\lambda(t)\,X(t) - \mu\,X(t) \\
\frac{dY_U(t)}{dt} &= c(1+i)\,\lambda(t)\,X(t) + g\,Y_T(t) - \sigma\,Y_U(t) - \mu\,Y_U(t) - \nu_U\,Y_U(t), \qquad (2.5.1) \\
\frac{dY_T(t)}{dt} &= \sigma\,Y_U(t) - g\,Y_T(t) - \mu\,Y_T(t) - \nu_T\,Y_T(t).
\end{aligned}$$

(continued)

Box 2.5 *(continued)*

TABLE 2.5.1
Parameters in HIV/AIDS model. Parameters in the model predicting HIV incidence following antiretroviral therapy (Blower et al. 2000). All rates are per year, and the model assumes that changes to the community are occurring continuously.

Parameter	Description	Value
π	The rate at which HIV– men join the gay community in SF	2133
μ	The rate at which gay men leave the sexually active community[a]	1/30
c	The number of partners per year with whom risky sex occurs (before ART)	1.7
$c\,(1 + i)$	The number of partners per year with whom risky sex occurs (after ART)	$1.7\,(1 + i)$
β_U	Chance of infection per HIV+ partner (untreated) with whom risky sex occurs	0.1
β_T	Chance of infection per HIV+ partner (treated) with whom risky sex occurs	0.025
$\lambda(t)$	Force of infection per partner; $\lambda(t) = \dfrac{\beta_U Y_U(t) + \beta_T Y_T(t)}{X(t) + Y_U(t) + Y_T(t)}$	Varies
σ	The rate at which untreated HIV+ men enter treatment	0.5[b]
g	The rate at which treated HIV+ men abandon treatment	0.05
ν_U	The death rate of untreated HIV+ men from AIDS (expected survival of 12 years following infection)	1/12
ν_T	The death rate of treated HIV+ men from AIDS (expected survival of 27 years following infection)	1/27

[a]By moving away, becoming sexually inactive, or dying for reasons unrelated to HIV/AIDS.
[b]The current rate in San Francisco.

In this model, it is assumed that uninfected individuals engage in behavior that puts them at risk of contracting HIV at a rate of $c(1 + i)$ per uninfected individual per year, where i equals zero before ART but rises to some unknown value after ART. Unlike the flu model (2.7), this rate is assumed to be a personal decision that does not depend on the number or density of possible sexual partners. That is, individuals don't just bump into each other randomly as assumed in a mass-action model; instead they actively seek out sexual partners at a particular rate $c(1 + i)$. For each sexual contact, the composition of the population determines the probability that the contact results in an infection. At time t, this probability is given by $\lambda(t)$ (the per capita "force of infection"), which incorporates the probability that a sexual partner is HIV+ times the probability of acquiring HIV from this partner during sex, β. Specifically, if $N(t)$ is the total number of potential partners, $N(t) = X(t) + Y_U(t) + Y_T(t)$, then the probability that a sexual partner is HIV+ but not undergoing treatment is $Y_U(t)/N(t)$; such partners tend to have a higher transmission

(continued)

Box 2.5 *(continued)*

probability β_U. Similarly, the probability that a sexual partner is HIV+ and undergoing treatment is $Y_T(t)/N(t)$; such partners tend to have a lower transmission probability β_T. Accounting for the possibility of contracting HIV from either type of individuals, the force of infection is

$$\lambda(t) = \beta_T \frac{Y_T(t)}{N(t)} + \beta_U \frac{Y_U(t)}{N(t)}. \tag{2.5.2}$$

Using the parameter values in Table 2.5.1, equations (2.5.1) were solved numerically to generate Figure 1.6 (as described in Chapter 4).

Box 2.6: The Relationship between Discrete-Time and Continuous-Time Models

Although discrete- and continuous-time models are different, they share several fundamental similarities. In fact, one can derive a continuous-time model directly from a discrete-time model by shrinking the length of the time unit down to zero. By describing this procedure, we gain a much clearer understanding of the relationship between discrete- and continuous-time models.

Consider the mouse model, which was derived using a day as the unit of time. What would happen in a shorter unit of time, Δt? In order for this procedure to work, we have to assume that the same set of events could occur in the same order in successively smaller time units. (If this does not make biological sense, e.g., if migration only happens at night, then we should not use a continuous-time model to describe the process.) In the mouse model, the first event that happened was that the cat ate a fraction d of the mouse population. Now, in half a day, we would expect the cat to eat half this amount, $d/2$. In general, in a shorter amount of time Δt, we would expect the cat to eat a fraction $d\,\Delta t$ of the mouse population. In this procedure, we assume that d retains the same value. Nevertheless, as the time interval shrinks, it is possible for d to take on larger and larger values without depleting the entire population of mice. For instance, in cutting the day in half, $d/2$ is the fraction of mice eaten in half a day, and this must still lie between zero and one. Now, however, d can lie anywhere between zero and two. Similarly, for smaller time increments we have the restriction that $d\Delta t$ must lie between zero and one and therefore that d must lie between zero and $1/\Delta t$. As Δt gets smaller and smaller, the maximum allowable value of d gets larger and larger. This reveals why parameters that describe flow are restricted to be less than one in discrete-time models but can have no upper limit in continuous-time models. The same argument applies to both births and migration events, of which we now expect $b\,\Delta t$ and $m\,\Delta t$ to occur in the time unit Δt. Thus, for a discrete-time model with time unit Δt, we would replace d, b, and m in Figure 2.2b with $d\,\Delta t$, $b\,\Delta t$, and $m\,\Delta t$.

(continued)

Box 2.6 *(continued)*

We then proceed, event by event, through the life cycle, using Recipe 2.1 to generate a recursion equation:

$$n'(t) = n(t) - d\,\Delta t\, n(t) \qquad \text{after predation by the cat,}$$

$$n''(t) = n'(t) + b\,\Delta t\, n'(t) \qquad \text{after births,}$$

$$n'''(t) = n''(t) + m\,\Delta t \qquad \text{after migration.}$$

Now $n'''(t)$ is $n(t+\Delta t)$, the number of mice after the time unit, Δt has passed. By plugging the first equation for n' into the second equation and the resulting equation for $n''(t)$ into the third equation, we get

$$n(t + \Delta t) = (1 + b\,\Delta t)\,(1 - d\,\Delta t)\,n(t) + m\,\Delta t. \tag{2.6.1}$$

Next, we can use the definition of a derivative to convert recursion (2.6.1) into a differential equation (Box 2.2). According to the definition of a derivative (Appendix 2),

$$\frac{\mathrm{d}n(t)}{\mathrm{d}t} \equiv \lim_{\Delta t \to 0}\left[\frac{n(t + \Delta t) - n(t)}{\Delta t}\right] \tag{2.6.2}$$

Here's how to read equation (2.6.2) in words: the derivative of n with respect to t is defined ("$\equiv$") as the change in n over a time interval (that is, $n(t + \Delta t) - n(t)$) divided by the length of the time interval (Δt), in the limit as the time interval shrinks to zero ("$\lim_{\Delta t \to 0}$").

We begin the conversion process by plugging (2.6.1) into the term in square brackets in (2.6.2):

$$\begin{aligned}\frac{n(t + \Delta t) - n(t)}{\Delta t} &= \frac{(1 + b\Delta t)(1 - d\,\Delta t)\,n(t) + m\Delta t - n(t)}{\Delta t} \\ &= b\,n(t) - d\,n(t) + m - b\,d\,n(t)\Delta t.\end{aligned} \tag{2.6.3}$$

Next, we let the time interval Δt go to zero, which causes the last term to drop out. We are left with the same differential equation (2.9) that we derived directly from the flow diagram for the continuous-time model. This procedure works for any discrete-time model as long as it is reasonable to allow each event to occur in successively smaller periods of time Δt.

As another example, consider the discrete-time flu model. In a short amount of time Δt, we expect that an infected person contacts any given susceptible person with probability $c\,\Delta t$. Therefore, the total number of contacts with susceptible individuals is $c\,\Delta t\, s(t)$ per infected individual in this short period of time. Again, while the parameter c must lie between zero and one in the discrete-time model, the maximum allowable value of this parameter now increases without bound as we shrink the time interval Δt to zero. Every time a contact occurs, whether in discrete or continuous time, the probability that the flu is transmitted to the healthy person is a. Given that there are $n(t)$ such infected individuals, we expect a total number of new flu cases in

(continued)

Box 2.6 *(continued)*

the time interval Δt to be $a\,c\,\Delta t\,n(t)\,s(t)$. Using Recipe 2.1, the number of flu cases after an interval of time Δt would be

$$n(t + \Delta t) = n(t) + a\,c\,\Delta t\,n(t)\,s(t). \tag{2.6.4}$$

Plugging (2.6.4) into (2.6.2) and taking the limit

$$\frac{\mathrm{d}n(t)}{\mathrm{d}t} \equiv \lim_{\Delta t \to 0}\left[\frac{a\,c\,\Delta t\,n(t)\,s(t)}{\Delta t}\right] = a\,c\,n(t)\,s(t), \tag{2.6.5}$$

we regain the differential equation (2.10a).

The above procedure illustrates that the way in which we scale down the flow rates as the time interval Δt decreases determines the restrictions on the parameters in the continuous-time models. For example, in deriving the continuous-time flu model from equation (2.6.4), we saw that the maximum allowable value of c increased to infinity. But the parameter a, which represents the probability of transmission per contact, retains the restriction of having to lie between zero and one from the discrete-time model because we did not scale this parameter at all when shrinking the time interval Δt to zero.

Because it is a considerable source of confusion for new (and experienced) modelers, it is also worth clarifying the difference between discrete- and continuous-time models in terms of the units of the parameters. In our discrete-time mouse model, d was a *fraction* (and therefore was constrained to lie between zero and one) whereas such parameters are referred to as *rates* in continuous-time models. But how and why did a fraction become a rate when moving to continuous time, and what is the relationship between the two? One way to understand this is to notice that the fraction of mice eaten in any time interval can always be written as a rate of consumption per unit time, d, multiplied by the length of the time interval in question, Δt. Of course, as we have seen, we must ensure that d and Δt are chosen so that $d\,\Delta t$ lies between zero and one because this represents the *fraction* eaten. In a discrete-time model, the unit of time is arbitrarily assigned the value of $\Delta t = 1$, so that we can view d as really being $d\,\Delta t$, where d remains a rate and $d\,\Delta t$ remains a fraction. This would, however, make the discrete-time equations harder to read, so that it is much clearer to refer to $d\,\Delta t$ as simply d. In deriving the continuous-time model, we put Δt explicitly back into the discrete-time model. But when we applied equation (2.6.2) to obtain a differential equation, we divided our expression for the change in $n(t)$ over the time interval by the length of the interval Δt. As a result, the continuous-time equations involve d alone on the right-hand side of the equation (rather than $d\,\Delta t$), which is a consumption *rate* (i.e., consumption per unit time) rather than a fraction.

Finally, just because we can derive continuous-time equations from discrete-time equations does not guarantee that they will behave in the same way. In fact, we will see some spectacular differences between these two types of models in Chapter 4. By deriving continuous-time equations from discrete-time equations, however, we gain some insight into when and why the dynamics of these models should differ. In the above, we shrunk the time interval, from one to 1/2 to Δt, which we then allowed to shrink to zero. This procedure changes how often the variables in

(continued)

Box 2.6 *(continued)*

the model are updated, from once to twice to $1/\Delta t$ times per original time unit. Every time we update the variables, we allow the changes that occur within one time interval to impact the changes that occur within the next time interval. If none of the variables change by much within one unit of time, then updating the variables will make little difference. If, however, the variables undergo large changes within a time unit (i.e., changes by more than just a few percent), then it will matter whether we fix the value of the variables to their initial values within a time unit or update these variables after every small interval of time Δt. Continuous-time models represent the extreme case where the variables are continuously updated over time. We will return to this issue in Chapters 4 and 6, but for now we conclude with the following important point: *The behavior of discrete-time and continuous-time models will be similar if each variable changes little over the time unit considered in the discrete-time model.*

When changes over a time unit are not small, all bets are off, and the discrete-time and continuous-time models can behave quite differently. In this case, discrete-time models are also quite sensitive to the ordering of events within a time unit (see the luncheon discussion in section 2.5.1). Unless there is a good biological reason to believe in one type of model (discrete-time versus continuous-time) and one type of ordering, then you should be careful before placing too much stock in any predictions from a model in which large changes can occur over a time unit. When such large changes are possible, it might be worthwhile deriving both a discrete-time and a continuous-time version of the model to see how sensitive the results are to the way in which the problem is modeled.

If you are just starting to model, the number of different mathematical techniques that are available is daunting. Keep in mind that even the best mathematicians do not know them all. Any modeler knows only a subset of possible techniques. It helps to remember this—not only because it's easier to manage learning math when you don't feel that you have to learn everything, but also because it is important to recognize that you should always keep an eye out for useful new techniques to add to your mathematical toolbox. We can always learn (and develop!) more techniques. A good idea is to read papers in the area that interests you to decide which mathematical techniques to learn first. While no one person can master all mathematical techniques, knowing the basic steps of modeling can allow you to collaborate effectively with modelers who do know the techniques that you need.

2.7 Checks and Balances

The process of mathematical modeling is rarely smooth. Rarely does the first set of equations that you write down end up being the final set. Generally, modeling is an iterative procedure. First of all, everybody makes mistakes. This means that it is critical to check your equations and analyses thoroughly and to start over again whenever you discover a mistake. The most obvious way to

check for mistakes is to rederive everything. Oddly, many mistakes are not caught this way, probably because our minds are likely to make the same error twice. Therefore, it is a good idea to get into the habit of checking your results using other pieces of information.

If there are any constraints on the variables, make sure that your results obey these constraints. For example, if you are modeling the proportion of females within a population and the proportion of males, then the sum of these proportions should equal one. If you are modeling the number of mice within a population, you should stop the model as soon as the number becomes negative, which means that the mice have gone extinct. If the biological processes considered only add to a variable, that variable should never decrease over time.

Similarly, make sure that each equation has the right units or "dimensionality"—if you are modeling the number of individuals in a population, your answer should have the dimensions of a number, not a number squared. Plus, the units of the right-hand side of an equation should equal the units of the left-hand side. For example, in the cat-mouse model, equation (2.4) has units of number of mice on the left, $n(t + 1)$. On the right, we have $(1 + b)(1 - d)$ $n(t) + m$. The term $(1 + b)$ has units "number of mice per mouse." The term $(1 - d)$ measures the fraction of surviving mice and therefore has no units. Thus, $(1 + b)(1 - d)\, n(t)$ correctly has units of "number of mice," as does m, the number of (migrant) individuals. In the flu model, equation (2.7a) has units of number of infected individuals on the left, $n(t + 1)$. On the right, we have $n(t) + a\, c\, n(t)\, s(t)$. The first term, $n(t)$, has the units "number of infected individuals." At first, it might seem as if the second term $a\, c\, n(t)\, s(t)$, has the dimensions of "(number of infected individuals) × (number of susceptible individuals)" because it involves the product of $n(t)$ and $s(t)$. But c is the probability of any given infected individual contacting a susceptible individual per number of susceptible individuals in the population (see section 2.4.2) and therefore has units of "1/(number of susceptible individuals)," while a is a probability and has no units. Thus, $a\, c\, n(t)\, s(t)$ also has the units "number of infected individuals."

Another way to check your results is to look at special cases where you know what should happen. For example, in the mice model, if predation and immigration are absent ($d = 0$ and $m = 0$), the equations should describe the same growth process as the tree-branching model (e.g., the recursion equation (2.4) becomes $n(t + 1) = (1 + b)\, n(t)$, which is identical to equation (2.3)). Therefore, any results that you obtain for the mice model with $d = 0$ and $m = 0$ should be the same as those for the branching model. Another good idea is to check results against simulations, which represent a special case where all parameters and starting conditions are specified.

Conversely, you can save a lot of effort if you notice that your model is a special case of another model or can be written in the same form as another model. This is helpful because you can then apply the known results of the other model to your own problem. For example, the equations for the branching model have the same form as the equations describing exponential growth (Chapter 3). The exponential growth model has been well studied, and we

know a lot about its behavior. Realizing this, we can apply our knowledge about exponential growth (e.g., Figure 4.1 and equation (4.1)) to our tree branching model.

Finally, but most importantly, you should check your results by seeing if they make sense. What did you expect to happen? Do the results match your expectations? If the results match, they are more likely to be correct. If they don't match, then either the results are wrong or your intuition is wrong. If, after extensive checking, you cannot find any errors in the math, then try to figure out why your intuition was wrong. This is often an extremely valuable exercise, allowing you to correct and refine your understanding of the biological system. For example, the results of the model studied by Phillips (1996) led the research community to reevaluate what forces were driving HIV dynamics after infection (see Box 2.4 and section 1.3.1). Phillips' model showed that it was possible for viral loads to rise rapidly and then decline without the immune system kicking in. In hindsight, this result makes sense, even if it was difficult to foresee before modeling the problem.

Another reason why your final model might be different from the initial model is that you can get through the entire process and realize that your initial model was too simple or too complex. Sometimes, your results will indicate that your model was not exactly what you intended. For example, because the tree-branching model has the same solution as the model of exponential growth, the number of branches will grow exponentially (see Figure 4.1). This makes sense, because we did not include anything in the model to slow growth once the tree gets large. Realizing this, we might want to redo our model and include the possibility that, as the tree grows, it experiences more competition and shading from nearby branches. One way to do this would be to let the growth rate b decline as the number of branches grows. We shall talk about an extension to the exponential growth model that does exactly this in Chapter 3 (the logistic growth model). Similarly, as we have modeled the flu, susceptible individuals get infected until everybody has the flu. But in the real world there is never a time when everybody has the flu, because people recover from the flu and can also become resistant (see the SIR model developed in Chapter 3).

It is also easy to make your initial model too complex by including too many variables and parameters. With an overly complex model, you are much more likely to run into a brick wall in the analysis. Always consider whether every variable and parameter is necessary for you to address the biological question. Sometimes the answer will be "yes," in which case explore the model as best you can. At other times, certain details that, upon reflection, are less important can be dropped. For example, in the cat and mouse model, you might initially keep track of both the number of male mice and the number of female mice. But after running some simulations of the model, you might realize that it is only the number of female mice that matters to the dynamics, except when there are not enough males to fertilize the females. At this point, you might then decide to reduce the complexity of the model (following the principle of parsimony) and focus only on the total number of female mice.

Part of the art of mathematical modeling is learning how models can be simplified. Sometimes, parameters can be grouped together to reduce the total

number of parameters in a model. For example, in the flu model (2.7), the contact rate c and the probability of infection per contact, a, always enter into the equations as the product $a\,c$. Thus, we can define a new variable $\beta = a\,c$, which measures the infectivity of the flu. Replacing a and c with β allows us to reduce the number of parameters by one. Doing so also allows us to see that increasing the contact rate or the probability of transmission per contact should have equivalent effects on the spread of the flu.

It is also sometimes possible to reduce the number of variables in a model. For example, in the flu model, the total number of individuals, $N = n(t) + s(t)$, remains constant. (You can show this for the discrete model by adding together $n(t + 1) + s(t + 1)$ and showing that the sum is the same as in the previous generation.) Therefore, you can rewrite equation (2.7a) by substituting $s(t) = N - n(t)$ to get an equation that involves only one variable ($n(t)$):

$$\begin{aligned} n(t + 1) &= n(t) + a\,c\,n(t)\,(N - n(t)) \\ &= n(t) + \beta\,n(t)\,(N - n(t)). \end{aligned} \tag{2.11}$$

Reducing the number of parameters and variables makes a model more elegant, but more importantly it can make the model easier to analyze, as we shall see throughout the book. We will describe some methods to reduce the number of parameters and variables in a model in Chapter 9. But the best advice is to keep an eye out for features of a model (e.g., that the number of individuals remains constant) that might help to describe a model in the simplest and most elegant terms.

2.8 Relate the Results Back to the Question

You might be tempted to think that you are done once you have analyzed your model. Modeling biological processes, however, is worthwhile only if the mathematical results are related back to biological problems. At its best, theory is closely tied to empirical observations and tests. For example, empirical observations suggest a theoretical model, which generates an empirical test, which suggests that the model needs to be refined in particular ways. This interplay is extremely fruitful.

Even in the absence of such a tight interplay between modeling and empirical research, it is always desirable to go beyond the mathematical analysis and determine the broader biological insights that can be gained. How do the results alter the way scientists should think about a problem? What predictions can be made based on the model? What experiments could test these predictions? Are there any data that can be explained or better understood in light of the model? For example, the models that we described in the last chapter on HIV dynamics were interesting, not because of their mathematical equations, but because they helped us better understand how HIV might replicate and spread. They changed the way we thought about HIV, made counterintuitive predictions, and suggested empirical tests. Essentially, the difference between an important and widely read model and an irrelevant and obscure model lies,

not in the modeling steps described above, but in this final step: describing how the model helps us better understand and interpret the biological world around us.

2.9 Concluding Message

In this chapter we have introduced the process of model construction by decomposing the task into a series of seven steps (Box 2.1). We have illustrated these steps with a series of toy examples as well as with models from the literature on HIV described in Chapter 1 (Boxes 2.4 and 2.5). In the next chapter we apply these same steps to derive some classic models from ecology, evolution, and epidemiology.

Problems

Problem 2.1: In Phillips' model of HIV dynamics within the body (Box 2.4), what parameter would you alter to incorporate an immune response to HIV particles? In three or four sentences, say how you might expand the model to incorporate this immune response.

Problem 2.2: Ground squirrels engage in alarm calls to alert their fellow squirrels that a predator may be present. Upon hearing a call, silent squirrels may start calling. Over time, calling squirrels may stop calling if the danger has not materialized. Draw a flow diagram with two circles representing the number of silent and calling ground squirrels over time. Place flow rates above each arrow in your diagram and describe in words what the flow represents.

Problem 2.3: The genome of any organism consists of a number of purine nucleotides (adenine and guanine) and pyrimidine nucleotides (cytosine and thymine). During DNA replication, however, mutations occasionally occur, causing a purine to be incorrectly replaced by a pyrimidine or vice versa. Figure 2.5 illustrates a flow diagram for this mutation process.

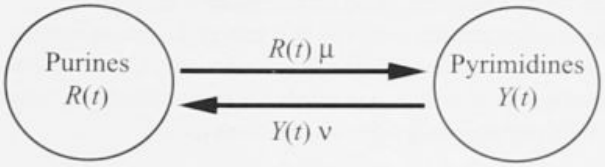

Figure 2.5: Two-state mutation model

(a) Based on the flow diagram, write down discrete-time equations for the number of purines, $R(t)$, and pyrimidines, $Y(t)$. You may choose to write either recursion or difference equations, but you should specify which type of equation you have chosen. (b) Write down continuous-time equations for the number of purines, $R(t)$, and pyrimidines, $Y(t)$.

Problem 2.4: Yeast and bacterial cells can be grown so that they divide continually using a "chemostat." Chemostats are tanks carrying a complete medium with all of the sugars and essential elements necessary for microbial growth, as illustrated in Figure 2.6. New medium is added to the tank via a constant drip (inflow), while

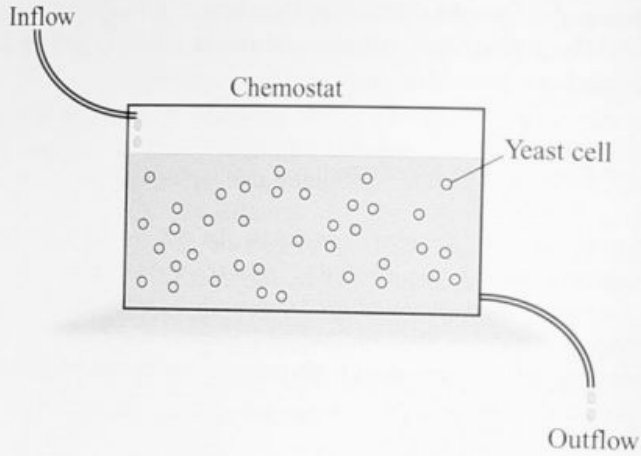

Figure 2.6: Yeast in a chemostat

used medium and cells exit via an effluent tube (outflow). To model the dynamics of a yeast population grown in a chemostat, (a) list all of the variables that you would want to include, (b) list all of the parameters that you think might be relevant, (c) describe the type of model that you are considering (discrete or continuous), and (d) specify any restrictions on the variables and parameters (e.g., $x(t)$ must be positive). Don't forget to describe the units for the variables that you choose (e.g., "number of individuals" or "density per milliliter") and the parameters (e.g., "rate of loss per cell per minute").

Problem 2.5: In Chapter 8 we will analyze a model for disease transmission based on the following equations: $dS/dt = \theta - d\,S - \beta\,S\,I + \gamma\,I$ and $dI/dt = \beta\,S\,I - (d + \nu + \gamma)I$. The variables S and I denote the number of susceptible and infected individuals. (a) Draw and label a flow diagram for these two variables. (b) Suggest a plausible biological interpretation of the parameters γ and ν.

Problem 2.6: Suppose that after contracting the flu, people are initially resistant to reinfection, but this immunity eventually wanes. (a) Alter the flow diagram for the flu model in Figure 2.4c to include a "recovered and immune" class with these properties. (b) Suppose that immune individuals have a constant per capita rate of losing immunity. What are the continuous-time equations (2.10) for this modified flu model?

Problem 2.7: There are six different possible orderings of events in the mouse model of the text. There are, however, only four different recursion equations, because some equations are compatible with more than one ordering of events. Match the orders of events (a)–(f) to their corresponding recursion equations (i)–(iv):

(a) Census, births, predation, migration
(b) Census, births, migration, predation
(c) Census, predation, births, migration
(d) Census, predation, migration, births
(e) Census, migration, births, predation
(f) Census, migration, predation, births

(i) $n(t + 1) = (1 + b)(n(t)\,(1 - d) + m)$
(ii) $n(t + 1) = (1 - d)(n(t)\,(1 + b) + m)$
(iii) $n(t + 1) = (1 - d)\,(1 + b)\,(n(t) + m)$
(iv) $n(t + 1) = (1 - d)\,(1 + b)\,n(t) + m$

Problem 2.8: The flow diagram in Figure 2.7 might describe the dynamics of colonial animals (e.g., naked mole rats) with reproductive individuals, nonreproductive "workers," and a specialized group of workers ("soldiers") that defend the colony and recruit new soldiers from among the worker class. (a) Infer which variables

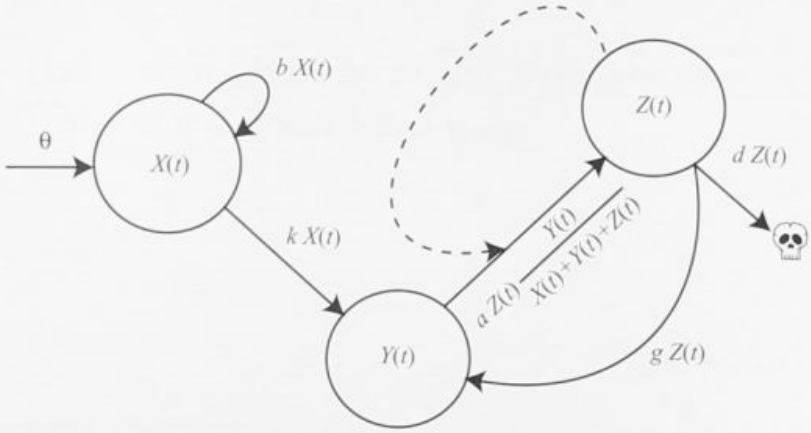

Figure 2.7: Class structure in a naked mole rat colony

correspond to the reproductive class, the soldier class, and the worker class. Specify what parts of the flow diagram you used to draw your inferences. (b) Derive continuous-time differential equations for the three variables $X(t)$, $Y(t)$, and $Z(t)$.

References

Blower, S. M., H. B. Gershengorn, and R. M. Grant. 2000. A tale of two futures: HIV and antiretroviral therapy in San Francisco. *Science* 287:650–654.

Phillips, A. N. 1996. Reduction of HIV concentration during acute infection: Independence from a specific immune response. *Science* 271:497–499.