EXCERPTED FROM

STEPHEN WOLFRAM
A NEW KIND OF SCIENCE

SECTION 9.3

Irreversibility and the Second Law of Thermodynamics

In some cases, the behavior is fairly simple, and the patterns obtained have simple repetitive or nested structures. But in many cases, even with simple initial conditions, the patterns produced are highly complex, and seem in many respects random.

The reversibility of the underlying rules has some obvious consequences, such as the presence of triangles pointing sideways but not down. But despite their reversibility, the rules still manage to produce the kinds of complex behavior that we have seen in cellular automata and many other systems throughout this book.

So what about localized structures?

The picture on the facing page demonstrates that these can also occur in reversible systems. There are some constraints on the details of the kinds of collisions that are possible, but reversible rules typically tend to work very much like ordinary ones.

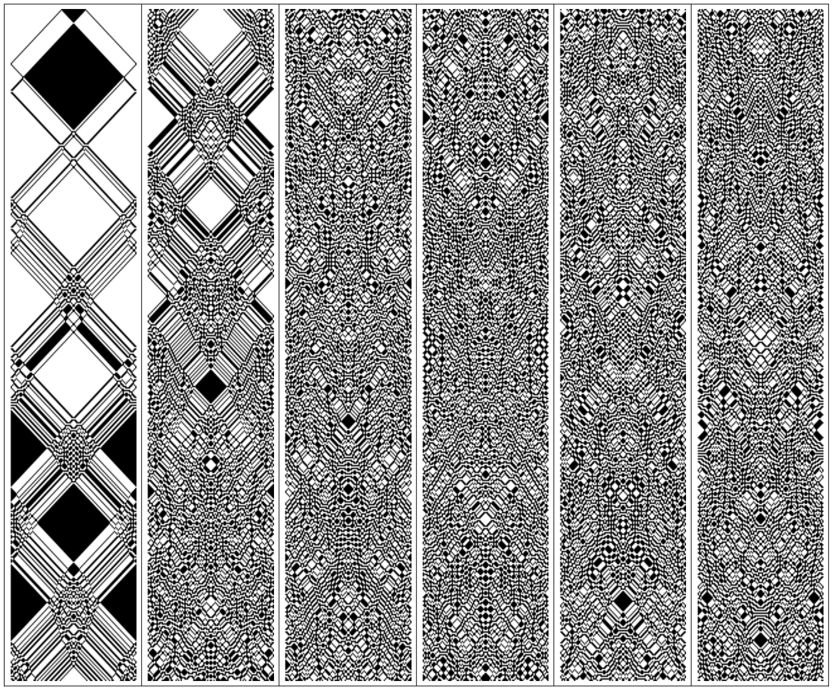
So in the end it seems that even though only a very small fraction of possible systems have the property of being reversible, such systems can still exhibit behavior just as complex as one sees anywhere else.

Irreversibility and the Second Law of Thermodynamics

All the evidence we have from particle physics and elsewhere suggests that at a fundamental level the laws of physics are precisely reversible. Yet our everyday experience is full of examples of seemingly irreversible phenomena. Most often, what happens is that a system which starts in a fairly regular or organized state becomes progressively more and more random and disorganized. And it turns out that this phenomenon can already be seen in many simple programs.

The picture at the top of the next page shows an example based on a reversible cellular automaton of the type discussed in the previous section. The black cells in this system act a little like particles which bounce around inside a box and interact with each other when they collide.

At the beginning the particles are placed in a simple arrangement at the center of the box. But over the course of time the picture shows that the arrangement of particles becomes progressively more random.

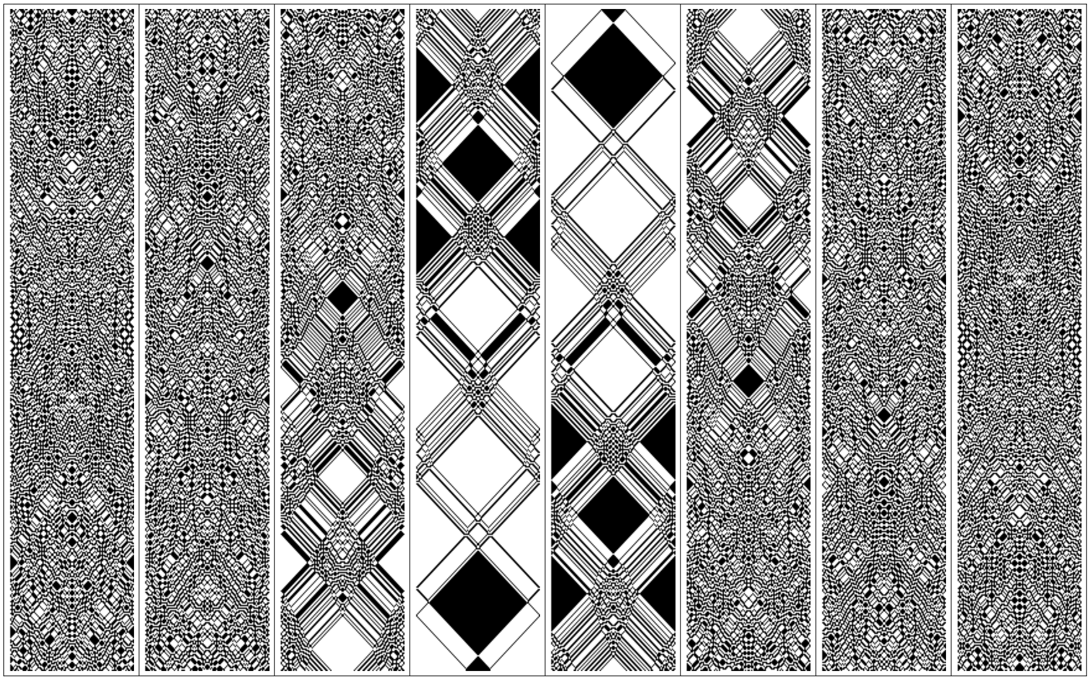


A reversible cellular automaton that exhibits seemingly irreversible behavior. Starting from an initial condition in which all black cells or particles lie at the center of a box, the distribution becomes progressively more random. Such behavior appears to be the central phenomenon responsible for the Second Law of Thermodynamics. The specific cellular automaton used here is rule 122R. The system is restricted to a region of size 100 cells.

Typical intuition from traditional science makes it difficult to understand how such randomness could possibly arise. But the discovery in this book that a wide range of systems can generate randomness even with very simple initial conditions makes it seem considerably less surprising.

But what about reversibility? The underlying rules for the cellular automaton used in the picture above are precisely reversible. Yet the picture itself does not at first appear to be at all reversible. For there appears to be an irreversible increase in randomness as one goes down successive panels on the page.

The resolution of this apparent conflict is however fairly straightforward. For as the picture on the facing page demonstrates, if the



An extended version of the picture on the facing page, in which the reversibility of the underlying cellular automaton is more clearly manifest. An initial condition is carefully constructed so that halfway through the evolution shown a simple arrangement of particles will be produced. If one starts with this arrangement, then the randomness of the system will effectively increase whether one goes forwards or backwards in time from that point.

simple arrangement of particles occurs in the middle of the evolution, then one can readily see that randomness increases in exactly the same way—whether one goes forwards or backwards from that point.

Yet there is still something of a mystery. For our everyday experience is full of examples in which randomness increases much as in the second half of the picture above. But we essentially never see the kind of systematic decrease in randomness that occurs in the first half.

By setting up the precise initial conditions that exist at the beginning of the whole picture it would certainly in principle be possible to get such behavior. But somehow it seems that initial conditions like these essentially never actually occur in practice.

There has in the past been considerable confusion about why this might be the case. But the key to understanding what is going on is simply to realize that one has to think not only about the systems one is studying, but also about the types of experiments and observations that one uses in the process of studying them.

The crucial point then turns out to be that practical experiments almost inevitably end up involving only initial conditions that are fairly simple for us to describe and construct. And with these types of initial conditions, systems like the one on the previous page always tend to exhibit increasing randomness.

But what exactly is it that determines the types of initial conditions that one can use in an experiment? It seems reasonable to suppose that in any meaningful experiment the process of setting up the experiment should somehow be simpler than the process that the experiment is intended to observe.

But how can one compare such processes? The answer that I will develop in considerable detail later in this book is to view all such processes as computations. The conclusion is then that the computation involved in setting up an experiment should be simpler than the computation involved in the evolution of the system that is to be studied by the experiment.

It is clear that by starting with a simple state and then tracing backwards through the actual evolution of a reversible system one can find initial conditions that will lead to decreasing randomness. But if one looks for example at the pictures on the last couple of pages the complexity of the behavior seems to preclude any less arduous way of finding such initial conditions. And indeed I will argue in Chapter 12 that the Principle of Computational Equivalence suggests that in general no such reduced procedure should exist.

The consequence of this is that no reasonable experiment can ever involve setting up the kind of initial conditions that will lead to decreases in randomness, and that therefore all practical experiments will tend to show only increases in randomness.

It is this basic argument that I believe explains the observed validity of what in physics is known as the Second Law of Thermodynamics. The law was first formulated more than a century ago, but despite many related technical results, the basic reasons for its validity have until now remained rather mysterious.

The field of thermodynamics is generally concerned with issues of heat and energy in physical systems. A fundamental fact known since the mid-1800s is that heat is a form of energy associated with the random microscopic motions of large numbers of atoms or other particles.

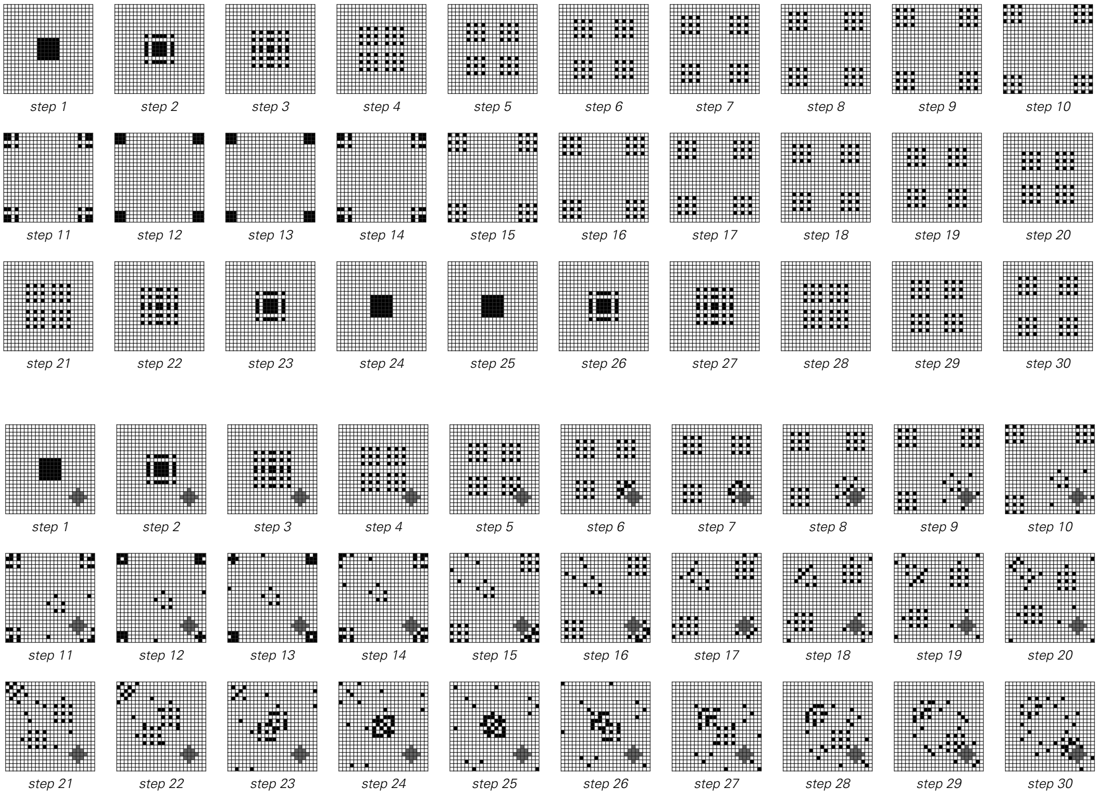
One formulation of the Second Law then states that any energy associated with organized motions of such particles tends to degrade irreversibly into heat. And the pictures at the beginning of this section show essentially just such a phenomenon. Initially there are particles which move in a fairly regular and organized way. But as time goes on, the motion that occurs becomes progressively more random.

There are several details of the cellular automaton used above that differ from actual physical systems of the kind usually studied in thermodynamics. But at the cost of some additional technical complication, it is fairly straightforward to set up a more realistic system.

The pictures on the next two pages show a particular two-dimensional cellular automaton in which black squares representing particles move around and collide with each other, essentially like particles in an ideal gas. This cellular automaton shares with the cellular automaton at the beginning of the section the property of being reversible. But it also has the additional feature that in every collision the total number of particles in it remains unchanged. And since each particle can be thought of as having a certain energy, it follows that the total energy of the system is therefore conserved.

In the first case shown, the particles are taken to bounce around in an empty square box. And it turns out that in this particular case only very simple repetitive behavior is ever obtained. But almost any change destroys this simplicity.

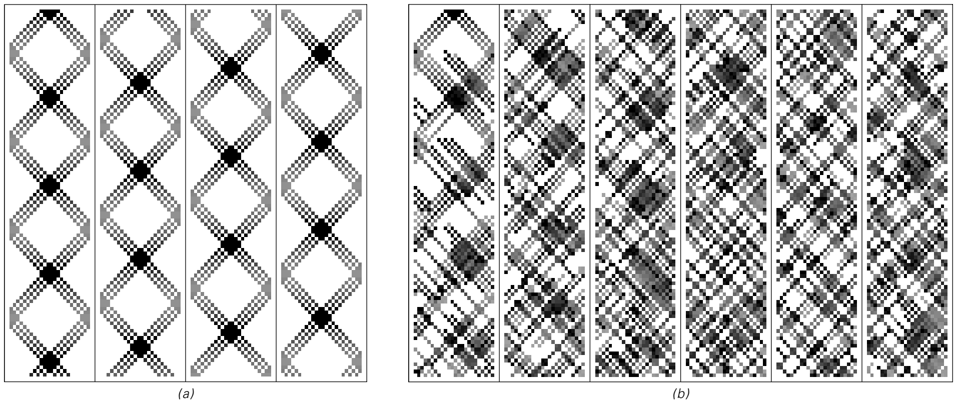
And in the second case, for example, the presence of a small fixed obstacle leads to rapid randomization in the arrangement of particles—very much like the randomization we saw in the one-dimensional cellular automaton that we discussed earlier in this section.



The behavior of a simple two-dimensional cellular automaton that emulates an ideal gas of particles. In the top group of pictures, the particles bounce around in an empty square box. In the bottom group of pictures, the box contains a small fixed obstacle. In the top group of pictures, the arrangement of particles shows simple repetitive behavior. In the bottom group, however, it becomes progressively more random with time. The underlying rules for the cellular automaton used here are reversible, and conserve the total number of particles. The specific rules are based on 2 × 2 blocks—a two-dimensional generalization of the block cellular automata to be discussed in the next section. For each 2 × 2 block the configuration of particles is taken to remain the same at a particular step unless there are exactly two particles arranged diagonally within the block, in which case the particles move to the opposite diagonal.

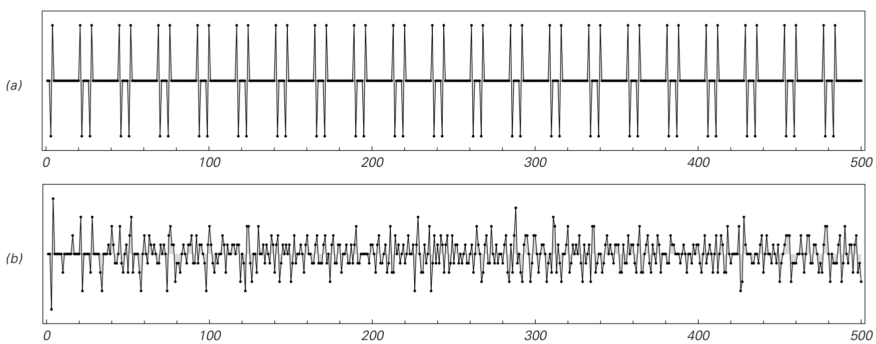
So even though the total of the energy of all particles remains the same, the distribution of this energy becomes progressively more random, just as the usual Second Law implies.

An important practical consequence of this is that it becomes increasingly difficult to extract energy from the system in the form of systematic mechanical work. At an idealized level one might imagine trying to do this by inserting into the system some kind of paddle which would experience force as a result of impacts from particles.



Time histories of the cellular automata from the facing page. In each case a slice is taken through the midline of the box. Black cells that are further from the midline are shown in progressively lighter shades of gray. Case (a) corresponds to an empty square box, and shows simple repetitive behavior. Case (b) corresponds to a box containing a fixed obstacle, and in this case rapid randomization is seen. Each panel corresponds to 100 steps in the evolution of the system; the box is 24 cells across.

The pictures below show how such force might vary with time in cases (a) and (b) above. In case (a), where no randomization occurs, the force can readily be predicted, and it is easy to imagine harnessing it to produce systematic mechanical work. But in case (b), the force quickly randomizes, and there is no obvious way to obtain systematic mechanical work from it.



The force on an idealized paddle placed on the midline of the systems shown above. The force reflects an imbalance in the number of particles at each step arriving at the midline from above and below. In case (a) this imbalance is readily predictable. In case (b), however, it rapidly becomes for most practical purposes random. This randomness is essentially what makes it impossible to build a physical perpetual motion machine which continually turns heat into mechanical work.

One might nevertheless imagine that it would be possible to devise a complicated machine, perhaps with an elaborate arrangement of paddles, that would still be able to extract systematic mechanical work even from an apparently random distribution of particles. But it turns out that in order to do this the machine would effectively have to be able to predict where every particle would be at every step in time.

And as we shall discuss in Chapter 12, this would mean that the machine would have to perform computations that are as sophisticated as those that correspond to the actual evolution of the system itself. The result is that in practice it is never possible to build perpetual motion machines that continually take energy in the form of heat—or randomized particle motions—and convert it into useful mechanical work.

The impossibility of such perpetual motion machines is one common statement of the Second Law of Thermodynamics. Another is that a quantity known as entropy tends to increase with time.

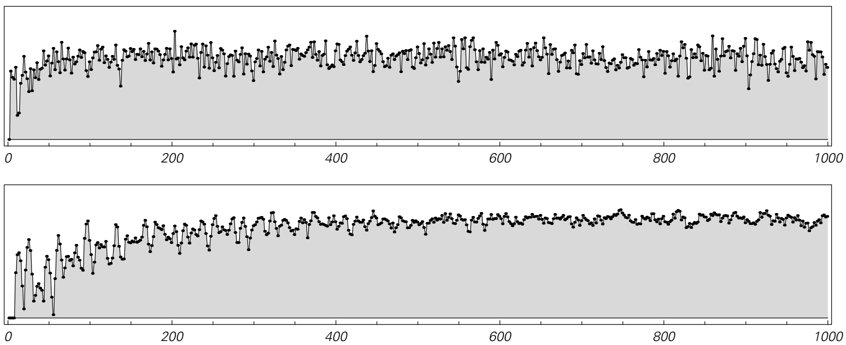
Entropy is defined as the amount of information about a system that is still unknown after one has made a certain set of measurements on the system. The specific value of the entropy will depend on what measurements one makes, but the content of the Second Law is that if one repeats the same measurements at different times, then the entropy deduced from them will tend to increase with time.

If one managed to find the positions and properties of all the particles in the system, then no information about the system would remain unknown, and the entropy of the system would just be zero. But in a practical experiment, one cannot expect to be able to make anything like such complete measurements.

And more realistically, the measurements one makes might for example give the total numbers of particles in certain regions inside the box. There are then a large number of possible detailed arrangements of particles that are all consistent with the results of such measurements. The entropy is defined as the amount of additional information that would be needed in order to pick out the specific arrangement that actually occurs.

We will discuss in more detail in Chapter 10 the notion of amount of information. But here we can imagine numbering all the possible arrangements of particles that are consistent with the results of our measurements, so that the amount of information needed to pick out a single arrangement is essentially the length in digits of one such number.

The pictures below show the behavior of the entropy calculated in this way for systems like the one discussed above. And what we see is that the entropy does indeed tend to increase, just as the Second Law implies.



The entropy as a function of time for systems of the type shown in case (b) from page 447. The top plot is exactly for case (b); the bottom one is for a system three times larger in size. The entropy is found in each case by working out how many possible configurations of particles are consistent with measurements of the total numbers of particles in a 6 × 6 grid of regions within the system. Just as the Second Law of Thermodynamics suggests, the entropy tends to increase with time. Note that the plots above would be exactly symmetrical if they were continued to the left: the entropy would increase in the same way going both forwards and backwards from the simple initial conditions used.

In effect what is going on is that the measurements we make represent an attempt to determine the state of the system. But as the arrangement of particles in the system becomes more random, this attempt becomes less and less successful.

One might imagine that there could be a more elaborate set of measurements that would somehow avoid these problems, and would not lead to increasing entropy. But as we shall discuss in Chapter 12, it again turns out that setting up such measurements would have to involve the same level of computational effort as the actual evolution of the system itself. And as a result, one concludes that the entropy associated with measurements done in practical experiments will always tend to increase, as the Second Law suggests.

In Chapter 12 we will discuss in more detail some of the key ideas involved in coming to this conclusion. But the basic point is that the phenomenon of entropy increase implied by the Second Law is a more or less direct consequence of the phenomenon discovered in this book that even with simple initial conditions many systems can produce complex and seemingly random behavior.

One aspect of the generation of randomness that we have noted several times in earlier chapters is that once significant randomness has been produced in a system, the overall properties of that system tend to become largely independent of the details of its initial conditions.

In any system that is reversible it must always be the case that different initial conditions lead to at least slightly different states—otherwise there would be no unique way of going backwards. But the point is that even though the outcomes from different initial conditions differ in detail, their overall properties can still be very much the same.

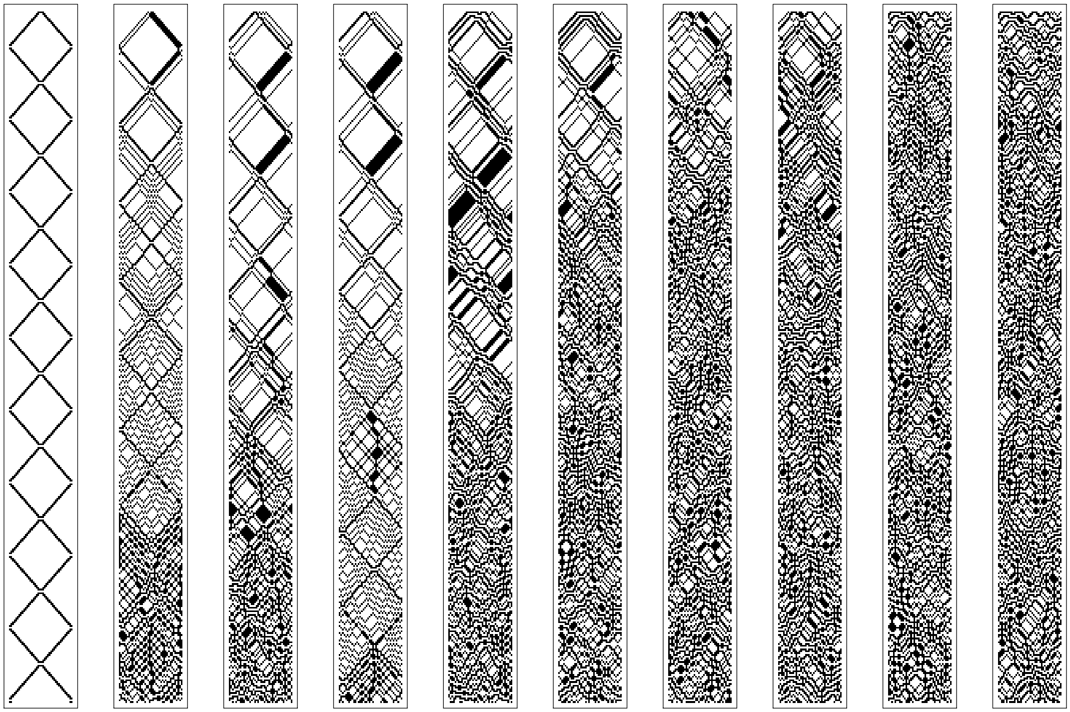
The pictures on the facing page show an example of what can happen. Every individual picture has different initial conditions. But whenever randomness is produced the overall patterns that are obtained look in the end almost indistinguishable.

The reversibility of the underlying rules implies that at some level it must be possible to recognize outcomes from different kinds of initial conditions. But the point is that to do so would require a computation far more sophisticated than any that could meaningfully be done as part of a practical measurement process.

So this means that if a system generates sufficient randomness, one can think of it as evolving towards a unique equilibrium whose properties are for practical purposes independent of its initial conditions.

This fact turns out in a sense to be implicit in many everyday applications of physics. For it is what allows us to characterize all sorts of physical systems by just specifying a few parameters such as temperature and chemical composition—and avoids us always having to know the details of the initial conditions and history of each system.

The existence of a unique equilibrium to which any particular system tends to evolve is also a common statement of the Second Law of



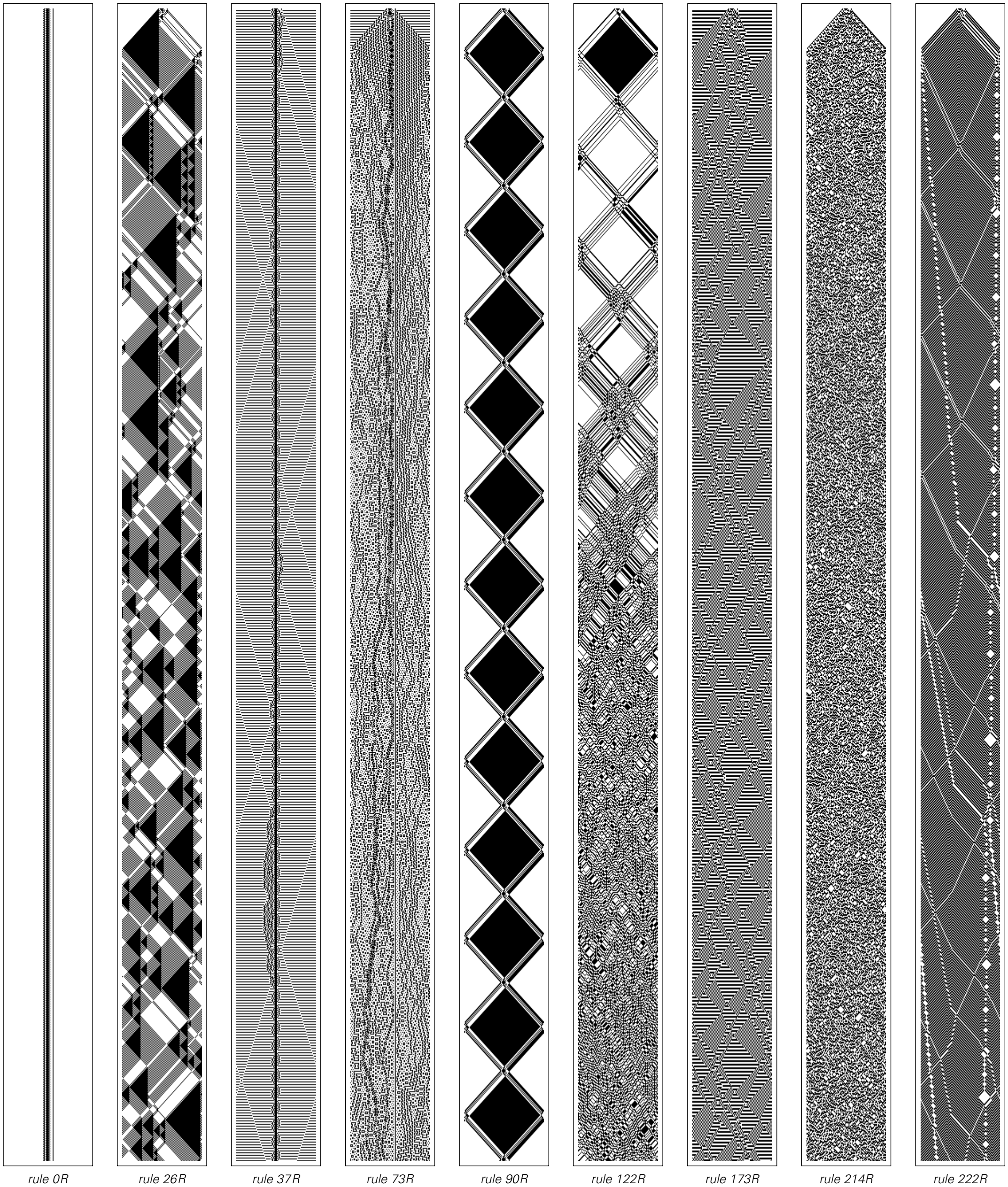
The approach to equilibrium in a reversible cellular automaton with a variety of different initial conditions. Apart from exceptional cases where no randomization occurs, the behavior obtained with different initial conditions is eventually quite indistinguishable in its overall properties. Because the underlying rule is reversible, however, the details with different initial conditions are always at least slightly different—otherwise it would not be possible to go backwards in a unique way. The rule used here is 122R. Successive pairs of pictures have initial conditions that differ only in the color of a single cell at the center.

Thermodynamics. And once again, therefore, we find that the Second Law is associated with basic phenomena that we already saw early in this book.

But just how general is the Second Law? And does it really apply to all of the various kinds of systems that we see in nature?

Starting nearly a century ago it came to be widely believed that the Second Law is an almost universal principle. But in reality there is surprisingly little evidence for this.

Indeed, almost all of the detailed applications ever made of the full Second Law have been concerned with just one specific area: the behavior of gases. By now there is therefore good evidence that gases obey the Second Law—just as the idealized model earlier in this section suggests. But what about other kinds of systems?



Examples of reversible cellular automata with various rules. Some quickly randomize, as the Second Law of Thermodynamics would suggest. But others do not—and thus in effect do not obey the Second Law of Thermodynamics.

The pictures on the facing page show examples of various reversible cellular automata. And what we see immediately from these pictures is that while some systems exhibit exactly the kind of randomization implied by the Second Law, others do not.

The most obvious exceptions are cases like rule 0R and rule 90R, where the behavior that is produced has only a very simple fixed or repetitive form. And existing mathematical studies have indeed identified these simple exceptions to the Second Law. But they have somehow implicitly assumed that no other kinds of exceptions can exist.

The picture on the next page, however, shows the behavior of rule 37R over the course of many steps. And in looking at this picture, we see a remarkable phenomenon: there is neither a systematic trend towards increasing randomness, nor any form of simple predictable behavior. Indeed, it seems that the system just never settles down, but rather continues to fluctuate forever, sometimes becoming less orderly, and sometimes more so.

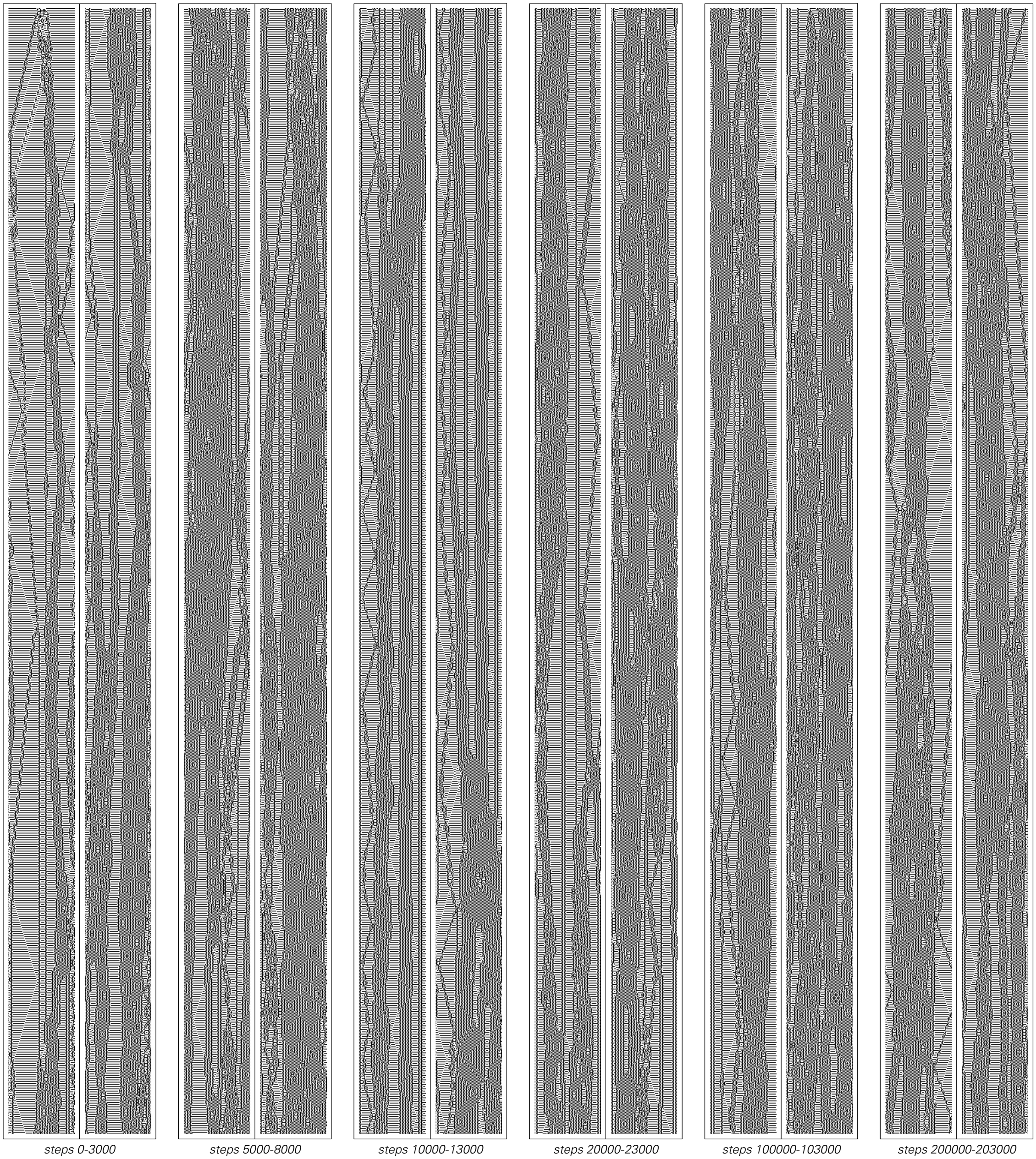
So how can such behavior be understood in the context of the Second Law? There is, I believe, no choice but to conclude that for practical purposes rule 37R simply does not obey the Second Law.

And as it turns out, what happens in rule 37R is not so different from what seems to happen in many systems in nature. If the Second Law was always obeyed, then one might expect that by now every part of our universe would have evolved to completely random equilibrium.

Yet it is quite obvious that this has not happened. And indeed there are many kinds of systems, notably biological ones, that seem to show, at least temporarily, a trend towards increasing order rather than increasing randomness.

How do such systems work? A common feature appears to be the presence of some kind of partitioning: the systems effectively break up into parts that evolve at least somewhat independently for long periods of time.

The picture on page 456 shows what happens if one starts rule 37R with a single small region of randomness. And for a while what one sees is that the randomness that has been inserted persists. But eventually the system instead seems to organize itself to yield just a small number of simple repetitive structures.



More steps in the evolution of the reversible cellular automaton with rule 37R. This system is an example of one that does not in any meaningful way obey the Second Law of Thermodynamics. Instead of exhibiting progressively more random behavior, it appears to fluctuate between quite ordered and quite disordered states.

This kind of self-organization is quite opposite to what one would expect from the Second Law. And at first it also seems inconsistent with the reversibility of the system. For if all that is left at the end are a few simple structures, how can there be enough information to go backwards and reconstruct the initial conditions?

The answer is that one has to consider not only the stationary structures that stay in the middle of the system, but also all various small structures that were emitted in the course of the evolution. To go backwards one would need to set things up so that one absorbs exactly the sequence of structures that were emitted going forwards.

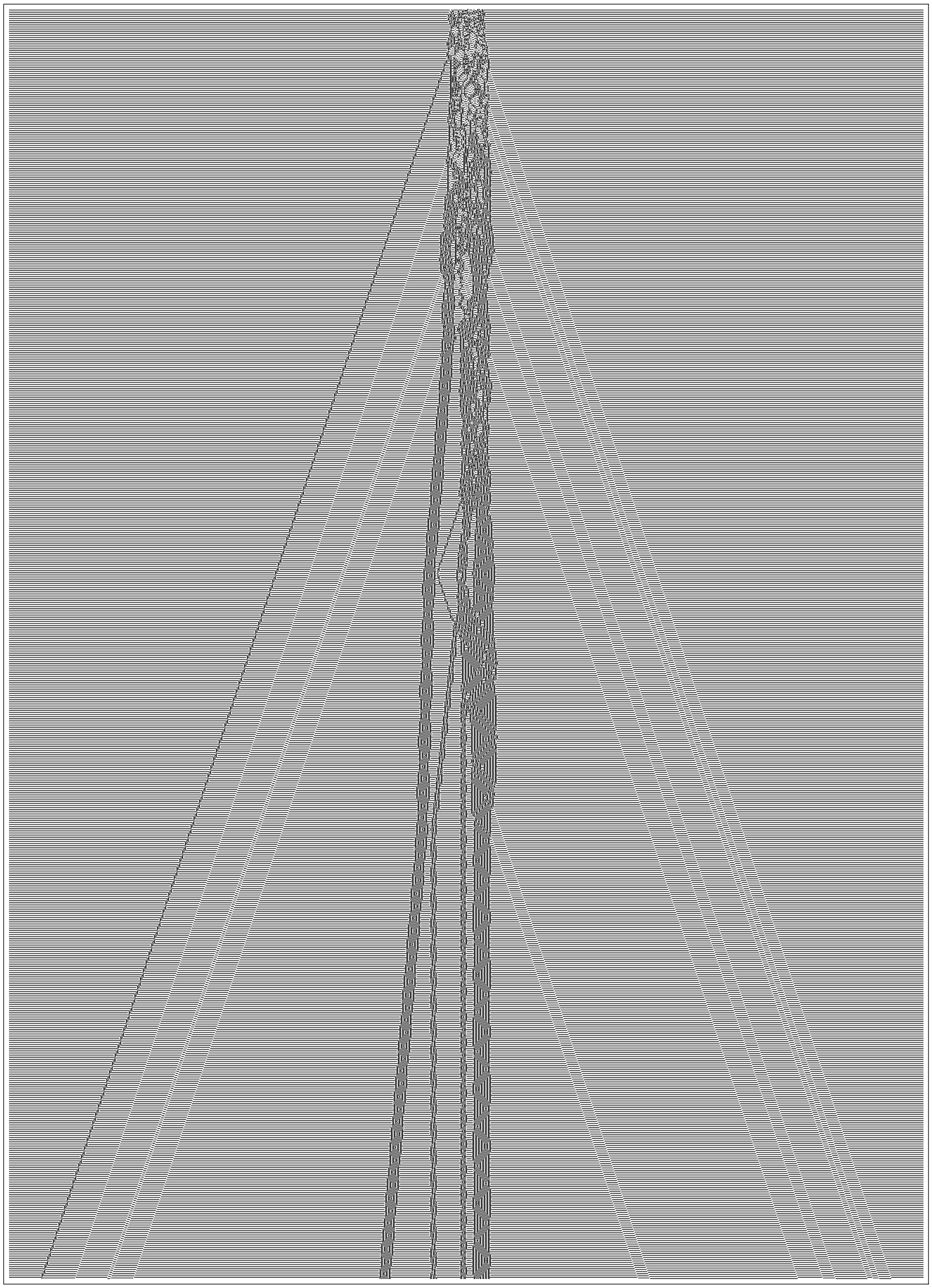
If, however, one just lets the emitted structures escape, and never absorbs any other structures, then one is effectively losing information. The result is that the evolution one sees can be intrinsically not reversible, so that all of the various forms of self-organization that we saw earlier in this book in cellular automata that do not have reversible rules can potentially occur.

If we look at the universe on a large scale, then it turns out that in a certain sense there is more radiation emitted than absorbed. Indeed, this is related to the fact that the night sky appears dark, rather than having bright starlight coming from every direction. But ultimately the asymmetry between emission and absorption is a consequence of the fact that the universe is expanding, rather than contracting, with time.

The result is that it is possible for regions of the universe to become progressively more organized, despite the Second Law, and despite the reversibility of their underlying rules. And this is a large part of the reason that organized galaxies, stars and planets can form.

Allowing information to escape is a rather straightforward way to evade the Second Law. But what the pictures on the facing page demonstrate is that even in a completely closed system, where no information at all is allowed to escape, a system like rule 37R still does not follow the uniform trend towards increasing randomness that is suggested by the Second Law.

What instead happens is that kinds of membranes form between different regions of the system, and within each region orderly behavior can then occur, at least while the membrane survives.



An example of evolution according to rule 37R from an initial condition containing a fairly random region. Even though the system is reversible, this region tends to organize itself so as to take on a much simpler form. Information on the initial conditions ends up being carried by localized structures which radiate outwards.

This basic mechanism may well be the main one at work in many biological systems: each cell or each organism becomes separated from others, and while it survives, it can exhibit organized behavior.

But looking at the pictures of rule 37R on page 454 one may ask whether perhaps the effects we see are just transients, and that if we waited long enough something different would happen.

It is an inevitable feature of having a closed system of limited size that in the end the behavior one gets must repeat itself. And in rules like 0R and 90R shown on page 452 the period of repetition is always very short. But for rule 37R it usually turns out to be rather long. Indeed, for the specific example shown on page 454, the period is 293,216,266.

In general, however, the maximum possible period for a system containing a certain number of cells can be achieved only if the evolution of the system from any initial condition eventually visits all the possible states of the system, as discussed on page 258. And if this in fact happens, then at least eventually the system will inevitably spend most of its time in states that seem quite random.

But in rule 37R there is no such ergodicity. And instead, starting from any particular initial condition, the system will only ever visit a tiny fraction of all possible states. Yet since the total number of states is astronomically large—about 10^{60} for size 100—the number of states visited by rule 37R, and therefore the repetition period, can still be extremely long.

There are various subtleties involved in making a formal study of the limiting behavior of rule 37R after a very long time. But irrespective of these subtleties, the basic fact remains that so far as I can tell, rule 37R simply does not follow the predictions of the Second Law.

And indeed I strongly suspect that there are many systems in nature which behave in more or less the same way. The Second Law is an important and quite general principle—but it is not universally valid. And by thinking in terms of simple programs we have thus been able in this section not only to understand why the Second Law is often true, but also to see some of its limitations.