

Chapter Title: Agent-Based and Other Computational Models for Complex Systems

Book Title: Dynamic Models in Biology

Book Author(s): Stephen P. Ellner and John Guckenheimer

Published by: Princeton University Press

Stable URL: <http://www.jstor.com/stable/j.ctvc4h1q.13>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



Princeton University Press is collaborating with JSTOR to digitize, preserve and extend access to *Dynamic Models in Biology*

JSTOR

8 Agent-Based and Other Computational Models for Complex Systems

So far we have stuck to our general theme: simple models, using tried-and-true approaches that have a mature mathematical theory to help us understand how the models behave. An understanding of simple models and hands-on experience with them are essential prerequisites for working effectively with complicated models.

The purpose of this chapter is to round out the picture by examining some models whose complexity puts them beyond the reach of mathematical analysis. These models are based on simulating each individual unit in a system, generically called an “agent.” In Chapter 4 we briefly considered systems biology—large systems of differential equations for the interactions among numerous genes, signal transduction pathways, and other within-cell processes. The only thing special about those models is their size, dozens to hundreds of state variables. Agent-based models, in contrast, represent a totally different modeling paradigm. Each agent is represented explicitly, and the model consists of context- and state-dependent rules for agents’ possible actions and the consequences that result from them. The term “agents” brings to mind individual organisms, but agent-based models have also been developed for molecules, genes, cells, and tissue segments. The goal of agent-based models is to explain or predict the dynamics of “macroscopic” properties from the rules that operate at the “microscopic” level of system components interacting with their local environment and other agents.

We have already seen one simple agent-based model: discrete-event simulations of an epidemic in a finite population (Chapter 6). In that model each agent is characterized by a single state variable—disease state—and has a simple decision rule for state transitions. The appeal of agent-based models is that it is not much harder to program situations where agents are characterized by many state variables and have more elaborate decision trees for choosing actions. Based on a review of agent-based models for human systems, Bonabeau (2002) identified some properties that favor agent-based modeling:

1. Agents have discontinuous “either-or” responses to other agents’ behaviors.
2. Agents adapt, learn, or otherwise modify their behavior based on past experience.
3. Agents are heterogeneous, either inherently or due to learning and adaptation.
4. Interactions are localized—for example, with spatial neighbors or a small network of social contacts—but the pattern of interactions is dynamic.

Peck (2004) offers a more concise summary:

When an analytic model can be used, it should be used. They are simpler, clearer and usually can be fit to data to make their interpretation much easier. However, some processes are inescapably complex. This is the domain of the complex simulation model.

It is important to realize that “complex” in this quote does not just mean “complicated.” It refers to phenomena that occur with a large number of interacting components or processes, but not with a small collection of the same type of entities. In this sense, discrete-event epidemic models (Chapter 6) are not complex simulation models: in that case the new phenomena arise when there are few enough agents that coin-tossing stochasticity becomes important, while large numbers of agents are well described by a differential equation model. In either case, phenomena beyond the present reach of mathematical analysis can nonetheless be studied by simulation.

The argument against elaborate computational models, discussed in Chapter 4, is that they replace a complex biological system that we don’t understand with a computer model that we don’t understand either. So after presenting examples of agent-based models, we consider two approaches that facilitate the process of understanding a complex model’s properties and predictions.

1. A computer model can be studied “experimentally” in much less time than the real system, with no practical obstacles to changing parameters or modifying system properties. When done *successfully*, this puts computational models on an equal footing with simpler models whose properties can be studied analytically. But success requires really exploring the full range of possibilities, which may be hard for large models with many parameters and assumptions. We describe methods for identifying which parameters and assumptions have the largest effects on the behavior of a model, so that attention can focus on the most important ones.
2. A complementary approach is to derive an approximate model which is easier to understand. Although this is often done in an *ad hoc* way, there are also some systematic approaches that can be tried.

Exercise 8.1. Based on the considerations discussed above, suggest a biological question that could profitably be addressed by an agent-based model, and explain why an agent-based approach is reasonable. Be original: do a literature search to make

sure that your proposed study has not already been published, and describe how you conducted your search. Don't just do just a web search—use Science Citation Index or some other appropriate bibliographic resource for the area of biology.

8.1 Individual-Based Models in Ecology

In ecology, agent-based models are known as “individual-based” models (DeAngelis and Gross 1992). The agents are single organisms, each characterized by a set of “*i*-state variables” specifying attributes that can differ between individuals and change over time. The total number of state variables in the model is then the product of the number of individuals, and the number of *i*-state variables per individual.

8.1.1 Size-Dependent Predation

Individual-based models need not be complicated, because it may be very easy to describe assumptions at the individual level. An example is the model by Rice et al. (1993) for survival of larval fish in the presence of size-dependent predation. The model was developed to explore how individual variation in growth rate might affect the fraction of individuals surviving the larval stage and becoming a juvenile. The model has a daily time step and follows a cohort of larvae over 60 days, with individual size (length) as the only *i*-state variable.

- Fish *i* has initial length 12 mm and growth rate $g_i = (G + z_i)$ mm/day, where $G = 0.2, 0.4, \text{ or } 0.6$, and z_i is chosen from a normal distribution with mean 0, and standard deviation $\sigma = 0.04, 0.08, \text{ or } 0.16$.
- Each day, each fish has a 20% chance of encountering a predator.
- On encounter with a predator, the chance of a fish being eaten depends on the size ratio: $P(\text{capture}) = -0.33 + 0.15 \times (\text{predator length}/\text{prey length})$, truncated onto the interval $[0, 1]$. This equation was based on experiments for 90 mm alewife preying on larval bloater. It assumes that predators are mature fish, all roughly the same size, and not changing in size over the time period being modeled.
- At the end of each day, the size of each surviving fish is increased by its growth rate.

The assumption of constant growth rates set at birth is unrealistic, so Rice et al. (1993) considered a second model with dynamic random growth rates:

$$g_i(t + 1) = g_i(t) + d_i(t),$$

where $d_i(t)$ was drawn each day from a Gaussian distribution with zero mean and standard deviation $\sigma_d = 0.08$. In this model there is no intrinsic tendency for growth rate to go up or down (the mean change is 0), but as time goes on individuals become more variable in growth rate.

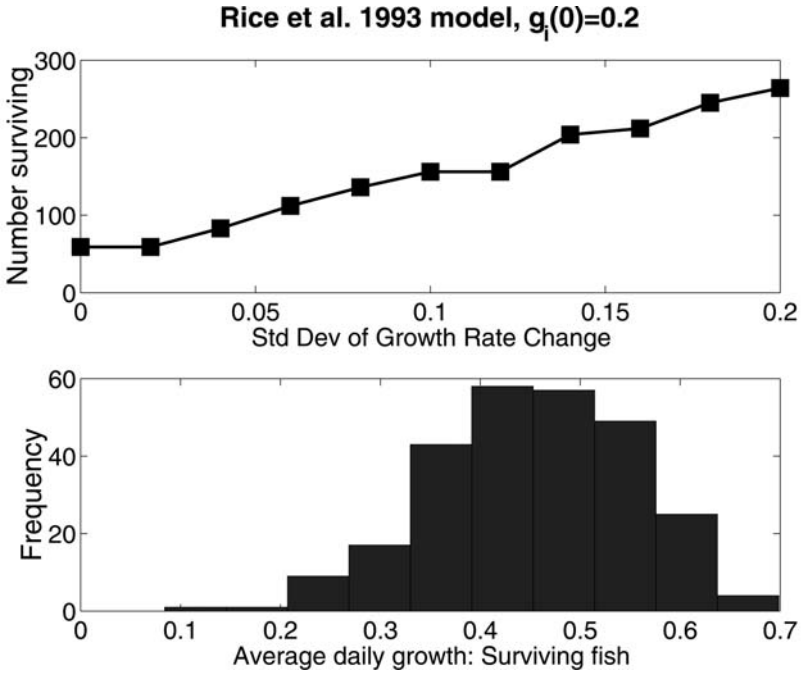


Figure 8.1 Results from simulations of the Rice et al. (1993) individual-based model for larval fish growth and survival in the presence of size-dependent predation. Individuals were started at 12 mm with $g_i(0) = 0.2$ mm/d, and were constrained to have growth rate between 0 and 0.8 mm/day, as in Rice et al. (1993).

The main result from this model was that increases in either the mean or the variance of growth rate increased the probability of survival. The effect of the mean is no surprise. The effect of the variance says that survivors are mostly the lucky few individuals who achieve above-average growth rates. In simulations of the second model, up to 90% of the surviving individuals had actual average growth rates above the initial mean growth rate, depending on model parameters. Growth rate variability could more than double the number of survivors, with most survivors having above-average growth rates (Figure 8.1).

One could use a matrix model based on size rather than an agent-based model. For example, the “stages” could be size-classes 0–0.2mm, 0.2–0.4mm, 0.4–0.6mm, and so on. However, the model would actually require a two-dimensional classification, by size and growth rate, because the probability of size-class transitions depends on growth rate. So a matrix model would need to have a lot of stages, and a large and complicated matrix to describe the odds of moving between categories. In contrast, coding up the individual-based model is a simple exercise (Table 8.1).

This model is *theoretical* even though it is *computational*. Larval growth is a highly simplified abstraction, predation risk is constant even though prey abun-

```

nfish=5000; ndays=50; %5000 fish for 50 days
sizes=zeros(ndays,nfish); growth=zeros(ndays,nfish);
alive=ones(1,nfish);
sigmag=0.08; % standard deviation of changes in growth

% initial sizes and growth rates
sizes(1,1:nfish)=0.4*ones(1,nfish);
growth(1,1:nfish)=0.4;

%iterate the model
for jday=1:(ndays-1);
    % Predation. For simplicity of coding, the dead can die again.
    meetpred=rand(1,nfish)<0.2; %does a predator find me?
    pdie=-0.33 + 0.15*(90./sizes(jday,:)); %if so....
    eaten=rand(1,nfish)<pdie;
    m=find(meetpred.*eaten>0); % do I live or die?
    alive(m)=0;
    % find sizes and growth rates for tomorrow
    sizes(jday+1,:)=sizes(jday,:)+growth(jday,:);
    growth(jday+1,:)=growth(jday,:)+sigmag*randn(1,nfish);
    growth(find(growth<0))=0;
    growth(find(growth>0.8))=0.8;
end;

% plot growth rate of survivors
m=find(alive>0); growthrate=(sizes(ndays,m)-sizes(1,m))/ndays;
hist(growthrate);

```

Table 8.1 MATLAB code for the Rice et al. (1993) model with random changes in growth rate over time.

dance is changing rapidly and predators are changing in size, the causes of growth variation are left unspecified, and so on. In order to focus on the main theoretical question—interactions between variation in growth rate and size-dependent predation—everything else is represented simply or ignored.

8.1.2 Swarm

Enquist and Niklas (2001) used an individual-based model to explain the constancy of size-frequency distributions in forests. The empirical pattern is that the number of trees of trunk diameter x in a given area is proportional to x^{-2} . To

explain this pattern they developed a simulation model in which trees compete with each other for space and light and follow size-dependent rules for allocating assimilated energy. Enquist and Niklas (2001) wrote their model in Swarm, an open-source toolkit for developing agent-based models. The Swarm Development Group (wiki.swarm.org) claimed 300–500 users as of 2002, and lists applications to cell biology, animal behavior, ecology, economics, sociology, and military strategy.

Enquist and Niklas's (2001) simulations begin by scattering a user-specified number of seeds across a finite area. These may be all of one species, or else are individually assigned a species-specific biomass, light requirement, and dispersal range. Each tree's total energy intake is partitioned into stem, leaf, and reproduction, and converted into biomass of each type according to "allometric" allocation rules. These rules specify the relationship between the total biomass of the individual and its parts, and were derived from a theoretical model for such scaling relationships in vascular plants. A tree's stem, leaf, and reproductive biomass determine its height, canopy diameter, and seed production, respectively. Some of the available reproductive biomass may be reserved for future use. Propagules are dispersed randomly over a circle whose radius depends on the height of the parent and the weight of the seed. Individuals growing under the canopies of other, larger individuals receive less light, and therefore grow more slowly, or may die, depending on their light requirements (differing among species). There is also ongoing mortality unrelated to light limitation.

Simulations of the model replicate the size-frequency spectrum of natural forest ecosystems (frequency \sim size⁻²). They were also reported to replicate another pattern in real forests, that the total biomass per area at steady state is roughly constant—-independent of latitude, elevation, and the number of species present. Enquist and Niklas (2001) concluded that "the invariant properties identified for real plant communities emerge from the allometric rules that influence the behaviour of individual plants competing for space and limited resources."

8.1.3 Individual-Based Modeling of Extinction Risk

VORTEX (Lacy et al. 2003) is an agent-based model for population viability analysis (PVA)—quantitative analysis to estimate the risk of extinction for a population, group of populations, or species, over a finite time horizon (typically 10 to 100 years), or to compare possible management options for reducing the risk of extinction. VORTEX takes an individual-based approach to modeling the extinction process in structured populations, and incorporates genetic processes in addition to the demographic processes included in matrix models (Chapter 2).

There are two motivations for an agent-based approach. The first is that extinction necessarily occurs in a small population where the randomness of individual "coin tossing" is likely to be important. The second is that, as in Rice

et al. (1993), individuals are characterized by so many state variables that a matrix model would be unwieldy. The first two are age and sex: a user specifies a maximum possible age, and the age at first breeding for females and males. Juvenile (i.e., nonbreeding) individuals have an age-specific survivorship. Survival of adults is independent of age until they reach the maximum, at which point death is certain. Adult females have an annual probability of breeding, which may be constant or vary randomly, and a user-specified probability distribution for the number of offspring when breeding occurs (i.e., a breeding female has j offspring with probability $a_j, j = 1, 2, \dots, M$). Adult males have an annual probability of mating; the mate for each mating female is chosen at random from the pool of mating males for that year. If the user chooses a monogamous mating system, then each male is only allowed to mate with one female. The pairing up of mating individuals is made independently each year—the model does not allow long-term pair bonds.

The third state variable is genotype, in order to model inbreeding depression. As a population shrinks there will be a loss of genetic variability and a higher rate of matings between closely related individuals (“inbreeding”). “Inbreeding depression” refers to the fact that matings between close relatives tend to produce fewer viable offspring. This may result from deleterious recessive alleles, which are expressed more often in inbred individuals due to their greater chance of getting the same allele from both parents, or from a general advantage of heterozygosity. VORTEX accounts for both of these possible effects. The program tracks the kinship between all individuals, and offspring of closely related parents have a decreased chance of survival in their first year. Each individual in the starting population is assigned ten unique recessive lethal alleles—two each at five hypothetical loci—and the model tracks the genotype of each individual at these loci in subsequent generations, using random numbers to simulate the genotype of offspring based on those of their parents. Offspring having identical alleles at any of the five loci are killed immediately upon birth.

As additional complications, the birth and survival rates can be affected by population density, environmental variation, and catastrophes, in ways determined by user-specified parameters. The population may be subdivided into up to fifty distinct subpopulations, with the user specifying different parameter values for each subpopulation and the degree of migration between subpopulations. The flexibility of VORTEX has allowed it to be used for a wide range of species; a recent program manual (Miller and Lacy 2003) lists over 100 publications or reports using VORTEX; see Table 8.2. The price of flexibility is that a user has to specify lots of parameters (Table 8.3), or else accept default values that may not apply to their species.

The value of VORTEX hinges on its predictions being accurate enough for real-world decision making, even in the typical case where many parameter estimates are imprecise or simply lacking. Results on this count are mixed. For example,

Birds	Other vertebrates
Humbolt penguin	Asiatic lion
Whooping crane	Baird's tapir
Capricorn silvereye	Florida panther
Bearded vulture	Giant panda
Red-cockaded woodpecker	Chinese river dolphin
Kirtland's warbler	Tree kangaroo
Attwater's prairie chicken	Lion-tailed macaque
Mississippi sandhill crane	Leadbeater's opossum
Javan hawk-eagle	Northern white rhinoceros
White-winged wood duck	Sumatran tiger
Hihi	Hawaiian monk seal

Table 8.2 Some of the animal species for which PVA has been carried out using Vortex, based on Appendix III of Miller and Lacy (2003)

Required
Age at first breeding α , males and females
Mating system
Maximum longevity
Mean % of adults breeding each year, males and females
Variance in % breeding each year, males and females
Maximum litter size M
Litter size distribution a_j , $j = 0, 1, \dots, M$
Age-specific mortality p_x , $x = 0, 1, \dots$, males and females
Magnitude of environmental variability, in survival and in fecundity
Correlation of environmental variability between survival and fecundity
% inbreeding depression due to lethals
Optional to create more complex models
Frequency of catastrophes
Effect of catastrophes on survival and on reproduction
Number of populations
Which sexes disperse?
Which ages disperse?
Survival during dispersal
% of individuals dispersing
Effects of population density on survival, fecundity, dispersal, etc.

Table 8.3 Partial list of the parameters involved in specifying a VORTEX model

Lindenmayer et al. (2000) tested VORTEX by using it to predict the abundance of three species of arboreal marsupials across 39 remnant patches of native *Eucalyptus* forest in southeastern Australia, embedded in a landscape dominated by introduced pine forest. Life-history parameters were available for all three species, but little was known about their dispersal, so Lindenmayer et al. (2000) considered five different plausible scenarios:

1. No migration
2. Migration to all patches equally likely
3. Migration only to the nearest patch
4. Migration to each other patch proportional to patch size divided by distance (or distance²) from the patch of origin

Good agreement between observed and predicted abundances could be obtained, but only for one of the five dispersal scenarios (size/distance²), specific values of the overall migration rate, and the mortality rate during dispersal, and if differences in patch quality were taken into account (patch quality was based on how well the available vegetation matched the dietary preference of the species). Lindenmayer and Lacy (2002) similarly found, for two small mammals in the same habitat, that actual patterns of patch occupancy and abundance were obtained in only a fraction of the plausible scenarios. In contrast, Penn et al. (2000) found very good agreement between observed population trends in two koala populations, and the trends predicted by VORTEX using independent data to estimate parameters.

Brook et al. (1999) compared extinction forecasts for whooping crane from Vortex and five other commonly used programs. The whooping crane has been reduced to a single wild flock of roughly 150 individuals, and about 130 captive individuals (Brook et al. 1999). The whooping crane has been monitored since it was recognized as endangered in 1939, giving 57 years of demographic data that could be used to estimate model parameters. The models differed as to which processes were included, with the following included in only some of the packages: catastrophes, inbreeding depression, breeding structure (whooping crane is monogamous), correlation between survival and recruitment variation, sex ratio and mate availability, and the ability to specify an arbitrary probability distribution for litter size. Each model was used to generate 500 replicate projections for 50 years starting from a population of 18 individuals.

All models were parameterized from exactly the same information but they produced significantly different predictions. Predicted mean size after 50 years ranged from 66 to 129, and the risk of extinction from 1% to 7.6%. Moreover, there were significant differences between two versions of VORTEX. In version 5 males are first paired with females, and it is then determined whether or not the female will produce a litter that year. This version predicted a mean final

size of 35 individuals, and a 9.4% risk of extinction over 50 years. Starting from version 6 males are only paired with females who will produce a litter that year if they have a mate. This change doubled the mean final population size, and halved the predicted extinction risk. As a control for their comparisons Brook et al. (1999) developed a simplified crane life history that could be represented in all the programs. In these “standardized” simulations the models produced very similar results. Thus the different predictions were due to differences in basic assumptions rather than the simulation methods used in the programs.

Even VORTEX does not include everything. For example, infectious diseases play a significant role in regulating many animal populations (Hudson et al. 2002) and have been implicated in population declines of a number of threatened species including African wild dogs, Arctic foxes, mountain gorilla, and rainforest toads (Cleaveland et al. 2002). Diseases also appear to pose an increasing threat to marine species (e.g., Harvell et al. 1999; Ward and Lafferty 2004). Using an agent-based model Hess (1996) showed that increased opportunities for individual movement between remnant habitat patches—which reduces extinction risk in the absence of disease—could instead increase a small population’s risk of extinction in the presence of disease.

Individual-based modeling remains controversial in ecology. Proponents argue that individual-based models can be based on assumptions that are derived directly from observations of individual behavior and its consequences, and can incorporate all important variables (e.g., Huston et al. 1988; deAngelis et al. 1998). This is counterbalanced by the need for detailed information on all important variables and the potential for seemingly harmless assumptions to have large consequences, such as the mating rules in Vortex.

Exercise 8.2. Download Vortex (currently at vortex9.org) and use it to replicate the koala PVA in Penn et al. (2000) using the model parameters listed in Table 3 of that paper. Explore how sensitive the predicted extinction risk is to the assumed frequency of catastrophic years (1 year in 20, in the published results).

8.2 Artificial Life

Artificial life refers to the construction of artificial entities, mechanical or computational, that share some properties of biological life. Here we consider only one aspect: the use of agent-based simulations as models of long-term evolution.

For many basic questions the history of life on Earth constitutes a single unreplicated data point. As a result there is considerable controversy about the role of chance versus necessity—if the process were started again from scratch, would the second run look like the first (Gould 1990)? For example,

- Would life necessarily emerge from the primordial soup?

- The universality of the genetic code argues for a single origin. Of the code's many features—using DNA to store information, use of base-pair triplets as the units to specify amino acids, and so on—which if any are essential for life to proliferate and diversify, and which are just happenstance?
- Would multicellular life necessarily emerge? If so, would it necessarily diversify into a suite of taxa similar to what we now have and a similar spectrum of ecological roles (primary producers, herbivores, predators, parasites, and so on)?

Chance has almost certainly played a significant role, for example the importance ascribed to meteor impacts in mass extinction events. Proponents of chance point to such events and argue that the history of life would be very different if the process were started from scratch a second time. One purpose of artificial life is to provide additional instances of evolution under similar sets of rules. We can then see how widely outcomes vary when a set of rules is run multiple times. And by comparing models we can try to identify which properties of the underlying rules determine the consistent features of the “organisms” and “ecosystems” that evolve.

8.2.1 *Tierra*

Tierra (Ray 1991, 1992) is the best-known example of long-term evolution in digital organisms. Tierra simulates competition between bit segments that strive to self-replicate within a computer's memory. Allowing these “organisms” to compete for a limiting resource (CPU time) and to reproduce with random mutations yields a digital analog of organic evolution.

A Tierran organism is a sequence of 0/1 bits occupying a linear block of memory. Each five-bit segment of the organism corresponds to a meaningful instruction, a structure analogous to the actual genetic code. Tierra is seeded with a few individuals who replicate and spread through memory. “Mutations”—random copying errors—occur during replication, creating genetic diversity. There are also “somatic” mutations (random 0 ↔ 1 bit flips within existing organisms) that occur every time step. Once memory fills to some specified level, organisms begin to die and their space in memory is deallocated. Mutations conferring faster reproduction or lower mortality increase in frequency.

The limiting resource in Tierra is time. Organisms take turns getting a slice of the computer's CPU time allocated to the instructions in their “genome.” Organisms that use fewer instructions to replicate propagate faster, so there is selection for having a short genome without superfluous instructions. Mortality is age dependent: individuals are “born” at the bottom of a queue that advances toward the “reaper” over time. Individuals are also bumped up or down in the queue according to whether or not they successfully execute certain specific instructions.

In Ray's initial experiments (Ray 1992), Tierra was seeded with a single self-replicating "ancestor" organism of 80 instructions, which was designed to be a minimal self-replicating algorithm with no other functionality. The ancestor examines itself to determine where in memory it begins and ends, indicated by specific bit strings. It then calculates its size, allocates a memory block for an offspring, and uses a copy procedure embedded in its own code to replicate its genome into the offspring memory block. It then registers the daughter as a new organism, puts it at the bottom of the CPU time and "reaper" queues, and starts trying to breed again.

According to Ray (1992) the first ancestor was created as a test case for debugging Tierra, and was expected to just spread and replicate. Instead, it initiated a complex evolutionary arms race.

- *Parasites* evolved that lacked the copy procedure and therefore were shorter and faster-replicating. Instead, they located and used the copy procedure from other nearby individuals. Their analog in real life would be viruses that exploit the replication machinery of the cells that they infect. Parasites usually emerged within the first few million instructions in a run.
- Organisms with *resistance to parasitism* then emerge, and may exclude parasites for long periods of time.
- Parasites could also be excluded by *hyperparasites*. When the parasite tries to use them for self-replication, the hyperparasite resets some of the parasite's memory registers so that the parasite replicates the hyperparasite instead of itself. These drive the parasite to extinction and become dominant in the population.
- Such conditions of high genetic uniformity favor the evolution of sociality. *Social hyperparasites* emerge, which are short and rapidly breeding but can only reproduce by exploiting nearby organisms that are genetically similar to themselves (e.g., their "tail" and part of a neighbor's "head" are needed for reproduction).
- This situation was then exploited by *cheaters*, who intercept the CPU pointer identifying the block of code that is being executed as it passes from one hyperparasite to another, and redirect it to their own genome.

This evolutionary sequence required about 1 billion instructions. In the process the creatures spontaneously developed sexual reproduction. "Sexual" organisms were producing offspring that were partial copies of themselves and partial copies of others in the population—a blending akin to sexual reproduction.

Exercise 8.3. Based on the results from Tierra, we might predict that if multicellular life has evolved on other planets, it will suffer from viruses—diseases that replicate themselves using the host's cellular replication mechanism. Eventually Star Fleet will come back with the data. What might you do now, on planet Earth, to challenge or support this prediction?

8.2.2 *Microbes in Tierra*

Yedid and Bell (2001) used *Tierra* to model short-term adaptation in microbes, motivated by two hypotheses for the mode of evolutionary adaptation in microbial cultures. The dominant paradigm is “periodic selection”: most of the time the culture has a single dominant genotype, but occasionally a beneficial mutation of the dominant type arises and then replaces its ancestor. The genealogy of successive dominant genotypes is then predicted to be a direct line of descent. In a new culture representing a novel environment for the organism, many different mutations will be helpful and adaptation will be rapid. Once those have become fixed, the rate of beneficial mutations will dwindle and the rate of adaptation is expected to diminish.

Yedid and Bell (2001) argued that some experimental results were inconsistent with this hypothesis: the rate of gene substitutions was too high to occur in a genetically uniform population, and did not diminish over time. They hypothesized that the population remained genetically diverse, and that each new dominant could descend from any of the types present.

Tierra was used to test these hypotheses because it allowed them to monitor genealogies in full detail, which is not yet possible with real organisms. They seeded *Tierra* with a single minimal genotype, with parameters set so that organisms could only read, write, and execute instructions in their own genome. Experiments were run at three mutation rates: 0.002, 0.01, and 0.2 per genome per generation, which cover the range of natural spontaneous mutation rates. Relatively small populations were used, about 500 individuals, because of the need to record each genotype and its genealogical relations.

About 75% of the genotypes that arose during the experiments were stable replicators who bred true. The remainder consisted of inviable mutants (about 17%) and mutators, who consistently produced offspring unlike themselves. At low and intermediate mutation rates, they found that evolution followed the periodic selection hypothesis. With few exceptions each dominant was a direct descendant of the previous dominant, and most dominants achieved frequencies of 80% or higher in the population (Figure 8.2). But at high mutation rates (0.2/genome/generation) dominants generally “emerge from the fog of rare genotypes that collectively constitute a large fraction of the population” (Yedid and Bell 2001, p. 479).

Yedid and Bell (2001) interpreted these results as indicating that different patterns of adaptation may occur, depending on mutation rate. At lower mutation rates such as those believed to hold in microbes, the classical periodic selection hypothesis holds, with occasional selective sweeps as a new type evolves and replaces the current dominant. At higher genomic mutation rates, perhaps typical of multicellular eukaryotes, there are periods of dominance by a single dominant, but also long periods when multiple quasineutral types coexist, any of which could be the ancestor of the next dominant.

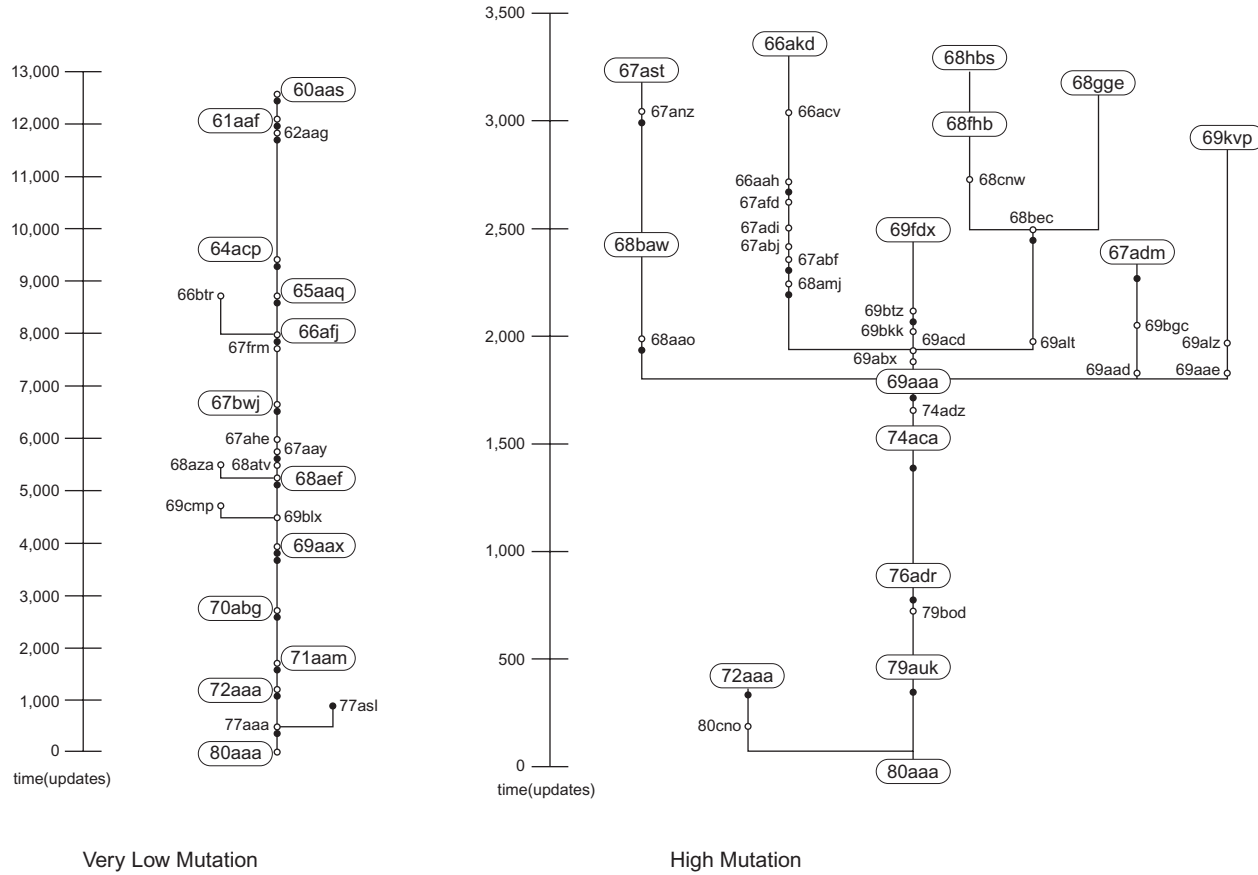


Figure 8.2 Genealogies of successive dominant genotypes under low and high mutation rates (from Yedid and Bell 2001). More common genotypes are coded by length and sequence of appearance.

8.2.3 *Avida*

The *Avida* program (Adami 2002; Ofria and Wilke 2004) is a modification of *Tierra* with two-dimensional space; *Avidans* “live” on a two-dimensional grid of cells with each individual occupying a single cell. Individuals interact through competition for space, which occurs at birth: when a parent reproduces, one of its neighbors is removed and replaced by the newborn. Reproduction is limited by CPU time as in *Tierra*, creating selection for short genomes. Individuals can also be given extra CPU time when they execute certain specific instructions, which allows selection for traits other than genome length. *Avida* uses a simpler genome than *Tierra*, consisting of a list of instructions such as “add,” “copy,” and “search in the forward direction.” There are about 50 potential instructions, and the user specifies which of these are available in a run of the program.

Avida has been used to explore a number of issues in evolutionary biology, including effects of mutation rate (Wilke et al. 2001) and the processes whereby stepwise evolutionary change can produce complex adaptations (Lenski et al. 2003); see Wilke and Adami (2002) and O’Neill (2003) for reviews.

Chow et al. (2004) conducted *Avida* experiments on the relationship between productivity and species diversity in ecosystems. Species diversity is often (but not universally) observed to peak at intermediate levels of productivity, and a number of possible factors have been invoked to explain this pattern, including spatial heterogeneity and effects of predators. Chow et al. (2004) used *Avida* to test whether these factors are necessary to explain the observed pattern. They simulated a spatially homogeneous ecosystem without predators, in which individuals compete for a number of different “resources.” Each resource was a CPU time reward for executing certain logical operations—an analog of obtaining food by carrying out a successful hunt, with different hunting behaviors required to obtain different types of food. Competition was modeled by having a finite supply rate of the reward for each operation, that was partially consumed each time an organism executed that operation. When a resource type becomes scarce because most of it has been consumed, the reward for the corresponding instruction goes down—“hunting” for that food item is less rewarding per unit of time.

In each experiment a grid of 300 sites was initiated with a founding genotype that could replicate but did not execute any of the rewarded operations, and evolution was allowed to proceed. Twenty-five replicates were run at each of seven values of the resource supply rate, which was interpreted as a proxy for productivity since it determined the rate at which individuals could reproduce. The results showed a clear pattern of maximum species diversity at intermediate productivity (Figure 8.3). Similar results were obtained with a generalist founder that could perform all rewarded operations. These results show that the observed qualitative relationship between diversity and productivity could result simply from evolutionary diversification in resource usage: the only requirement is a pool of different resources that organisms can exploit.

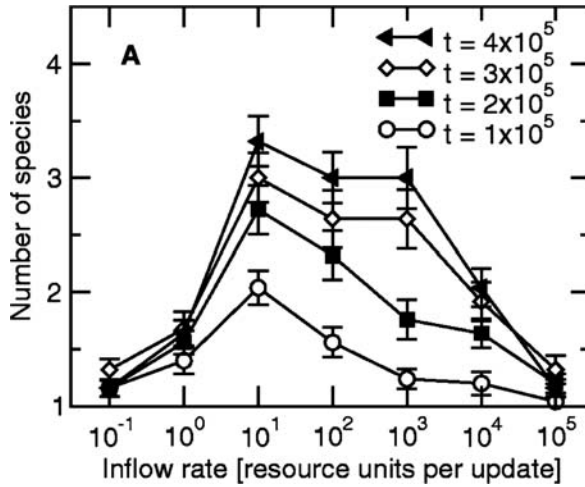


Figure 8.3 Relationship between resource supply rate and the number of distinct species descended from the ancestor, at different times in the simulation (time t is measured in “updates”, corresponding roughly to the CPU time required for 30 instructions to be executed). From Chow et al. (2004)

The relevance of experiments in digital evolution depends on the analogy between digital and organic evolution. In some ways they are similar—being driven by mutation, selection, drift, and limited resources—but in other respects they can be quite different. Evolution in *Tierra* is driven by the benefits of a short genome. Yedid and Bell (2001) found that genome length most often changed via mutants that miscalculate their genome size and produce offspring shorter than themselves. Sometimes those offspring were still able to replicate and therefore had higher fitness than the ancestor. Consequently, adaptation in *Tierra* was dominated by processes very different from those typical of actual organisms.

Exercise 8.4. The online Supplementary Information for Waide et al. (1999) provides summary information on published studies of the relationship between productivity and species diversity. Considering only those studies showing a unimodal relationship between productivity and diversity, evaluate (and explain how you evaluated) the *quantitative* correspondence between Figure 8.3 and empirical studies. How should your findings affect our interpretation of the results in Chow et al. (2004)? What additional simulation experiments do your findings suggest?

Exercise 8.5. Read up on current theories for the origin of sexual reproduction, and conduct some experiments in *Tierra* or *Avida* to test one of them. Publish your results in *Science* or *Nature*, and send us a copy. If your paper is rejected, read up on current theories for the origin of cooperation and try again.

8.3 The Immune System and the Flu

The National Institutes of Health recommends yearly influenza vaccination, especially for certain groups. The influenza virus evolves rapidly, so each year a new vaccine is produced to provide protection against the strains predicted to be prevalent in the next flu season. Smith et al. (1999) studied how the efficacy of a vaccine depends on previous vaccinations. Historical studies of this question reached different conclusions as to whether repeated vaccination was effective, so Smith et al. (1999) developed an agent-based model to explore the efficacy of repeated vaccination.

Each individual has a vast number of B-lymphocytes, cells that produce antibodies that bind to viral and bacterial antigens and destroy them. Some antibodies are incorporated as receptors into the membranes of B-cells, and binding of these to antigen stimulates the proliferation of a clone of immune cells. Some of these progeny cells secrete antibodies while others are long-lived memory cells that allow for a faster and more vigorous response to similar antigens in the future. Vaccines stimulate the immune system to produce antibodies to antigens from disease organisms. Their effectiveness depends upon an individual's history of exposure to similar antigens.

Exposure to one strain of flu can confer partial or complete immunity to closely related strains. Also, vaccination against one strain affects the response to subsequent vaccination against similar strains. If an individual is vaccinated in two successive years with similar strains, the second vaccination produces a smaller immune response because the immune system already has produced circulating antibodies that bind to the antigens of the second vaccine. This may decrease the effectiveness of the second vaccination, especially against strains of flu that are more similar to the specific target of the second vaccine than the first.

The model of Smith et al. (1999) represented the specificity of each B-cell receptor, antibody, and antigen as a 20-letter word with a 4-letter alphabet. Thus, their model allows for 4^{20} different antibodies and antigens. They use the *Hamming distance* to define the specificity of antibody and antigen: the interactions between an antibody and an antigen are a function of the number of letters that are the same in the two. For example the words

ABCDABCDBCDADDCABAB

ABCDABCDBCDADDCABBAB

have distance 2 because the letters in positions 16 and 17 differ. Cells and antigens whose distance is 7 or smaller were allowed to interact, the strength of the reaction decreasing with distance.

Consider a person who was vaccinated against strain 1 last year and strain 2 this year, but has just been infected with strain 3—because the circulating strain is not exactly the one predicted when the vaccine was produced. Is their immune response to strain 3 stronger or weaker as a result of the previous year's vaccination? Assume the distance from strain 1 to strain 2 is 2, and that the distance from strain 3 to strain 1 is 3. The distance from strain 3 to strain 2 can be as small as 1 or as large as 5. The immune system produced lots of antibody to strain 1 last year, some of which is still present in the body, and smaller amounts of antibody to strain 2 this year. If strain 3 is closer to strain 2 than to strain 1, then there may be a weaker response than if no vaccination occurred last year. On the other hand, if strain 3 is closer to strain 1, then there may be a stronger response than would have occurred with a single vaccination this year to strain 2.

The model studied the interactions between repeated vaccinations by simulating a collection of 10^7 B-cell clones. Rules were established for the proliferation and death of B-cells when challenged with antigen at different antigenic distances. Rules were also established for the proliferation of viruses in the system that are bound and unbound by antibody. Simulations were performed that corresponded to no vaccination, vaccination “last year” only, vaccination “this year” only, and vaccination in both years. In each case, the system was challenged with attack by a flu virus two months after the time of vaccination this year. Simulations were run for 450 days to observe the response of the system over two subsequent flu seasons.

Simulations of the model clearly showed both positive and negative interference effects of repeated vaccination. With repeated vaccination, the response of the system depended upon the antigenic distance of the virus to both the first and second vaccines. In all simulations, vaccination in the current year conferred greater immunity than no vaccination. However, in some cases there was greater immunity this year if no vaccination occurred last year.

The study implies a definite policy recommendation for the production of flu vaccines. Each year's vaccine targets strains that are predicted to be prevalent, but there is a range of uncertainty as to which strain will predominate. Within the range of the predicted strains, the authors recommend that vaccine be produced to those antigenically farthest from the vaccine of the previous year. This will produce the largest response of individuals who were vaccinated last year to vaccination this year.

8.4 What Can We Learn from Agent-Based Models?

The studies reviewed above illustrate that the things we can learn from complex, agent-based models are no different from what we can learn from any other kind of theoretical dynamic model. The models provide the link between process and

pattern: how mechanistic assumptions about key variables and processes lead to predictions of observable phenomena. Yedid and Bell (2001) predicted how mutation rate affects the dynamics of adaptation; Chow et al. (2004) showed that evolutionary diversification in ecological niches could produce a hump-shaped relationship between productivity and diversity; Enquist and Niklas (2001) found that allometric resource allocation rules could generate observed size-frequency scalings in forests. In principle these are no different from (for example) our simple models for gene regulation networks, in which certain specific types of interactions were shown to induce certain kinds of temporal patterns in the state variables.

However, the unavoidable complexity of agent-based models can make it hard to know *why* the observed outcomes occurred, as critics of simulation models often complain. The Enquist and Niklas (2001) model reproduces the observed size-frequency scaling. But is that really due to their agent-level allocation rules, as they claim, or is it a consequence of competition for space and light that would hold for other allocation rules? If so, what is the set of allocation rules that would produce results consistent with the observed size-frequency scaling? Given the number of assumptions in a realistic agent-level model, it may not be feasible to just vary each assumption and see what happens. Four options for each of five assumptions yields 1024 alternative models, each needing to be explored through multiple simulations across a range of parameter values—and it is a rare agent-based model that only has five assumptions.

In the rest of this chapter, we consider two approaches for addressing this issue: structured computational methods for understanding the predictions of a complex model, and the formulation of simpler dynamic models where the link between assumptions and outcomes is more easily determined. Used in this way, analytic and computational models can be mutually reinforcing approaches to understanding a complex system, rather than polar alternatives.

In a sense these two approaches are analogous to doing a bifurcation analysis on a classical dynamical systems model with many parameters (Chapter 5). Bifurcation analysis identifies the special locations in parameter space where qualitative changes occur. The overall behavior of the model is then sketched out by locating the bifurcation curves or surfaces and identifying the qualitative changes that occur at each of them. For computational models we similarly want to start by figuring out where the action is happening, and then focus our attention there.

8.5 Sensitivity Analysis

The goal of sensitivity analysis is to evade the curse of dimensionality by identifying a subset of parameters and assumptions that most strongly affect a model's behavior, at much lower computational cost than a full exploration of parameter

space. Ranking parameters by their impact on model predictions is also important information for planning real-world experiments, because it identifies which additional information will be most beneficial for narrowing down the range of model predictions.

The benchmark for computational cost is the exponential dependence on the number of parameters in a brute force approach: to check all possible combinations of k values for d parameters requires

$$k^d = e^{d \log(k)} \quad [8.1]$$

model runs (if a model is stochastic, then some quantities have to be estimated by doing repeated simulations at the same parameter values—for simplicity we will also refer to this as a “model run”).

The simplest form of sensitivity analysis is a computational version of the local sensitivity analysis for eigenvalues of matrix models (Chapter 2). “Local” means that we only consider small perturbations from a reference set of parameter values. Let Y denote a quantity of interest calculated from the output of a model run, and $Y(p_i)$ its value as a function of the i th parameter p_i , with all other parameters held fixed at their reference value. The sensitivity of output Y to parameter p_i is defined as the percentage change in Y relative to the percentage change in p_i , for a small change in p_i . Typically 5% or 10% changes are used with a centered difference estimate of the response. With a 5% change, the fractional change in Y is $(Y(1.05p_i) - Y(0.95p_i))/Y(p_i)$ and the fractional change in p_i is 0.1, so the sensitivity is computed as

$$s_i = \frac{Y(1.05p_i) - Y(0.95p_i)}{0.1 \times Y(p_i)}. \quad [8.2]$$

Note that [8.2] is what we called the elasticity in Chapter 2. Ecologists and economists use the term elasticity for [8.2] but in the simulation modeling literature it is called the sensitivity. You might as well get used to this—it’s not going to change any time soon.

Local sensitivity analysis is computationally cheap, requiring $2d$ additional runs for a model with d parameters. Even with $k = 2$ (the absolute minimum) the computational cost relative to [8.1] is reduced over 50-fold with 10 parameters, and more than 25000-fold with 20 parameters. However it has some significant limitations:

1. It characterizes the model’s response only near the reference parameter set. This has limited relevance in the typical situation where many or most parameter estimates have low precision.
2. It takes no account of interactions between parameters. If a small increase in parameter 1 has no direct effect on Y , but it makes Y much more sensitive to changes in parameter 2, local sensitivity analysis will incorrectly say that parameter 1 is unimportant.

A *global sensitivity analysis* removes these limitations by making simultaneous large changes in multiple parameters. The goal is to identify the most important parameters by doing model runs at a relatively small but well-chosen set of parameter combinations. The model is viewed as a “black box” that transforms a probability distribution of inputs—parameters, assumptions, and initial conditions—into a probability distribution of model outputs. The results from model runs are treated as experimental data and analyzed statistically to determine which inputs have the most effect on the output distribution. The results of this analysis are meaningful so long as the input-output relationship satisfies the assumptions of the statistical procedures used to analyze the “data.”

Three main types of statistical procedures have been used:

1. *Response surface* methods posit that the relationship between parameters and model outputs Y can be approximated by an assumed functional form that is simple enough to be estimated well from a small number of model runs. Response surface methods are simple and easy to implement using a statistics package. However their reliability depends on how well the assumed functional form approximates the input-output relationship, which is difficult to assess.
2. *Correlation* methods estimate the strength and direction of the statistical association between each parameter and Y . Correlation coefficients can be misleading when the parameter-output relationship is nonmonotonic or strongly nonlinear. For example, the correlation between p and $Y(p) = 1 - p^2$ is exactly 0 if p is chosen at random from the interval $[-1, 1]$. However, these properties are easy to examine by plotting Y as a function of each parameter.
3. *Variance decomposition* methods decompose the variance in Y across model runs into contributions from each parameter, alone and in combination with other parameters. These methods are applicable in principle to any parameter-output relationship with finite variance of the output.

Selection of parameter combinations for a global sensitivity analysis is usually based on a set of *input distributions* $f_i(p_i)$ for each parameter p_i . Each f_i is a probability distribution representing the range and relative likelihood of different values for the i th parameter. If parameters were estimated from data and interest focuses on parameters representing the experimental setting, a reasonable choice of f_i might be Gaussian distributions corresponding to confidence intervals on each parameter (see Chapter 9)—if $[a, b]$ is the 95% confidence interval, the corresponding Gaussian parameters are $\mu = (a + b)/2$, $\sigma = (b - a)/3.92$. But in general f_i can represent any parameter range over which you want to characterize the behavior of the model. If the model output of interest depends on the model’s initial conditions, the initial value of each state variable can be treated as an additional parameter, or the analysis can be repeated for different sets of initial conditions.

8.5.1 Correlation Methods

Latin hypercube sampling (LHS, McKay et al. 1979) is way of sampling parameter space that for many purposes gives much more precise results than a random sample of the same size [Blower and Dowlatabadi (1994) review comparisons that have been made between LHS and other sampling schemes]. For LHS with sample size k , the range for each parameter is divided into k intervals having equal probability under the input distribution f_i . A value in each interval is selected, possibly at random. This assures that the full range of each parameter is represented in the sample. The outcomes from this process are assembled into a matrix \mathbf{P} whose (i, j) th entry is the j th selected value of parameter i . For example, suppose there are two parameters a and b , and the $k = 3$ selected values are $a = \{1, 2, 3\}$, $b = \{10, 20, 30\}$. Then

$$\mathbf{P} = \begin{bmatrix} 1 & 2 & 3 \\ 10 & 20 & 30 \end{bmatrix}. \tag{8.3}$$

Each column of \mathbf{P} is then a complete parameter vector for the model. These vectors are far from random: reading across each row the values are monotonically increasing. To destroy this correlation, each row of \mathbf{P} is shuffled into random order. The columns of the resulting matrix $\tilde{\mathbf{P}}$ are then a set of k parameter vectors for the model, such that each selected value of each parameter occurs in exactly one of the parameter vectors, and there are no systematic cross-correlations between parameters. Two possible outcomes of row-shuffling [8.3] are

$$\tilde{\mathbf{P}}_1 = \begin{bmatrix} 3 & 1 & 2 \\ 20 & 30 & 10 \end{bmatrix}, \quad \tilde{\mathbf{P}}_2 = \begin{bmatrix} 2 & 3 & 1 \\ 30 & 10 & 20 \end{bmatrix}. \tag{8.4}$$

Matrix $\tilde{\mathbf{P}}_1$ says to do model runs with parameter vectors $(a, b) = (3, 20), (1, 30)$, and $(2, 10)$; $\tilde{\mathbf{P}}_2$ says to do runs with $(a, b) = (2, 30), (3, 10)$, and $(1, 20)$.

The model is run at all parameter vectors in one or more $\tilde{\mathbf{P}}$ matrices. Statistical measures of correlation are then used to summarize how strongly each parameter affects Y . Possible measures include the following:

1. The linear correlation coefficients between $p_{i,j}$ and Y_j , where $p_{i,j}$ is the value of p_i and Y_j is the value of Y for the j th model run. This measures the overall linear association between each parameter and Y .
2. The standardized linear regression coefficients. These measure the strength of each parameter-output association independent of the linear effects of other parameters. Their values can be obtained by doing multiple linear regression of Y on all parameters, with all variables scaled relative to their standard deviations.
3. Rank correlation coefficients. These can be used instead of linear correlations. The parameter values $p_{i,j}$ are replaced by their ranks $r_{i,j}$ across the parameter samples—the smallest value of $p_{1,j}$ is replaced by 1, the second smallest is replaced by 2, and so on. The same is done to the Y values, and then the linear correlation

Parameter	Optimistic			Pessimistic		
	Year 1	Year 5	Year 10	Year 1	Year 5	Year 10
Fraction of drug-sensitive cases treated	0.99	0.99	0.99	0.97	0.97	0.95
Transmission coefficient of drug-sensitive treated infection	-0.18	-0.65	-0.84	-0.10	-0.43	-0.60
Average survival time of drug-sensitive treated individuals	0.86	0.85	0.78	0.52	0.35	0.26
Increase in risky behavior				-0.86	-0.89	-0.90

Table 8.4 Sensitivity coefficients for key parameters in the Blower et al. (2000) model for antiretroviral treatment of HIV/AIDS, taken from Table 1 of that paper. The model output variable was the number of AIDS deaths averted by treatment, and the tabulated values are partial rank correlation coefficients between parameters and output where the correlation was above 0.5 in magnitude. “Optimistic” and “pessimistic” scenarios differ in their assumptions about the fraction of treated infections where resistance develops and the possible rate at which risky behaviors increase.

coefficient between rank values is computed. This statistic is called Spearman’s ρ , and is available in most statistics package without the user having to compute the ranks.

To see why standardized regression coefficients are recommended, consider the “model output” $Y = p_1 + 2p_2$ with input distributions $f_1 =$ uniform on $[-1, 1]$ and $f_2 =$ uniform on $[-2, 2]$. Because p_2 has twice the impact per unit change and twice the range of variation, the answer we want for global sensitivity is that p_2 is four times as important as p_1 . The linear regression coefficients are 1 and 2, but standardizing gives regression coefficients 1 and 4 (why?). The linear correlation coefficient includes this standardization automatically. Opinion is mixed about the merits of rank correlations. Some authors consider rank correlations more appropriate if parameter-output relationships are monotonic but strongly nonlinear. Others feel that transformation to ranks is too strong a distortion for the correlation coefficients to be meaningful.

Blower et al. (2000) used LHS with partial rank correlation coefficients to identify key parameters affecting the effectiveness of antiretroviral therapy for HIV/AIDS, using the model discussed in Chapter 6. The model has a total of 24 parameters, but four of these were identified as having particularly strong impacts on the numbers of AIDS deaths averted (Table 8.4). Note that the relative importance of parameters depends on the time horizon being considered, and also on whether the assumed range of possible parameter variation was “optimistic” or “pessimistic.” The results from a global sensitivity analysis always depend on the input distributions that define what “global” means.

Exercise 8.6. Consider the “model output” $Y(p_1, p_2, p_3) = e^{p_1+2p_2+3p_3}$. Write a script that uses LHS in combination with linear correlation and also with rank correlation to do a global sensitivity analysis with the input distributions f_i all being uniform distributions on $[-1, 1]$. Parameter sampling with these f_i is relatively easy, because equiprobable intervals are equal in length. For example with $k = 5$ the 6 interval endpoints are $\{-1, -0.6, -0.2, 0.2, 0.6, 1.0\}$ and the selected values can be the mid-points $p_i = \{-0.8, -0.4, 0, 0.4, 0.8\}$. Write your script so that a user can choose the value of k by editing just the first line of the script.

8.5.2 Variance Decomposition

Variance decomposition methods for global sensitivity analysis are based on the *analysis of variance* decomposition from theoretical statistics. The model output of interest can be decomposed as

$$Y(p_1, p_2, \dots, p_d) = Y_0 + \sum_{i=1}^d Y_i(p_i) + \sum_{1 \leq i < j \leq d} Y_{i,j}(p_i, p_j) + \dots + Y_{1,2,\dots,d}(p_1, p_2, \dots, p_d). \quad [8.5]$$

In statistical terminology, [8.5] decomposes Y into the “main” effect of each parameter, and interaction effects of higher and higher orders. If parameter values are chosen independently, this decomposition is unique under the constraints that the functions on the right-hand side of [8.5] all have expected value zero and are mutually uncorrelated. Let $\sigma_i^2, \sigma_{i,j}^2$, and so on denote the variances of the terms in [8.5]. Because the terms are uncorrelated, [8.5] implies a corresponding decomposition of the total variance in Y ,

$$\sigma^2 = \sum_{i=1}^n \sigma_i^2 + \sum_{1 \leq i < j \leq d} \sigma_{i,j}^2 + \dots + \sigma_{1,2,\dots,k}^2 \quad [8.6]$$

where σ^2 is the variance of $Y(p_1, p_2, \dots, p_d)$. Sobol’ (1993) proposed sensitivity indices defined by scaling [8.6] relative to the total variance:

$$1 = \sum_{i=1}^n M_i + \sum_{1 \leq i < j \leq d} C_{i,j} + \dots + C_{1,2,\dots,k}. \quad [8.7]$$

The “main” sensitivity index M_i is the fraction of the total variance in Y accounted for directly by parameter i . The individual C ’s are not particularly significant, but they can be combined into a measure of each parameter’s total contribution to the variance,

$$T_i = M_i + \sum_{j=1}^d C_{i,j} + \sum_{1 \leq j < k \leq d} C_{i,j,k} + \dots + C_{1,2,\dots,k}. \quad [8.8]$$

This “total” sensitivity index for parameter i is the sum of all terms in [8.7] that involve parameter i . Fortunately T_i can be estimated without estimating each individual term in [8.8]. For any i , we can split the parameter vector into p_i and

$p_{\sim i}$, the latter being a vector of all parameters besides p_i . Consider the quantity

$$Y(p_i, p_{\sim i}) - Y(p_i, p'_{\sim i}) \tag{8.9}$$

where p_i is a random draw from the input distribution on parameter i , and $p_{\sim i}, p'_{\sim i}$ are two independent random draws of all other parameters from their input distributions. Because both terms in [8.9] use the same value of p_i , their difference lacks the $Y_i(p_i)$ term in the decomposition [8.5], but there are two sets of all other terms. As a result,

$$\frac{1}{2}E[Y(p_i, p_{\sim i}) - Y(p_i, p'_{\sim i})]^2 = \sigma^2 - \sigma_i^2$$

and therefore

$$M_i = 1 - \frac{1}{2\sigma^2}E[Y(p_i, p_{\sim i}) - Y(p_i, p'_{\sim i})]^2 \tag{8.10}$$

(Chan et al. 2000), where E denotes the expected value with respect to the parameter input distributions. Similarly in $Y(p_i, p_{\sim i}) - Y(p'_i, p_{\sim i})$ the main effects of all parameters other than p_i cancel out while any term including p_i appears twice, so

$$T_i = \frac{1}{2\sigma^2}E[Y(p_i, p_{\sim i}) - Y(p'_i, p_{\sim i})]^2 \tag{8.11}$$

where p_i, p'_i are two independent random draws from the input distribution on parameter i (Chan et al. 2000).

Using these formulas the complete set of main and total sensitivity indices can be estimating efficiently using “winding stairs” parameter sampling (Jansen et al. 1994). The winding stairs method iteratively generates a matrix of parameter vectors and corresponding Y values with d columns. Let $p_{i,j}$ denote the j th sample of parameter i , drawn at random from the input distribution f_i . With $d = 3$ parameters in the model, the form of the matrix is

$$W = \begin{bmatrix} Y(p_{1,1}, p_{2,1}, p_{3,1}) & Y(p_{1,1}, p_{2,2}, p_{3,1}) & Y(p_{1,1}, p_{2,2}, p_{3,2}) \\ Y(p_{1,2}, p_{2,2}, p_{3,2}) & Y(p_{1,2}, p_{2,3}, p_{3,2}) & Y(p_{1,2}, p_{2,3}, p_{3,3}) \\ Y(p_{1,3}, p_{2,3}, p_{3,3}) & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}. \tag{8.12}$$

Construction of W starts at the top left with the first sampled value of each parameter. Then proceeding to the right in the top row, the second column “updates” the value of parameter 2, and the third column “updates” the value of the parameter 3, in each case leaving the previous values of the other parameters unchanged. Then dropping down to the next row, the entry in the first column updates the value of parameter 1. This process continues until r rows have been filled in.

The quantities [8.10] and [8.11] can all be estimated from the output matrix W . For example, to estimate M_1 we need to average squared differences of Y values

Parameter group	M_i	T_i
Parr per spawner	0.38	0.39
Parr to smolt survival	0.01	0.09
Smolt to adult survival	0.52	0.54
Proportional spawning at age 4	0.01	0.001

Table 8.5 Main and total Sobol’ sensitivity indices for stage-specific parameter groups in a stochastic age-structured model for salmon populations, using parameter input distributions corresponding to the uncertainty in parameter estimates (Ellner and Fieberg 2003). The response variable Y is the predicted long-term population growth rate, estimated from 5000 model runs at each parameter set. The exact values of the indices must satisfy $M_i < T_i$, but this may not hold for estimated values that are computed from a finite sample of parameters. From Ellner and Fieberg (2003).

where only the value of p_1 is the same for both runs. The values in the first and last columns of W fit that description, so the estimate of M_1 is

$$\hat{M}_1 = 1 - \frac{1}{2r\hat{\sigma}^2} \sum_{j=1}^r (W_{j,1} - W_{j,d})^2.$$

For T_1 we need model runs where p_1 has changed but all other parameters are the same. Pairs like that occur in the construction of W every time one row is finished and the next row is started. The estimate is therefore

$$\hat{T}_1 = \frac{1}{2(r-1)\hat{\sigma}^2} \sum_{j=2}^r (W_{j,1} - W_{j-1,d})^2.$$

Similar choices of entries in W can be used to estimate the sensitivity indices for all parameters. The Y value pairs needed for estimating T_i are adjacent in the sequence used to compute W (only one parameter changes), and those for computing M_i are $d - 1$ steps apart (all parameters but one have changed).

Table 8.5 shows the main and total sensitivity indices for stage-specific parameter groups in a stage-structured model for salmon populations (Ellner and Fieberg 2003). The model response variable Y is the long-term population growth rate. The purpose of the sensitivity analysis was to identify which parameters contributed most to uncertainty in the predicted growth rate, so the input distributions reflected the ranges of uncertainty in parameter estimates. Because the parameters for each particular life stage were estimated from the same data, the parameter uncertainty distributions have within-stage correlations whereas the methods for computing sensitivity indices assume independence. However the parameters p_i can be vectors rather than single numbers, so we computed the sensitivity indices to the four independent groups of parameters characterizing different life stages. The results were striking: two of the four life stages account for virtually all of the uncertainty.

Winding stairs sampling makes the Sobol’ sensitivity indices a general and easy to implement method for global sensitivity analysis. The principal limitation is the requirement that parameters be generated independently. For dimension reduction—finding parameters that have limited effect on the model—that requirement can be met by using independent random draws at each step of constructing the W matrix. However, for quantifying the impact of parameter uncertainty, correlated distributions of parameter uncertainty are inevitable unless each parameter comes from a separate data set.

Exercise 8.7. In Table 8.5 three of the life stages had $M_i \approx T_i$. What does that tell us about the impact of the corresponding parameter groups on population growth? For the fourth life stage (parr to smolt) T_i is larger than M_i by nearly an order of magnitude—what does that tell us?

Exercise 8.8. Consider again the “model output” $Y(p_1, p_2, p_3) = e^{p_1+2p_2+3p_3}$. Write a script that uses winding stairs to compute the main and total sensitivity indices for each parameter, with the input distributions f_i for all parameters being Gaussian with mean 0 and standard deviation 1. Write your script with so that a user can set the value of r (the number of rows in the W matrix) by editing only the first line. Report and interpret your results for $r = 1000$.

Exercise 8.9. Derive [8.10] and [8.11] using the analysis of variance decomposition [8.5] and the property that

$$\int Y_{i,\dots,j,\dots,k}(p_i, \dots, p_j, \dots, p_k) f_j(p_j) dp_j = 0$$

for any fixed values of the variables other than p_j (in words: the expectation of each $Y_{i,\dots,j,\dots,k}$ conditional on the values of any strict subset of its arguments is 0).

8.6 Simplifying Computational Models

A second way of trying to understand a complicated model is to try to find a simpler, approximating model that captures its essential features. This is sometimes done by trial and error, or on the basis of sensitivity analysis. However there are also some approaches that derive approximations directly from the model’s dynamic equations.

8.6.1 Separation of Time Scales

Biological systems often include processes that happen on very different time scales. When this occurs, a model for the system can sometimes be simplified by assuming that the fast processes happen infinitely fast relative to the slow processes. We have seen examples of this in Chapters 1 (enzyme kinetics), 3

(Morris-Lecar model), and 4 (gene regulation). Recall the equations for mRNA m and repressor protein p in the repressilator model:

$$\begin{aligned} \dot{m}_i &= -m_i + \frac{\alpha}{1 + p_i^n} + \alpha_0 \\ \dot{p}_i &= -\beta(p_i - m_i). \end{aligned} \tag{8.13}$$

If β is very large, then p_i changes more rapidly than m_i , and we can imagine that the second equation in [8.13] goes all the way to its asymptotic state while m_i remains at its current value—implying that $p_i \approx m_i$ always holds. The model can therefore be reduced to the equations for mRNA:

$$\dot{m}_i = -m_i + \frac{\alpha}{1 + m_i^n} + \alpha_0. \tag{8.14}$$

When β is very large, the repressilator system is an example of a *fast-slow* or *singularly perturbed* vector field. Abstractly, we regard the system as having the form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{y}) \\ \epsilon \dot{\mathbf{y}} &= \mathbf{g}(\mathbf{x}, \mathbf{y}), \end{aligned} \tag{8.15}$$

where we have split the state vector into the “slow variables” \mathbf{x} and the “fast variables” \mathbf{y} . If ϵ is small, the fast variables may approach an equilibrium with $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ before \mathbf{x} changes appreciably. The *singular limit* of the system [8.15] is then a system of *differential algebraic equations*

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{y}) \\ 0 &= \mathbf{g}(\mathbf{x}, \mathbf{y}). \end{aligned} \tag{8.16}$$

If we can solve the equations $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ in the form $\mathbf{y} = \mathbf{h}(\mathbf{x})$, then we can reduce the system to $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{h}(\mathbf{x}))$. This is what we have done in each of the examples cited above. The following *quasi-steady-state hypotheses* must hold in order for this reduction to be possible in this explicit fashion:

1. There must be a clear separation of time scales for the variation of \mathbf{x} and \mathbf{y} .
2. There must be a function $\mathbf{y} = \mathbf{h}(\mathbf{x})$ that we can compute so that equations $\mathbf{g}(\mathbf{x}, \mathbf{h}(\mathbf{x})) = 0$ for all \mathbf{x} in the region of interest.
3. For each fixed \mathbf{x} in the region of interest, the system $\dot{\mathbf{y}} = \mathbf{g}(\mathbf{x}, \mathbf{y})$ must have a unique stable equilibrium at $\mathbf{y} = \mathbf{h}(\mathbf{x})$.

When the fast equations do not have a unique stable equilibrium, a fast-slow system can have solutions that alternate between two different kinds of behavior:

- Periods when it behaves like the reduced slow system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{h}_i(\mathbf{x})), \quad \mathbf{y} = \mathbf{h}_i(\mathbf{x})$$

where $\mathbf{h}_i(\mathbf{x})$ is one of the stable solutions of $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$.

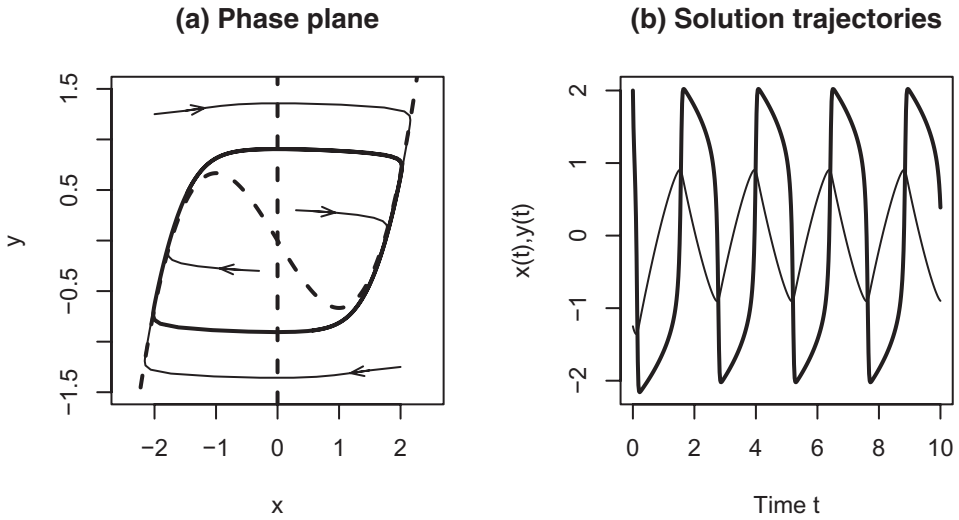


Figure 8.4 Relaxation oscillations in the van Der Pol model [8.17] with $C = 20$. In the phase portrait (a) the nullclines are drawn as dashed lines, and four solution trajectories as solid lines. All trajectories converge very quickly onto the limit cycle, drawn in bold. Panel (b) shows the state variables $x(t)$ (bold) and $y(t)$.

- Periods when the fast variables rapidly jump from one stable solution to another.

A simple example is given by the van der Pol system. Proposed as a model for heartbeat, the model can be rescaled into the form

$$\begin{aligned} \dot{x} &= C(y - f(x)) & \text{where } f(x) &= -x + x^3/3 \\ \dot{y} &= -x. \end{aligned} \tag{8.17}$$

The y -nullcline is the line $x = 0$. The x -nullcline is the curve $y = f(x) = -x + x^3/3$. The only equilibrium is their intersection at $(0, 0)$, which is locally unstable for any $C > 0$. When C is large this is a fast-slow system. Starting from any point above the x nullcline [i.e., $y > f(x)$], $x(t)$ increases rapidly until the right branch of the nullcline $y = f(x)$ is reached (see Figure 8.4). The solution then remains on the nullcline, moving according to the slow equation. Eventually the solution “falls off” the bottom of the nullcline. At that point $y < f(x)$ so $x(t)$ decreases rapidly until the left branch of the nullcline is reached—and so on, leading to a strongly stable limit cycle.

The periodic orbits computed in Figure 8.4 are called *relaxation oscillations*. As $C \rightarrow \infty$ these orbits come closer and closer to closed curves formed from segments parallel to the x -axis that join the local minimum and maximum of the x -nullcline to the opposite branch of the nullcline, and the parts of the nullcline that connect these segments.

Exactly the same behavior can occur in the Morris-Lecar model as studied in Chapter 5. When the parameter ϕ is small, the typical rate of change of w is slow compared to the rate of change of v . So we would like to solve for v as a function of w and reduce the system to an equation just for w . But this can't be done because the v -nullcline is not monotonic. As in the van der Pol system it has a local maximum and minimum that separate the nullcline into three branches. The left and right branches are stable equilibria for the fast \dot{v} equation and the middle branch is unstable. If we select parameters so that the only intersection of the nullclines occurs on the middle branch, we obtain the same type of oscillations as in the van der Pol system. Trajectories follow the left and right branches of the nullcline until they reach the local minimum or maximum, and then rapidly jump to the opposite branch.

Exercise 8.10. Plot nullclines and trajectories of the Morris-Lecar model in the (v, w) phase plane using the first parameter set of Table 5.1 but setting $\phi = 0.001$ and $i = 110$. In addition, plot solutions $v(t)$ and $w(t)$ as functions of time, observing the rapid jumps in v that alternate with periods of slower variation. What happens to w during the jumps in v ? Note: be careful with the numerical integration and choice of time steps in doing this exercise, because the system is stiff as discussed in Section 5.7.

Exercise 8.11. Continuing the previous exercise, explore how the shape of the periodic orbits changes as $\phi \rightarrow 0$.

8.6.2 Simplifying Spatial Models

The formation of spatial patterns as a result of decisions and movements by individual organisms has recently been a major application area for agent-based models. Dieckmann et al. (2000), Parrish et al. (2002), and Chowdhury et al. (2004) summarize many such applications including bacterial colonies, ant trails, fish schools, and the walking paths followed by humans in crowds. Agent-based models are a natural starting point for phenomena that seem to arise from decisions by discrete agents responding to the behaviors of other discrete agents. Helbing et al. (2000) model escape panics, such as attempts to rapidly exit a room where a fire has started. Everyone's desire to exit rapidly leads to congestion and injuries at pileups that reduce the average exit rate. The model treats agents as circles of fixed diameter moving on a surface, that try to achieve a particular direction and velocity of movement but also try to avoid getting too close to other agents or the walls. Additional rules govern what happens when collisions occur. Simulations of the model replicate the counterproductive effects of panic, suggesting that the model could help design structures for rapid escape during emergencies. Paradoxically, a well-placed obstacle might actually increase the rate of escape during an emergency (Figure 8.5). The key features of the system—agents heading for an exit but bumping into each other and obstacles—would be

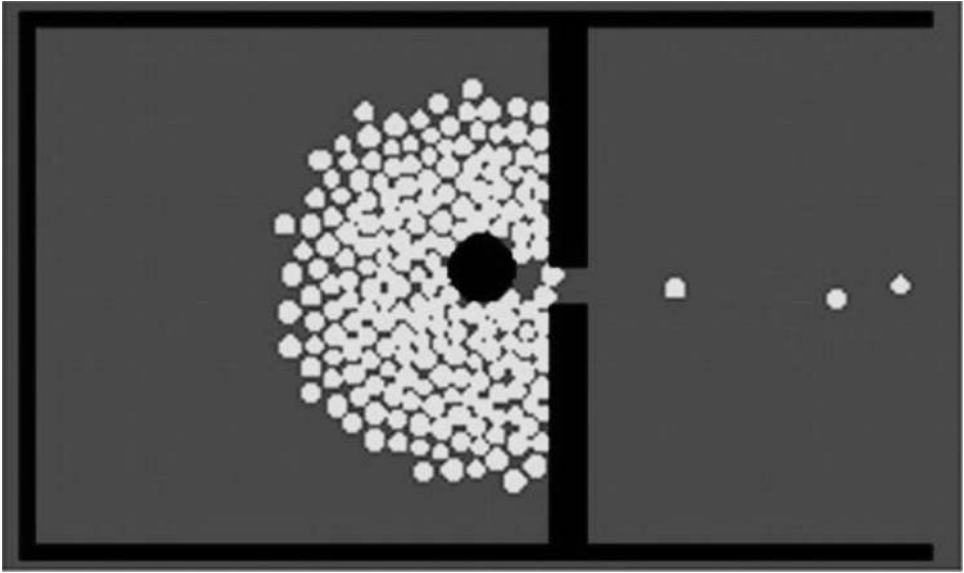


Figure 8.5 A simulation of the Helbing et al. (2000) model for exit panic with a column placed asymmetrically in front of the exit (from Bonabeau 2002). The presence of the column causes the agents to organize their movements in a way that decreases the number of injuries and increases the rate of exit.

more difficult to represent in a partial differential equation model for the “concentration” of pedestrians.

In this section we will briefly describe some methods that can sometimes lead to useful simplifications of spatial agent-based models: mean field equations, hydrodynamic limits, and moment equations. These methods are most easily introduced for spatial “lattice models” such as Avida. Space is represented by a regular discrete grid of sites—such as the points (x, y) in the plane where x and y are both integers. Each site is in one of a finite set of possible states. These models are an import from theoretical physics, where the agents might be elementary particles sitting at each vertex of a regular two- or three-dimensional lattice, with possible states “spin up” and “spin down.”

For a simple biological example (based on Hiebeler 2000), consider a two-dimensional lattice such that each site can be suitable or unsuitable for occupancy (with this distinction being permanent), and each suitable site can be either empty or occupied by a single agent. We posit the following simple rules for what can happen between times t and $t + \Delta t$:

1. Each agent has probability $\mu \times \Delta t$ of dying.
2. Each agent has probability $b \times \Delta t$ of producing an offspring, which is dispersed at random into one of the four sites immediately adjacent to the parent (up, down, left, right). If the site is occupied or unsuitable, the offspring dies. If the cell is suitable and vacant, the offspring takes it and the cell is then occupied.

If all sites are suitable, these assumptions specify the classical “contact process” model for population spread. Hiebeler (2000) introduced unsuitable cells in order to model habitat loss in a fragmented landscape, and to ask how the amount and spatial arrangement of suitable habitat affects species persistence.

For what values of b and μ will the population persist? By rescaling time we can see that the answer will involve only the ratio b/μ , but beyond that it is a difficult question and the exact answer is not known.

One way of simplifying the model is to pretend that the sites around a given agent are drawn independently at random from the entire “landscape.” This is called the mean field approximation, and is derived in general by asking how agents would behave if their local environments were statistically homogenized. Then assuming “many” agents, coin-tossing randomness is replaced by expected transition rates (as in Chapter 3). The result is typically a system of differential equations for the frequencies of agents in different possible states.

For the Hiebeler model let $O(t)$, $E(t)$, and U be the numbers of occupied, empty but suitable, and unsuitable sites with $O(t) + E(t) + U \equiv N$ (constant total number of sites). Since $O(t) + E(t) \equiv N - U$ we only need one state variable, $O(t)$. New occupied sites are created by births: offspring are created at rate $bO(t)$ per unit time, and the fraction that live to occupy a site is given by the fraction of sites that are suitable but unoccupied, $E(t)/N = (N - U - O(t))/N$. This calculation is where spatial structure is ignored: sites near a parent are more likely to be occupied than a “typical” site, so the fraction of offspring that survive is actually lower than $E(t)/N$. Occupied sites are vacated by deaths, which occur at total rate $\mu O(t)$ per unit time. Assuming enough sites that the coin-tossing randomness averages out, we then have

$$\dot{O} = bO(N - U - O)/N - \mu O$$

or equivalently

$$\dot{p} = bp(s - p) - \mu p \tag{8.18}$$

where $p(t) = O(t)/N$ is the fraction of occupied sites and $s = (N - U)/N$ is the fraction of suitable sites. The condition for population persistence is easy to find by linear stability analysis of the equilibrium $\dot{p} = 0$.

The mean field approximation does not totally ignore space. Agents are still viewed as interacting with their local neighbors, whose states are the outcome of random processes. These considerations can affect the expected rate at which sites change state. For example, suppose we assume that an agent’s probability of death per unit time is $\mu(1 + C^2)$ where C is the number of occupied neighboring sites. If a given site has z neighboring sites, then in the mean field approximation C is a Binomial(z, p) random variable. The per agent death rate is then proportional to the expected value of $1 + C^2$, $E[1 + C^2] = 1 + (zp)^2 + zp(1 - p)$. The term $zp(1 - p)$ reflects the stochastic variation among the local neighborhoods of dif-

ferent agents. If we took the approximation one step further and pretended that each local neighborhood contained exactly $E[C] = zp$ occupied sites, this term would be eliminated. But because the mortality rate is a concave-up function of the number of occupied neighbor sites, the effect of local stochasticity is to increase the average mortality rate.

Sometimes the mean field approximation works—in the sense of preserving the model’s qualitative properties—and sometimes it doesn’t. We will see below that the Hiebeler model is a “doesn’t,” despite its simplicity. In contrast, Moorcroft et al. (2001) found that the mean field approximation for a very complicated and computationally intensive agent-based forest simulation model was very accurate, in the sense that it replicated the average trajectory of the model over many simulations. The mean field approximation in that case was a system of partial differential equations, describing the frequency distribution of trees classified by species, the amount of biomass in several compartments, and the time since a local disturbance of their site. The mean field equations also had to be solved numerically, but they could be scaled up to much larger land areas than could be simulated tree by tree. This allowed Hurtt et al. (2002) to use the mean field model for studying the carbon budget of the entire continental United States in response to past and projected future land use patterns.

For some types of models on a homogeneous lattice, Durrett and Levin (1994) conjectured that the model’s behavior can often be predicted from the behavior of the mean-field approximate model:

- When the mean field model has a single fixed point with all state densities positive, the lattice model converges to a steady state with all state densities positive.
- When the mean field model has several fixed points, its long-term behavior depends on initial conditions. In contrast, the lattice model predictably converges to a steady state corresponding to one of the fixed points.
- When the mean field model has a limit cycle, the lattice model converges to a steady state with the same states present, and spatial average densities on intermediate spatial scales exhibit persistent cyclic oscillations.

Durrett (1999) reviews evidence in support of this conjecture, and some cases where mean field behavior is misleading. However, it is generally true (and proved for some models) that as the range of interagent interactions is increased, the model converges to the mean-field approximation. In the Hiebeler model this occurs as the range of offspring dispersal is increased so that offspring can land on any site within distance r of the parent, $r \gg 1$. When $r = \infty$ the site where an offspring lands really is a random draw from the entire landscape, so the model behaves just like the mean field approximation.

Exercise 8.12. Derive the complete mean field approximation to the Hiebeler model under the modified assumptions that in a short time interval of length Δt

- (a) each agent has probability $\mu(1 + \alpha_1 C + \alpha_2 C^2)\Delta t$ of dying, and
- (b) each agent has probability $b(1 + \beta_1 C + \beta_2 C^2)\Delta t$ of producing an offspring,

where C is the agent's number of occupied neighboring sites. Depending on their signs the α and β coefficients could model local competition or cooperation.

8.6.3 Improving the Mean Field Approximation

Mean-field equations obliterate the correlation between agents and their neighborhoods: a kindergarten class with one case of flu will soon have many, which limits the rate at which each infective child generates new cases. To retain these local correlations we need approximations that do not totally homogenize space.

The mean field approximation can be viewed as the limit of "mixing" the system more and more rapidly. Suppose that in each time interval of length $\Delta t \ll 1$ each pair of neighboring sites swap their contents with probability $m\Delta t$. Then if m is large enough the system approaches the mean-field property that each agent's neighbors are a random draw from the entire set of agents. Some local spatial structure can be preserved by simultaneously shrinking the size of lattice cells. As in the convergence of random walk to diffusion (Chapter 8), mixing and shrinking rates are linked so that agents' mean square displacement as a function of time approaches a limit. In physics this is called a "hydrodynamic limit." For some types of model the hydrodynamic limit is valid, meaning that an infinite-lattice model with rapid enough mixing has the same qualitative behavior. In many cases there is a simple recipe for the hydrodynamic limit: a mean field model of the form $du/dt = f(u)$ is replaced by a reaction-diffusion model of the form

$$\frac{\partial u}{\partial t} = f(u) + \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2},$$

and similarly for models with more state variables (Durrett 1999).

Another way of improving the mean field is to expand the scale at which independence is assumed. Where mean field theory treats all sites as independent, *pair approximation* includes correlations between neighboring sites. Let ρ_i denote the fraction of sites in state i , and ρ_{ij} the fraction of neighboring site pairs in state (i, j) . Assumptions of between-pair independence are then used to derive equations for the pair frequencies ρ_{ij} . In the Appendix to this chapter we show how approximate pair-frequency equations are derived for a simple model.

The pair approximation for the Hiebeler model is a system of three differential equations. There are nine different pair types, but only three state variables are needed because ρ_2 and ρ_{22} are constant by assumption, we have the symmetries $\rho_{ij} = \rho_{ji}$, and the pair frequencies sum to 1. Analysis of the pair approximation equations leads to the prediction that the fraction of suitable sites that are occupied depends on the conditional probability that the neighbor of a suitable site

is unsuitable, not on the total fraction of suitable sites. The mean field approximation [8.18] makes exactly the opposite prediction. Simulations show that the pair approximation prediction is very close to the truth (Hiebeler 2000).

More accurate approximations can be obtained by using larger building blocks, but this is rarely done. A more important extension has been the use of similar ideas when interactions are defined by an interaction network rather than spatial proximity, for example, a network of social contacts allowing disease transmission (e.g., Rand 1999; van Baalen 2000; Eames and Keeling 2003).

Pair approximation is one example of *moment closure* methods, in which approximate equations are written for a limited number of statistics (“moments”) characterizing the state of the system. For pair approximation these are the site-pair frequencies; in other cases the moments are means, variances, and covariances (in space and/or time) of agents of different types or agents in different states. These are more complicated than pair approximation but do not require artificially gridding up space into discrete cells. Bolker et al. (2000) review these methods and some of their applications.

8.7 Conclusions

Agent-based models are likely to remain an important approach in computational biology, in part because the agent-based paradigm has become a part of how we build models. It is increasingly common for classical differential equation models to be constructed by making assumptions at the level of agents, and then explicitly deriving the model as a mean-field or some other approximation to the “exact” agent-based model. This leads to deterministic models whose equations reflect the local stochastic nature of interactions among agents, which in the past typically would have been ignored.

We remind readers that this is the chapter where we have allowed ourselves to stray off the beaten path toward the “bleeding edge.” It will probably become obsolete faster than the rest. We also caution readers that the sections on sensitivity analysis and model simplification are more a prescription for the future than a description of current practice. We believe that these methods have a lot of potential and deserve more widespread use, but it is too early to say if they will live up to their promise. In particular, it is not clear whether the methods to improve on mean-field approximations in spatial models will be useful for other kinds of agent-based simulations, or even for more realistic spatial models. The central message of those sections is that structured approaches, grounded in statistical and mathematical theory, can be enormously more efficient than trial and error for figuring out the behavior of a complicated computational model.

8.8 Appendix: Derivation of Pair Approximation

Consider the continuous-time Hiebeler model with (for simplicity) all sites suitable, and let states 0 and 1 indicate empty and occupied sites. This model is known as the “contact process.” Each pair of neighboring cells is then in one of the four states (0, 0), (1, 1), (1, 0), (0, 1). Although there are four types of site pairs, we only need two state variables because of the constraints

$$\sum_{i,j=0}^1 \rho_{ij} = 1, \quad \rho_{10} = \rho_{01}. \tag{8.19}$$

It is convenient to use ρ_1 and ρ_{11} as the state variables. Because $\rho_{01} + \rho_{11} = \rho_1$ we have $\rho_{01} = \rho_{10} = \rho_1 - \rho_{11}$. Occupied sites are lost by deaths at total rate $\mu\rho_1$. They are gained by offspring birth and survival at total rate $b\rho_1q_{0/1}$ where $q_{0/1}$ is the conditional probability that a randomly chosen neighbor to an occupied site is vacant. So ignoring the coin-tossing randomness, we have

$$\dot{\rho}_1 = b\rho_1q_{0/1} - \mu\rho_1.$$

By the conditional probability formula $P(A|B) = P(A \cap B)/P(B)$ we have

$$q_{0/1} = \rho_{01}/\rho_1 = (\rho_1 - \rho_{11})/\rho_1$$

and therefore

$$\dot{\rho}_1 = b(\rho_1 - \rho_{11}) - \mu\rho_1. \tag{8.20}$$

Next we need the equation for ρ_{11} . A (1, 1) site pair is lost whenever one of the pair dies, so the total loss rate is $2\mu\rho_{11}$. (1, 1) pairs are gained when the vacant site in a (0, 1) or (1, 0) pair becomes occupied. Each occupied neighbor of the vacant site sends in offspring at rate $b/4$; one neighbor is occupied with probability 1, the other three with probability $q_{1/01}$, the conditional probability of state 1 in a randomly chosen neighbor of the 0 in a (0, 1) pair, $q_{1/01} = \rho_{101}/\rho_{01}$. The expected total rate of (0, 1) \rightarrow (1, 1) transitions is therefore

$$\rho_{01} \frac{b}{4} (1 + 3q_{1/01}).$$

(1, 0) \rightarrow (1, 1) transitions occur at the same rate so we have

$$\dot{\rho}_{11} = \rho_{01} \frac{b}{2} (1 + 3q_{1/01}) - 2\mu\rho_{11}. \tag{8.21}$$

The equation for ρ_{11} involves $q_{1/01} = \rho_{101}/\rho_{01}$, and the trio density ρ_{101} cannot be computed from our two state variables. This problem cannot be solved by adding more state variables, because the pattern continues—dynamic equations for site trio densities involve the densities of site quartets, and so on. At some point we have to somehow truncate this process. Pair approximation stops at site pairs by making the approximation $q_{1/01} = q_{1/0}$. The 0 site in a (0, 1) pair isn’t any old vacant site—it’s a vacant site with an occupied neighbor—but we

ignore this and hope for the best. The total rate of $(0, 1) \rightarrow (1, 1)$ transitions is then approximated as

$$\rho_{01} \frac{b}{2} (1 + 3q_{1/0}).$$

The result, with a bit of algebra, is

$$\dot{\rho}_{11} = \frac{b}{2} (\rho_1 - \rho_{11}) \left(1 + 3 \frac{\rho_1 - \rho_{11}}{1 - \rho_1} \right) - 2\mu\rho_{11}. \quad [8.22]$$

Equations [8.20] and [8.22] are now a closed system of equations that can be analyzed or solved numerically. For example, linear stability analysis of the equilibrium $(0, 0)$ gives population persistence for $b/\mu > 4/3$. The corresponding prediction of the mean-field model is persistence for $b/\mu > 1$. The actual critical value of b/μ (estimated from simulations) is about 1.65, so pair approximation has reduced by half the error of the mean-field approximation. But numerical accuracy is not really the point—you can do better by simulation. Pair and other moment approximations are valuable when they generate qualitative predictions that can be confirmed in simulations but would have been hard to discover inductively from simulations (e.g., Bolker and Pacala 1999; Keeling et al. 2000; Snyder and Chesson 2004).

8.9 References

- Adami, C. 2002. Ab initio modeling of ecosystems with artificial life. *Natural Resource Modeling* 15: 133–145.
- Blower, S. M., and H. Dowlatabadi. 1994. Sensitivity and uncertainty analysis of complex models of disease transmission: An HIV model, as an example. *International Statistical Review* 2: 229–243.
- 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.
- Bolker, B. M., and S. W. Pacala. 1999. Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. *American Naturalist* 153: 575–602.
- Bolker, B. M., S. W. Pacala, and S. A. Levin. 2000. Moment methods for stochastic processes in continuous space and time. Pages 388–411 in Dieckmann et al. (2000).
- Bonabeau, E. 2002. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences, U.S.A.* 99, Suppl. 3: 7280–7287.
- Brook, B. W., J. R. Cannon, R. C. Lacy, C. Mirande, and R. Frankham. 1999. Comparison of the population viability analysis packages GAPPs, INMAT, RAMAS, and VORTEX for the whooping crane (*Grus americana*). *Animal Conservation* 2: 23–31.
- Chan, K., A. Saltelli, and S. Tarantola. 2000. Winding Stairs: A sampling tool to compute sensitivity indices. *Statistics and Computing* 10: 187–196.
- Chow, S. S., C. O. Wilke, C. Ofria, R. E. Lenski, and C. Adami. 2004. Adaptive radiation from resource competition in digital organisms. *Science* 305: 84–86.
- Chowdhury, D., K. Nishinari, and A. Schadschneider. 2004. Self-organized patterns and traffic flow in colonies of organisms: from bacteria and social insects to vertebrates. *Phase Transitions* 77: 601–624.
- Cleaveland, S., G. R. Hess, A. P. Dobson, M. K. Laursen, H. I. McCallum, M. G. Roberts, and R. Woodroffe. 2002. The role of pathogens in biological conservation. Pages 139–150 in Hudson et al. (2002).

- DeAngelis, D. L., and L. J. Gross (eds.). 1992. *Individual-Based Models and Approaches in Ecology*. Chapman and Hall, New York.
- DeAngelis, D. L., L. J. Gross, M. A. Huston, W. F. Wolff, D. M. Fleming, E. J. Comiskey, and S. M. Sylvester. 1998. Landscape modeling for everglades ecosystem restoration. *Ecosystems* 1: 64–75.
- Dieckmann, U., R. Law, and J.A.J. Metz. 2000. *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge University Press, Cambridge, U.K.
- Durrett, R. 1999. Stochastic spatial models. *SIAM Review* 41: 677–718.
- Durrett, R., and S. A. Levin. 1994. The importance of being discrete (and spatial). *Theoretical Population Biology* 35: 252–283.
- Eames, K.T.D. and M. Keeling. 2003. Contact tracing and disease control. *Proceedings of the Royal Society of London, Series B* 270: 2565–2571.
- Ellner, S. P., and John Fieberg. 2003. Using PVA for management despite uncertainty: Effects of habitat, hatcheries, and harvest on salmon. *Ecology* 84: 1359–1369.
- Enquist, B. J., and K. J. Niklas. 2001. Invariant scaling relations across tree-dominated communities. *Nature* 410: 655–660.
- Gould, S. J. 1990. *Wonderful Life: The Burgess Shale and The Nature of History*. W. W. Norton and Company, New York.
- Harvell, C. D., K. Kim, J. M. Burkholder, R. R. Colwell, P. R. Epstein, J. Grimes, E. E. Hofmann, E. K. Lipp, A.D.M.E. Osterhaus, R. Overstreet, J. W. Porter, G. W. Smith, and G. R. Vasta. 1999. Emerging marine diseases: Climate links and anthropogenic factors. *Science* 285: 1505–1510.
- Helbing, D., I. J. Farkas, and T. Vicsek. 2000. Simulating dynamical features of escape panic. *Nature* 407: 487–490.
- Hess, G. 1996. Disease in metapopulation models: Implications for conservation. *Ecology* 77: 1617–1632.
- Hiebeler, D. 2000. Populations on fragmented landscapes with spatially structured heterogeneities: Landscape generation and local dispersal. *Ecology* 81: 1629–1641.
- Hudson, P. J., A. Rizzoli, B. T. Grenfell, H. Heesterbeek, and A. P. Dobson (eds.). 2002. *The Ecology of Wildlife Diseases*. Oxford University Press, Oxford, U.K.
- Hurttt, G. C., S. W. Pacala, P. R. Moorcroft, J. P. Caspersen, E. Shevliakova, and B. Moore. 2002. Projecting the future of the US Carbon sink. *Proceedings of the National Academy of Sciences U.S.A.* 99: 1389–1394.
- Huston, M., D. deAngelis, and W. Post. 1988. New computer models unify ecological theory. *Bio-Science* 38: 682–691.
- Jansen, M.J.W., W.A.H. Rossing, and R. A. Daamen. 1994. Monte Carlo estimation of uncertainty contributions from several independent multivariate sources. Pages 334–343 in J. Gasman and G. van Straten (eds.), *Predictability and Nonlinear Modelling in Natural Sciences and Economics*. Kluwer Academic, Dordrecht.
- Keeling, M. J., H. B. Wilson, and S. W. Pacala. 2000. Reinterpreting space, time lags, and functional responses in ecological models. *Science* 290: 1758–1761.
- Lacy, R. C., M. Borbat, and J. P. Pollak. 2003. VORTEX: A stochastic simulation of the extinction process, version 9. Chicago Zoological Society, Brookfield, IL. Online at <http://www.vortex9.org>
- Lenski, R. E., C. Ofria, R. T. Pennock, and C. Adami. 2003. The evolutionary origin of complex features. *Nature* 423: 139–144.
- Lindenmayer, D. B., R. C. Lacy, and M. L. Pope. 2000. Testing a simulation model for population viability analysis. *Ecological Applications* 10: 580–597.
- Lindenmayer, D. B., and R. C. Lacy. 2002. Small mammals, habitat patches and PVA models: a field test of model predictive ability. *Biological Conservation* 103: 247–265.
- McKay, M. D. 1992. Latin Hypercube Sampling as a tool in uncertainty analysis of computer models. Pages 557–564 in J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson (eds.), *Proceedings of the 1992 Winter Simulation Conference*. Association for Computing Machinery, Arlington, VA. Online at portal.acm.org
- McKay, M. D., W. J. Conover, and R. J. Beckman. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 22: 239–245.

- Miller, P. S., and R. C. Lacy. 2003. VORTEX: A stochastic simulation of the extinction process, version 9.21 user's manual. Conservation Breeding Specialist Group (SSC/IUCN), Apple Valley, MN. Online at <http://www.vortex9.org/v921manual.pdf>
- Moorcroft, P. R., G. C. Hurtt, and S. W. Pacala. 2001. A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs* 74: 557–586
- Ofria, C., and C. O. Wilke. 2004. Avida: A software platform for research in computational evolutionary biology. *Artificial Life* 10: 191–229.
- O'Neill, B. 2003. Digital evolution. *PLOS Biology* 1: 11–14.
- Parrish, J. K., S. V. Viscido, and D. Grünbaum. 2002. Self-organized fish schools: An examination of emergent properties. *Biological Bulletin* 202: 296–305.
- Peck, S. L. 2004. Simulation as experiment: A philosophical reassessment for biological modeling. *Trends in Ecology and Evolution* 19: 530–534.
- Penn, A. M., W. B. Sherwin, G. Gordon, D. Lunney, A. Melzer, and R. C. Lacy. 2000. Demographic forecasting in koala conservation. *Conservation Biology* 14: 629–638.
- Rand, D. A. 1999. Correlation equations for spatial ecologies. Pages 99–143 in J. McGlade (ed.) *Advanced Ecological Theory*. Blackwell Scientific Publishing, London.
- Ray, T. S. 1991. An approach to the synthesis of life. Pages 371–408 in C. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (eds.), *Artificial Life II*. Santa Fe Institute Studies in the Sciences of Complexity vol. X. Addison-Wesley, Redwood City, CA.
- Ray, T. S. 1992. Evolution, ecology and optimization of digital organisms. Santa Fe Institute Working Paper 92-08-042.
- Rice, J. A., T. J. Miller, K. A. Rose, L. B. Crowder, E. A. Marschall, A. S. Trebitz, and D. L. DeAngelis. 1993. Growth rate variation and larval survival: Inferences from an individual-based size-dependent predation model. *Canadian Journal of Fisheries and Aquatic Sciences* 50: 133–142.
- Smith, D. J., S. Forrest, D. H. Ackley, and A. S. Perelson. 1999. Variable efficacy of repeated annual influenza vaccination. *Proceedings of the National Academy of Sciences USA* 96: 14001–14006.
- Snyder, R. E., and P. Chesson. 2004. How the spatial scales of dispersal, competition, and environmental heterogeneity interact to affect coexistence. *American Naturalist* 164: 633–650.
- Sobol', I. M. 1993. Sensitivity analysis for nonlinear mathematical models. *Mathematical Modeling and Computational Experiment* 1: 407–414 [English translation from *Matematicheskoe Modelirovanie* 2: 112–118 (1990)].
- van Balen, M. 2000. Pair approximations for different spatial geometries. Pages 359–387 in U. Dieckman, R. Law, and J.A.J. Metz (eds.), *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge University Press, Cambridge, U.K.
- Waide, R. B., M. R. Willig, C. F. Steiner, G. Mittelbach, L. Gough, S. I. Dodson, G. P. Juday, and R. Parmenter. 1999. The relationship between productivity and species richness. *Annual Review of Ecology and Systematics* 30: 257–300.
- Ward, J. R., and K. D. Lafferty. 2004. The elusive baseline of marine disease: Are diseases in ocean ecosystems increasing? *PLOS Biology* 2: 542–547.
- Wilke, C. O., J. L. Wang, C. Ofria, R. E. Lenski, and Christoph Adami. 2001. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* 412: 331–333.
- Wilke, C. O., and C. Adami. 2002. The biology of digital organisms. *Trends in Ecology and Evolution* 17: 528–532.
- Yedid, G., and G. Bell. 2001. Evolution in an electronic microcosm. *American Naturalist* 157: 465–487.

This page intentionally left blank