PREPRINT (July 19, 2006). Published in *Complexity* (12, 1), Sep/Oct 2006

Emergence Explained: Abstractions

Getting epiphenomena to do real work

Russ Abbott

Department of Computer Science, California State University, Los Angeles
and
The Aerospace Corporation
Russ.Abbott@GMail.com

Abstract. Emergence—macro-level effects from micro-level causes—illustrates the fundamental dilemma of science and is at the heart of the conflict between reductionism and functionalism: how can there be autonomous higher level laws of nature (the functionalist claim) if everything can be reduced to the fundamental forces of physics (the reductionist position)? In this, the first of two papers, we conclude the following. (a) What functionalism calls the special sciences (sciences other than physics) do indeed study autonomous laws. (b) These laws pertain to real higher level abstractions (discussed in this paper) and entities (discussed in the second paper). (c) Higher level interactions are epiphenomenal in that they can always be reduced to fundamental physical forces. (d) Since higher-level models are simultaneously both real and reducible we cannot avoid multiscalar systems. (e) Multiscalar systems are downward entailing and not upward predicting.

1 Introduction

Although the field of complex systems is relatively young, the sense of the term *emergence* that is commonly associated with it—that micro phenomena often give rise to macro phenomena[1]—has been in use for well over a century. The article on *Emergent Properties* in the Stanford Encyclopedia of Philosophy [1] begins as follows.

> Emergence [has been] a notorious philosophical term of art [since 1875]. … We might roughly characterize [its] meaning thus: emergent entities (properties or substances) 'arise' out of more fundamental entities and yet are 'novel' or 'irreducible' with respect to them. … Each of the quoted terms is slippery in its own right.

In a 1998 book-length perspective on his life's work [2], John Holland, the inventor of genetic algorithms and one of the founders of the field of complex systems, offered an admirably honest account of the state of our understanding of emergence at the time.

> It is unlikely that a topic as complicated as emergence will submit meekly to a concise definition, and I have no such definition to offer.

In a review of Holland's book, Shalizi [3] wrote the following.

> Someplace … where quantum field theory meets general relativity and atoms and void merge into one another, we may take "the rules of the game" to be given. But the rest of the observable, exploitable order in the universe—benzene molecules, *PV = nRT,* snowflakes, cyclonic storms, kittens, cats, young love, middle-aged remorse, financial euphoria accompanied with acute gullibility, prevaricating candidates for public office, tapeworms, jet-lag, and unfolding cherry blossoms—where do all these regularities come from? Call this

[1] Recently the term *multiscale* has gained favor as a less mysterious-sounding way to refer to this macro-micro interplay.

> *emergence* if you like. It's a fine-sounding word, and brings to mind southwestern creation myths in an oddly apt way.

The preceding is a poetic echo of the position expressed in a landmark paper [4] by Philip Anderson when he argued against what he called the *constructionist hypothesis,* namely, "[the] ability to reduce everything to simple fundamental laws … implies the ability to start from those laws and reconstruct the universe." Anderson explained as follows.

> At each level of complexity entirely new properties appear. … [O]ne may array the sciences roughly linearly in [a] hierarchy [in which] the elementary entities of [the science at level n+1] obey the laws of [the science at level n]: elementary particle physics, solid state (or many body) physics, chemistry, molecular biology, cell biology, …, psychology, social sciences. But this hierarchy does not imply that science [n+1] is 'just applied [science n].' At each [level] entirely new laws, concepts, and generalization are necessary. … Psychology is not applied biology, nor is biology applied chemistry. … The whole becomes not only more than but very different from the sum of its parts.

Although not so labeled, the preceding provides a good summary of the position known as functionalism—developed at about the same time—which argues that autonomous laws of nature appear at many levels.

Anderson thought that the position he was taking was radical enough that he included a reaffirmation of his adherence to reductionism.

> [The] workings of all the animate and inanimate matter of which we have any detailed knowledge are all … controlled by the same set of fundamental laws [of physics]. … [W]e must all start with reductionism, which I fully accept.

In the rest of this paper, we elaborate and extend the position that Anderson introduced. We claim to offer a coherent explanation for how nature can be simultaneously reductive and non-reductive. Much of our approach is derived from concepts borrowed from Computer Science—which more than any other human endeavor has tackled the job of building detailed, rigorous, and formal models of how we think. [5]

This is the first of two papers. In the second we apply the notions developed here to topics including: the nature of entities, the fundamental importance of interactions between entities and the environment, the central (and often ignored) role of energy (especially in computer science), the aggregation of complexity, and the limitations of modeling.

2 Background and foundations

To contrast reductionism and functionalism we use papers written by Steven Weinberg, the Nobel-prize winning physicist and an articulate defender of reductionism, and Jerry Fodor, one of the founders of the functionalist school of philosophy.

Functionalism holds [6] that there are so-called 'special sciences' (in fact, all sciences other than physics and perhaps chemistry) which study regularities in nature that are in some sense autonomous of physics. In [7] Fodor wrote the following reaffirmation.

> The *very existence* of the special sciences testifies to the reliable macrolevel regularities that are realized by mechanisms whose physical substance is quite typically heterogeneous. Does anybody really doubt

> that mountains are made of all sorts of stuff? Does anybody really think that, since they are, generalization about mountains-as-such won't continue to serve geology in good stead? Damn near everything we know about the world suggests that unimaginably complicated to-ings and fro-ings of bits and pieces at the extreme *micro*level manage somehow to converge on stable *macrolevel* properties.

Although Fodor does not use the term, the phenomena studied by the special sciences are the same sort of phenomena that we now call multiscale, i.e., emergent. Why is there emergence? Fodor continues as follows.

> [T]he 'somehow' [of the preceding extract] really is entirely mysterious … . *Why is there anything except physics?* … Well, I admit that I don't know why. I don't even know how to think about why. I expect to figure out why there is anything except physics the day before I figure out why there is anything at all … .

On the other side Weinberg distinguishes [8] between grand and petty reductionism.

> Grand reductionism is … the view that all of nature is the way it is (with certain qualifications about initial conditions and historical accidents) because of simple universal laws, to which all other scientific laws may in some sense be reduced. Petty reductionism is the much less interesting doctrine that things behave the way they do because of the properties of their constituents: for instance, a diamond is hard because the carbon atoms of which it is composed can fit together neatly. …
>
> Petty reductionism is not worth a fierce defense. … In fact, petty reductionism in physics has probably run its course. Just as it doesn't make sense to talk about the hardness or temperature or intelligence of individual "elementary" particles, it is also not possible to give a precise meaning to statements about particles being composed of other particles. We do speak loosely of a proton as being composed of three quarks, but if you look very closely at a quark you will find it surrounded with a cloud of quarks and anti-quarks and other particles, occasionally bound into protons; so at least for a brief moment we could say that the quark is made of protons.

Weinberg uses the weather to illustrate grand reductionism.

> [T]he reductionist regards the general theories governing air and water and radiation as being at a deeper level than theories about cold fronts or thunderstorms, not in the sense that they are more useful, but only in the sense that the latter can in principle be understood as mathematical consequences of the former. The reductionist program of physics is the search for the common source of all explanations. …
>
> Reductionism … provides the necessary insight that there are no autonomous laws of weather that are logically independent of the principles of physics. … We don't know the final laws of nature, but we know that they are not expressed in terms of cold fronts or thunderstorms. …
>
> Every field of science operates by formulating and testing generalizations that are sometimes dignified by being called principles or laws. … But there are no principles of chemistry that simply stand on their own, without needing to be explained reductively from the properties of electrons and atomic nuclei, and in the same way there are no principles of psychology that are free-standing, in the sense that they do not need ultimately to be understood through the study of the human brain, which in turn must ultimately be understood on the basis of physics and chemistry.

Thus the battle is joined: can all the laws of the special sciences be derived from physics?

3 Epiphenomena and Emergence

If one doesn't already have a sense of what it means, the term *epiphenomenon* is quite difficult to understand. The WordNet definition [9] is representative.

> A secondary phenomenon that is a by-product of another phenomenon.

It is not clear that this definition pins much down. It's especially troublesome because the terms *secondary* and *by-product* should not be interpreted to mean that an epiphenomenon is separate from and a consequence of the state of affairs characterized by the "other" phenomenon.

We suggest that a better way to think of an epiphenomenon is as an alternative way of apprehending or perceiving a given state of affairs. Consider Brownian motion, which appears to be motion that very small particles of non-organic materials apparently engage in on their own. Before Einstein, Brownian motion was a mystery. How could inanimate matter move on its own? We now know that Brownian motion is *an epiphenomenon of* collisions of atoms or molecules with the visibly moving particles.

The key is that we observed and described a phenomenon—the motion of visible inorganic macro-particles—without knowing what brought it about. We later found out that this phenomenon is an epiphenomenon of the underlying reality—the collision of micro-sized atoms or molecules with the visible macro particles. With this example as a guide we define the term *epiphenomenon* as follows.

> *Epiphenomenon*. A phenomenon that can be described independently of the underlying phenomena that bring it about.

We define *emergent* as synonymous with *epiphenomenal*. A phenomenon is emergent if it may be characterized independently of its implementation.[2]

Defined in this way, *emergence* is synonymous with concepts familiar from Systems Engineering and Computer Science. System requirements and software specifications are by intention written in terms that do not depend on the design or implementation of the systems that realize them. System requirements are written before systems are designed, and software specifications are intended to be implementation-independent. Thus system requirements and software specifications describe properties that are intended to be emergent once the specified system or software is implemented.

[Sidebar] Four simple examples of emergence

Even very simple systems may exhibit emergence. Here are four examples.

1. Consider a satellite in geosynchronous orbit. It has the property that it is fixed with respect to the earth as a reference frame. This property is emergent because it may be specified independently of how it is brought about. A satellite teth-

[2] In the second paper, we extend this notion when applied to entities. What will become important is how the entity acts in/on its environment, i.e., its functionality, independently of the mechanism that implements it. A Turing Machine, understood as a collection of tuples, acts on its environment, its tape, according to the function it computes. In his talk at the 2006 Understanding Complex Systems Symposium Eric Jakobsson made the point that biology must be equally concerned with what organisms do in their worlds and the mechanisms that allow them to do it.

ered to the ground by a long cable—like a balloon, were that possible—would also be fixed with respect to the earth as a reference frame.

Of course that's not how geosynchronicity works. A satellite in geosynchronous orbit circles the earth at the equator with a period that matches the earth's period of rotation. It is the combination of two independently produced phenomena (the satellite's orbit and the earth's rotation) that results in geosynchronicity.[3] If emergence is considered a defining characteristic of complex systems, this two-element system is probably as simple a complex system as one can imagine.

2. Consider the following code snippet.

```
temp := x;
x := y;
y := temp;
```

   This familiar idiom exchanges the values of x and y. Since this property may be specified independently of the code—there are many ways to exchange two variables—the exchange of x and y is an emergent effect of running this code. Here the coordinated combination of three independent actions produces an emergent result.

3. The bacterium *E. coli* produces an enzyme which digests lactose. But it produces it only when lactose is present. How does the bacterium "sense" the presence of lactose and control the production of the enzyme? A molecule that is normally bound to the bacterium's DNA blocks RNA transcription of the gene that codes for the enzyme. Lactose, when present, binds to the blocking molecule, pulling it off the DNA and allowing transcription and production to proceed. The way the bacterium acts in its environment may be described independently of what may be a surprising (but very clever) mechanism[4] (another combination of independent actions) for bringing that activity about.

4. Giuseppe Arcimboldo's paintings (see figure 1) illustrate emergence in a somewhat different way. What is most striking about Arcimboldo's paintings is that they are not optical illusions that include the outlines of different figures depending on how one looks at them. They consist of nothing but separate fruits and vegetables—which together make a face.

[Subhead in Sidebar] Emergence and surprise

We tend to reserve the term *emergent* for properties that appear in systems that are not explicitly designed by human engineers to have them. Emergence sometimes seems like a magic trick: we see that it happens but we didn't anticipate it, and we don't—at least initially—understand how it's done. This may be why emergence is sometimes associated with surprise. We suggest that it is wrong to rely on surprise as a character-

[3] Jonathan von Post (private communication) tells the story of how Arthur C. Clarke once applied for a British patent for geosynchronous orbits. It was rejected as impractical—but not as unpatentable. Imagine the lost royalties!

[4] We now know that this mechanism, the control of gene expression, is central to how biological organisms function.

istic of emergence. Emergence is a property of something in the world. Whether an observer is surprised has nothing to do with how we understand phenomena in the world.

3.1 Supervenience

A term from the philosophical literature that is closely related to *emergence* is *supervenience*. The intended use of this term is to relate a presumably higher level set of predicates (call such a set H for *higher*) to a presumably lower level set of predicates (call such a set L for *lower*). The predicates in H and L are all presumed to be applicable to some common domain of discourse. H and L are each ways of characterizing the state of affairs of the underlying domain. One says that H supervenes on (or over) L if it is never[5] the case that two states of affairs will assign the same configuration of values to the elements of L but different configuration of values to the elements of H.

Consider the following example. Let the domain be a sequence of n bits. Let L be the statements: bit 1 is on; bit 2 is on; etc. Let H be statements of the sort: exactly 5 bits are on; an even number of bits are on; no two successive bits are on; the bits that are on form the initial values in the Fibonacci sequence; etc. H supervenes over L since any configuration of truth values of the L statements determines the truth values of the H statements.

However, if we remove one of the statements from L, e.g., we don't include in L a statement about bit 3, but we leave the statements in H alone, then H does not supervene over L. To see why, consider the H statement

An even number of bits is on. (h_1)

For concreteness, let's assume that there are exactly 5 bits. Assume first, as in the first line of Figure 2, that all the bits are on except bit 3, the one for which there is no L statement. Since 4 of the 5 bits are on, h_1 is true. Since there is no L statement about bit 3, all the L statements are true even though bit 3 is off. Now, assume that bit 3 is on as in the second line of Figure 2. All the L statements are still true. But since 5 bits are now on, h_1 is now false. Since there is an H statement that has two different truth values for a single configuration of truth values of the L statements, H does not supervene over L.

The notion of supervenience captures the relationship between epiphenomena and their underlying phenomena.[6] Epiphenomena supervene over underlying phenomena: distinct epiphenomena must be associated with distinct underlying phenomena. Note that the reverse is not true. Two different states of the underlying phenomena may result in the same epiphenomena. In our bit example, there are many different ways in which an even number of bits may be on.

3.2 Supervenience, strong emergence, and causation

Returning to Weinberg and Fodor, presumably both would agree that phenomena of the special sciences supervene over phenomena in physics. A given set

[5] Some definitions require that not only is it never the case, it never can be the case. It does make a formal difference whether we base supervenience on a logical impossibility or on empirical facts. We finesse that distinction by adopting the rule of thumb of fundamental particle physicists: if something can happen it will.

[6] As stated on his website (http://cscs.umich.edu/~crshalizi/notebooks/emergent-properties.html) Shalizi's definition of emergence amounts to supervenience plus efficiency in an information-theoretic sense.

of phenomena at the level of fundamental physics is associated with no more than one set of phenomena at the level of any of the special sciences. This is Weinberg's petty reductionism, a case he makes sarcastically.

> Henry Bergson and Darth Vader notwithstanding, there is no life force. This is [the] invaluable negative perspective that is provided by reductionism.

What Weinberg is presumably getting at is that the standard model of physics postulates four elementary forces: the strong force, the weak force, the electromagnetic force, and gravity. I doubt that Fodor would disagree. Weinberg's sarcastic reference to a life force is an implicit criticism of an obsolete strain of thinking about emergence. The notion of vitalism—the emergence of life from lifeless chemicals—postulates a new force of nature that appears at the level of biology and is not reducible to lower level phenomena. Emergence of this sort is what Bedau [10] has labeled *strong emergence*. But as Bedau also points out, no one takes this kind of emergence seriously.[7]

If one dismisses the possibility of strong emergence and agrees that the only forces of nature are the fundamental forces of physics, then Fodor must also agree (no doubt he would) that any force-like construct postulated by any of the special sciences must be strictly reducible to the fundamental forces of physics. This is a stark choice: strict reductionism with respect to forces or strong emergence. This leads to an important conclusion. Any cause-like effect that results from a force-like phenomenon in the domain of any of the special sciences must be epiphenomenal.[8, 9] Weinberg backed away from petty reductionism when conceptualized in terms of matter. When conceptualized in terms of forces we see this as a conclusive argument in its favor.

3.3 Supervenience, reductionism, and emergence

It would appear that the relationship defined by supervenience will be useful in analyzing multiscale phenomena—especially if one want to "reduce" H statements to L statements or to show how H statements "emerge" from L statements. To some extent this is the case. But supervenience is not as useful as one might have hoped. One reason is the difficulty one encounters when using supervenience in a universe that obeys a higher-order regularity. Suppose that in our bit world H includes this statement.

> The prime bits are on, and the non-prime bits are off. (h_2)

Let's assume that h_2 is true—that it expresses a regularity about our bit world. Clearly h_2 supervenes over L. But knowing that h_2 supervenes over L doesn't help us if we want either to reduce h_2 to L or to show how h_2 emerges from L.

---

[7] Even were evidence of strong emergence to be found, science would carry on. Dark energy, the apparently extra force that seems to be pushing the Universe to expand may be a new force of nature. Furthermore, even if other (spooky) forces of nature like vitalism were (mysteriously) to appear at various levels of complexity, science would continue. We would do our best to measure and characterize them. After all, the known primitive forces just seemed to pop up out of nowhere, and science has taken them in stride.

[8] Kim [11] used the term *epiphenomenal causation* to refer to interactions of this sort.

[9] Compare this with the conclusion Hume reached [12] in his considerations of causality—that when one looks at any allegedly direct causal connection one always finds intermediary links. Since Hume did not presume a bottom level of fundamental physical forces, he dismissed the notion of causality entirely.

Our H statement expresses a regularity about our bit world that isn't explicit in our collection of L statements. A good scientist would ask why h_2 is true. The difficulty is that we haven't said why H is true. Yet there is nothing in the nature of supervenience as a relationship between sets of statements that provides a vehicle for explicating higher level regularities. Supervenience on its own is an insufficient framework for formulating either emergence or reduction.

4 Emergence in the Game of Life

The Game of Life[10] [13] is a totalistic[11] two-dimensional cellular automaton. The Game of Life grid is assumed to be unbounded in each direction. Each cell is either "alive" or "dead"—or more simply on or off. The 8 surrounding cells are a cell's neighbors. At each time step a cell determines whether it will be alive or dead at the next time step as follows. A live cell with two or three live neighbors stays alive; otherwise it dies. A dead cell with exactly three live neighbors becomes alive. All cells update simultaneously.

It is useful to think of the Game of Life in the following three ways.

1. As an agent-based model—of something, perhaps life and death phenomena. For our purposes it doesn't matter that the Game of Life isn't a realistic model—of anything. Many agent-based models are at the same time both simple and revealing.
2. As a trivial physical universe. Recall Shalizi, "Someplace … where quantum field theory meets general relativity … we may take 'the rules of the game' to be given." The Game of Life rules will be those "rules of the game." The rules that determine how cells turn on and off will be taken as the primitive operations of the physics of the Game of Life universe.[12] The reductionist agenda within a Game of Life universe would be to reduce every higher level phenomenon to the Game of Life rules.
3. As a programming platform.

Although these three perspectives reflect different emphases, it will always be the Game of Life rules that determine whether cells turn on or off.

4.1 *Epiphenomenal gliders*

Figure 3 shows a sequence of 5 time steps in a Game of Life run. The dark cells are "alive;" the light cells are

[10] The Game of Life is a popular example in discussions of emergence. Bedau [10] uses it as his primary example. Dennett [14] uses the fact that a Turing Machine may be implemented in terms of Game of Life patterns to argue that the position he takes in *The Intentional Stance* [15] falls midway along a spectrum of positions ranging from what he calls "Industrial strength Realism" to eliminative materialism, i.e., that beliefs are nothing but convenient fictions. Dennett also notes that when compared with the work required to compute the equivalent results in terms of primitive forces, one gets a "stupendous" "scale of compression" when one adopts his notion of an intentional stance. Although [14] doesn't spell out the link explicitly, Dennett's position appears to be that because of that intellectual advantage, one should treat the ontologies offered by the intentional stance as what he calls "mildly real"—although he doesn't spell out in any detail what regarding something as "mildly real" involves. We go further than Dennett in that we claim (below) that higher level abstractions are real in an objective sense—even though higher level interactions remain epiphenomenal. Our focus also differs from Dennett's in that we are concerned with the nature of regularities—whether or not those regularities are the subject matter of anyone's beliefs.

[11] The action taken by a cell depends on the number of neighbors in certain states, not which cells are in which states.

[12] This is the basis of what is sometimes called "digital physics." See [16], [17], and [18] which attempt to understand nature in terms of cellular automata.

"dead." One can apply the rules manually and satisfy oneself that they produce the sequence as shown. Notice that the fifth configuration shows the same pattern as the first offset by one cell to the right and down. If there are no other live cells on the grid, this process could be repeated indefinitely, producing a glider-like effect. Such a glider is an epiphenomenon of the Game of Life rules. A glider can be described—independently of the rules—as a sequence of patterns that traverses the grid.

What are the consequences of gliders from our three perspectives?

- When looked at from an *agent-based modeling perspective*, gliders may represent epidemics or waves of births and deaths. If one thinks about it—and forgets that one already knows that the Game of Life can produce gliders—gliders are quite amazing. A pattern sequence that traverses the grid arises from very simple (and local) rules for turning cells on and off. There is nothing in the rules about waves of cells sweeping across the grid. If one were attempting to demonstrate that such waves could be generated by simple agent-agent interactions, one might be quite pleased by this result.

- From our *physics perspective*, we note that the rules are the only forces in our Game of Life universe. Being epiphenomenal, gliders are causally powerless.[13] A glider neither changes how the rules operate nor determines which cells will be switched on and off. Gliders may be emergent, but they do not represent a new force of nature in the Game of Life universe.

  It may appear to us as observers that a glider moves across the grid and turns cells on. But that's not true. It's only the rules that turn cells on and off. A glider doesn't "go to an cell and turn it on." There is no glider "life force." Things happen only as a result of the lowest level forces of nature, the rules.

- From a *programming perspective* gliders are trivial. Once we know how to build a glider, it's a simple matter to make as many as we want. As a programming platform—imagine that we are kids fooling around with a new toy—we might experiment to see whether we can make other sorts of patterns. If we find some, which we will, we might want to see what happens when patterns crash into each other. After a while, we might compile a library of Game of Life patterns and their interactions.[14] It has even been shown [19] that by suitably arranging Game of Life patterns, one can implement Turing Machines.[15]

4.2 Using epiphenomena to do real (functional) work

What did we just say? What does it mean to say that epiphenomenal gliders and other epiphenomenal patterns can implement a Turing Machine? How can it mean anything? Neither the patterns nor the Turing Machine are real; they are

[13] All epiphenomena are causally powerless. Glider effects illustrate epiphenomenal causation.

[14] Since its introduction three decades ago, an online community of Game of Life programmers has developed. That community has created such libraries. A good place to start is Paul Callahan's "What is the Game of Life?" at http://www.math.com/students/wonders/life/life.html. See the appendix of this paper for a formalization of how such a pattern library may be produced.

[15] In [5] we refer to the generation of Game of Life gliders and Turing Machines as "non-algorithmic programming."

all epiphenomenal. Furthermore, the interactions between patterns aren't real either. They're also epiphenomenal; the only real action is at the level of the Game of Life rules. No matter how real the patterns look, interaction among them is always epiphenomenal.

What does one do to show that a Game of Life implementation of a Turing machine is correct? One treats the patterns and their interactions, i.e., the design itself, as an abstraction. One then shows two things: that the design implements a Turing Machine and that the patterns and their interactions can be implemented by Game of Life rules.

In other words, we build a Turing Machine on two levels of emergence. Both the pattern library and the Turing Machine are specified independently of their implementations. The Turing Machine is implemented by elements from the pattern library, and the pattern library is implemented by Game of Life rules. The use of emergent patterns and their epiphenomenal interactions to implement a Turing Machine—which can then be used to do real computations—illustrates the use of epiphenomena to do real work.

[Sidebar] Game of Life anthropologists

Let's pretend that we are anthropologists and that a previously unknown tribe has been discovered on a remote island. It is reported that their grid-like faces are made up of cells that blink on and off. We get a grant to study them. We travel to their far-off village and learn their language. They can't seem to explain what makes their cells blink on and off; we have to figure that out for ourselves.

After months of study, we come up with the Game of Life rules as an explanation for how the grid cells are controlled. Every single member of the tribe operates in a way that is consistent with those rules. The rules even explain the unusual patterns we observe—some of them, glider-like, traverse the entire grid. Pleased with our analysis, we return home and publish our results.

But something continues to nag. One of the teenage girls—she calls herself Hacka—has a pattern of activities on her grid that seems somehow more complex than the others. The Game of Life rules fully explain every light that goes on and every light that goes off on Hacka's pretty face. But that explanation just doesn't seem to capture everything that's going on. Did we miss something?

To make a long story short, it turns out that the tribe was not as isolated as we had thought. In fact they have an Internet connection. Hacka had learned not only that she was a Game of Life system but that the Game of Life can implement a Turing Machine. She had decided to program herself to do just that—and she used her Turning Machine implementation to solve problems that no one else in the tribe could approach. Her parents disapproved, but girls just want to have fun.

No wonder we felt uncertain about our results. Even though the Game of Life rules explained every light that went on and off on Hacka's face, it said nothing about the functionality implemented by Hacka's Turing Machine implementation. The rules explained everything about how the system worked; they said nothing about what the system did. The rules simply have no way to talk about Turing Machines. A Turing machine is an autonomous abstraction that Hacka built on top of the Game of Life rules.

4.3 Downward entailment

Recall Weinberg's statement: there are no autonomous laws of weather that are logically independent of the principles of physics. Clearly there are lots of autonomous "laws" of Turing Machines (namely computability theory), and they are all logically independent of the rules of the Game of Life. The fact that one can implement a Turing Machine on a Game of Life platform tells us nothing about Turing Machines.

An implementation of a Turing Machine on a Game of Life platform is an example of what might be called a non-reductive regularity. The Turing Machine and its implementation is certainly a kind of regularity, but it is a regularity that is not a logical consequence of—is not reducible to and cannot be derived from—the Game of Life rules. The theorems of computability theory are derived *de novo*. That Turing Machines can be realized using Game of Life rules tells us nothing about computability theory.

On the other hand, the fact that a Turing Machine can be implemented using the Game of Life rules does tell us something about the Game of Life—namely that the results of computability theory can be applied to the Game of Life. The property of being Turing complete applies to the Game of Life precisely because a Turing Machine can be shown to be one of its possible epiphenomena. Similarly we can conclude that the halting problem for the Game of Life—which we can define as determining whether a Game of Life run ever reaches a stable (unchanging or repeating) configuration—is unsolvable because we know that the halting problem for Turing Machines is unsolvable.

In other words, epiphenomena are downward entailing. Properties of epiphenomena are also properties of the phenomena from which they spring. This is not quite as striking as downward causation[16] would be, but it is a powerful intellectual tool. Let's consider in a bit more detail how we would conclude that the Game of Life halting problem is unsolvable. Because we can implement Turing Machines using the Game of Life, we know that we can reduce the halting problem for Turing Machines to the halting problem for the Game of Life: if we could solve the Game of Life halting problem, we could solve the Turing Machine halting problem. But we already knew that the Turing Machine halting problem is unsolvable. Therefore the Game of Life halting problem is unsolvable. Thus a consequence of downward entailment is that reducibility cuts both ways. One can conclude that if something is impossible at a higher level it must also be impossible at the lower level. We reach that conclusion by reasoning about the higher level as an independent abstraction and then reconnecting that abstraction to the lower level.

Earlier, we dismissed the notion that a glider may be said to "go to a cell and turn it on." The only things that turn on Game of Life cells are the Game of Life rules. But because of downward entailment, there is hope for talk of this sort. Once we established that a Turing Machine can be implemented on a Game of Life platform, we were able to apply results about Turing Machines to the Game of Life. We can do the same thing with gliders. We can establish a domain of discourse about gliders as abstract entities. Within that domain of discourse we can reason about gliders; in particular we can reason about how fast and in

[16] See, for example [20] for a number of sophisticated discussions of downward causation.

which directions they will move. Having developed facts and rules about gliders as abstractions, we can use the fact that gliders are epiphenomena of the Game of Life and apply those facts and rules to the Game of Life cells that gliders traverse. It is then reasonable to say that a glider goes to a cell and turns it on.

5 Science and higher level abstractions

In a recent book [21], Laughlin argues for what he calls *collective principles of organization*, which he finds to be at least as important as reductionist principles. In discussing Newton's laws, for example, he concludes that since

> these [otherwise] overwhelmingly successful laws … make profoundly wrong predictions at [the quantum] scale … Newton's legendary laws [are] emergent. They are not fundamental at all but a consequence of the aggregation of quantum matter into macroscopic fluids and solids. … [M]any physicists remain in denial. [They] routinely speak about Newton's laws being an 'approximation' for quantum mechanics, valid when the system is large—even though no legitimate approximation scheme has ever been found.

Laughlin also uses as an example the solid state of matter, which may be characterized as a three dimensional lattice of components held together by forces acting among those components. Once one has defined an abstract structure of this sort, one can derive properties of matter having this structure without knowing anything more about either (a) the particular elements at the lattice nodes or (b) how the binding forces are implemented. All one needs to know are the strengths of the forces and the shape of the lattice.

From our perspective, both Newton's laws and the solid state of matter are abstractions that nature implements under certain conditions. They apply in much the same way as the Turing Machine abstraction applies to certain cell configurations in the Game of Life. Laughlin calls the implementation of such an abstraction a *protectorate*.

Laughlin points out that protectorates tend to have feasibility ranges, which are often characterized by size, speed, and temperature. A few molecules of H_2O won't have the usual properties of ice. And ice melts when heated to the point at which the attractive forces are no longer able to preserve the lattice configuration of the elements. Similarly Newton's laws fail at the quantum level. The existence of such feasibility ranges does not reduce the importance of either the solid matter abstraction or the Newtonian physics abstraction. They just limit the conditions under which nature implements them.

The more general point is that nature implements a great many such abstractions. As is the case with computability theory, there are often sophisticated theories that characterize the properties of such naturally occurring abstractions. These theories may have nothing to do with how the abstract designs are implemented. They are theories that apply to the abstractions themselves. To apply such theories to real phenomena all one needs are physical examples in which the abstraction is implemented.

Furthermore and perhaps more importantly, abstractions of this sort are neither derivable from nor logical consequences of their implementations—i.e., grand reductionism fails. Abstractions and the theories built on them are new and creative constructs and are not derivable as consequences of the properties of the platform on which they are implemented. The Game of Life doesn't include the concept of a Turing machine, and quantum physics doesn't include the concept of a Newtonian solid.

The point of all this is to support Laughlin's position: when nature implements an abstraction, the epiphenomena described by that abstraction become just as real any other phenomena, and the abstraction that describes them is just as valid a description of that aspect of nature as any other description of any other aspect of nature. That much of nature is best understood as implementations of abstractions suggests that much of science is best expressed at two levels: (1) the abstraction itself, i.e., how it works as an abstraction—what Weinberg denigratingly refers to as the principles of the higher level science—and (2) how and under what conditions nature implements that abstraction.

5.1 Phase transitions

Since nature often implementers abstract designs only within feasibility regions, there will almost always be borderline situations in which the implementation of an abstract design is on the verge of breaking down. These borderline situations frequently manifest as what we call phase transitions—regions or points (related to a parameter such as size, speed, temperature, and pressure) where multiple distinct and incompatible abstractions may to be implemented.

Newton's laws fail at both the quantum level and at relativistic speeds. If as Laughlin suggests, the Newtonian abstraction is not an approximation of quantum theory, phase transitions should appear as one approaches the quantum realm. As explained by Sachdev [22], the transition from a Newtonian gas to a Boise-Einstein condensate (such as super-fluid liquid helium) illustrates such a phase transition.

> At room temperature, a gas such as helium consists of rapidly moving atoms, and can be visualized as classical billiard balls which collide with the walls of the container and occasionally with each other.
>
> As the temperature is lowered, the atoms slow down [and] their quantum-mechanical characteristics become important. Now we have to think of the atoms as occupying specific quantum states which extend across the entire volume of the container. … [Since helium] atoms are 'bosons' … an arbitrary number of them can occupy any single quantum state. … If the temperature is low enough … *every* atom will occupy the *same* lowest energy … quantum state.

On the other hand, since Newton's laws are an approximation of relativistic physics, there are no Newtonian-related phase transitions as one approaches relativistic speeds.

These considerations suggest that whenever data suggests a phase transition, one should look for two or more abstractions with overlapping or adjacent feasibility regions.

5.2 Static and dynamic emergence

Abstractions may be implemented in two ways. Static abstraction (or static emergence) comes about as a result of energy wells. Solids are an example; they exist in energy wells. To convert a solid into a liquid or a gas, one must add energy.

Abstractions may also be created when energy flows though an open but constrained system. Dynamic abstraction (or dynamic emergence) produces what is famously known as far-from-equilibrium systems. The global weather system is an example. Meteorological regularities occur when energy flows through the environment. Prigogine [23] refers to systems of this sort as dissipative structures. Dynamic abstractions are particularly important for biological and social entities, which we discuss in the second paper.

5.3 The reality of higher level abstractions

Are higher level abstractions objectively real? Our answer is "yes" on two grounds.

1. **Lower entropy.** When higher level abstractions are implemented, entropy is reduced. Solids have lower entropy than liquids and gasses, and the global weather system has less entropy than the elements that compose it would have were they are not implementing weather abstractions.
2. **Either more or less mass.** When organized in terms of higher level abstractions matter has either more or less mass than that same matter when not so organized. When organized as a solid, matter has less mass (by a negligible but real amount) than when not so organized. Similarly, the matter that makes up the global weather system along with the energy flowing through it has more mass (by a negligible but real amount) than that same matter without the energy.

Because higher level abstractions are physically identifiable from both an entropy and mass perspective, we feel justified in asserting that they are objectively real.

Humankind has intuitively recognized the reality of higher level abstractions for a long time. The fundamental dilemma of science has been to reconcile the reality of higher level abstractions with the epiphenomenal nature of higher level interactions. The dilemma is resolved when one realizes that (a) the subject matter of the higher level sciences are abstractions; (b) those abstractions are instantiated as physically real when implemented by lower level phenomena; yet (c) interaction among those abstractions is epiphenomenal and may always be reduced to the fundamental forces of physics.

Furthermore, downward entailment may be applied to conclusions drawn about higher level abstractions to derive results about the elements that implement them. But the principles that govern a higher level abstraction are generally not derivable from the principles that govern the implementing mechanisms.

A consequence of all this is that multiscalar systems are inevitable. Recall the nursery rhyme.

For want of a nail, a shoe was lost.
For want of a shoe, a horse was lost.
For want of a horse, a rider was lost.
For want of a rider, a message was lost.
For want of a message, a battle was lost.
For want of a battle, a kingdom was lost.
All for want of a nail.
- George Herbert (1593-1632)

Abstractions at all levels are central to how we look at the world, but one must always be aware of the feasibility ranges within which an abstraction is implemented. A tragic example is the case of the O-rings on Challenger. They failed when they were used outside the temperature feasible range for which they functioned as sealants.

[Sidebar] Implications for Modeling

The perspective we have described yields two major implications for modeling. We refer to them as the *difficulty of* *looking downwards* and *the difficulty of looking upwards*. In both cases, the problem is that it is very difficult to model significant creativity—notwithstanding the fact that surprises appear in some of our models.

Modeling: the difficulty of looking downward

Strict reductionism, our conclusion that all forces and actions are epiphenomenal over forces and actions at the fundamental level of physics, implies that it is impossible to find a non-arbitrary base level for models. One never knows what unexpected effects one may be leaving out by defining a model in which interactions occur at some non-fundamental level.

Consider a model of computer security. Suppose that by analyzing the model one could guarantee that a communication line uses essentially unbreakable encryption technology. Still it is possible for someone inside to transmit information to someone outside.

How? By sending messages in which the content of the message is ignored but the frequency of transmission carries the information, e.g., by using Morse code. The problem is that the model didn't include that level of detail. This is the problem of looking downward.

A further illustration of this difficulty is that there are no good models of biological arms races. (There don't seem to be any good models of significant co-evolution at all.) There certainly are models of population size effects in predator-prey simulations. But by biological arms races we are talking about not just population sizes but actual evolutionary changes.

Imagine a situation in which a plant species comes under attack from an insect species. In natural evolution the plant may "figure out" how to grow bark. Can we build a computer model in which this solution would emerge? It is very unlikely. To do so would require that our model have built into it enough information about plant biochemistry to enable it to find a way to modify that biochemistry to produce bark, which itself is defined implicitly in terms of a surface that the insect cannot penetrate. Evolving bark would require an enormous amount of information—especially if we don't want to prejudice the solution the plant comes up with.

The next step, of course, is for the insect to figure out how to bore through bark. Can our model come up with something like that? Unlikely. What about the plant's next step: "figuring out" how to produce a compound that is toxic to the insect? That requires that the model include information about both plant and insect biochemistry—and how the plant can produce a compound that interferes with the insect's internal processes. This would be followed by the development by the insect of an anti-toxin defense.

To simulate this sort of evolutionary process would require an enormous amount of low level detail—again especially if we don't want to prejudice the solution in advance.

Other than Tierra (see [Ray]) and its successors, which seem to lack the richness to get very far off the ground, as far as we know, there are no good computer models of biological arms races. A seemingly promising approach would be an agent-based system in which each agent ran its own internal genetic programming model. But we are unaware of any such work.[17]

Finally, consider the fact that geckos climb walls by taking advantage of the Van der Walls "force." (We put *force* in quotation marks because there is no Van der Walls force. It is an epiphenomenon of relatively rarely occurring quantum phenomena.) To build a model of evolution in which creatures evolve to use the Van der Walls force to climb walls would require that we build quantum physics into what is presumably intended to be a relatively high-level biological model in which macro geckos climb macro walls

It's worth noting that the use of the Van der Walls force was apparently not an extension of some other gecko process. Yet the gecko somehow found a way to reach directly down to a quantum-level effect to find a way to climb walls.

The moral is that any base level that we select for our models will be arbitrary, and by choosing that base level, we may miss important possibilities. Another moral is that models used when doing computer security or terrorism analysis—or virtually anything else that includes the possibility of creative adaptation—will always be incomplete. We will only be able to model effects on the levels for which our models are defined. The imaginations of any agents that we model will be limited to the capabilities built into the model.

Modeling: the difficulty of looking upward

We noted earlier that when a glider appears in the Game of Life, it has no effect on the how the system behaves. The agents don't see a glider coming and duck. More significantly we don't know how to build systems so that agents will be able to notice gliders and duck.

It would be an extraordinary achievement in artificial intelligence to build a modeling system that could notice emergent phenomena and see how they could be exploited. Yet we as human beings do this all the time. The dynamism of a free-market economy depends on our ability to notice newly emergent patterns and to find ways to exploit them.

Al Qaeda noticed that our commercial airlines system can be seen as a network of flying bombs. Yet no model of terrorism that doesn't have something like that built into it will be able to make that sort of creative leap. Our models are blind to emergence even as it occurs within them.

Notice that this is not the same as the difficulty of looking downward. In the Al Qaeda example one may assume that one's model of the airline system includes the information that an airplane when loaded with fuel will explode when it crashes. The creative leap is to notice that one can use that phenomenon for new purposes. This is easier than the problem of looking downward. But it is still a very difficult problem.

[17] Genetic programming is relevant because we are assuming that the agent has an arbitrarily detailed description of how the it functions and how elements in its environment function.

Notice how difficult it would be implement such a system. The agent's internal model of the environment would have to be updated continually as the environment changed. That requires a means to perceive the environment and to model changes in it. Clearly that's extraordinarily sophisticated. Although one could describe such a system without recourse to the word *consciousness*, the term does come to mind.

Nature's approach is much simpler: change during reproduction and see what happens. If the result is unsuccessful, it dies out; if it is successful it persists and reproduces. Of course that requires an entire generation for each new idea.

The moral is the same as before. Models will always be incomplete. We will only be able to model effects on the levels for which our models are defined. The imaginations of any agents that we model will be limited to the capabilities built into the model.

6 It's a tie

In the debate between reductionism and functionalism, the final score is 1-1.

- Petty reductionism gets a point with respect to causation. Forces and interactions can always be reduced to the fundamental forces of physics. There is no life force.
- Grand reductionism loses a point with respect to derivability. Just as the laws governing Turning machines are not derivable from the rules of the Game of Life, the laws governing higher level abstractions are not in general derivable from the fundamental laws of physics—even when as in the case of solids and Newtonian physics nature implements those abstractions without our help.

To paraphrase Weinberg, the goal of science is to find simple universal laws that explain why nature is the way it is. When understood in this way, mathematics, computer science, and engineering—all of which create and study conceptual structures that need not exist—are not science. Fortunately for us, neither is nature. Nature produces results which need not exist. Evolution may be a blind watchmaker; nature in general is a blind engineer; and engineers implement abstractions.

Isn't there something unreal about explaining nature at least in part as implementations of abstractions? Not necessarily. Is there a better way to understand the origins of Newtonian mechanics, the solid state of matter, and phase transitions? These phenomena are part of the offerings that nature sets before us. More importantly, they are evidence that emergence, i.e., the creation and implementation of abstractions, is not only a fundamental aspect of nature but a fundamental principle of science.

The perspectives developed in this paper reflect those of computer science. Is this parochialism? It's difficult to tell from so close. One thing is clear. Because computer science has wrestled—with some success—with many serious philosophical challenges,[18] it is not unreasonable to hope that the field may contribute something to the broader philosophical community. In this paper, it is the computer science notion of an abstract software specification that is most significant. Software abstractions are both conceptual and real. An implemented software specification combines the formality and abstraction of mathematics with the reality of nature. Computers are reification devices, capable of making the abstract concrete. (See [5].)

Nature's abstractions differ from software abstractions in that they are not conceptual; they are always implemented by something. The combination of higher level abstractions along with the epiphenomenal and range-limited nature of higher level causality makes multiscale systems unavoidable—a problem which typically doesn't plague abstractions implemented in software.[19]

---

[18] It has been suggested that Computer Science is applied philosophy. Fred Thompson, an early mentor, is Emeritus Professor of Applied Philosophy and Computer Science at Caltech.

[19] Software developers face feasibility issues in such areas as scalability and numerical range and precision limits.

One may be tempted to look for a new mathematics that explains in closed-form how all phenomena arise from some set of primitive elements. We won't find one. The abstractions of science are downward—not upward—entailing. We will never develop a mathematics that maps the motion of electrons to the functionality of the software a computer is running. The design of the computer (at all levels) along with the software itself is that mapping. Nothing simpler will do.

Nature allows for the creation and implementation of new abstractions, i.e., emergence. Are there simple universal laws of emergence? Are there necessary and sufficient conditions under which emergence occurs? Answers to these questions would help explain why nature is the way it is.

7 Acknowledgement

We are grateful for numerous enjoyable and insightful discussions with Debora Shuger during which many of the ideas in this paper were developed and refined.

References

[1] O'Connor, T. and Hong Yu Wong, "Emergent Properties", *The Stanford Encyclopedia of Philosophy* (Summer 2005 Edition), Edward N. Zalta (ed.), URL: http://plato.stanford.edu/archives/sum2005/entries/properties-emergent/.

[2] Holland, J. *Emergence: From Chaos to Order*, Addison-Wesley, 1997.

[3] Shalizi, C., "Review of Emergence from Chaos to Order," *The Bactra Review*, URL as of 6/2005: http://cscs.umich.edu/~crshalizi/reviews/holland-on-emergence/

[4] Anderson, P.W., "More is Different," *Science*, 177 393-396, 1972.

[5] Abbott, R. "If a Tree Casts a Shadow, is it Telling the Time?" to appear in *Proceedings of the Fifth International Conference on Unconventional Computation*, Lecture Notes in Computer Science, Springer, 2006.

[6] Fodor, J. A., "Special Sciences (or the disunity of science as a working hypothesis)," *Synthese* 28: 97-115. 1974.

[7] Fodor, J.A., "Special Sciences; Still Autonomous after All These Years," *Philosophical Perspectives, 11, Mind, Causation, and World,* pp 149-163, 1998.

[8] Weinberg, S., "Reductionism Redux," *The New York Review of Books*, October 5, 1995. Reprinted in Weinberg, S., *Facing Up*, Harvard University Press, 2001.

[9] WordNet 2.0, URL as of 6/2005: www.cogsci.princeton.edu/cgi-bin/webwn.

[10] Bedau, M.A., "Downward causation and the autonomy of weak emergence". *Principia* 6 (2002): 5-50.

[11] Kim, J. "Multiple realization and the metaphysics of reduction," *Philosophy and Phenomenological Research*, v 52, 1992.

[12] Hume, D. *An Enquiry Concerning Human Understanding*, Vol. XXXVII, Part 3. The Harvard Classics. New York: P.F. Collier & Son, 1909–14; Bartleby.com, 2001.

[13] Gardner, M., Mathematical Games: "The fantastic combinations of John Conway's new solitaire game 'life'," *Scientific American*, October, November, December, 1970, February 1971.

[14] Dennett, D. C. "Real Patterns," *The Journal of Philosophy*, (88, 1), 1991.

[15] Dennett, D. C., *The Intentional Stance*, MIT Press/Bradford Books, 1987.

[16] Zuse, K., "Rechnender Raum" (*Vieweg, Braunschweig*, 1969); translated as *Calculating Space*, MIT Technical Translation AZT-70-164-GEMIT, MIT (Project MAC), Cambridge, Mass. 02139, Feb. 1970.

[17] Fredkin, E., "Digital Mechanics", Physica D, (1990) 254-270, North-Holland.

[18] Wolfram, S., *A New Kind of Science*, Wolfram Media, 2002.

[19] Rendell, Paul, "Turing Universality in the Game of Life," in Adamatzky, Andrew (ed.), *Collision-Based Computing,* Springer, 2002.

[20] Emmeche, C, S. Køppe and F. Stjernfelt, "Levels, Emergence, and Three Versions of Downward Causation," in Andersen, P.B., Emmeche, C., N. O. Finnemann and P. V. Christiansen, eds. (2000): *Downward Causation. Minds, Bodies and Matter*. Århus: Aarhus University Press.

[21] Laughlin, R.B., *A Different Universe*, Basic Books, 2005.

[22] Sachdev, S, "Quantum Phase Transitions," in *The New Physics*, (ed. G. Fraser), Cambridge University Press, 2006.

[23] Prigogine, Ilya and Dilip Kondepudi, *Modern Thermodynamics: from Heat Engines to Dissipative Structures*, John Wiley & Sons, N.Y., 1997.

Appendix. Game of Life Patterns

Intuitively, a Game of Life pattern is the step-by-step time and space progression on a grid of a discernable collection of inter-related live cells. We formalize that notion in three steps.

1. First we define a static construct called the live cell group. This will be a group of functionally isolated but internally interconnected cells.
2. Then we define Game of Life *basic patterns* as temporal sequences of live cell groups. The Game of Life *glider* and *still-life* patterns are examples
3. Finally we extend the set of patterns to include combinations of basic patterns. The more sophisticated Game of Life patterns, such the *glider gun*, are examples.

Live cell groups

The fundamental construct upon which we will build the notion of a pattern is what we shall call a *live cell group*.

A live cell group is a collection of live and dead cells that have two properties.

1. They are functionally isolated from other live cells.
2. They are functionally related to each other.

More formally, we define cells c_0 and c_n in a Game of Life grid to be *connected* if there are cells c_1, c_2, …, c_{n-1} such that for all i in 0 .. n-1

1. c_i and c_{i+1} are *neighbors*, as defined by Game of Life, and
2. either c_i or c_{i+1} (or both) are *alive*, as defined by Game of Life.

Connectedness is clearly an equivalence relation (reflexive, symmetric, and transitive), which partitions a Game of Life board into equivalence classes of cells. Every dead cell that is not adjacent to a live cell (does not have a live cell as a Game of Life neighbor) becomes a singleton class.

Consider only those connectedness equivalence classes that include at least one live cell. Call such an equivalence class a *live cell group* or *LCG*.

Define the *state* of an LCG as the specific configuration of live and dead cells in it. Thus, each LCG has a state.

No limitation is placed on the size of an LCG. Therefore, if one does not limit the size of the Game of Life grid, the number of LCGs is unbounded.

Intuitively, an LCG is a functionally isolated group of live and dead cells, contained within a boundary of dead cells. Each cell in an LCG is a neighbor to at least one live cell within that LCG.

As a consequence of this definition, each live cell group consists of an "inside," which contains all its live cells (possibly along with some dead cells), plus a "surface" or "boundary" of dead cells. (The surface or boundary is also considered part of the LCG.)

Basic patterns: temporal sequences of live cell groups

Given this definition, we can now build temporal sequences of LCGs. These will be the Game of Life basic patterns.

The Game of Life rules define transitions for the cells in a LCG. Since an LCG is functionally isolated from other live cells, the new states of the cells in an LCG are determined only by other cells in the same LCG.[20]

Suppose that an LCG contains the only live cells on a Game of Life grid. Consider what the mapping of that LCG by the Game of Life rules will produce. There are three possibilities.

1. The live cells may all die.
2. The successor live cells may consist of a single LCG—as in a *glider* or *still life*.
3. The successor live cells may partition into multiple LCGs—as in the so-called *bhepto* pattern, which starts as a single LCG and eventually stabilizes as 4 *still life* LCGs and two *glider* LCGs.

In other words, the live cells generated when the Game of Life rules are applied to an LCG will consist of 0, 1, or multiple successor LCGs.

More formally, if ℓ is an LCG, let Game of Life(ℓ) be the set of LCGs that are formed by applying the Game of Life rules to the cells in ℓ. For any particular ℓ, Game of Life(ℓ) may be empty; it may be contain a single element; or it may contain multiple elements. If ℓ' is a member of Game of Life(ℓ) write ℓ -> ℓ'.

For any LCG ℓ_0, consider a sequence of successor LCGs generated in this manner:

ℓ_0 -> ℓ_1 -> ℓ_2 -> ℓ_3 -> … .

Extend such a sequence until one of three conditions occurs.

1. There are no successor LCGs, i.e., Game of Life(ℓ_i) is empty—all the live cells in the final LCG die. Call these terminating sequences.
2. There is a single successor LCG, i.e., Game of Life(ℓ_i) = {ℓ_k}, but that successor LCG is in the same state as an LCG earlier in the sequence, i.e., $\ell_k = \ell_j$, $j < k$. Call these repeating sequences.
3. The set Game of Life(ℓ_i) of successor LCGs contains more than one LCG, i.e., the LCG branches into two or more LCGs. Call these branching sequences.

Note that some LCG sequences may never terminate. They may simply produce larger and larger LCGs. The so-called *spacefiller* pattern, which actually consists of multiple interacting LCGs, one of which fills the entire grid with a single LCG as it expands,[21] is an amazing example of such a pattern. I do not know if there is an LCG that expands without limit on its own. If any such exist, call these infinite sequences.

For any LCG ℓ_0, if the sequence

[20] In particular, no LCG cells have live neighbors that are outside the LCG. Thus no cells outside the LCG need be considered when determining the GoL transitions of the cells in an LCG. A dead boundary cell may become live at the next time-step, but it will do so only if three of its neighbors within the LCG are live. Its neighbors outside the LCG are guaranteed to be dead.

If a boundary cell does become live, the next-state LCG of which it is a member will include cells that were not part of its predecessor LCG.

[21] See the spacefiller pattern on http://www.math.com/students/wonders/life/life.html or http://www.ibiblio.org/lifepatterns.

$\ell_0 \text{->} \ell_1 \text{->} \ell_2 \text{->} \ell_3 \text{->} \ldots$.

is finite, terminating in one of the three ways described above, let seq(ℓ_0) be that sequence along with a description of how it terminates. If

$\ell_0 \text{->} \ell_1 \text{->} \ell_2 \text{->} \ell_3 \text{->} \ldots$.

is infinite, then seq(ℓ_0) is undefined.

Let BP (for Basic Patterns) be the set of finite non-branching sequences as defined above. That is,

$$\text{BP} = \{\text{seq}(\ell_0) \mid \ell_0 \text{ is an LCG}\}$$

Note that it is not necessary to extend these sequences backwards. For any LCG ℓ_0, one could define the pre-image of ℓ_0 under the Game of Life rules. Game of Life^{-1}(ℓ) is the set of LCGs ℓ' such that Game of Life(ℓ') = ℓ.

For any chain seq(ℓ_0) in BP, one could add all the chains constructed by prefixing to seq(ℓ_0) each of the predecessors ℓ' of ℓ_0 ℓ' as long as ℓ' does not appear in seq(ℓ'). But augmenting BP in this way would add nothing to BP since by definition seq(ℓ') is already defined to be in BP for each ℓ'.

We noted above that we do not know if there are unboundedly long sequences of LCGs beginning with a particular ℓ_0. With respect to unboundedly long predecessor chains, it is known that such unbounded predecessor chains (of unboundedly large LCGs) exist. The so-called *fuse* and *wick* patterns are LCG sequences that can be extended arbitrarily far backwards.[22] When run forward such fuse or wick LCGs converge to a single LCG. Yet given the original definition of BP even these LCG sequences are included in it. Each of these unbounded predecessor chains is included in BP starting at each predecessor LCG.

Clearly BP as defined includes many redundant pattern descriptions. No attempt is made to minimize BP either for symmetries or for overlapping patterns in which one pattern is a suffix of another—as in the fuse patterns. In a computer program that generated BP, such efficiencies would be important.

BP is recursively enumerable

The set BP of basic Game of Life patterns may be constructed through a formal iterative process. The technique employed is that used for the construction of many recursively enumerable sets.

1. Generate the LCGs in sequence.
2. As each new LCG is generated, generate the next step in each of the sequences starting at each of the LCGs generated so far.
3. Whenever an LCG sequence terminates according to the BP criteria, add it to BP.

The process sketched above will effectively generate all members of BP. Although theoretically possible, such a procedure will be so inefficient that it is useless for any practical purpose.[23] The

[22] A simple fuse pattern is a diagonal configuration of live cells. At each time step, the two end cells die; the remaining cells remain alive. A simple fuse pattern may be augmented by adding more complex features at one end, thereby building a pattern that becomes active when the fuse exhausts itself. Such a pattern can be built with an arbitrarily long fuse.

[23] Many much more practical and efficient programs have been written to search for patterns in the GoL and related cellular automata. See http://www.ics.uci.edu/~eppstein/ca/search.html for a list of such programs.

only reason to mention it here is to establish that BP is recursively enumerable. Whether BP is recursive depends on whether one can in general establish for any LCG ℓ_0 whether seq(ℓ_0) will terminate.[24]

Game of Life patterns: combinations of basic patterns

Many of the interesting Game of Life patterns arise from interactions between and among basic patterns. For example, the first pattern that generated an unlimited number of live cells, the glider gun, is a series of interactions among combinations of multiple basic patterns that cyclically generate gliders.

To characterize these more complex patterns it is necessary to keep track of how basic patterns interact. In particular, for each element in BP, augment its description with information describing

a) its velocity (rate, possibly zero, and direction) across the grid,

b) if it cycles, how it repeats, i.e., which states comprise its cycle, and

c) if it branches, what the offspring elements are and where they appear relative to final position of the terminating sequence.

Two or more distinct members of BP that at time step i are moving relative to each other may interact to produce one or more members of BP at time step i+1. The result of such a BP "collision" will generally depend on the relative positions of the interacting basic patterns. Even though the set BP of basic patterns is infinite, since each LCG is finite, by using a technique similar to that used for generating BP itself, one can (very tediously) enumerate all the possible BP interactions.

More formally, let $\mathcal{P}_f$(BP) be the set of all finite subsets of BP. For each member of $\mathcal{P}_f$(BP) consider all possible (still only a finite number) relative configurations of its members on the grid so that there will be some interaction among them at the next time step. One can then record all the possible interactions among finite subsets of BP.

These interactions would be equivalent to the APIs for the basic patterns. We could call a listing of them BP-API. Since BP is itself infinite, BP-API would also be infinite. But BP-API would be effectively searchable. Given a set of elements in BP, one could retrieve all the interactions among those elements. BP-API would then provide a documented starting point for using the Game of Life as a programming language.

As in traditional programming languages, as more complex interactions are developed, they too could be documented and made public for others to use.

[24] This is not the same question as that which asks whether any Game of Life configuration will terminate. We know that is undecidable.

Figures

Figure 1. Autumn by Giuseppe Arcimboldo.

From: http://commons.wikimedia.org/wiki/Image:Giuseppe_Arcimboldo_-_Autumn,_1573.jpg

The image is public domain.

Figure 2. Bit 3 off and on.

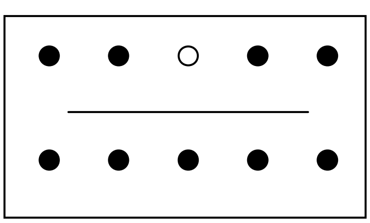

Figure 3. A glider.