II

Life and Death on the Sugarscape

In this chapter the simplest version of our artificial world is described. A single population of agents gathers a renewable resource from its environment. We investigate the distribution of wealth that arises among the agents and find that it is highly skewed. It is argued that such distributions are *emergent* structures. Other emergent phenomena associated with mass agent migrations are then studied. Social networks among neighboring agents are illustrated and their significance is discussed. Finally, it is argued that artificial societies can serve as laboratories for social science research.

In the Beginning There Was Sugar

Events unfold on a "sugarscape." This is simply a spatial distribution, or topography, of "sugar," a generalized resource that agents must eat to survive. The space is a two-dimensional coordinate grid or lattice. At every point (x, y) on the lattice, there is both a sugar level and a sugar capacity, the capacity being the maximum value the sugar level can take at that point. Some points might have no sugar (a level of zero) and low capacity, others might have no sugar but large capacity—as when agents have just harvested all the sugar—while other sites might be rich in sugar and near capacity.

The Sugarscape software system (that is, the computer program proper) permits one to specify a variety of spatial distributions of levels and capacities. But let us begin with the particular sugarscape shown in figure II-1, which consists of 2500 locations arranged on a 50 x 50 lattice with the sugar level at every site initially at its capacity value.

The sugar score is highest at the peaks in the northeast and southwest quadrants of the grid—where the color is most yellow—and falls off in a series of terraces.[1] The sugar scores range from some maximum—

1. Terms like "peak" or "mountain" are not used to suggest physical elevation, but to denote regions of high sugar level.

Figure II-1. A Sugarscape

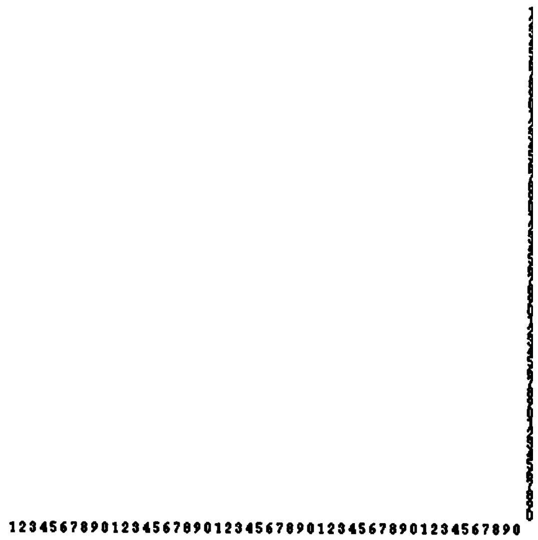
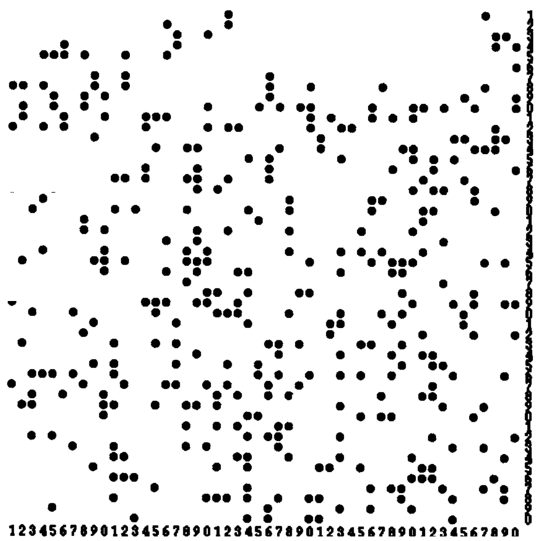


Figure II-2. Sugarscape with Agents



here 4—at the peaks to zero at the extreme periphery. The sugarscape wraps around from right to left (that is, were you to walk off the screen to the right, you would reappear at the left) and from top to bottom, forming a doughnut—technically, a torus.

Simple Local Rules for the Environment

In our model, autonomous agents inhabit this sugarscape and constantly collect and consume sugar. We therefore need to postulate a rule for how the sugar regenerates—how it grows back after it is harvested by the agents.

Various rules are possible.[2] For instance, sugar could grow back instantly to its capacity. Or it could grow back at a rate of one unit per time step. Or it could grow back at different rates in different regions of the sugarscape. Or the growback rate might be made to depend on the sugar level of neighboring sites. We will examine several of these possibilities. To begin, however, we stipulate that at *each lattice point* the sugarscape obeys the following simple rule:

---

2. The main constraint we impose on ourselves in constructing such rules throughout this book is to make them as simple as possible. This has two main implications, one theoretical and one practical. Theoretically, rule simplicity suggests that the agents use only local information. Practically, we want to be able to state a particular rule in just a few lines of code.

Sugarscape growback rule $\mathbf{G}_\alpha$: At each lattice position, sugar grows back at a rate of α units per time interval up to the capacity at that position.[3]

With the sugarscape described, we now "flesh out" what we mean by "agents."

The Agents

Just as there is an initial distribution of sugar, there is also an initial population of "agents." We want to give these agents the ability to move around the sugarscape performing various tasks. In this chapter they simply gather sugar and eat it.[4] In later chapters their behavioral repertoire expands to include sex, cultural exchange, combat, trade, disease transmission, and so on. These actions require that each agent have internal states and behavioral rules.[5] We describe these in turn.

Agent States

Each agent is characterized by a set of fixed and variable states. For a particular agent, its genetic characteristics are fixed for life while its wealth, for instance, will vary over time.

One state of each agent is its location on the sugarscape. At every time each agent has a position given by an ordered pair (x, y) of horizontal and vertical lattice coordinates, respectively. Two agents are not allowed to occupy the same position. Some agents are born high on the sugarscape near the peaks of the sugar mountains shown in figure II-1. Others start out in the sugar "badlands" where sugar capacities are very low. One might think of an agent's initial position as its "environmental endowment." We shall first investigate a random distribution of 400 agents, as shown in figure II-2.

Each agent has a genetic endowment consisting of a sugar metabolism

---

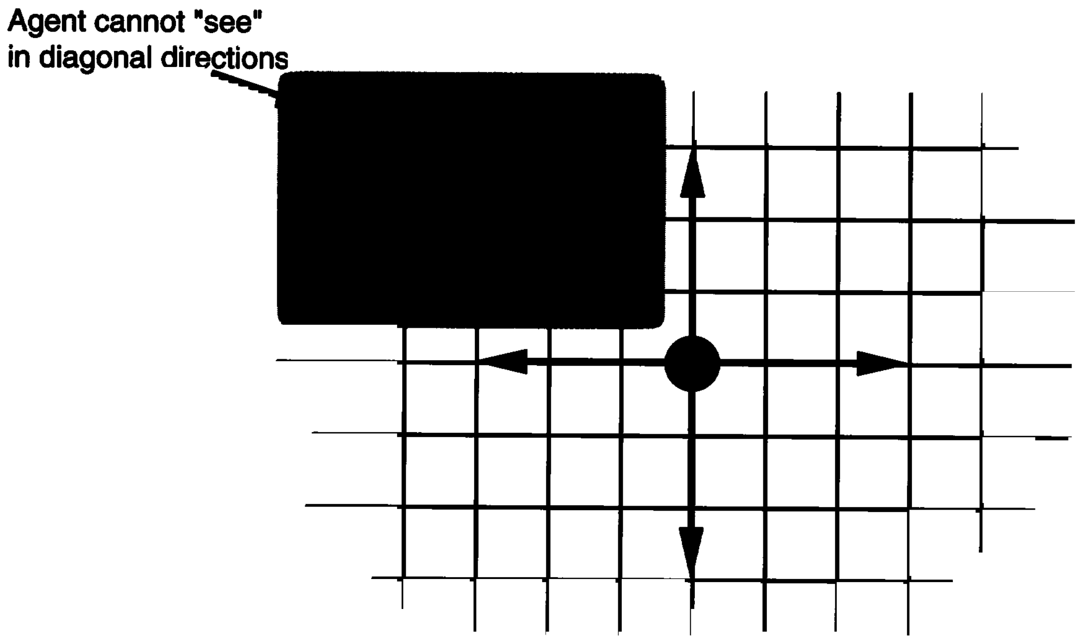
3. The rule can be stated formally. Call the current resource (sugar) level r^t and the capacity c. Then the new resource level, r^{t+1}, is given by

$$r^{t+1} = min(r^t+\alpha,c).$$

4. For a similar model, see Packard's [1989] artificial ecology.

5. As noted in Chapter I, each agent is implemented as an "object"; its internal states are its "instance variables," while its behavioral rules are specified by its "methods." Technically, the states of an agent are data while its behavioral rules are procedures (or subroutines).

Figure II-3. Agent Vision



and a level of vision. Agents have different values for these genetic attributes; thus the agent population is *heterogeneous*.[6]

The agent's metabolism is simply the amount of sugar it burns per time step, or iteration. Metabolisms are randomly distributed across agents. For the runs of the model described immediately below, metabolism is uniformly distributed with a minimum of 1 and a maximum of 4.

Agent vision is also randomly distributed. Agents with vision v can see v units in the four principal lattice directions: north, south, east, and west. Agents have no diagonal vision. This lack of diagonal vision is a form of imperfect information and functions to bound the agents' "rationality," as it were. The nature of agent vision is illustrated in figure II-3. An agent with vision 3 can look out 3 units in the principal lattice directions. In what follows vision is initially distributed uniformly across agents with values ranging from 1 to 6, unless stated otherwise.

All agents are given some initial endowment of sugar, which they carry with them as they move about the sugarscape. Sugar collected but not eaten—what an agent gathers beyond its metabolism—is added to the agent's sugar holdings.[7] There is no limit to how much sugar an individual agent may accumulate.

6. When the number of genetic attributes is large it may even be the case that no two agents are genetically identical.

7. Agent holdings do not decay over time.

Simple Local Rules for the Agents

The agents are also given a movement rule. Movement rules process local information about the sugarscape and return rank orderings of the sites according to some criterion. Such rules are called "movement rules" since each agent moves to the site it ranks highest. As with the sugarscape growback rule, we require that agent movement be governed by a *simple* rule.

A natural way to order the sites is by the amount of sugar present at each site within an agent's vision. This results in the following movement rule, which is a kind of gradient search algorithm:

Agent movement rule **M**:

- Look out as far as vision permits in the four principal lattice directions and identify the unoccupied site(s) having the most sugar;[8]
- If the greatest sugar value appears on multiple sites then select the nearest one;[9]
- Move to this site;[10]
- Collect all the sugar at this new position.

Succinctly, rule **M** amounts to this: From all lattice positions within one's vision, find the nearest unoccupied position of maximum sugar, go there and collect the sugar.[11]

At this point the agent's accumulated sugar wealth is *incremented* by the sugar collected and *decremented* by the agent's metabolic rate. If at any time the agent's sugar wealth falls to zero or below—that is, it has been unable to accumulate enough sugar to satisfy its metabolic demands—then we say that the agent has starved to death and it is removed from the sugarscape. If an agent does not starve it lives forever.

8. The order in which each agent searches the four directions is random.

9. That is, if the largest sugar within an agent's vision is four, but the value occurs twice, once at a lattice position two units away and again at a site three units away, the former is chosen. If it appears at multiple sites the same distance away, the first site encountered is selected (the site search order being random).

10. Notice that there is no distinction between how far an agent can move and how far it can see. So, if vision equals 5, the agent can move up to 5 lattice positions north, south, east, or west.

11. Since all agents follow this behavioral rule, there is a sense in which they are quite homogeneous. However, recalling that vision is randomly distributed in the agent population, two distinct agents placed in identical environments will not generally respond (behave) in the same way, that is, move to the same location.

Each agent is permitted to move once during each artificial time period. The order in which agents move is randomized each time period.[12]

Artificial Society on the Sugarscape

All the ingredients are now in hand. We have a sugarscape and an initial population of agents, each of whom comes into the world with an environmental and genetic endowment, and we have simple behavioral rules for the sugarscape and the agents. Initially there will be only one rule for the agents and one for the sugarscape, but in subsequent chapters both the environment and the agents will execute multiple rules.[13] So a notation is needed to describe the rules being executed for any particular run of the model. Call **E** the set of rules that the environment executes, and let **A** be the set of rules the agents follow. Then the ordered pair (**E**, **A**) is the complete set of rules.

For the first run of the model, the sugarscape will follow an instance of the general rule, $\mathbf{G}_\alpha$, that we call the "immediate growback rule."

Sugarscape rule $\mathbf{G}_\infty$: Grow back to full capacity immediately.[14]

This rule says that no matter what the current sugar level is at a site, replace it with that site's sugar capacity. The agents will all execute movement rule **M**. Thus the complete set of rules being executed is $(\{\mathbf{G}_\infty\}, \{\mathbf{M}\})$.

Can you guess what will happen for these rules? Will the agents all clump together atop the sugar mountains? Will agent motion persist indefinitely? Actual dynamics are shown in animation II-1.[15]

What is striking to us is the way the agents ultimately "stick" to the

12. All results reported here have been produced by running the model on a serial computer; therefore only one agent is "active" at any instant. In principle, the model could be run on parallel hardware, permitting agents to move simultaneously (although **M** would have to be supplemented with a conflict resolution rule to handle cases in which two or more agents simultaneously decide to inhabit the same site). Whenever one simulates on a serial machine processes that occur in parallel, it is important to randomize the agent order periodically to ensure against the production of simulation artifacts [Huberman and Glance, 1993].

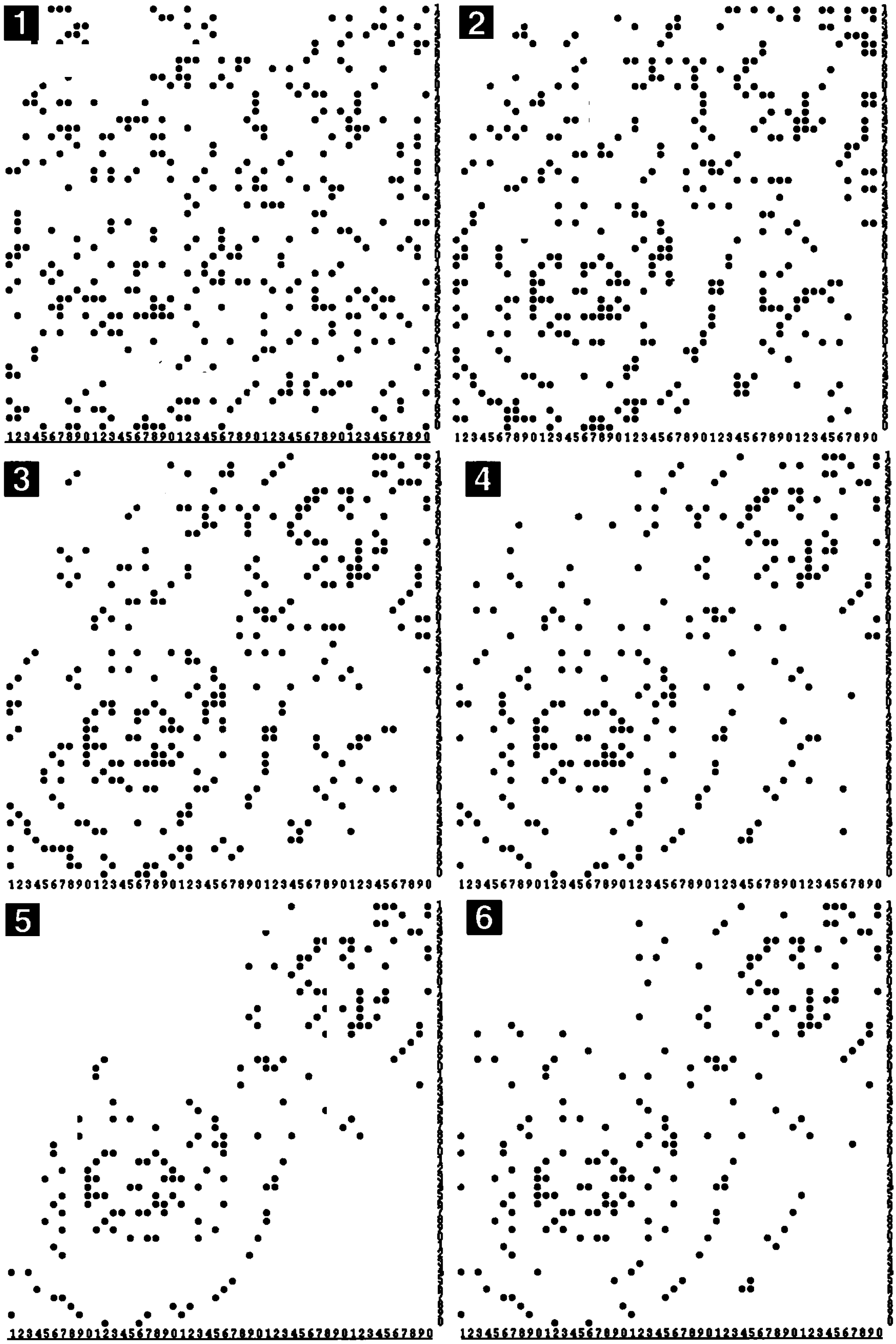
13. Appendix B presents a summary statement of all rules used, in their most general form.

14. Under the definition of G_α, G_∞ ensures that sites grow back instantly to capacity, since

$$r^{t+1} = \min(\infty, c) = c.$$

15. Users wishing to view animations should consult the README file on the CD-ROM for instructions.

Animation II-1. Societal Evolution under Rules ({G_∞}, {**M**}) from a Random Initial Distribution of Agents



ridges of the terraced sugarscape. With immediate growback to capacity, the agents' limited vision explains this behavior. Specifically, suppose you are an agent with vision 2 and you are born on the terrace of sugar height 2, just one unit south of the sugar terrace of level 3. With vision 2, you can see that the nearest maximum sugar position is on the ridge of the sugar terrace of height 3, so, obeying rule **M**, you go there and collect the sugar. Since there is instant growback, no point *on* the level 3 sugar terrace is an *improvement*; and with vision of only 2, you cannot see the higher terrace of sugar level 4. So you stick on the ridge.

Also notice that some agents die. For those with high metabolism *and* low vision, life is particularly hard. This run of the model reaches a steady-state configuration once these unfortunates have died and the rest have attained the best positions they can find.[16] Much richer dynamics result if we slow down the rate at which the sugarscape regenerates, as shown in the next run of our artificial society.

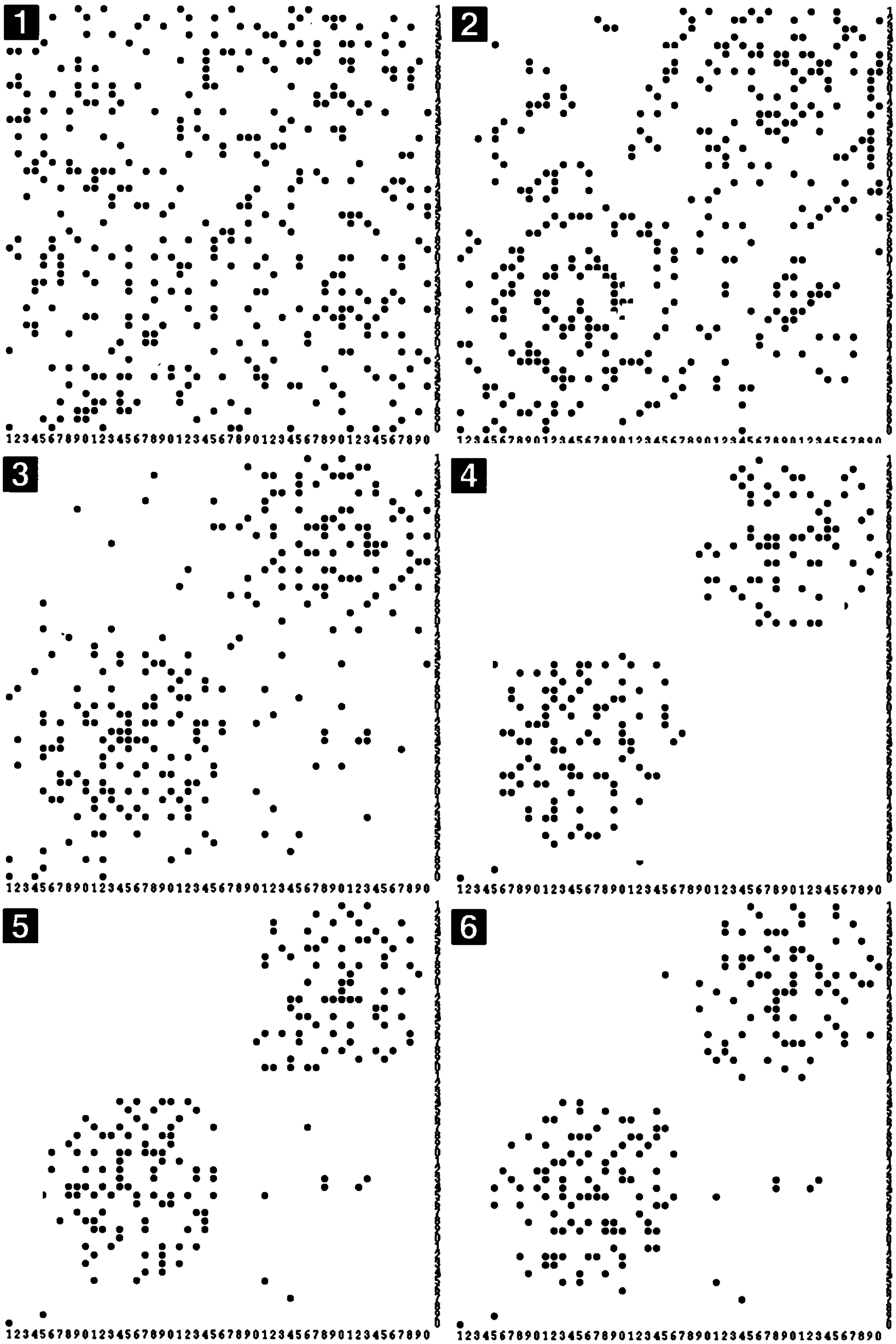
For this second run we again take the initial population to be 400 agents arranged in a random spatial distribution. Each agent again executes rule **M**. But now let us change the sugarscape rule to $\mathbf{G}_1$: Every site whose level is less than its capacity grows back at 1 unit per time period. The complete rules are then $(\{\mathbf{G}_1\}, \{\mathbf{M}\})$. The evolution is shown in animation II-2.

What first catches the eye in this animation is the continuous buzz of activity; it reminds one of "hiving." But it is both purposeful and efficient. It is purposeful in that the agents concentrate their activities on the sugar peaks. Indeed, two "colonies" seem to form, one on each mountain. If the intervening desert (low sugar zone) between the main

---

16. Each time the model is run under rules $(\{\mathbf{G}_\infty\}, \{\mathbf{M}\})$, results qualitatively similar to those of animation II-1 are produced. However, because the initial population of agents is random—each agent has genetics and initial location drawn from certain probability distribution functions—runs made with different streams of random numbers will generally be completely different microscopically, that is, at the level of the agents. (Distinct random number streams are created from run to run either by using distinct seeds in a fixed random number generator (RNG) or by employing altogether different RNGs. A common way to make successive random number seeds uncorrelated in consecutive runs of a model is to tie them to something independent of the model, such as the actual time at which the user starts the run.) Because the search direction in rule **M** is stochastic, even runs having identical populations of agents will differ at the micro-level when, for example, a RNG is re-seeded in the course of a run. All this said, however, let us emphasize that any particular run of the model is completely reproducible. That is, when the sequence of random numbers is specified *ex ante* the model is deterministic. Stated yet another way, model output is invariant from run to run when all aspects of the model are kept constant, including the stream of random numbers.

Animation II-2. Societal Evolution under Rules ({G_1}, {**M**}) from a Random Initial Distribution of Agents



sugar mountains were widened, the spatial segregation seen here would be even more pronounced.

The agents are also efficient grazers. Focusing your attention on a particular sugar location atop one of the sugar peaks, you will see that once it attains some level near its capacity value, it is struck. Then the agent moves away, leaving a white site. Once the site grows back, some agent will zip over and hit it, and so it goes.

An alternative view of rule **M** is that it is a *decentralized* harvesting rule. Specifically, imagine yourself to be the owner of the sugarscape resources and that your goal is to harvest as much sugar as possible. You could give each of your agents explicit instructions as to which site to harvest at what time. Such a harvesting program could turn out to be very complicated indeed, especially when the differential capabilities of the agents are taken into account. But **M** is also a harvesting program, a highly decentralized one.

Carrying Capacity

This simulation illustrates one of the fundamental ideas in ecology and environmental studies—the idea of a *carrying capacity*: A given environment will not support an indefinite population of agents.[17] In this case, although 400 agents begin the simulation, a carrying capacity of approximately 224 is eventually reached. This is revealed in the time series of agent population given in figure II-4.

We can systematically study the dependence of the carrying capacity on the genetic composition of the agent population. To do this one simply specifies particular distributions of vision and metabolism among the agents and lets the model evolve until the asymptotic population level—the carrying capacity—is reached. For a given set of distributions, each run of the model will produce a somewhat different population value, due to stochastic variations, hence, multiple runs must be performed. Figure II-5 gives the dependence of carrying capacity on inital mean vision, parameterized by initial mean metabolism, <m>, starting with 500 agents.[18]

As agent vision increases each agent can see more of the sugarscape and is therefore a more efficient harvester. Similarly, as agent metabolism decreases, each agent finds it somewhat easier to survive.

---

17. For a comprehensive and considered inquiry into the question of Earth's human carrying capacity, see Cohen [1995].

18. Each data point represents the mean value of 10 runs.

Figure II-4. Time Series of Population under Rules ($\{G_1\}$, $\{M\}$) from a Random Initial Distribution of Agents; Asymptotic Approach to the Environmental Carrying Capacity of 224

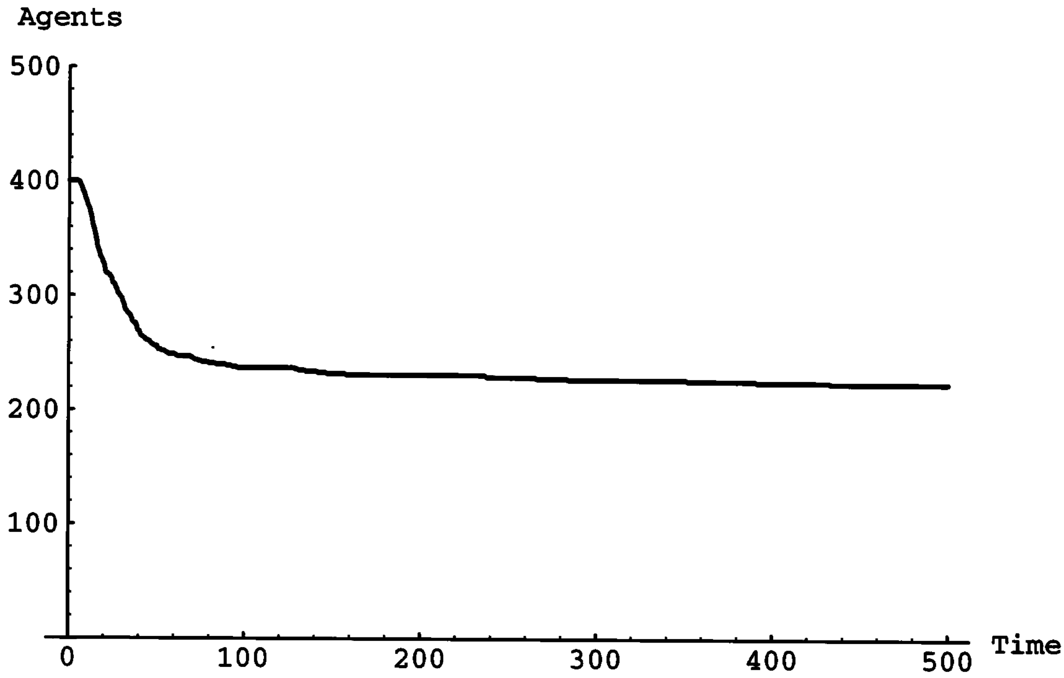
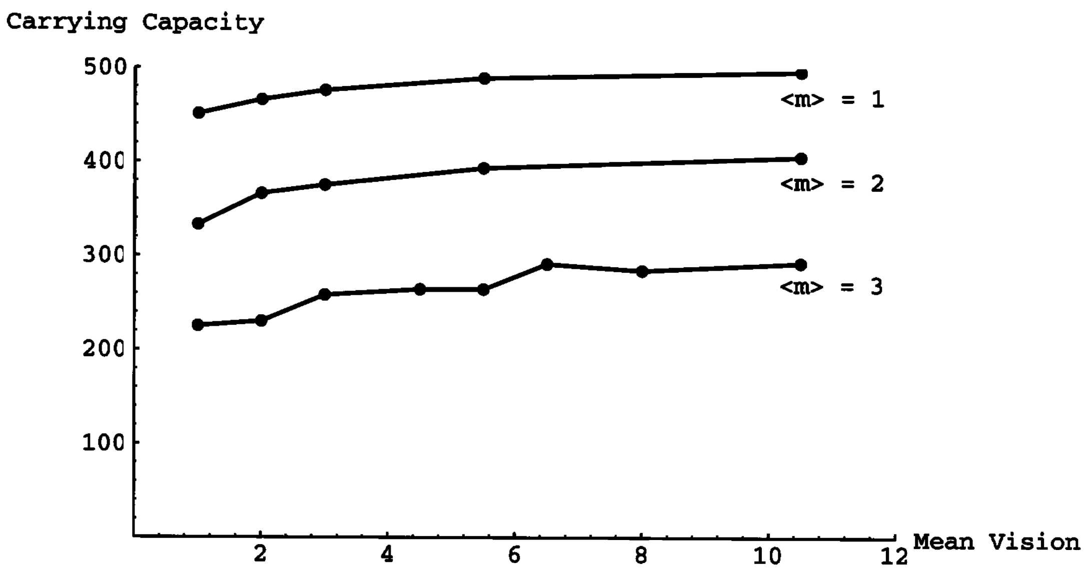


Figure II-5. Carrying Capacities as a Function of Mean Agent Vision, Parameterized by Metabolism, under Rules ($\{G_1\}$, $\{M\}$) from a Random Initial Distribution of Agents



Selection without Sex

In a primitive form our artificial world also illustrates a central idea of evolutionary theory, that of *selection*. As mentioned above, at the outset metabolism and vision are randomly distributed across agents, with each varying between some minimum and maximum value. However, by the time the carrying capacity is attained, the population is skewed in favor of agents with low metabolism and high vision. These agents enjoy a selective advantage over the high metabolism, low vision agents. And as we shall see in the next chapter, when we add sexual reproduction to the agents' behavioral repertoire, this process becomes accretive generationally, producing much stronger tendencies toward agents who are increasingly "fit." Even without sex, selection pressures can be substantial. In the run depicted in animation II-2, the initial mean vision and metabolism were 3.5 and 2.5, respectively. After 500 time periods, selection had increased mean vision to 4.1 and reduced mean metabolism to 1.8.

Wealth and Its Distribution in the Agent Population

All the while in our artificial world agents are accumulating wealth (measured, of course, in sugar). And so, at any time, there is a distribution of wealth in society. The topic of wealth distribution has always interested economists. To study the distribution of wealth in our artificial society we need to modify the previous run in two related ways. First, if agents are permitted to live forever then no stationary wealth distribution ever obtains—the agents simply accumulate indefinitely. Since death is indisputably a fact of life, it is only realistic to insist on finite agent lifetimes. So we set each agent's maximum achievable age—beyond which it cannot live—to a random number drawn from some interval [a,b]. Of course, agents can still die of starvation, as before.

Given that agents are to have finite lifetimes, the second modification that must be implemented is a rule of agent replacement. One can imagine many such rules; for example, a fixed number of new agents could be added each period. However, to ensure a stationary wealth distribution it is desirable to use a replacement rule that produces a constant population. The following replacement rule achieves this goal.

> <u>Agent replacement rule $\mathbf{R}_{[a,b]}$</u>: When an agent dies it is replaced by an agent of age 0 having random genetic attributes, random

position on the sugarscape, random initial endowment, and a maximum age randomly selected from the range [a,b].

To study the actual evolution of the distribution of wealth on the sugarscape we place 250 agents—approximately the carrying capacity—randomly about the sugarscape and let them move and accumulate sugar as before (agent movement rule **M**). Replacement rule $\mathbf{R}_{[60,\ 100]}$ is in effect.[19] Initial agent endowments are uniformly distributed between 5 and 25. The sugarscape grows back at unit rate (environment rule $\mathbf{G}_1$). Now, since we want to track the distribution of wealth, not the spatial distribution of agents, we show a histogram of wealth animated over time. In animation II-3, the horizontal axis gives the range of individual wealth in society, divided into ten "bins." The vertical axis gives the number of agents falling into the various bins. How does the distribution evolve?

While initially quite symmetrical, the distribution ends up highly skewed.[20] Such skewed wealth distributions are produced for wide ranges of agent and environment specifications. They seem to be characteristic of heterogeneous agents extracting resources from a landscape of fixed capacity. By contrast, the distribution of *income*, defined as the amount harvested per period less metabolism, is much less skewed.[21]

Emergence

In the sciences of complexity, we would call this skewed distribution an *emergent structure*, a stable macroscopic or aggregate pattern induced by the local interaction of the agents. Since it emerged "from the bottom up," we point to it as an example of *self-organization*. Left to their own, strictly local, devices the agents achieve a *collective* structuring of some sort. This distribution is our first example of a so-called *emergent structure*.

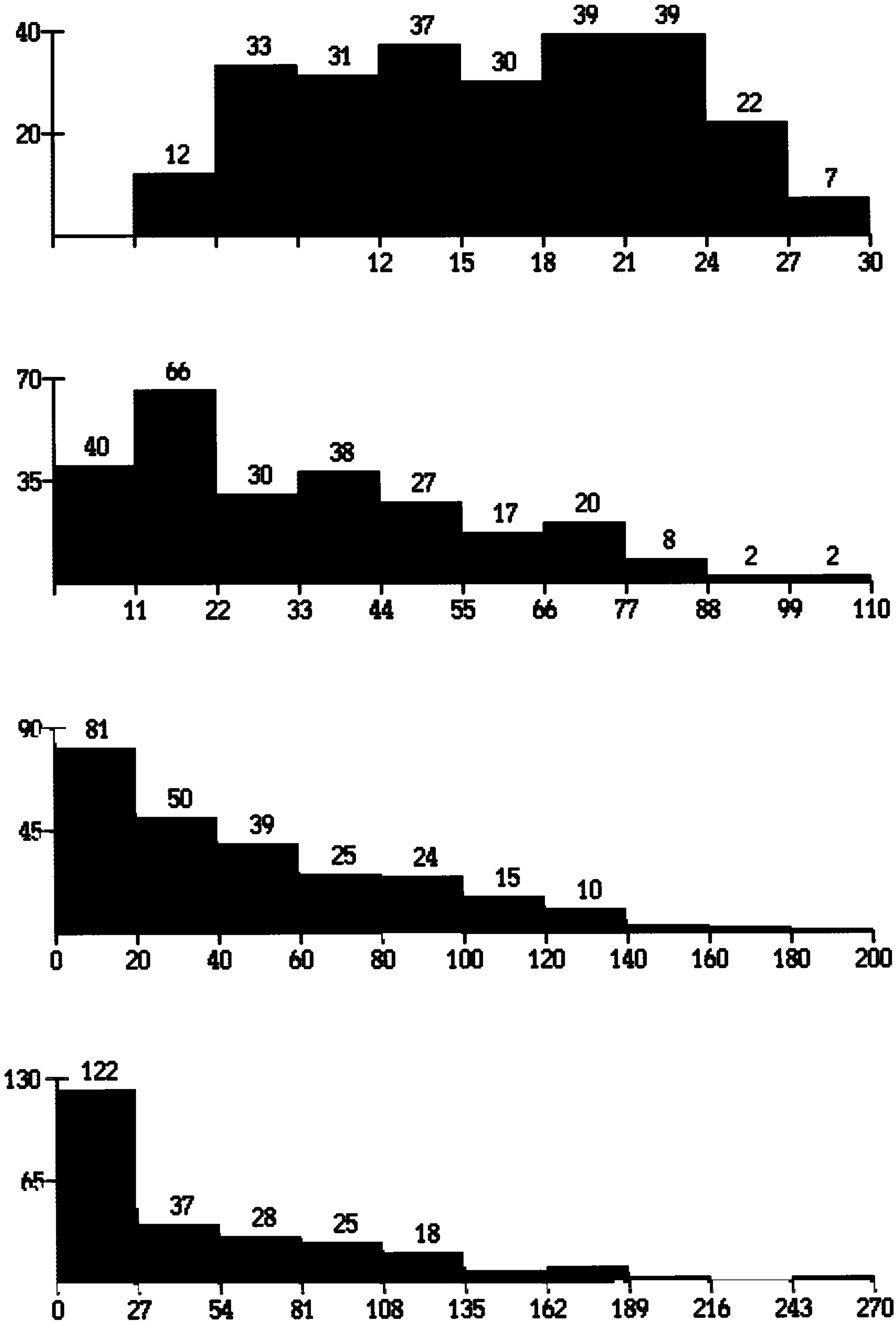
The term "emergence" appears in certain areas of complexity theory, distributed artificial intelligence, and philosophy. It is used in a variety of ways to describe situations in which the interaction of many autonomous individual components produces some kind of coherent, systematic behavior involving multiple agents. To our knowledge, no

19. Note that the mean death age will be 80 when few agents die of starvation.

20. Agents having wealth above the mean frequently have both high vision and low metabolism. In order to become one of the very wealthiest agents one must also be born high on the sugarscape and live a long life.

21. The maximum income possible is 3, since the maximum sugar level is 4 and the minimum metabolism is 1.

Animation II-3. Wealth Histogram Evolution under Rules ({$\mathbf{G_1}$}, {$\mathbf{M}$, $\mathbf{R_{[60,100]}}$}) from a Random Initial Distribution of Agents



completely satisfactory *formal* theory of "emergence" has been given.[22] A particularly loose usage of "emergent" simply equates it with "surprising" or "unexpected," as when researchers are unprepared for the kind of systematic behavior that emanates from their computers.[23] A less subjective usage applies the term to group behaviors that are qualitatively different from the behaviors of individuals composing the group.

We use the term "emergent" to denote *stable macroscopic patterns arising from the local interaction of agents.* One example is the skewed wealth distribution; here, the emergent structure is statistical in nature. There is a qualitatively different type of emergent phenomenon that we also observe. An example of this, described below, occurs when a wave of agents moves collectively in a diagonal direction on the sugarscape, this even though individual agents can move only north, south, east, or west. That is, the group adopts a heading unavailable to any individual! While both the highly skewed wealth distribution and the collective wave satisfy our definition of emergence, they differ in a fundamental respect. We know what it would *mean* for an individual agent to travel on a diagonal; the local rule simply prohibits it. By contrast, we do not know what it would *mean* for an individual to have a wealth distribution; at a given time, only groups can have distributions.[24]

Understanding how simple local rules give rise to collective structure is a central goal of the sciences of complexity. As we will frequently observe, such understanding would have fundamental implications for policy. For instance, we might be able to distinguish conditions (on information or spatial heterogeneity, for example) conducive to the emergence of efficient markets from conditions making their emergence highly unlikely. We might then be better equipped to answer the following sort of question: Is it reasonable to base policy on the assumption that if central authorities "just get out of the way" then efficient markets will self-organize in Russia? Clearly, implicit assumptions on seemingly

---

22. Interesting efforts are under way, however; see Baas [1994].

23. This usage obviously begs the question, "Surprising to whom?"

24. To formalize things somewhat, let A denote an agent and C denote a collection of agents. Let P(A) denote the proposition "A has property P," and likewise for P(C). Then there are at least two types of emergence:

1. P(A) and P(C) are both meaningful, but only P(C) is observed (for example, collective diagonal waves);
2. Only P(C) is meaningful and it is observed (for example, the skewed wealth distribution).

abstract questions of "emergence" drive policy at fundamental levels.

Returning to the wealth distribution of animation II-3, some have argued, especially in the context of the so-called Pareto "law," that highly skewed distributions of income and wealth represent some sort of "natural order," a kind of immutable "law of nature."[25] Artificial social systems let us explore just how immutable such distributions are. We can adjust *local* rules—like those concerning inheritance and taxation—and see if the same *global* pattern in fact emerges.

Measures of Economic Inequality: The Gini Coefficient

It is possible to fit the wealth distribution of animation II-3, or its cumulative distribution counterpart, to any number of empirically significant distribution functions. Such distributions—the Pareto-Levy distribution being perhaps the best known—are typically characterized by one or two parameters, and it might be informative to compare the parameter values obtained from our artificial society with those from real societies. The point of such exercises is to compress information on whole distributions into just a few parameters. This not only facilitates comparison with real economic data but also provides a basis for describing the results of simulations in summary terms. For example, if rules (**E**, **A**) yield a wealth distribution statistic S while rules (**E**, **A**') result in S' > S, it can unambiguously be said that changing agent rules from **A** to **A**' causes S to increase.

In particular, we are interested in summary statistics that can be interpreted as measures of *inequity*. There are a variety of ways to accomplish this when the distribution function to be fit is specified. For example, the exponent in the Pareto distribution is a measure of the inequality of the distribution. However, its interpretation is far from unambiguous.[26] One summary statistic relating to inequality of income or wealth is the so-called Gini coefficient. It has the desirable property that it does not depend on an underlying distribution; that is, it is a "distribution-free" statistic.

The nature of the Gini coefficient or ratio is conveniently explained by reference to the so-called Lorenz curve. This is a plot of the fraction of total social wealth (or income) owned by a given poorest fraction of the

---

25. A large literature surrounds the Pareto "law." See, for example, Kirman [1990] and Persky [1992].

26. See Steindl [1990].

population. Any unequal distribution of wealth produces a Lorenz curve that lies below the 45° line—the poorest X percent of the population controls less than X percent of the society's total wealth. The Gini ratio is a measure of how much the Lorenz curve departs from the 45° line. If everyone has the same amount of wealth the Gini ratio is zero, while if a single individual owns everything then the Gini ratio is one. As the Gini coefficient increases society becomes less egalitarian.[27]

To construct a Lorenz curve for wealth, one first ranks the agents from poorest to wealthiest. Each agent's ranking determines its position along the horizontal axis. Then, for a given agent (abscissa) an ordinate is plotted having a value equal to the total wealth held by the agent and all agents poorer than the agent. The first image in animation II-4 is a Lorenz curve for the initial distribution of wealth on the sugarscape for the run described in animation II-3. When the animation is run, one observes a monotone increase in the curvature of the Lorenz curve—it progressively "bows" outward as inequality grows.

The animation also displays a real-time computation of the Gini coefficient. Note that it starts out quite small (about 0.230) and ends up fairly large (0.503). This Gini ratio, approximately constant for long-time evolutions of the society, is much lower than that seen in industrial societies. In subsequent chapters we shall augment the agents' rules of behavior to include, for example, inheritance, trade, and so on. The Gini ratios of the artificial societies that result then begin to resemble those of developed economies.

The ability to alter agent interaction rules and noiselessly compute the effect on the Gini ratio and other summary statistics is one of the most powerful features of artificial societies. They are "laboratories" for the study of social systems.

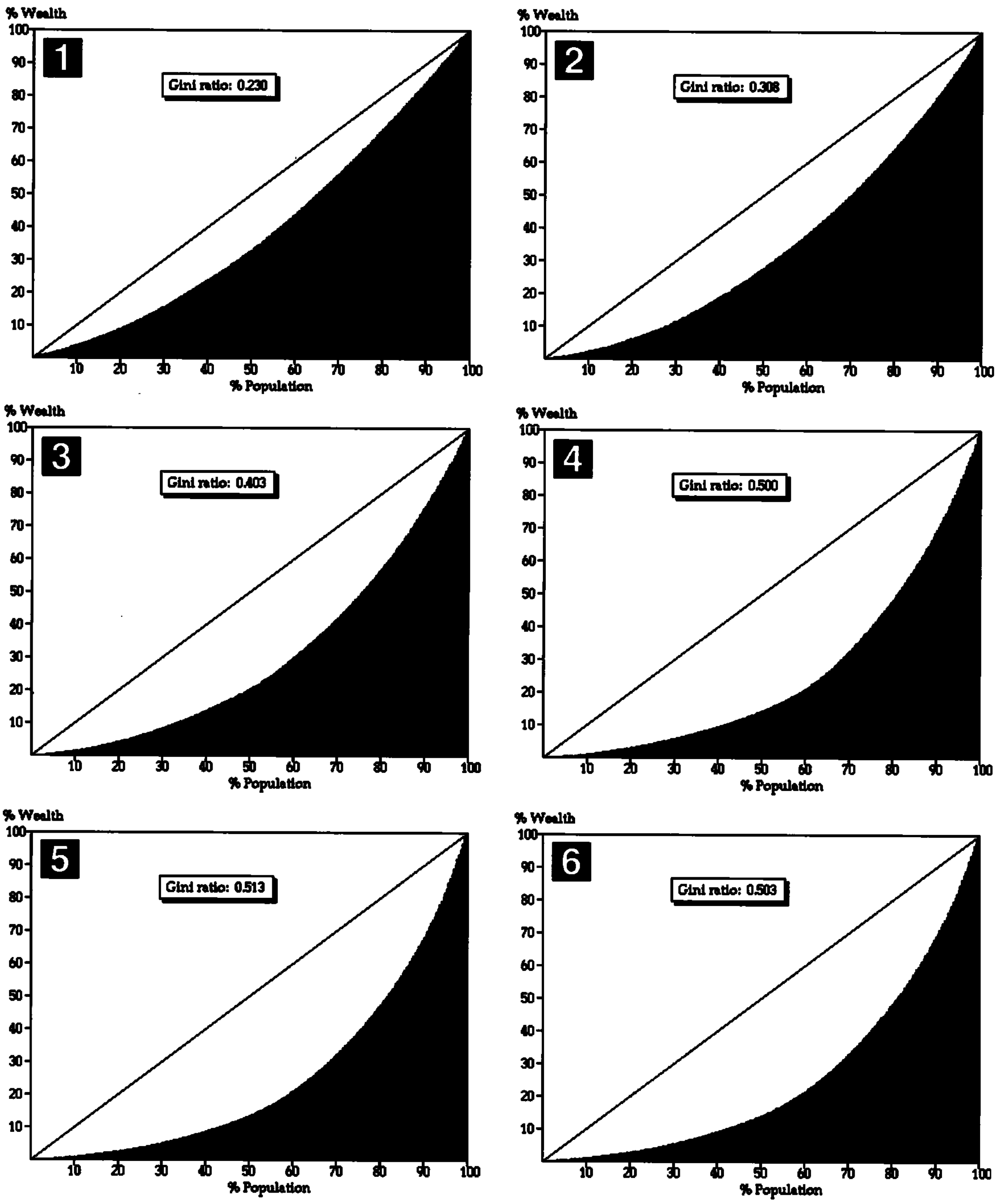
Social Networks of Neighbors

As described in Chapter I, we study various agent connection networks in this book. The first of these will be relatively straightforward, adding insight to the basic picture of "hiving" on the sugarscape. Specifically, we want to keep track of each agent's "neighbors."

One might define the term "neighbor" in a variety of ways. Since our

27. For a more detailed exposition of the Lorenz curve, see Kakwani [1990]. A concise description of the Gini ratio is Dagum [1990].

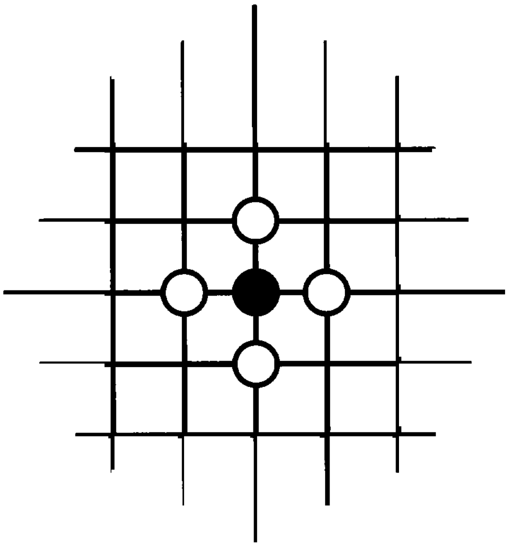
Animation II-4. Evolution of the Lorenz Curve and the Gini Coefficient under Rules ({$\mathbf{G}_1$}, {$\mathbf{M}$, $\mathbf{R}_{[60,100]}$})



agents live on a rectangular lattice it is natural to use the so-called von Neumann neighborhood, defined to be the set of sites immediately to the north, south, east, and west of a particular site. A von Neumann neighborhood is depicted in figure II-6.

An alternative is the Moore neighborhood, which includes all four sites of the von Neumann neighborhood as well as the four sites along

Figure II-6. An Agent and Its von Neumann Neighborhood



the diagonals. Thus there are eight Moore neighbors as shown in figure II-7.

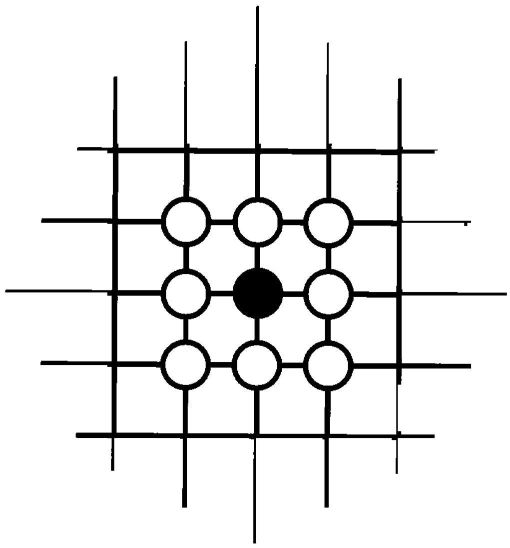
In what follows we shall always employ the von Neumann neighborhood.[28] When an agent moves to a new position on the sugarscape it has from zero to four neighbors. Each agent keeps track of these neighboring agents internally until it moves again, when it replaces its old neighbors with its new neighbors.

The neighbor connection network is a directed graph with agents as the nodes and edges drawn to the agents who have been their neighbors; it is constructed as follows. Imagine that agents are positioned on the sugarscape and that none has moved. The first agent now executes **M**, moves to a new site, and *then* builds a list of its von Neumann neighbors, which it maintains until its next move. The second agent then moves and builds its list of (post-move) neighbors. The third agent moves and builds its list, and so on until all agents have moved. At this point, lines are drawn from each agent to all agents on its list. The resulting graph—a social network of neighbors—is redrawn after every cycle through the agent population.[29] What is most interesting about such

28. In the Sugarscape software system that produced the animations in this book and CD-ROM, one can specify that either a von Neumann or a Moore neighborhood be used.

29. Note that agent-neighbor connections may be asymmetrical (that is, agent i is on agent k's list but not conversely) *and* may extend beyond an agent's von Neumann neighborhood. To see this, imagine that agent i moves into agent k's neighborhood and, accord-

Figure II-7. An Agent and Its Moore Neighborhood



graphs, or networks, is that they change over time as agents (the nodes) move around on the sugarscape. Animation II-5 depicts the development of agent connection networks under rules ({$\mathbf{G}_1$}, {**M**}), the same rules that produced animation II-2. Note that some of the neighbor graphs are simple, while others are elaborate webs containing cycles and other structures. Clearly, a rich variety arises.

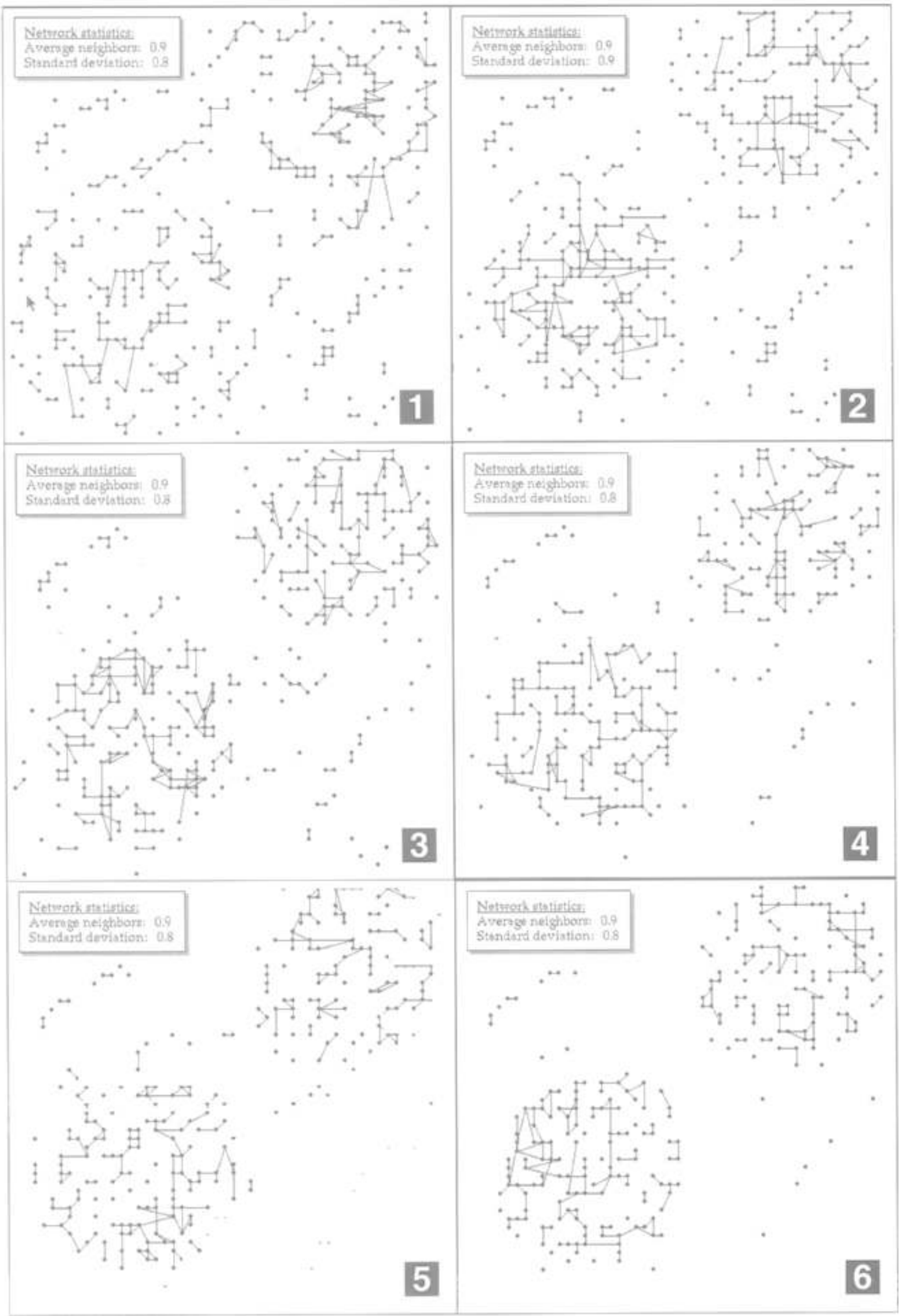
The connection network reveals something not visible in the earlier animations of agents on the sugarscape. If, for instance, message-passing is permitted only between neighbors, what is the chance that a message could make its way across the entire sugarscape? If we whisper it in the ear of a southwestern agent, will a northeastern one ever hear it? If the world is divided into two spatially separated and noninteracting networks, then neighborwise communication will be limited, and information may be localized in a concrete sense, a phenomenon with important implications in a number of spheres. When agent *interaction* (for example, trade) occurs over such networks, the term "connection network" seems less apt than the term "social network."[30] In essence, the connections describe a topology of social interactions.

---

ingly, puts k on its list. Then, when it's k's turn to move, it hops out of i's neighborhood, so when it builds its (post-move) list, i is not on it. In the resulting graph, then, the edge from i to k will go beyond i's neighborhood, and i will not be on k's list (asymmetry).

30. The literature on social networks is large; Scott [1992] and Kochen [1989] are good introductions. Recent work especially relevant to the dynamic networks presented here

Animation II-5. Evolution of Social Networks of Neighbors under Rules ({$\mathbf{G}_1$}, {**M**})



Migration

The skewed wealth distributions above are examples of emergent structures. We turn now to a different kind of emergent structure, this one spatial in nature. To "grow" it we need to give the agents a maximum vision of ten rather than the value of six used above. Now, instead of the random initial distribution of agents on the sugarscape used earlier, suppose they are initially clustered in the dense block shown in the first frame of animation II-6. In all other respects the agents and sugarscape are exactly as in animation II-2. How will this block start affect the dynamics?

A succession of coherent waves results, a phenomenon we did not expect. Reflecting on the local rule, however, the behavior is understandable. Agents in the leading edge proceed to the best unoccupied sugar site within their vision. This leaves a "bald zone" where they had been. The agents behind them must wait until the bald spot grows back under $\mathbf{G}_1$ before they have any incentive, under rule **M**, to move to those lattice points, and so on for the agents behind them. Hence, the series of waves.[31]

While these waves seem to qualify as emergent structures, the diagonal *direction* in which they propagate is perhaps even more interesting. Recall that during a single application of **M** the *individual* agent can only move north, south, east, or west. Yet the collective wave is clearly moving northeast—a heading unavailable to individuals! On closer examination, the collective northeast direction results from a complex interweaving of agents, none of whom can move in this direction. This is shown in figure II-8. Here, the local rule *precludes* individual behavior mimicking the collective behavior.[32]

---

includes Banks and Carley [1994a, 1994b], Sanil, Banks, and Carley [1994], and Carley *et al.* [1994].

31. In pure cellular automata (CA) models, waves are phenomena of significant interest. Recently Sato and Iwasa [1993] have produced these in a CA model of forest ecology. Recent attempts by mathematical biologists to model the wavelike movement of certain mammal herds include Gueron and Levin [1993, 1994] and Gueron, Levin, and Rubenstein [1993]. For an economic model of "herding," see Kirman [1993].

32. Thus emergence, in this case, is the opposite of self-similarity, in which a given pattern is observed on all scales (that is, all orders of magnification) as in fractals.

Animation II-6. Emergent Diagonal Waves of Migrators under Rules ({G_1}, {**M**}) from an Initial Distribution of Agents in a Block

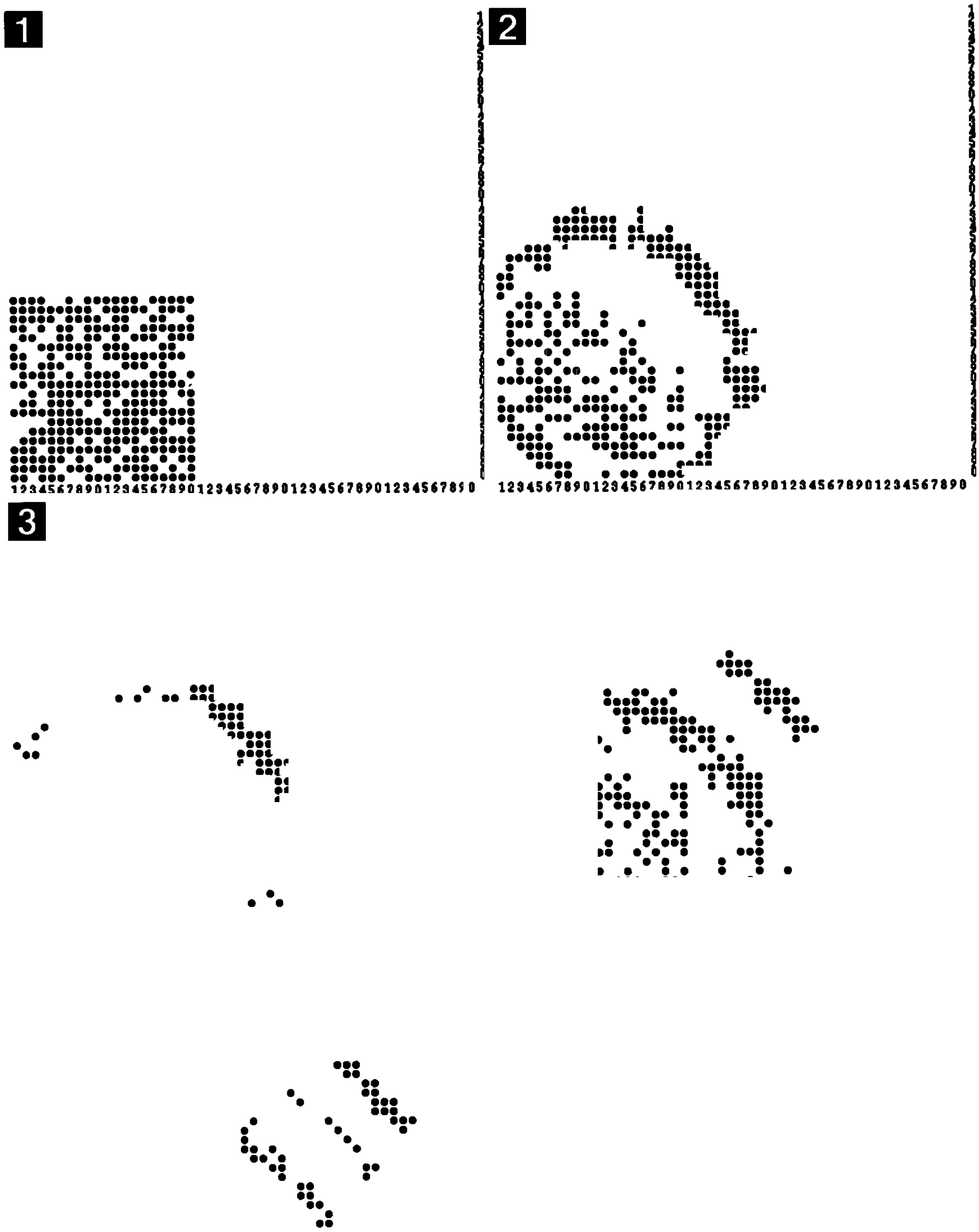
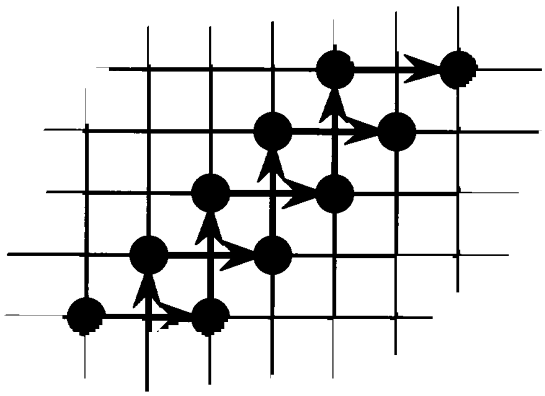


Figure II-8. Interweaving Action of Agents



Seasonal Migration

As another example of macrobehavior patterns arising from simple local rules, let us see if our agents can migrate with the seasons. First, to create artificial seasons, we split the familiar sugarscape into a north and a south by drawing an imaginary equator, a horizontal line cutting the sugarscape in half. For the opening season, the sugarscape grows back at unit rate in the north and at one-eighth that rate in the south; it is "bloom" season in the north and "drought" in the south. Then, after fifty time periods, the situation is reversed; the seasons change. The south grows back at unit rate and the north regenerates at one-eighth that rate. And so it goes, season after season. The general rule can be stated as follows:

> Sugarscape seasonal growback rule $\mathbf{S}_{\alpha\beta\gamma}$: Initially it is *summer* in the top half of the sugarscape and *winter* in the bottom half. Then, every γ time periods the seasons flip—in the region where it was summer it becomes winter and vice versa. For each site, if the season is summer then sugar grows back at a rate of α units per time interval; if the season is winter then the growback rate is α units per β time intervals.[33]

33. This rule can be stated formally. Noting the time by t, for the sites in the top half (bottom half) of the sugarscape if the value of the quantity

$$\frac{t \bmod (2\gamma)}{\gamma}$$

is less than 1, then the season is summer (winter). Otherwise the season is winter (summer). If the season is summer then the growback rate is α units per time interval. If it is winter it is α units per β time intervals. Here the "mod" operator symbolizes the remainder resulting from an integer division operation; x mod y is the remainder upon division

The question we ask now is this: If the simulation is begun with the same simple agents randomly distributed on the sugarscape, will they migrate back and forth with the seasons? Animation II-7 gives the answer.

Again, the agents, operating under the same simple local rule **M**, exhibit collective behavior far more complex—and far more realistic—than we had expected. Yes, we get migrators. But, we also get "hibernators"! The high vision ("bird-like") creatures migrate. The low vision–low metabolism ("bear-like") creatures hibernate. Agents with low vision and high metabolism generally die; they are selected against.

Notice, however, that a hibernator born in the south rarely goes north, and a hibernator born in the north rarely goes south. Northern and southern hibernators, in short, would rarely meet and, hence, would rarely *mate*. They would form, in effect, separate mating pools and, in evolutionary time, "speciation" would occur.

Pollution: A Negative Externality

So far, in simply grazing the sugarscape, agents have been interacting with one another indirectly. That is, agents move on the basis of what they find in their local environment, and what they find is the result of the actions of other agents.[34] Such indirect interactions are a kind of externality.[35] Externalities can be positive or negative. Pollution is an example of the latter type. A polluter degrades the environment in which other agents live and in so doing reduces the welfare of other agents, and possibly its own welfare as well.

There are many ways in which pollution can be added to the sugarscape. It might be produced by agent movement, agent gathering activities, agent sugar consumption, sugar growback, or some other mechanism. There might be many types of pollutants, each produced at different rates. Pollutants may get transported to other sites at various rates and could possess a natural growth or decay rate. And in order for the pollution to be a negative externality it must affect the agent

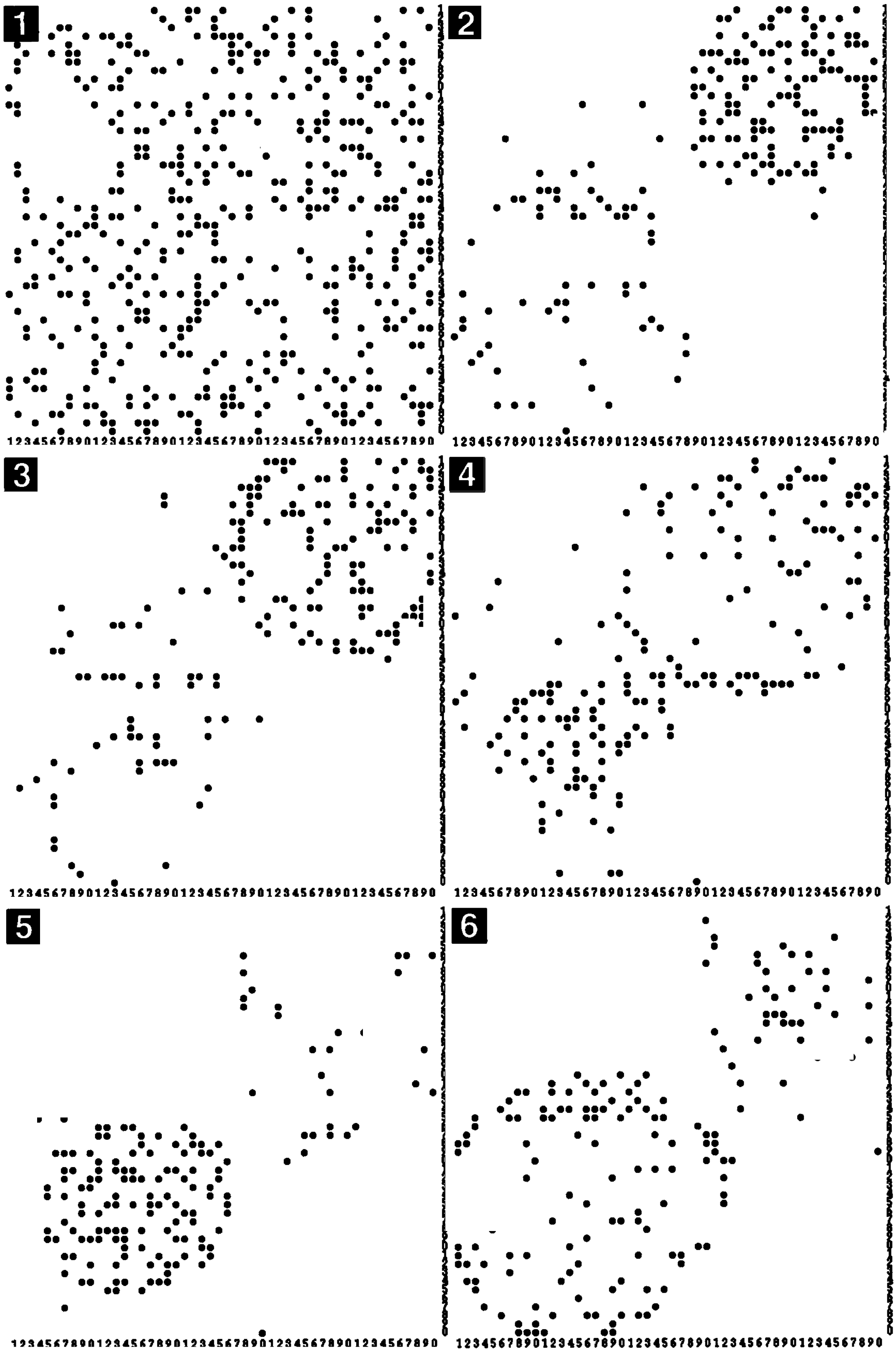
---

of x by y. For example, 5 mod 2 = 1. Note that for γ larger than the duration of a run the seasons never change.

34. But the agents do not interact *directly*. In subsequent chapters the agents interact with one another, through behaviors such as sex and trade.

35. When the action of one agent affects the welfare—here, the sugar wealth—of a second agent and is not constrained socially (through a market, for instance), then an externality exists [Campbell, 1987: 57].

Animation II-7. Migration and Hibernation Resulting from Rules $(\{\mathbf{S}_{1,8,50}\}, \{\mathbf{M}\})$ and a Random Initial Distribution of Agents



adversely. It could enter the agents' bodies and degrade their vision, say, or increase their metabolisms, as if it made them sick. Or it could simply be a negative amenity—the agents could just dislike it and so try to avoid it whenever possible. In that case it would be as if a second commodity had been added to the sugarscape, an economic "bad."

We have chosen a very simple pollution formation rule. There is one type of pollutant. It is produced by both gathering and consumption activities, in proportion to the amount of sugar gathered and consumed, respectively. It accumulates on the sites at which the gathering and consumption activities occur. Stated formally, the rule is:

> Pollution formation rule $\mathbf{P}_{\alpha\beta}$: When sugar quantity s is gathered from the sugarscape, an amount of production pollution is generated in quantity αs. When sugar amount m is consumed (metabolized), consumption pollution is generated according to βm. The total pollution on a site at time t, p^t, is the sum of the pollution present at the previous time, plus the pollution resulting from production and consumption activities, that is, $p^t = p^{t-1} + \alpha s + \beta m$.[36]

Pollution affects the agents in a very simple way: it has negative amenity value. That is, they just do not like it! The simplest way to incorporate this is to modify the agent movement rule somewhat, to let the pollution devalue—in the agents' eyes—the sites where it is present. Instead of moving to the site of maximum sugar, we now specify that the agents select the site having the maximum sugar to pollution ratio.[37] That is, those sites with high sugar levels and low pollution levels are the most attractive. The modified agent movement rule now reads (with the changes to the previous rule italicized):

> Agent movement rule **M**, modified for pollution:
> - Look out as far as vision permits in the four principal lattice directions and identify the unoccupied site(s) having the maximum *sugar to pollution ratio;*

---

36. The Sugarscape software system offers a more general pollution formation rule than this. There can be multiple types of pollutants, each produced at different rates. In Chapter IV, when investigating the effect of pollution on prices and economic trade activity, we shall make use of the general pollution formation rule. This more general rule is described fully in Appendix B, along with the general forms of all other rules.

37. To be precise, the ratio computed is actually $s/(1+p)$ to preclude division by zero in the no pollution case.

- If the maximum *sugar to pollution ratio* appears on multiple sites, then select the nearest one;
- Move to this site;
- Collect all the sugar at this new position.

The final ingredient to add is some form of pollution transport. Without transport or dissipation pollution simply accumulates without bound at the sites where it is produced. Perhaps the simplest form of transport is diffusion. Diffusion on a lattice like the sugarscape is simply implemented as a local averaging procedure. That is, diffusion transports pollution from sites of high levels to sites of low levels. This can be stated algorithmically as:

Pollution diffusion rule $\mathbf{D}_\alpha$:
- Each α time periods and at each site, compute the pollution flux—the average pollution level over all von Neumann neighboring sites;
- Each site's flux becomes its new pollution level.

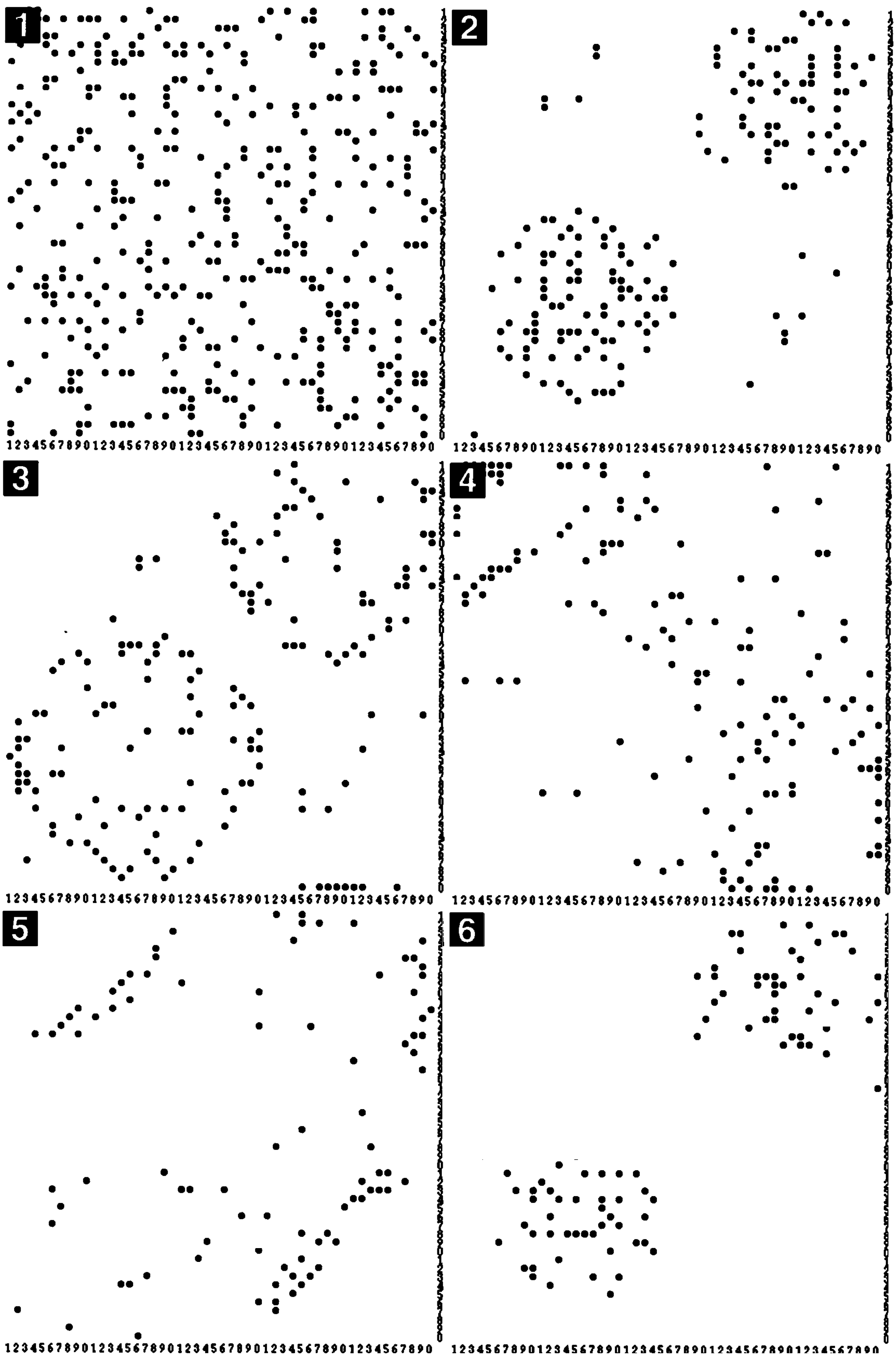
The reader with a knowledge of cellular automata (CA) will notice that this rule, which relates the pollution on any site to that on other sites, makes the sugarscape a true CA.[38] Note that as α is increased the rate of diffusion is decreased, so $\mathbf{D}_1$ is the fastest diffusion possible.

These simple rules, taken together, prove sufficient to "grow" a reasonable story of agent response to an agent polluted environment. In animation II-8 agents execute the modified movement rule, **M**. The sugarscape grows back according to $\mathbf{G}_1$. At $t = 50$ pollution begins (rule $\mathbf{P}_{11}$ is turned on). Then, at $t = 100$, diffusion begins (rule $\mathbf{D}_1$ is switched on).

At first the agents are merrily hiving the sugar hills, as usual. Pollution levels are low, and the behavior produced by the modified movement rule is not much different from that produced by the original movement rule. Eventually, however, pollution levels build up and the agents progressively abandon the polluted zone. They are forced off the sugar

38. The simple growback rule $\mathbf{G}_\alpha$ does this only trivially since growback rates on any site are independent of either growback rates or sugar levels on neighboring sites. We have experimented with growback rules that do, in fact, have this dependence. One such rule is as follows: if the level of sugar is not 0, then apply $\mathbf{G}_\alpha$; however, if the sugar level is 0, then grow back only if some neighboring site has a nonzero sugar level. It is as if each barren site must be "seeded" by neighbors.

Animation II-8. Agent Migration in a Polluted Environment; Rule System $(\{\mathbf{G}_1, \mathbf{D}_1\}, \{\mathbf{M}, \mathbf{P}_{11}\})$



peaks, migrating into relatively pristine areas where no pollution from sugar production has accumulated. However, as agents continue to eat from their personal accumulations, they progressively despoil even this area through consumption pollution. Because there is little food for the agents out in this resource-poor hinterland, competition is intense. Many die of starvation—the carrying capacity of the polluted environment is lower.

Subsequently, when diffusion is turned on, the pollution quickly spreads more or less uniformly around the landscape and many agents move back to the regions of highest sugar. As they continue to gather and metabolize sugar, pollution increases while diffusing over the entire landscape. There is a kind of rising "red tide" that diminishes the welfare of all agents still alive on the sugarscape.[39]

It turns out that the same type of dynamic pattern appears when the externality involved is positive rather than negative. Positive externalities—increasing returns or network externalities, for example—give agents reasons to associate with one another, to spatially cluster. Of course, the two types of externalities can be combined: there may be positive externalities associated with production but negative externalities associated with consumption. Are cities the "balance points" between these opposed effects?[40]

A Social Interpretation

It is possible to give a social intrepretation to the migratory dynamics just discussed. When seasons change or pollution levels rise, large numbers of agents migrate to particular regions of the sugarscape. In effect, they

39. It is at this point that a clear "tragedy of the commons" interpretation of life on the sugarscape manifests itself. Metabolisms are constant and so, for a fixed agent population, sugar consumption and the pollution it generates are fixed; thus, the only way pollution levels can be decreased is to reduce the amount of pollution generated through production (harvesting) activities. If, for instance, those agents who harvest more sugar than they consume in following **M** were to follow some alternative rule, harvesting only, say, half the sugar they find beyond their metabolic needs, then overall pollution levels would fall. While this behavioral rule would make these agents worse off, in comparison to **M**, by lowering their income, perhaps all agents could be made better off through side payments. Alternative rules—institutions—for managing such common property resource problems in general are investigated at length by Ostrom [1990] and Ostrom, Gardner, and Walker [1994].

40. For a very general analysis along these lines, see Papageorgiou and Smith [1993]; see also Krugman [1996].

are environmental refugees; an environmental catastrophe has struck their zone and they flood into better areas. In Chapter III, we introduce combat. Its intensity can grow when competition for resources becomes severe. An influx of environmental refugees suddenly boosts the agent density in the receiving zone and, naturally, competition for sugar intensifies dramatically. The model suggests, therefore, that *environmental degradation can have serious security implications.*[41]

Summary

These exercises make clear that a wide range of collective structures and collective patterns of behavior can emerge from the spatio-temporal interaction of agents operating, individually, under simple local rules. For example, only one agent rule, **M**, has been used, and it is about as primitive a rule as we could construct. Paraphrasing, it amounts to the instruction: "Look around for the best free site; go there and harvest the sugar." And yet, all sorts of unexpected things emerge from the *interaction* of these agents: basic principles like the existence of environmental carrying capacities; skewed distributions of wealth; coherent group structures like waves that move in directions unavailable to individuals; and biological processes like hibernation and migration (refugees). And that strikes us as surprising. The nature of the surprise is worth discussing.

The Surprising Sufficiency of Simple Rules

We have succeeded in "growing" a number of quite familiar collective behaviors, such as migration, and familiar macroscopic structures, such as skewed wealth distributions. And we grow many more familiar macroscopic entities below. Now, upon first exposure to these familiar social, or macroscopic, structures—be they migrations, skewed wealth distributions, or the like—some people say, "Yes, that looks familiar. But, I've seen it before. What's the surprise?"

The surprise consists precisely in the emergence of familiar macrostructures *from the bottom up*—from simple local rules that out-

41. The connection between environmental change and security is the subject of several recent studies by Homer-Dixon [1991, 1994]. The mathematical structure of spatial patterns resulting from conflicts has been studied by Vickers, Hutson, and Budd [1993].

wardly appear quite remote from the social, or collective, phenomena they generate. In short, *it is not the emergent macroscopic object per se that is surprising, but the generative sufficiency of the simple local rules.*

Of course, for the model to be of practical use to social scientists, a minimum requirement is that it generate familiar phenomena with some fidelity. If the model cannot generate the familiar world as a base case, then how can we use it to examine the effects of various policies, for example?

Furthermore, there may be familiar and important social phenomena that are hard to study with standard tools. For instance, we can do more than turn pollution on and off in our model; we can track its effect on prices (see Chapter IV). We find that a pollution-induced shortage of one good increases its price, an effect described in standard economics texts. But when we then diffuse the pollutant, relieving the shortage, relative prices do not return to equilibrium instantly—on the contrary, the adjustment may take a long time. And adjustment *dynamics* are difficult to model within the standard equilibrium framework. Moreover, had we been unable to get the familiar result (that is, the "right" price response to shortage), this lag in adjustment would not be credible.

The main point, however, is that, when—in subsequent chapters—we grow a familiar macrostructure, it is the sufficiency of the local rules that is surprising.

Artificial Social Systems as Laboratories

Of course, in this exposition, we presented the rules *before* carrying out any simulations. We might have proceeded differently. Imagine that we had begun the entire discussion by simply running animation II-2, which shows a buzz of agents "hiving" the sugar mountains, and that we had then bluntly asked, "What's happening here?" Would you have guessed that the agents are all following rule **M**? We do not think we would have been able to divine it. But that really is all that is happening. Isn't it just possible that something comparably simple is "all that is happening" in other complex systems, such as stock markets or political systems? As social scientists, this is the problem we normally confront. We observe the complex collective—already emerged—behavior, and we seek simple local rules of individual behavior (for example, maximize profit) that could generate it.

The Sugarscape model can function as a kind of *laboratory* where we "grow" fundamental social structures *in silico*, thereby learning which micromechanisms are sufficient to generate macrostructures of interest. Such experiments can lead to hypotheses of social concern that may subsequently be tested statistically against data.

In Chapter III we expand the behavioral repertoire of our agents, allowing us to study more complex social phenomena.