What is Life?
First Principles, Basic Functions, and the Origin of Life

By Mark White, MD
Copyright 2007, Rafiki, Inc.

To Erwin Schrödinger and Charles Darwin
# Table of Contents

I. One Man’s Approach to Describing Life ................................................................. 3
II. What is Information? .............................................................................................. 9
III. What is Space and Time? .................................................................................... 16
IV. What is Symmetry? ............................................................................................ 18
V. What is Complexity? ............................................................................................. 24
VI. What is Evolution? ............................................................................................. 27
VII. What is a Crystal? .............................................................................................. 33
VIII. What is Biochemistry? .................................................................................... 38
IX. What Are Boundaries? ....................................................................................... 42
X. What is a Code? .................................................................................................... 45
XI. What is the Genetic Code? ................................................................................ 52
XII. What are the broad patterns of life? .................................................................. 85
XIII. How could life begin? ...................................................................................... 106
XIV. What is life? ....................................................................................................... 107
I. One Man’s Approach to Describing Life

Life is a crystal. Of course, there is a little bit more to it than that, but I didn’t want to be accused of the ol’ bait and switch, so I thought I’d start out by answering the question asked by the title: What is Life? Life is a crystal.

The title itself is partially borrowed from Erwin Schrödinger. He wrote a dandy little book in the early part of the fifth decade of the twentieth century. Schrödinger was a Nobel Prize-winning physicist who contributed greatly to the creation of the discipline known as quantum physics. He was also an accomplished writer who liked to speculate broadly on a number of topics, and it is lucky for us that he decided to turn his attention to biology. As far as I’m concerned, Erwin got it just about right when he asked and partially answered the question: What is Life? Of course, that was so long ago that we knew virtually none of the details that we know today. People then were limited to broad speculation. We still are today, but the speculation now can revolve around so much more molecular detail. Schrödinger applied to biology the abstract skills that he learned as a quantum physicist, and he concluded that life must be some form of a molecular information system. He was exactly right - because it is. Of course, at the time, there were no digital computers and we also knew nothing of the fantastic details involved in the many complex phenomena involving DNA and protein. But, at bottom, Erwin nailed it. He said that life would use a molecular “code script” that would guide ontogeny. He also knew that this general scenario would surely fly in the face of everything he had worked so hard to understand about the behavior of insentient atoms. He was relatively unfazed by this, and so he logically assumed that there was much to be learned. He knew that there was much that he could not yet possibly understand. Again, he was right, and again it is still true today.

This book is dedicated to Erwin Schrödinger and Charles Darwin. I think that the two combined perhaps could have figured out the basic riddles of Life if only they had the opportunity, the time, and the vast number of details that are accumulating by the minute today. I do not mean to imply that I have solved the riddle – I have not – but I am blessed with the vantage point not afforded to either of these two. From here, I am able to realize just how important and insightful both of them actually were. I also think they were thinking in basically the same terms and had surprisingly similar ideas about Life, but that is just one man’s opinion. They were perhaps not men before their time, but time did not allow them to answer the interesting questions that they both so intensely tried to answer.

Schrödinger’s work is lauded, as it should be, but it is unfortunately too often overlooked, misinterpreted and sometimes altogether forgotten. However, the most instructive thing about his work can be found in his general approach, his unhesitant willingness to explore areas out of his formal expertise, and his openness to the idea that he was lacking knowledge, or perhaps even being misled by things of which he felt certain. He approached the question of life simultaneously from above and below. He looked at the picture from the smallest and largest scales, and this is exactly what is required here. Life is tremendously complex and consists of many patterns on many scales.
Schrödinger knew that at bottom Life is an atomic phenomenon. He knew that atoms generally obey certain rules of behavior, yet when he examined the actual atoms of life he found them inexplicably disobeying the most fundamental rules to he thought he knew. He sensed that his ideas should require a basic grounding, yet the most basic grounding for atoms is to first ask if they are involved in a phase of gas, liquid or solid. He reasoned that life must be a solid, and this solid must have a particular structure. Generally speaking, solids with a definite structure are called crystals, and so Erwin supposed that the code script of life should be a crystal. He also intuited the informative nature of life’s code script, and he knew that most crystals require a regularly repeating pattern of atoms. If Life’s code script did this, however, it would be hard to imagine how it could contain any meaningful information, so he proposed that it was an aperiodic crystal. In other words, he felt sensed that the atomic arrangement must be regular yet it could not be slavishly repeating. This insight is partially right, but Schrödinger would have been stunned to know the truth at that time. The molecular structure about which he was speculating proved to be DNA, and DNA has a repeating structure employing an almost non-repeating sequence of component parts. We can definitely call DNA a crystal, but the issue of its periodicity will be left for a later section.

Perhaps the most insightful thing in Schrödinger’s book is the way he addresses the problems presented by the second law of thermodynamics with respect to the obvious organizational properties of life. The second law tells us that atoms and molecules will be driven by thermodynamics, which means heat motion. Heat motion is random in every way, yet the molecules of Life are behaving in a decidedly non-random fashion. Schrödinger knew that this contradiction must somehow be resolved if we are ever to understand Life.

A gross measure of the randomness of a molecular system is called entropy. Since life seemed to move in the opposite direction from entropy, Erwin decided that whatever it was that life did, or however it was that it did it, life should be marked by negentropy. He admitted of no greater understanding of this phenomenon than this simple description, but once again, he was exactly right. He said that whatever negentropy actually is, life “drinks it up.” He said that ordinary things, like periodic crystals, created order from disorder. In other words, when a disordered atomic system entered a period of favorable conditions, such as a decrease in temperature, it could self-organize into a configuration with more order. However, life seemed to be doing something more, something inherently different, which he described as a process of creating more order from existing order, and he called this negentropy. It is a good word and a good idea that is currently misunderstood. We will, of course, address it more later.

Professor Schrödinger’s book was merely one man’s idiosyncratic view of an extremely complex phenomenon. This book merely represents my idiosyncratic view of that same phenomenon. Schrödinger was greatly limited by the existing knowledge of his time. Although I have access to vastly more information and ideas than did he, I am still greatly limited. I have personal limitations in what I am able to know and understand, but I am also limited by a severe lack of data. Even though the data available today dwarf’s that available in the mid-twentieth century, it is still a demonstrably incomplete set of data. Consequently, my view of Life could never be complete. I
seriously question whether any human will ever have a vantage point that allows for a complete view of life.

It is grossly presumptuous to speak of Life in terms of what it is and how it behaves, but this does not appear to have stopped anybody from ever doing it, least of all me. I am not an expert in any field, but that is an advantage in many respects. Any understanding of Life will necessarily be a multi-disciplinary affair. Surely it could never be limited to the thoughts and opinions of biologists or chemists. It is best viewed from a broad platform not a narrow one. So, from my broad, highly idiosyncratic platform, I shall give you my view of life. It is necessarily a very general view, but I’m sure that I will include enough specifics to annoy experts in virtually every related field. I want to keep this accessible to a general reader; whereas the biochemists have constructed huge barriers to entry in this field. I believe that we will benefit by tearing down many of those barriers and allow linguists, mathematicians, puzzle-solvers, cryptologists, crystallographers, philosophers, and yes, even theologians to flood the field. The riddle of Life is too big for any one man, and far to interesting to be restricted to any one field. I suspect that in the future, if it is to be restricted to a single area of expertise, that area will be information technology. Call me crazy. Everybody else does.

My view is necessarily incomplete, but there is another more significant difficulty involved with any description of Life. Life is a metaphor. It is not just that our description of Life is a metaphor it is that Life itself is a metaphor. At the highest level of comparison in the universe, there are only two choices: Life and Not Life. Science will obviously have much better luck at comprehensively describing inorganic systems, but then again, inorganic systems are the only things we can use for comparison to decide that any system is organic in the first place. In this book, I will describe the inorganic system of minerals so that we can use it to compare and contrast with Life. Minerals have definite rules of formation and molecular behavior. They are simple rules by comparison and markedly different from organic molecular systems, yet there are definable first principles of molecular behavior that are clearly shared by inorganic and organic systems. It is only by starting with the similarities that we might finally and fully appreciate the differences.

Life at bottom is a molecular phenomenon. It involves complex forms of molecular information, but all molecular information is based on the specifics of molecular structure. Organic systems are distinguished by their extreme complexity, and this complexity plays out in time and space. Organic molecular systems are manifest by complex molecular languages. It is the language of life that leads to it being defined by events and not merely by the molecules involved. Life is a process not a thing or a single event. The scope of Life’s many processes is vast in terms of numbers, variety, and scales of both time and space. Although Life is defined by distinct boundaries within the system, there is no distinct boundary within Life on this planet. It is a single, vast, complex crystal system on Earth.

A metaphor is an implied meaning between things within a specific context. This is exactly how molecular meaning is derived for all molecules within any organic system. No molecule has any inherent “meaning” outside of the specific context in which it exists. The meaning of every organic molecule is derived from the logical structural
relationships with all of its fellow molecules. This determines not just the quality but also the quantity of molecular information represented by any single molecule. The metaphorical nature of molecular information makes it recursive and unavoidably complex. In addition to the metaphorical nature of molecules, we must also accept the metaphorical nature of the human languages that we might hope to use in describing Life. So, we have the unavoidable situation of applying metaphor to metaphor, which surely can only lead us further than hoped from our desired destination: understanding Life.

We essentially do not “know” the fundamental nature of the universe, only some of its gross patterns of behavior. Life is the most complex manifestation of the universe, so it will be the most difficult to comprehend. Perhaps we will never be able to comprehend it in any way that we can find satisfying. This point could never be made any better than it was by Evelyn Fox Keller in her book Making Sense of Life:

“But by what mandate is the world obliged to make sense to us? Is such an assumption even plausible? I would say no, and on a priori grounds. One need invoke neither divine intervention nor unknown forces in order to doubt our ability to make rational sense of all that we encounter in the natural world. The human mind does not encompass the world; rather, it is itself a part of that world, and no amount of self-reflection provides an escape from that limitation. Most of us would agree that the mind – along with its capacity to make rational sense – is itself a biological phenomenon and hence a product of evolution, brought into existence by the forces of natural selection. The selective advantages accruing from the ability to make sense of one’s immediate and even remote environment are obvious. Yet an evolutionary process that could give rise to a mental apparatus with an unlimited capacity for making sense, however desirable such a capacity might seem, is difficult to imagine. And conceiving of a process that would produce minds capable of fully comprehending themselves is more difficult still. I can imagine neither a design nor an associated selective advantage that would be adequate to such a capacity.

Once we grant that our access to the natural world is not unlimited, we might also recognize that other biological marvels crafted by natural selection could be equally elusive. I see nothing counterintuitive in the possibility that there are phenomena in the natural world extending beyond the grasp of human comprehension – if only by virtue of their sheer complexity. Embryonic development might very well be one of these.”

Nonetheless, we can try. And since this is an idiosyncratic exercise in metaphor creation - we will be going all the way here. I will attempt a systematic reconstruction of many basic terms in the English language so that they can be applied to a new understanding of Life. Perhaps the reader has already noted one of my quirky lexical conventions in that I consistently capitalize Life. This reflects the difference in my mind between the singular process in the universe, or at least here on earth, that is Life and the more ordinary process involved in a single life, like yours or mine. Beyond this I will introduce many new conventions to the language that I hope to use in describing Life. In the process, I intend to ignore some old standards in biology and biochemistry. I feel they have become comically dysfunctional in some cases. Some of the words and their meanings have become counter productive. I will try to ignore the vast confusion that naturally accompanies many of the necessary naming conventions in biochemistry. As I
previously stated, I hope to keep this book accessible to a more general audience. However, I will not be able to avoid some basic mathematical terminology, because this – Life in general - is an area dominated by information, and information is an area dominated by numbers. Furthermore, this is an area where much of the information consists of spatial structures, which is the natural domain of geometry, again a decidedly mathematical endeavor. I will try to introduce graphics where I can, because discussions of molecules and molecular information are discussions where ordinary words tend to fail. It is quite difficult to convey some of the more important yet perhaps subtle points with mere words alone. Therefore, I will sprinkle this book liberally with diagrams, icons and visual enhancements. We are entering a world of insentient molecules, so we need to begin to imagine how an insentient molecule finds its way in the world. It is not an easy task under the best of circumstances.

In some places we cannot avoid math, physics and chemistry, but I will limit them where possible. There are literally trillions of possible distinct biomolecules, and they each and every one can and should be named. Understandably, this will tend toward an onerous and confusing lexicon; therefore, I will avoid it where I can. In some instances, I will rename molecules and their processes in simpler, more descriptive ways. This will undoubtedly annoy those who have put in the time to master the conventional system, but if they bear with me, they may find new ideas useful to their old ways of thinking.

Since this book deals exclusively with the topic of Life, and since it addresses the issue of Life’s origin specifically, it is bound to ruffle more than a few feathers. I am well-aware of this and I try to be sensitive to people’s sensitivities to the extent that I am a remotely sensitive person. But what we are really concerned about here, when we speak of sensitivities, is beliefs. Everybody has them whether they acknowledge it or not, and everybody has strong beliefs when the subject is Life in general. I am no different. I have a strong set of beliefs about Life and its possible origins, and those beliefs will be front and center throughout this book. It is inevitable that my beliefs about life will stray into the domain of other’s beliefs about religion. I feel it necessary to say then that I agree with Einstein when he said that science and religion are merely two branches of the same tree. I believe that both are essential parts of the human experience and they both deal in the same business of questions and answers. Science offers nothing but questions because every answer in science merely leads to one thousand new questions. Religion tends to fare better in the business of answers, but sometimes the questions can be quite a bit trickier. Regardless, I see no reason that there cannot be room at this table for both science and religion, if that is what you seek. At the end of the day, any view of Life will require a huge measure of faith. There simply is no way of answering some of the more difficult questions that will inevitably arise, and perhaps it will always necessarily be just this way.
This book is one long argument designed to not only describe but also convince the reader to see things my way. No system of thought or structure for arguing can avoid being axiomatic at some level, and this one is no different. I intend to argue about the nature of Life axiomatically on several points, and so here they are:

- Life is singular
- Life is universal
- Life is informative
- Life is mineralogic

I will briefly explain each of my axioms now and allow the details to emerge from the rest of the book. I have briefly already touched on the first. I believe that Life is a singular process on Earth. In other words, for all intents and purposes, there is a single form of Life on Earth today. The second axiom, Life is universal, is a statement of my belief of how the universe is constructed and not about my view of how widespread life may be in the universe. I believe that the universe is constructed in such a way that Life becomes inevitable. Life in this universe is not an outlier or a fluke, it is the natural extension of universal processes operating under normal conditions. Life on earth perhaps represents an extreme instance of those processes, but it is nonetheless a product of the inherent logic of our universe. The third axiom, Life is informative, is a statement of my belief that the universe, and by extension Life, is a manifestation of a purely informative universe. In other words, all universal and organic processes are at bottom a manifestation of information processes. The last axiom is perhaps the most offensive. Life is mineralogic means that the information in organic molecules is fundamentally no different than the information in inorganic molecules. All molecules think alike. All molecular information and all molecular behavior is, at bottom, based on the informative properties and logical relationships between molecular structures.

Let’s get to it.
II. What is Information?

The concept of information is absolutely essential to understanding Life because Life is informative. Information is real. Consider the concept of money, where monetary value is abstract, yet money itself is real. For instance, the copper token known as a penny is worth one cent, which is one-hundredth part of one dollar. The penny is real even though the monetary value attached to it is merely relative to many other complex things in a complex human economy. However, from the standpoint of physical information, the penny can exist flat on a table in one of two states: heads or tails. You can think of a penny as a two-sided dice. There are two choices for the state of a penny so a penny is an ideal binary information device. Binary is a number system of base two. One binary digit is called a bit and it quantifies one choice from a set of two choices. By convention, the two digits in most binary systems are “0” and “1” so we can call tails 0 and heads 1. Flipping a coin, in this case a penny, is a good random binary number generator.

Consider a cloth bag of one hundred pennies. Take them out one at a time and flip them randomly in the air. Flip all one hundred pennies, and as they land, arrange them into a line of pennies that are randomly heads and tails. You have now generated a random binary string of one hundred bits of information. This information is real and it is recorded in the physical form of a sequence of pennies. Take a fine black paint brush and paint a 0 on each tails, and take a fine white paint brush and paint a 1 on each heads. Now we have a random string of 100 binary digits that perfectly represents a real world phenomenon. Collect all of the pennies from the table and put them back into the bag, and then use your black brush to label the bag “100.” This bag now contains exactly one hundred bits of real information. We can weigh the bag and divide the weight by one hundred and determine the information density of this bag of pennies.

The bag is labeled “100” because it contains exactly one hundred recorded random pennies, but let us consider a case where a bag contains one hundred two-headed pennies. If we flip the pennies in this new bag we know that every one will come up heads. We now paint them all with White 1s and put them back in the bag, which we must now label with a big black “0” because the bag contains no information. The weight of the bag is the same but the information density is zero. This is because information, like money, has a value. The value of information is determined by the number of possible choices and the probability of each choice. When the probability of a single binary choice goes to 1 then the information value goes to 0. The higher the probability of each choice the lower the information value of that choice. This is known as entropy in the parlance of information science. So, if the bag contained one half fair choices and one half two-headed pennies, the entropy or the information value of the entire bag would fall out at around 88. This is because we could “guess” the result of any toss with better than random accuracy. The information of anything is related to the number of possible things and the probability of each thing.

Now let us take our initial bag labeled “100” over to a different table and dump out all the pennies. We can orient all of the pennies as either heads or tails as they originally came up, and in this way we have translated information from one table to another. However, we have lost quite a bit of information about the original event. We
have no way of recreating the sequence of coins in the order in which they fell. Information has been lost in translation.

So, let’s start with a new bag of fair pennies and flip them all on the original table. Label them as we did before, and arrange them into a line. Now place a red string across the line and dab a dollop of hot-melt glue on top of the string in the middle of each penny. Allow it to cool briefly (kids, get your parents to help.) and place the string of pennies back in the bag. What number do we now use to label this bag? The answer is 2096. In other words, the string has added a tremendous amount of information to the bag and the information density has gone way up. The string that bonds the pennies into a sequence has added a tremendous amount of time information to the process of reconstructing the timeline for our original random penny tossing. When we take the bag over to the second table, we can now not only recreate precisely the outcome of each toss but also the order of each penny for all tosses. But now our curious little minds must wonder: where exactly did the pennies actually fall on the table?

Back we must go to the first table. Toss and label the pennies as before, but this time take a thin wire of spring steel and kink it with needle-nose pliers so that it traces a path from the first penny to the second, and on down the entire line. Gather up the spring-loaded string of pennies and compress them into the bag. Now how do we label this bag? It depends. It depends on how accurate the line of pennies is and how accurate we need it to be. However, we can now take the bag over to the second table and recreate most of the penny tossing event in time and space. However, we still only know the order of each toss and not the amount of time between tosses. Back to table number one.

Lay out a grid on the first table and get out your stop watch. Go back to using string to bond the pennies together, but add three binary numbers to the bond between each penny. The first number is the x-coordinate on the grid where the penny has landed and the second number is the y-coordinate. The third number records the time interval between tosses. We can perhaps make all three numbers of fewer digits by making the coordinates relative to the previous penny. Now each penny relative to the last penny is recorded in time and space by a series of black and white stripes painted on red string that bonds any two pennies. Now we are set, for when we place the pennies into the bag we have an easy way to measure the information and label the bag with its total information content. The majority of the information in the bag is now located in the bonds, so now we can return to table number two and efficiently recreate the random events in space and time.

So what? The main point of this exercise up to now is that information is real, it is stored in physical objects, it can represent random events in space and time, and it can always be translated into binary systems. It is an informative universe, and Life takes advantage of this fact in becoming life. We will now use this simple exercise to demonstrate how information can be used to translate things into other things.

Take a bag full of regular six-sided red dice and label each point of each die with a sequence of three black or white dots. We now have a simple spatial method for translating our string of pennies into the spatial information in dice. Every sequence of three pennies will have a corresponding point on a cube. Now, one cube is translated into three pennies and vice versa. We can think of thousands of clever ways to translate the pennies, their bonds and their timelines into a set of dice by merely rotating the dice and using the rotation as another form of information. We could then take the dice to the
second table and quickly recreate the random penny events with perfect fidelity. In other words, information can easily be translated from one form to another and back again.

Another way to do this, perhaps, is to translate the pennies into half-dollars. Let us pick four dates from a set of all possible dates for half-dollars: 1961, 1962, 1963 and 1964 – all good years, or so I’m told. Now, two bits of penny information equal one half-dollar. We can conceive of a system of flipping, rotating, bonding and labeling the half-dollars so that we efficiently capture all of the penny string information within a string of half-dollars where the bond information can be represented by physical storage in quantized rotations of the half-dollars. In other words, the original penny information can be stored in half-dollars, and the bag of pennies will have the same numeric label as the bag of half-dollars. However, in some sense, the information can be compressed into the system of half-dollars, perhaps, because there are fewer components and less markings. This is achieved by biasing the set of coins through careful selection of a small set of coins. Regardless, we can take the half-dollars over to the second table and decode them into the exact penny events that were recorded and stored in the pennies, translated into half-dollars, and then re-translated back into pennies. It is a symmetrical system of translation from pennies to half-dollars and back to pennies.

One thing definitely needs to be noted here: this system depends on us. The information system depends on an operator to execute the system. We have developed algorithms for executing the random penny system. An algorithm is merely a set of steps to achieve a particular function. I can record this penny algorithm and transmit it to you, and you can read it and perform the algorithm. This is another important form of information contained in the system. The code of penny tossing is itself quantifiable in terms of real information. More importantly, the numeric labels on the bags also depend entirely on the information in this code. Without the code, the information in each bag is zero. Depending on the precision and detail of the code, the information content of each bag can go way up or way down. Pennies and algorithms are interdependent for their information content, practical applications and their “meaning”.

The random penny exercise is just a convenient way to understand computer systems in the simplest possible terms. In this analogy, you the operator are the computer. This creates a computer metaphor for our thought experiment. You have basic logic circuits wired into your brain that allow you to interpret codes and perform simple functions. The original penny sequences represent raw data, or input that can be operated upon. The penny sequences represent output, or interpreted data. The algorithms and the data are mutually dependent. In other words, without coins there are no codes, and without codes there are no coins. The entire system is real, and it is recorded in and executed by real physical entities in this universe. The primary logic is hard-wired into the fabric of these objects, or more simply it is hardwired into the fabric of the universe itself. We have yet to determine that exact logic in the simplest possible terms, but that is precisely our goal. However, there are an infinite number of possible systems that could execute these basic functions. We are merely describing one idiosyncratic instance of this system. We should have no trouble imagining others. In fact, it is instructive to try to imagine more robust and efficient versions of this same system.

How might we do more with less? That is the question that Life always asks. The answer that Life comes up with time and again is Life itself. We can try to do so
here by applying a little brain power and creativity. For instance, the stopwatch analogy is somewhat of a kluge. We can create a more elegant substitute for time. Consider that our time precision, beyond the basic timeline itself, is perhaps not too important to us. How can we embed a simple timing mechanism into the objects themselves? Suppose that the original bag of pennies include various dates, but the dates are not evenly distributed. Suppose that there are twice as many even dates as there are odd dates. Suppose that we include a form of sequence symmetry in our penny pulling algorithm so that the sequence must include consecutive even dates followed by a single odd. Now it will take us some quantifiable amount of time to pull pennies and check dates to get an even or an odd. Over the random course of pulling many pennies, a time symmetry will develop that corresponds with the sequence symmetry. When we interpret the data later, we can know that each odd penny took twice as long to add as did each even penny. This will efficiently store the time information in the identity of the penny itself. It is but a single bit of time information that can differentiate a fast bond from a slow one. Through a clever use of limited parts, we have further evolved into a more efficient algorithm of fewer parts for storing more compressed information.

Let’s add a twist to the exercise. Let’s add a dimension of height to the table on which we flip our pennies. The table now contains yet more spatial information that relates to three axes and not just two, or it is now a robust spatial topography. If the table is a rough, hilly terrain instead of a smooth grid, perhaps we can store yet more information in our system. It seems likely that pennies will come to rest in low places more frequently than high places, and yet this simple fact represents more information about the system that can be stored by the system. More importantly, we can encode this information, somehow, in the sequence of coins and the sequence of bonds in an infinite number of ways. Perhaps we can even find a way to record heights in a clever way with our sequences of timing. In other words, rare sequences represent higher topographer than do common sequences. Regardless, we can see that combinations of a few well-behaved components of an information system can quickly convert random events into quantifiable, recordable, translatable physical information. But what could we ever do with it? Quite a bit, actually.

What if we wanted to recreate the initial table? All we would need is a repeatable penny algorithm and a system to interpret it. We could produce bags of efficiently stored and labeled half-dollars that we could transport anywhere and replicate the original table onto which the original pennies fell. We could build the second table purely from the information contained in the bags of pennies. We could store these bags for as long as we cared to, and translate them into any material we felt was necessary. Heck, we’ve got all day. In fact, we could run the penny flipping portion of the program endlessly, and then select only those sequences that efficiently map the table in time and space. By selecting only the “best” sequences for translation, we will further compress the data for efficient storage. We also will further bias the data into a group of highly selected bags. But remember, the bags of half-dollars already contain a highly biased set of coins because we selected only four dates to begin with. This initial bias represents a huge amount of information in the total system, because if we did not “know” the dates of the original coins, then a huge amount of additional information would be required to operate the system. Bias is a powerful way to compress and store information. Our bias is recorded in small sets of coins and small sets of bags.
Suppose we did not want to build a single table but an entire fleet of tables. We would need a system to replicate the bags so that many tables could be built at the same time. Suppose we got a set of magnets so that two magnets perfectly fit together and the other two did as well. We could glue one magnet to 1961 and its mate to 1962; likewise we could mate 1963 with 1964. Now, the process of replicating strings of half-dollars would be a simple process of pulling strings through bins of free half-dollars. We could make complementary strands that could also make complementary strands. The bonds would ensure that the information stays recorded in our coin sequence universe.

This exercise provides a decent analogy of information systems in general. It makes any discussion of computer systems highly metaphorical, but they always are anyway. Discussions of Life are highly metaphorical of information systems, as they need be. However, most conventional descriptions of Life leave out most of the important details of information systems, because our standard paradigm of Life is wholly inadequate. We can see from this that sequence, bond, code, selection, bias, translation all have new meaning here, and all of these ideas are essential to understanding life’s molecular information system. Our discussion of Life will be even more metaphorical because Life is a Molecular phenomenon and we are not individual molecules. We are also not computers, in the conventional sense, but an understanding of computers is essential toward an understanding of Life.

Consider the case of a digital camera. When you snap the shutter on your digital camera, it captures about five million tiny dots of color information on a rectangular grid. This represents an enormous bag of coins, and you would never need all of them to recreate that pattern of colors. The camera has algorithms for compressing that same information, more-or-less, into a much smaller bag of coins. That digital picture file can then be stored on your computer’s hard drive. You can copy it, send it, look at it, or print it as many times as you like. Life does something similar with molecules. It captures molecular information, stores it on a hard drive – DNA – copies it, sends it, looks at it, and prints it out as many times as it needs to in the form of other molecules.

Consider the case of the internet. The internet is not a single computer and it is not even a single computer network. It is a complex network of networks. Information exists in the form of files, applications, network systems, collections of network systems, and at the highest level it is a singular computer entity known as the internet. Life is very similar to this in that it is very hierarchical. There are individual molecules, individual codes, individual systems. There are codes of codes and systems of systems until finally at the highest level we see an individual entity that is Life itself.

We have been using computer metaphors for life ever since Erwin Schrödinger proposed that life operated on a code script. These metaphors have been stuck using a decidedly old parlance. Our concept and mastery of computers has greatly outpaced our concepts and mastery of biochemistry. The complexity of computers does not nearly rival that of Life, yet our computer metaphors for Life fail to even capture the complexity of computers, let alone the complexity of Life. We have upgraded our computers many times since we started using them as a metaphor for Life. It is now time that we upgrade our metaphors.

There is another simple way to conceptualize information. We can start by recognizing that information exists when there are boundaries for things and they produce
sets. We produced a date boundary for our half-dollars and thereby generated a set of four half-dollars. We produced spatial boundaries with strings and bags and thereby produced sets of strings and sets of bags. The table had its own spatial boundaries and we produced sets of tables. When you have sets you can begin counting and you can begin making choices. This produces probabilities and so those choices can be quantified and recorded. Information is produced and exists in physical form when these things happen. When you have sets, you can then have sets of sets. Here again, all of the previous operations can be repeated and yet more complex information is created. Information becomes hierarchical and yet more information emerges from this hierarchy. Naturally, logical relationships between the sets must exist, and information can travel between them, and translations can occur between sets. Any logical relationship between one set and another, or even logical relationships within sets, can be called a function. The logic of functions, and the combinations of logic between functions is called a code. Any algorithm that is used to perform a code is called a language. Therefore, all languages are codes and all codes are functions. There are many possible languages to execute any code and any function. Life has many molecular languages to execute its many codes consisting of many molecular functions. Life’s languages are designed to generate sets and logically relate them to each other. The languages themselves are highly metaphorical because all molecular meaning is derived from the relationship of one molecule to many others. No molecules would exist in the absence of molecular codes, and no codes would exist in the absence of these specific molecules. The codes exist as the collection of molecules. The molecules exist as functions of the codes. The information quality and quantity of each is mutually dependent. Life is highly metaphorical and so our description and understanding of it must be more so.

The universe itself is an informative structure. This point was made entirely clear to us with the creation of the discipline known as quantum physics. This is our basis of physics and chemistry, and no single scientific discipline has been more successful than quantum physics. It explains many observable phenomenon in our universe, and it is the basis for making and using the periodic table of elements.

Quantum physics provides the basis of our model of an atom. An atom is seen as a set of sets. Each atom has a set of protons, neutrons and electrons, which are themselves sets of sets. The periodic table is constructed mostly around the set of electrons in each atom. The first quantum physicists, scientists like Erwin Schrödinger, got started by quantizing all of the components of an atom. This meant that all of these components could be bounded in various ways, counted, and assigned whole numbers. Logical relationships were established between the sets. Every electron in an atomic set was defined by the set of whole numbers that describes its unique properties within particular bounds. The final property was known as electron spin. The final piece in the basic quantum puzzle is known as the Pauli exclusion principle, which states that no two electrons in an atom can share the exact same quantum numbers. This means that although all electrons are perceived as identical, no two electrons bounded by a single atom can have identical properties. This is a statement of the purely informative nature of the universe. After all, sets of atoms, or sets of quantized parts, are responsible for giving us the informative universe that we observe.
Our understanding of the universe depends on the belief that it is an informative universe, and our understanding of Life should be no different. We should strive to incorporate first principles of information into Life, and strive to utilize our most sophisticated understanding of information when we study life. We should search for the basic informative relationships within life so that we can conceptualize them and communicate them in ways reminiscent of the periodic table of the elements. We can do this.
III. What is Space and Time?

When we talk about Life we must put it into context. The context for Life is the universe. When we talk about the universe, at bottom, we are talking about two things: space and time. We have already seen how we might want to think of this universe as an informative universe, but what does this mean in terms of space and time?

Nobody yet has a comprehensive understanding of the universe and so nobody has a bulletproof notion of space and time. Our models of the universe are always imaginary because at some level we must always imagine space and time. Since we want to imagine a quantized universe that is purely informative, I prefer to imagine a quantized space and a quantized time. If we start at the very bottom, we can imagine the smallest universal unit of space and the smallest universal unit of time. At this level, I prefer to see space as composed of units of pure information and time composed of a single unit, or one universal tick of the clock. Time only moves forward, and the only impact of each tick of the universal clock is that all space must change. In other words, whatever the unit of information there is at every unit of space at any specific time, it must change with the next clock tick. The way that space must change is unknown to me, but I believe that the change in space will be dictated by the information contained only in neighboring space. In other words, there can be no force at a distance in the universe. Information can only propagate through space at a maximum rate dictated by the unit of space and the unit of time.

As we try to imagine the exact nature and configuration of space, we must start to think about geometry. I am not going to get into geometry to any level of sophistication, but there are a few basic things that we must think about. For instance, what is the shape of the smallest unit of space? I have no idea, but we conventionally think of this as a point in space. Most forms of geometry, which are always metaphorical human languages to communicate spatial concepts, are generated from the concept of a point. This is unfortunate because a point is seen as a dimensionless entity, so how can we fill space with them? I prefer to start with a plane and arbitrarily assign it a smallest unit of thickness. We could intersect three planes, and our unit of space would be a cube. However, we will never be able to perfectly build all other shapes if we start with a cube. We will never be able to build a single dodecahedron if we start with three planes. I know that I will need a purely informative dodecahedron later, so I merely start with the six planes of a dodecahedron, and I imagine that geometry is built up from there.

Regardless, once we have imagined space, we must imagine time. I have described how time works in general with regard to unit space, but how does it work with respect to all space? In this context, time becomes a measure of scale. Structures that arise in space are limited in their size and behavior by time. Structures can grow, move, evolve, but always on a particular scale in time. Time is the universal metric of scale in space in a purely informative universe.

Since this particular universe that we imagine is purely informative, it must have some rules for it to be animated and meaningful. We do not know these rules, but we must know that they are self-organizing. In other words, what ever those rules are, they must be sewn into the fabric of space and time, and all of the rules must be available to every unit of space during every unit of time. This, of course, paints a gross picture of
the universe as a singular, multi-processing computer, which I’m sure is unacceptable to most, but this is the picture being painted here. The result, however, is obvious. The universe will start with a structure and a set of rules, the programmer will set those rules and provide the starting state of the universe. Through time, structures will form and they will be bounded in space and time by the nature of the patterns that they form. More persistent patterns will be more tightly bounded and they will interact with other persistent patterns. Through time, larger structures and more hierarchical structures will emerge. The largest structures will take the longest time to form and will be the most complex. The universe itself is one gigantic pattern, but there is a seemingly endless number of patterns as we drill down into it. The universe is a fractal, but we’ll touch on this a bit later.

Our brief and idiosyncratic description of the universe here can be summarized quickly. The universe is purely informative, or more accurately it is quantized. It consists of pure information in terms of two things: space and time. Space and time are quantized and they define each other. They can be understood in terms of pure information, but their exact behaviors are too complex for us to now understand them in the kinds of details that we hope to know. However, from this platform we can turn our attention back to Life, because it is a function of the universe, and we can better understand the details of Life to which we are so far privy. The universe is constructed of laws and Life is a manifestation of those laws. Life is not an outlier or an arbitrary event. Life is a logical and consistent pattern in this universe. It is a manifestation of the broad pattern of the universe, but it is an intense manifestation of the most highly organized portion of that pattern. Life is the frilly lace at the fractal edge of a universal pattern.
IV. What is Symmetry?

Symmetry

**symmetry**
sym·me·try [simmatree]
(plural sym·me·tries)
noun
1. **property of sameness:** the property of being the same or corresponding on both sides of a central dividing line
2. **balanced proportions:** harmony or beauty of form that results from balanced proportions
3. **MATHEMATICS exact correspondence in position:** a correspondence in the position of pairs of points of a geometric object that are equally positioned about a point, line, or plane that bisects the object
4. **PHYSICS state of invariance:** a state of invariance shown by some phenomena when changes of orientation, charge, or parity are made

[Mid-16th century. Via Latin and Greek summetria, literally "similar measure," from, ultimately, metron "measure" (source of English meter).]

Microsoft® Encarta® Reference Library 2005. © 1993-2004 Microsoft Corporation. All rights reserved.

Symmetry is like art – we know it when we see it. However, some people are better at seeing it than others. I suppose I am really good at seeing it because I see it everywhere. I have provided the above dictionary definition of symmetry so that we will all have the same formal definition. Suffice it to say, I intend to go well beyond this definition of symmetry. We must.

Symmetry is sewn into the fabric of our universe. It is the first principle of organization in our universe. Without symmetry, there would be no organization in this universe at all. My one line definition of symmetry is that it is change without change. When things change yet stay the same, then that is symmetry. We could have millions of examples, but I will share just a few. Please note, however, that even though time moves in only one direction there is still a strong symmetry to time. The idiom, the more things change the more they stay the same, is a good summary of this idea. Things can change dramatically through time, yet first principles and basic functions can stay the same. Life is marked by a very strong time symmetry in this way.

Symmetry drives geometry, of course, but the way it does this is perhaps not so obvious. After all, there are many different forms of valid geometries, and they cannot all be built upon the exact same symmetries. The geometry with which most of us are familiar is based on simple and intuitive symmetries. It works well for most things, but when complex additional things are added to the system, simple geometries fail. For instance, an entirely new geometry was created to explain the symmetry between time and space that Einstein insisted upon to describe his model of the universe. In this exotic form we must allow that two parallel lines must at some point be allowed to cross. There are other geometries that are yet more extravagant than this. I mention this merely to point out that when we talk of symmetry and geometry, symmetry always comes first.
Besides geometry, there are many other kinds of systems that require symmetry. When systems include sequences, like human language, computer systems, and organic molecular languages, they too require symmetries. There are many different kinds of sequence symmetry. Sometimes these symmetries themselves contain meaning, like in the case of digital compression, or in the case of English palindromes. A palindrome is a sequence that can be read the same in one direction as the other. A good example of this is the panama palindrome:

A man, a plan, a canal – Panama!

The structure of this one palindrome can be used and expanded almost endlessly. Some folks have devoted the time to extend it to remarkable lengths, but I will share just one example found by Guy Steele when he extended it to 49 words:

A man, a plan, a canoe, pasta, heros, rajahs, a coloratura, maps, snipe, percale, macaroni, a gag, a banana bag, a tan, a tag, a banana bag again (or a camel), a crepe, pins, Spam, a rut, a Rolo, cash, a jar, sore hats, a peon, a canal – Panama!

I have examined a baby in the hospital whose name was a palindrome and the parents did not even know it. The point being, some languages are designed to contain palindromes, and English is not one of them, yet even this language cannot help but contain many elaborate palindromes. There could be yet more symmetry in a language than reading direction. Depending on the shape of the characters, a sentence could be flipped upside down or held to a mirror and still read the same. Some people have practiced the art of writing names and words with 180 degree rotational symmetry, and they have become skilled to the point of becoming commercially viable.

Sequences and languages are inherently symmetrical because at some level they are all built on inherent symmetries. However, it is usually only when these symmetries become broken that we recognize the utility of the symmetry in the first place. Using symmetry to create information and language is a simple two step process. Step one is to find a starting symmetry. Step two is to break the starting symmetry. Let me give you an example for geometry. Start with a tetrahedron because it has the simplest perfect symmetry of volume.

Figure 1.

All four points of a tetrahedron are identical. There is symmetry here but it is not yet broken and therefore we cannot use it as information. However, when we break the symmetry of the four points of a tetrahedron, we have created four things to count and a logical relationship between those things. The four points of a tetrahedron form a simple set of points bounded by a sphere, and those things can be counted, chosen and acted upon. We can see the symmetry in a tetrahedron by changing the tetrahedron through rotation. The points change but the tetrahedron itself and its inherent symmetry do not. It is change without change.

Figure 2.
Our selection of four colors and our selection of this particular arrangement of those four colors represents several different forms of information. There are twelve isomorphic rotations of this one tetrahedron, but there are also two different choices of these four colors. We chose this one, but we could have chosen the mirror twin. There also is a dual tetrahedron to this first tetrahedron, and we can pick up the information of the points of the dual by adding a single color, in this case we will add purple.

Figure 3.

There is a logical relationship between the points of a tetrahedron and the points of its dual, and all of these relationships can be represented by only five colors. However, a tetrahedron is not merely a set of points, it also consists of sets of faces and edges. The four points of the dual correspond to the four faces of the original, and vice versa. Of course, all of these additional elements are logically related to each other by the set of points. Furthermore, the set of points consisting of a tetrahedron and its dual form a cube, and a cube has faces and edges as well. So merely by breaking the symmetry of a single tetrahedral set of points, assigning colors for the points and their duals, we have invented functions, codes and an infinite number of possible languages form relating many sets of logically related things.

Figure 4.

Things quickly get better when we realize that the cube itself has a dual called an octahedron, so our system picks these elements free of charge. The octahedron has six points corresponding to the six faces of the cube.

Figure 5.

We start with a set of four points on a sphere, and because of perfect symmetry and our clever ability to break it, we generate a spatial language of immense power. The sphere is the most symmetrical object in space because every point is identical to every other. However, when we start arranging points on a sphere we begin to limit our options for creating perfect symmetry. We have built our way up to a cube rather easily, but there is only one option left open to use. We can create five cubes, one for each of our colors, and we can symmetrically arrange them on the sphere. This generates the highest level of perfect symmetry for points on a sphere, which is a dodecahedron. The dodecahedron is created from twenty symmetrical points on a sphere. It consists of five cubes that each share two points. Each of the shared points in the cube is taken from one of the dual tetrahedrons that made the cube. The dodecahedron itself has a perfect dual called an icosahedron, which has twelve points corresponding to the twelve faces of the dodecahedron.

Figure 6.
We have now generated all five possible perfect solids in the universe by starting with the four points of a tetrahedron. More importantly, we have merely broken the symmetry of those four points and built a comprehensive system of spatial information that relates all structures to all others, and all elements of all structures to all others.

Figure 7.

In an effort to make this information into a sequential language of shapes, we might find it helpful to produce maps of this information. Maps can be made in many forms, but humans prefer them in globes, or more often in flat sheets of paper. We can easily map the relationships of a tetrahedron, an octahedron and a cube in a flat format.

Figure 8.

Unfortunately, when we get to the level of the dodecahedron, this is no longer easily done. I have come up with a Rube Goldberg scheme for mapping the elements of a dodecahedron on an infinite plane. This is what is known as a tessalation. We cannot perfectly tessalate a dodecahedron because its faces are pentagons and pentagons do not perfectly fill a plane. We cannot tessalate an icosahedron because its faces are triangles and they join with four others, where they need to join with five others to fill the plane. We can, however, create a quixotic graph of dodecahedral elements that will fill a plane.

Figure 9.

This graph allows us to map out circuits of elements, but it does not allow us to do so with a single shape. It can, however, be wrapped around a cylinder and provide a crude but entertaining representation of a double helix.

Figure 10.

If we ever want to fully tile a plane with dodecahedrons we can only do it in fractal fashion. By making ever smaller symmetries of pentagons using the scaling function of phi, we can make whole multiples of successive pentagons and thereby fractalize a plane of pentagons.

Figure 11.

It is not such a big deal that we tessalate a dodecahedron, because we can always make sequences of dodecahedral elements using a globe, and if we are clever, we can do this using only four colors. This is exactly what Life did when it set up its language of protein synthesis called the genetic code.

Figure 12.

We will go into more detail about the genetic code later, but there are a few key concepts here with respect to symmetry. Symmetry is the basis of information but only
when symmetry is broken. There are many different kinds of symmetry - time, space, sequence to name a few – and they can be isometric. In other words, different forms of symmetry can share the same inherent symmetry properties and thereby information based on symmetry inherits the properties of the various forms. Spatial symmetry can give to sequence symmetry and sequence symmetry can give information back to spatial systems of information. Space and sequence can be isometric in their symmetry. This is an essential feature of any system that wants to translate information symmetrically between space and sequence through time. This is also an essential feature of any molecular system that wants to create sequences in space. This is a very important feature for living systems to not only function but to actually exist. We have already seen one example of how sequences can be used to compress and store large amounts of information about space and time. What we need to find is a mechanism for bounding spatial information in time and space. In other words, how do molecules ever come to make sequences?

Humans are quite adept at recognizing spatial symmetries wherever they might exist, less so at recognizing broader, more abstract forms of symmetry. These forms are no less real and no less important when it comes to understanding the universe. After all, the universe is principally built of complex symmetries. One human being in particular was extremely adept at recognizing abstract symmetry. He was a Frenchman named Evariste Galois, and he only graced the earth during a short period of time from 1811 to 1832. During this short time he accomplished a lot. He proved himself to be a man of great symmetry. Unfortunately, his ideas of symmetry in government landed him in jail, and his perception of the symmetry between men and women got him prematurely killed. Fortunately, before his senseless death, he taught his fellow and future man a good deal about symmetry in problem solving. Every problem implies a solution and every solution implies a problem. Galois recognized that many problems have many potential solutions. This is epitomized in the disturbing idiom, there is more than one way to skin a cat. (Dear reader, please do not skin any cats to prove the point.) However, if you were to skin many cats in many ways you would learn what Galois already knew: some sets of problems lead to sets of solutions. One particularly abstract problem is the problem of solving quintic equations. Galois knew that quintics would have multiple solutions and that these solutions would form a set of solutions. Galois proved that the sets would have symmetry, and he called this symmetry a group. Galois thereby invented the formal mathematical discipline of group theory, and group theory has come to be the formal language of symmetry in mathematics. This language has proven itself indispensable toward understanding the deep workings of our universe, and the same should be true of understanding the deep workings of Life in the universe. After all, Life is a product of the universe.

Beyond mathematics, there are many useful applications for symmetry. We use it constantly in storytelling, art, music, heck we use it for everything whether we know it or not. Most human cosmologies are founded on first principles of symmetry: heaven and earth, yin and yang, male and female, are all apt metaphors for practically everything familiar to us. Our basic problem with understanding Life is that we have heretofore ignored the many important forms and applications for symmetry. Take reproduction, for
instance. When a single cell divides, it creates two symmetrical versions of itself. This form of symmetry is maintained yet made far more complex and far more powerful when reproduction is carried out through sex. Your children are not replicants of you (hopefully) they are symmetrical solutions to the problem of reproduction that faced both you and your mate. The system of reproduction must be based on intense symmetries for it to ever function effectively in this manner. Your genes, cells, body parts, human groups, ecosystems must be highly symmetrical on all levels for this to be the chosen form of reproduction. But the object of the system in general is one based on symmetrical problem solving to begin with. After all, the goal of reproduction is not merely to build new units but also to simultaneously search for new units. If this were not the case, we would find perfect replicators in nature that perhaps would never die. There will be more on this later.

The point of this chapter is to introduce the concept of symmetry in its broadest possible form. Symmetry leads to pattern and our understanding of anything is based on our appreciation of and understanding of patterns. Patterns can be simple or complex. They can be small or large. They can exist in time and space. However, any pattern can be dissected down to first principles, and when it is we will usually find a basic symmetry lying at the bottom of it. Therefore, we must expand our thoughts to include all forms of symmetry and recognize them where they exist and in the functions for which they are most useful. The universe displays an enormous amount of symmetry because it is built of complex fractal symmetries. A fractal is any shape that displays the symmetry of scale. Our understanding of the universe must begin with an understanding of symmetry and scale, so fractals become a powerful metaphor for our understanding of the universe. There are symmetries of time, space, scale, sequence, thought, and symmetries of all reality. Our quest to understand Life can never be divorced from a solid understanding of symmetry of all kinds.
V. What is Complexity?

There is a useful, fundamental difference between the idea that something is complex versus something that is complicated. A complicated system is one of numerus parts, yet the whole is still roughly the sum of its parts. Conversely, a complex system is one with lots of parts, but the whole is not the sum of the properties of its parts. Basic properties of a large complex system only emerge from the interaction of large numbers of smaller individual agents. The numerous parts of a complex system must have rules of interaction to provide feedback to each other, and only then can a higher level of organization emerge from it. The adage for complexity is that more is not more; more is different.

Take for instance an automobile, it is a complicated system built of many sub-systems. There is a wheel base, a drive train, breaks, steering, etcetera. It is a complicated system designed from “above” for the purpose of transportation. We can continue to add sub-systems, like audio, AC and even add sub-sub-systems, like anti-lock brakes and such, but the sum of the parts continues to add up to an automobile. It is a complicated system of many parts but not a complex system.

In contrast, a living cell is truly complex, devised not from above but from below. It too is composed of numerous sub-systems, and sub-sub-systems, but there is nothing inherently “cell-like” about any of the component parts. The properties of an entire cell cannot be guessed by studying any, or for that matter all of its parts. Instead, the cell-like nature of a cell emerges from its inherent complexity. A cell is an entire system of adaptive self-organization. Similarly, an ant colony is not merely the sum of individual ants; the actions of a colony emerge from the actions of a large number of individual ant agents. There is a queen, but there is no “leader” ant commanding and controlling the behavior of an ant colony. The behavior pattern of an ant colony emerges from the collective interactive behaviors of individual ant agents. Similarly, a brain is not a complicated network of neurons it is a complex population of neurons from which “thought” emerges.

So, by comparison, an automobile of today can be quite complicated but probably not complex, not yet anyway. Even the ultimate doozy of a car will never begin to match the complexity of a living cell. However, a group of automobiles, the global transportation system they comprise, or the traffic patterns they might create, is complex. Automobiles can form a population of independent yet interrelated agents that provide feedback to one another. Yet one cannot truly anticipate or understand traffic patterns by studying the sub-systems of an automobile. Complexity results from a hierarchical interdependence of sets of things at evermore increasing scales in time and space.

A curious thing to note here about the difference between studying complex and complicated systems: the general approach to describing a complicated system is to utterly simplify it, but the general approach to a complex system is to generalize it, to find the global pattern of the system. Perhaps we might view this as finding the symmetry patterns within the complexity. After all, it is not the details that are complex it is the pattern that logically relates the details that leads to complexity. Symmetry is an abstract property of things that can remain independent of specific details. But pure simplification will usually destroy the sublime order of a complex self-organized
adaptive system. However, in order to simplify a system it is usually prudent to isolate each of the components and then simplify all of the mechanisms within them. In the case of a truly complex system, this attack through utter simplification or pure reductionism will typically lead to a poor understanding of the system in general. The trees will be well understood, perhaps, but the forest will be missed entirely.

The era in which we began to understand Life in terms of information was an era still dominated by a prejudice toward science via simplification. Although it was widely recognized that things could be very complicated, perhaps even complex, there was an unspoken faith that no system could at bottom defy a rigorous effort to reduce, simplify and thereby study it. This faith was ill-founded. The common sense approach to scientific investigation then was based on the premise that extremely complicated and complex systems, like a living cell, would best be understood in the light of utter simplification. In the brighter study light of complexity theory, any premise of simplicity is a demonstrably false premise. It turns out that broad, general paradigms of complexity are always more useful in the study of biology, at all levels, than are the narrow and specific paradigms of simplicity. Life is never simple. Complexity is the only platform from which we can begin to understand biology.

Chaos is a concept closely related to complexity. Chaos theory is the study of non-linear dynamic systems. A chaotic system is unpredictable in the general sense of the word, but a chaotic system is not random. The hallmark of a chaotic system is its exquisite sensitivity to initial conditions. The concept was made popular by the phrase ‘the butterfly effect’ (they even made a bad movie by that name). It means that small perturbations in initial conditions, like a butterfly flapping its wings, will lead to large consequences, like tornadoes, in complex weather patterns. A complicated system will usually follow linear dynamics whereas a complex system will follow non-linear dynamics.

These sorts of intellectual tools and languages were not available to the founders of our classic paradigms of molecular information, the most prominent being the genetic code. It was not an easy thing then to imagine that a basic molecular language could somehow follow the strict rules of a linear code, like morse code, yet still be non-linear in its overall operation. Consequently, a clearly chaotic natural process, like protein synthesis, was forced into a tiny “linear” box, and remarkably, it still resides unhappily cramped in our thoughts and textbooks in that ridiculous little box today. There is no butterfly effect in protein synthesis today because our poor, delicate butterfly is hidden in and suffocated by the linear black box of our classic linear paradigms. However, it appears empirically obvious now that both protein folding and sequencing - neither one - are a simple, linear handshake between DNA and protein. Protein synthesis in total is a complex and chaotic dance with multiple partners twirling back and forth across the dance floor. When taken as a whole, in a proper broad context, protein synthesis is a complex unpacking of emergent molecular information, not a simple lateral data dump. And thus the classic paradigm of the genetic code restricts our view of the sublime grace, beauty and elegance of the complex molecular dance that is being practiced right in front of our eyes.

Life itself is a supreme manifestation of complexity in a complex universe. It is entirely hierarchical, so there is no limit to the number of levels on which we might chose
to view it. We can drill down or we can pull back, and we will always be surprised and entertained by the complexity and the beauty of the patterns that will emerge. The universe is a fractal and so Life must be one as well. The most famous fractal is one discovered by the father of fractals, Benoit Mandelbrot. A quick exploration of the Mandelbrot set is quite illuminating toward an understanding of patterns, scales, and the symmetry of Life.

It is not necessary that we have a total understanding of the Mandelbrot set to make it useful to us here. We merely need to know that it is a set of numeric relationships that is generated by the iterative looping of a single formula. When this iteration is performed we produce a set that can be visualized on a graph. Here is the broad view of that set:

Figure 1.

We can see a broad symmetry in this set, but more than symmetry, there are examples of symmetry breaking everywhere we look. We can arbitrarily pick two places on this general curve, and we will witness the creation of distinct new symmetries within these curves.

Figure 2.

The most interesting thing about this curve, however, is in what it can teach us about scale. You see, there is no inherent scale anywhere in this curve. It represents an endless regress of complex shapes. We can pick a small area from one of the two areas we previously chose, and when we drill down on that area we see the creation of entirely new shapes with entirely unexpected detail. As far as I know, there is no bottom to this curve. It’s only known scale is a scale of self-reference. In other words, the Mandelbrot set is a shape metaphor where all shapes exist and have meaning only relative to all other shapes.

It seems to me that this set is not an invention of man but a wonderful discovery of a hidden truth and inner beauty of our universe. Perhaps it means to us that we can infinitely hone our attention on finer and finer detail, and for our trouble the universe will reveal to use ever-finer and more complex detail. The more answers we find the more questions we are obliged to ask. Our quest to find the bottom shape that is somehow generating the total curve will forever go unfulfilled. We are trapped in a curve with no bottom and no end to the complex ways it can be juxtaposed to itself. This too could epitomize our quest to define, describe and ultimately understand Life.

Figure 3.
VI. What is Evolution?

Evolution is the organization of information through time. This is the broadest possible definition that we can give, and it is useful in this generalized form. However, the term “evolution” packs quite a wallop for those who like to use and abuse it. Most people think only in terms of organic evolution, but there is no logical basis for doing so. Organic evolution is merely one form that differs from yet resembles all other forms of evolution in many important ways.

Evolution denotes change and change implies novelty. Therefore, evolution can rightly connote the creation of new things. Change requires time, and evolution is indeed always a function of time. We have already seen how things can generally be seen as basic compositions of information, so it is acceptable to view these things within a framework of organization, or perhaps at times we should view them within a framework of disorganization. Evolution, therefore, can be seen as the timeline of organization of information. Sometimes this timeline will trend towards decreasing organization, known as entropy in thermodynamic terms, and these things can be said to evolve into less and less distinct patterns. Sometimes the timeline is toward increasing organization, and these things can be said to evolve negentropy and involve more distinct patterns. Life is a clear example of the latter. Evolution creates patterns in both time and space. The patterns of Life are marked by increasing complexity and increasing information.

Imagine an alien landing on earth soon after its formation. Suppose this alien was tasked with taking a snapshot of earth for the purpose of recreating it more-or-less exactly as it is, but in another galaxy. This snapshot is nothing but pure data, just like a digital picture file. Think about how little space the snapshot would take up on the alien’s hard drive just after earth formed, and then think about how much more space the snapshot would take up today. The earth has evolved quite a bit since it first formed, and the molecular information content of earth has gone up tremendously through time. The vast majority of molecular information on earth can be found in a thin, mostly carbon, slushy crust at earth’s surface called Life. Life is a product of earth’s evolution.

Most people, when they hear the word evolution, think of Charles Darwin. However, Charles Darwin did not originate the concept of organic evolution. In fact, the concept of evolution was fairly well established before him. Charles Darwin originated a theory called natural selection that can neatly account for the mechanism that ensures that evolution will occur. The real beauty of natural selection is that in its most abstract form it can account for many kinds of evolution, and when it operates it guarantees that evolution will occur. Natural selection is a simple algorithm, and we can spell it out in four simple steps:

1. Choose
2. Record
3. Change
4. Repeat

Natural selection is a simple four-step loop for producing an organizational pattern through time within a set of things. First, we must have a set of things from
which to choose, and we must have a mechanism for making choices. Second, we must have a way to record the choices that we make. Third, we must have a mechanism for somehow changing the choices we have made. Lastly, we must have the ability to repeat the process endlessly. Darwin noted that the original set of things was the set of all living things on earth. He postulated that nature selected things from this set via a competition to reproduce. He knew that the winners in this competition would somehow be recorded in the generation of offspring, and yet each generation would somehow be slightly changed from the previous. He knew that over the long course of time, this cycle would get repeated many many times. In this scenario, the two steps of “record” and “change” are wedded in the process of reproduction. Darwin had no formal way to describe the molecular events of reproduction, but we have extremely fine details about it now. Darwin was exactly right about his simple algorithm.

Perhaps the most misunderstood step in the process is step number three: change. It is quite common to substitute the word “mutate” in for the word change; however, in organic systems the word “combine” is far more apt. Life creates change in future generations most effectively by making combinations of the recorded parts of past generations. Life is a highly symmetrical system of parts that combine and interact to form changes in future Life. Life constantly and most rapidly evolves via combinatorial processes. Sexual reproduction is the paradigm of combinatorial processes. Granted, mutations do cause changes, and mutations are an essential part of organic evolution, but combinations create far more changes, and they are a more logical and effective means to drive evolution of all kinds, but especially evolution of highly symmetrical information and symmetrical information systems. The more symmetrical a system, the more combinations that are possible. The more combinations, the more choices there will be. The more choices, the faster a “good” choice will be found.

Perhaps a mathematician will recognize this algorithm of natural selection as a basic fractal generator. This is true; it is a fractal algorithm. Mandelbrot originally described the fractal formula as involving an initiator and a generator, but we will use the terms x and y. We will use the symbol f() to stand as the formula or the function that operates on a variable. The basic pattern of a mathematical formula for producing a fractal curve is as follows:

1. \( f(x) = y \)
2. \( x = y \)
3. Repeat

In other words, to produce a fractal we start with a number, plug it into a function, and then use the output from that function as our input to the same function. This process might be repeated indefinitely, and it will produce a curve that has no inherent scale. It leads to self-same geometry. Here is a simple example called a Koch curve, where the initiator, or x, is a straight line. The function is to break the line into thirds and build an equilateral triangle above the middle third.
This process can be repeated indefinitely, where the generator, or y, from each round is used as the initiator for the next:

![Diagram of fractal generator](image)

The curve is a fractal because it has no inherent scale, and it has self-same geometry overall. However, we can now also see that natural selection is a fractal algorithm. The input is all life forms on earth. Natural selection is the function that operates on this input to create a new set of all life forms on earth. This new set is merely the input for the next round of natural selection.

1. Natural Selection (Life1) = Life2
2. Life1 = Life2
3. Repeat

Natural selection is not limited to organic systems. In fact, any system that can perform the four steps of the natural selection algorithm will necessarily evolve. It is a mechanism for self-organizing patterns of information of all types. The universe is informative, and it is capable of performing all four steps of the algorithm, so it does evolve. The fractal nature of the algorithm accounts for the fractal nature of the patterns in the universe in both time and space. It is only a matter of time before certain areas of the universe achieve enough organization that they can perform this algorithm on accelerating amounts of information. The results are cumulative, and therefore the more information plugged into the routine, the more information that is spit out. The algorithm organizes information at accelerating rates of increased precision, complexity and negentropy. The universe is a fractal pattern and Life represents the most embellished portion of the curve.

This should have answered for Schrödinger his question of how Life evades the second law of thermodynamics. Darwin was actually attacked with the second law, by a man named Mivart, even before there was any formal thermodynamics. Darwin used natural selection as his deft response to this attack:

“Assuming that an insect originally happened to resemble in some degree a dead twig or a decayed leaf, and that it varied slightly in many ways, then all the variations which rendered the insect at all more like any such object, and thus favoured its escape,
would be preserved, whilst other variations would be neglected and ultimately lost; or, if they rendered the insect at all less like the imitated object, they would be eliminated. There would indeed be force in Mr. Mivart’s objection, if we were to attempt to account for the above resemblances, independently of natural selection, through mere fluctuating variability; but as the case stands there is none.”

Darwin spells it out clearly here: Mother Nature has choices, she makes choices and she records her choices. The record of the choices she’s made in the past serve as the foundation of the choices she’ll have in the future. It would all be entropy if Mother Nature weren’t so hard at work making choices, but Mirvart apparently couldn’t see this obvious mechanism at work and where it would logically lead the entire system. This simple but time consuming work of having choices, making choices and recording them is Darwin’s algorithm of Natural Selection, and that’s exactly the way it works. Darwin did get many details wrong, but on this one crucial point he was entirely correct. It is obvious now, but he clearly pointed it out to us a century and a half ago when it was not or could not be so obvious. In the absence of natural selection all is entropy, but natural selection clearly drives the system in the direction away from entropy. This is also a perfect example of a simple description of any information system, pure and simple, have choices, make choices, record choices.

Remarkably, those who would attack Darwin and his ideas about evolution today still use the word ‘random’ as if it is a weapon that could slay Darwin. Nothing could be more false about Darwin, about what he said and what he believed, as the preceding passage perfectly illustrates. Unfortunately, the defenders of evolution today still seem to fail in their efforts to grasp Darwin’s central point. This is what opens the door for and actually invites the attacks on Darwin. He was arguing a view of negentropy and that view was based on natural selection. He knew that life must have choices, call them random if you must, but that once life made a choice, whatever randomness that might have existed has now clearly disappeared. Since life has the ability to make these choices and record these choices via the actual survival or continued existence of the choices made, these things would quite naturally serve as the platform of the next round of choices. This process would logically progress through vast amounts of space and time. Darwin knew what would inevitably become of this kind of system working in large sections of space through large sections of time. He knew that complexity and precision will inevitably lie at the end of this mechanistic road accruing to the benefit of logical choices made. He clearly described Mother Nature pointing time’s universal arrow against disorder with her primary force of natural selection:

“Although we have no good evidence of the existence in organic beings of an innate tendency towards progressive development, yet this necessarily follows, as I have attempted to show in the fourth chapter, through the continued action of natural selection. For the best definition which has ever been given of a high standard of organization, is the degree to which the parts have been specialized or differentiated, and natural selection tends towards this end, inasmuch as the parts are thus enabled to perform their functions more efficiently.”
“Although the belief that an organ so perfect as the eye could stagger any one; yet in the case of any organ, if we know of a long series of gradations in complexity, each good for its possessor, then, under changing conditions of life, there is no logical impossibility in the acquirement of any conceivable degree of perfection through natural selection.”

Darwin describes life essentially as a busily complex information system that is constantly in a competitive equilibrium and constantly pushing its point of equilibrium toward the accumulation of complexity, order, precision and, therefore, the accumulation of more information within the system. Darwin essentially told us to expect negentropy in the processes of life. This is exactly what Schrödinger was looking for, but apparently he didn’t know it. Schrödinger and Darwin made basically the same observation yet Darwin had already found a simple mechanism to explain it. That Schrödinger should seem to miss this is more ironic when one considers that he had just helped put the finishing touches on the quantum model that describes the universe as nothing but, what else, an information system.

Selection is perhaps the easiest part of the algorithm to understand. It is self-evident almost to the point of being tautological. It is frequently bastardized into the phrase “survival of the fittest,” which begs the question: what is the fittest? The answer, of course, is that which survives. Perhaps a better way to say it is persistence of pattern. In an informative universe, some patterns will persist longer than others. The patterns which persist are physically recorded in the universe itself. More importantly, patterns that persist are the only ones available to make new patterns. Protons, neutrons and electrons persist, for instance, and so they are the patterns available to combine into atoms. Atoms represent patterns that persist, so they are the ones that are available to combine into molecules. Some molecular patterns are better at persisting than are others. The ones that persist are the ones available to make new patterns. In this way it is not so much a matter of survival, it is more a matter of searching for patterns that can persist. In fact, natural selection is, at bottom, a search algorithm.

1. Build a generation of life forms.
2. Search the generation for persistent patterns.
3. Use the persistent patterns to build a new generation of life forms.

Life has two basic molecular functions: build and search. Life must have a mechanism to build molecules, and it must have a mechanism to search for new molecules to build. The algorithm of natural selection logically leads life to find both of these mechanisms. Life will favor those mechanisms that are most efficient at performing both of these functions. The most effective mechanisms are the ones that are most symmetrical both in terms of building and searching. The most symmetrical mechanisms are the ones that can simultaneously build and search. Life favors sexual reproduction because it is symmetrical in that it simultaneously builds and searches for new life forms. Death is a natural part of Life because it effectively speeds the search by pruning older life forms from the search. Life has found an effective mechanism to simultaneously build and search molecules, and this mechanism is called the genetic
code. The genetic code is very much like sexual reproduction because it is efficient at combining molecular components in building and searching those components. Codons are highly symmetrical sequences, and the assignments of codons form a highly symmetrical pattern. This pattern acts as an operating system for building proteins and a search engine for finding new proteins rapidly. Life itself is highly symmetrical in that all of these functions are tightly integrated and work well as a whole. Life does not operate on individual molecules, life forms, or species; Life operates on whole systems. In this way, Life becomes complex and symmetrical. The boundaries of sets are rarely ever clear. This was one of Darwin’s great insights. As a naturalist, he was tasked with drawing the lines between species, and he found that the lines did not exist. He knew that one species could be defined only relative to another, and sometimes that definition was not possible. Darwin knew that Life was complex, symmetrical and metaphorical. Three cheers for Darwin.
VII. What is a Crystal?

The word crystal has many meanings, so we must attempt to define it for later use. Let us start with the dictionary:

crystal

noun (plural crystals)
1. CHEMISTRY solid with repetitive internal structure: a solid containing an internal pattern of atoms, molecules, or ions that is regular, repeated, and geometrically arranged
2. MINERALS piece of crystal: a piece of a mineral in crystal form
3. MINERALS quartz: a clear colorless mineral, especially quartz
4. ELECTRONIC ENGINEERING electronic component: a crystalline substance that has semiconducting or piezoelectric properties and is used as an electronic component, or the electrical device using it
5. something like crystal: something that has the form of a crystal, for example, a frozen snowflake or a grain of salt
6. CRAFT heavy glass: a heavy transparent sparkling glass
7. HOUSEHOLD crystal glass objects: things made from crystal
8. glass over watch face: a transparent glass or plastic cover for the face of a clock or watch. Also called watch glass
   U.K. term watch glass n.1

adjective
very clear: clear and sparkling

[Pre-12th century. Via French cristal and Latin crystallum from Greek krustallos “ice” (the original sense in English). Ultimately, from an Indo-European word meaning “to freeze over” which is also the ancestor of English CRUST and cryo- .]

Dictionaries can be quite ironic. For instance, one who has taken an entry-level class in geology will note that the definition of a crystal contains a glaring contradiction. First, it is an orderly arrangement of atoms and then it is glass. Which is it, orderly or disorderly? After all, glass, from a molecular standpoint, is merely a highly viscous liquid. It is amorphous, or without form on an atomic level. This is one of the consequences of the inherent metaphorical nature of language. Things derive their meanings from comparisons to other things, and so sometimes the meanings must overlap. Sometimes the meanings can become bent back upon themselves, as in the above case with the definition of crystal. I suppose it is true that diamonds look like
glass, and a diamond is a crystal, so too then must be glass – right? Unfortunately, this will not do. After all, I have already told you that Life is a crystal, and Life does not look anything like diamond or glass – right? You might just be surprised.

When the going gets weird the weird turn pro, or so says Hunter S. Thompson. In this case the weird turn to the pros for help. I was never a professional mineralogist, but I was trained by one of the best, Cornelis Klein. He is my go-to-guy on crystals, and here is the definition of minerals and crystals that he was kind enough to give me:

“Minerals, with few exceptions, possess the internal, ordered atomic arrangement that is characteristic of crystalline solids (long range order). When conditions are favorable, they may be bounded by smooth plane surfaces and assume regular geometric forms known as crystals. Today, most scientists use the term crystal to describe any solid with an ordered internal structure, regardless of whether it possesses external faces. Because bounding faces are mostly an accident of growth and their absence in no way changes the fundamental properties of a crystal, this usage is reasonable. Thus a broader definition of a crystal becomes: a homogeneous solid possessing long range, three-dimensional order.

The crystalline state is determined by X-ray single-crystal techniques as well as high resolution transmission microscopy. Such techniques show whether a crystal is regular in its atomic structure throughout. It may show defects (as such it may lack some short range order) or it may be so poorly ordered that it is said to be amorphous.”

That should work well for us here. The bottom line is that minerals and the crystals that make them are phenomena that start at the atomic level. They are marked by order, and order is marked by pattern. Atomic patterns are frequently symmetric, and this leads to structures, small and large, that are characterized by symmetry. It has been mathematically demonstrated that symmetry is defined by groups of things, and there are only thirty-two possible symmetry groups in our universe. Atoms that arrange into ordered patterns have thirty-two choices for their basic symmetry patterns. Some groups have a small amount of symmetry, while others have the maximum amount possible. Life seems to always prefer the latter.

Cornelis Klein is the man who taught me mineralogy, and this is what he taught me several decades ago. He taught me that we cannot focus on macroscopic properties and gross behaviors of crystals in order to properly understand them. We must realize that atomic behavior can be dictated by many accidents in the environment that will lead to macroscopic variations, but at the atomic level, the rules are always the same. In other words, to understand the orderly behavior of atoms, we must first understand the rules that govern their behavior at bottom. That behavior is governed first and foremost by principles of symmetry.

When I was finally forced to take my first class in biochemistry – which was well after my experience with mineralogy - I quickly became confused. These wacky biochemists seemed to be wholly uninterested in the basic rules of atomic behavior that I had learned in mineralogy. Yet, they were describing some of the most well-behaved, orderly atoms on the planet. One of the first things I was given in biochemistry was a simple spreadsheet that ostensibly explained the orderly behavior of a large collection of highly specific atoms. I could count the atoms, and feel their pain. The spreadsheet was
easy for me to memorize, but what about the poor atoms? How is an atom supposed to see let alone read a spreadsheet? On what laws or first principles of atomic behavior is this spreadsheet based? Where in it shall we find those laws? After all, the orderliness of Life’s atomic behavior is marked by some of the most periodic patterns ever known to man. Why should the rules of atomic behavior be any different in Life than they are elsewhere in this universe?

Similarly, Erwin Schrödinger had trained himself to think in terms of order when describing the structure and behavior of any individual atom. This is the basis of quantum mechanics, and it gave us the periodic table of elements. Still, when atoms assemble in great numbers, the collective behavior is marked by disorder, and this was the only way Schrödinger had for understanding collective atomic behavior. It is a purely statistical phenomenon called thermodynamics. Random heat motion guides the mass behavior of atoms in the gas and liquid states. When atoms cool, they become solids. The most orderly patterns of solids are known as crystals, so Schrödinger was forced to see Life as a crystal - because it was obviously so orderly. He also knew that crystals are limited by their geometries, and he had no knowledge of any geometry that would lead to the patterns he was witnessing in Life. Order and disorder, as we’ve seen, are forms of information. Atoms derive their information from symmetry. They are ordered by it. Why should we not expect to find symmetry prominently observed in the middle of Life? We absolutely should, and fortunately we do.

Crystals are the manifestation of growth patterns of atoms when they become bounded. They are first and foremost bounded by the symmetry of atoms. They become bounded in time and space by several factors relating to the specific environment in which they grow. Gravity is the first environmental boundary that effects crystal growth on earth, and when gravity is removed from the equation, crystals of all types begin to grow “perfectly” with respect to their inherent symmetry. Chemical imperfections also are always a factor in crystal growth. The type and amount of chemical impurity will impact on crystal growth. Other physical parameters, such as the space in which a crystal can grow, will also play a significant role in crystal formation. Chemical concentrations, heat and pressure are also critical in the final form a crystal might assume. Time is of the essence. Crystal formation is ideally unbounded by time, but the reality is that all crystals form within time boundaries that are imposed by all of these other factors.
Given this new understanding of crystals, what else around us might be seen as a crystal in the most abstract terms? What about a computer, is it a crystal? It is a solid with an incredibly ordered arrangement, both at the level of atoms and above. The silicon and metal that go into a computer are certainly crystals. The macro-arrangement of these crystals is also quite orderly.

Figure 1.

A computer is a solid with a decidedly orderly arrangement of atoms. Man makes synthetic crystals of all types, so why should a digital computer be exempt from this definition? What feature of a computer disqualifies it from being considered a crystal?

How about a human brain; it is a computer of a sort. It is not synthetic. It is made of atoms that obviously have an orderly arrangement, so is it a crystal? If we study it on any level we find patterns that merely become more complex. It is fractal in its appearance and structure. Just because it is not “geometric” in the traditional sense, from the standpoint of gross appearances, does not exempt it from the basic rules that govern all atomic behavior. The physical structure leading to electro-chemical patterns in the brain – thought – must at least be considered crystalline at some level of abstraction. Perhaps thought itself can be considered as a form of complex crystallization.

At still higher levels of abstraction we must consider the notion that we are witnessing forms of complex crystalline behavior. Take human culture, for instance, it is governed by rules of order that merely compound into higher levels of organization that emerge from them. The United States Constitution might be seen as a fundamental symmetry group for human governance. It dictates the boundaries of law-making within a larger system of law making. Imperfections are introduced into the law-making environment and propagate throughout our system of laws, just as an imperfection naturally propagates through the atomic structure of an ordinary crystal. Similarly, our system of sciences behaves in a crystalline way. Laws of logic and systems of
observation are in place throughout the system. Seeds of ideas form, and larger ideas grow out around them. Structures of thought become interlaced and crystallized to form the superstructure we call knowledge.

When we turn toward Life, within this abstract framework and broad notion of crystals, it is not too hard to include Life within this mammoth tent. At the very bottom, Life is composed of atoms, and these atoms absolutely must obey the laws that are laid down for all atoms. Most of the atoms in living systems have aggregated into clusters of atoms that we now know as the “simple” building blocks of life. They are nucleotides, amino acids, sugars and lipids. These molecules are not crystals, per se, but they are built of the symmetry of atoms, and they themselves must have geometric properties. These building blocks are combined into orderly macro-arrangements that have repeatable patterns and assume consistent forms. They interact with each other in complex yet consistent ways. Biological macromolecules, such as DNA, RNA and protein are, in fact, tiny crystals and they do crystallize out from solution. Their atomic structures are regular, geometric and can be witnessed with X-ray techniques. These tiny crystals collaborate in intricate yet repeatable ways to form larger structures that are frequently defined by their geometric regularities, but need not be. They are the epitome of atomic order. An entire cell or even an entire life form must be seen as crystalline in this way. In fact, the entire collection of Life on this planet must be seen as a complex crystal forming system. It is the same basic system everywhere on earth and it can be seen to operate on the same basic rules, first principles and basic functions. At the very bottom, these rules should be seen as no different from the rules that govern ordinary crystals.
VIII. What is Biochemistry?

“Success has many fathers, but failure is an orphan.”
John F. Kennedy

That is a damn good question. Darwin would tell us that we are embarking on somewhat of a fool’s errand of human taxonomy. Biochemistry is whatever it is that biochemists do. Unfortunately, they rarely wear name tags or wear identifiable uniforms. However, they are organized into teams that exist in departments and compete for shelf space and grant money. So, to really get a foot hold on this definition, we need to revisit the trunk on the taxonomic tree. We can start by recognizing the natural boundaries of academic disciplines, where they might exist, and then draw the relationships between them. Start with logic, which leads to math, physics and then chemistry. From here we need to recognize the tug of other disciplines that fragment out the disciplines into sub-disciplines. There is biology, of course, which is the formal study of Life. There is also philosophy, linguistics, information science, and geology, which all must be considered as relevant to any formal study of Life. These things all tug on chemistry to pull it apart into organic chemistry, which is carbon based chemistry, physical chemistry, biochemistry and molecular biology. And don’t forget genetics!

I have always hated chemistry. I am simply no good at it. There are too many rules, and I can never keep them all straight. There are too many names, and I always get them confused; plus, the names are rarely descriptive in any useful way. When we journey into biochemistry, the situation merely gets worse because there are millions of known biochemicals, and they are each in their own way important. What is a poor boy to do?

When I started this project I intended to avoid anything chemical like the plague. I prefer the more abstract disciplines, like math, philosophy and information science. The path of science has veered so much toward a dependence on computers that an entirely new discipline, informatics, has become necessary. In the fields of biology, the information is accumulating so rapidly that a formal discipline, bioinformatics, absolutely is required also. Understanding Life requires a little bit of all these basic disciplines, yet nobody can be expected to master them all. Like I said, I tend to favor the computers, but no examination of Life can completely avoid the rest of it. As I progressed in my project and became more obnoxious and heretical, my local university tried to bounce me from department to department. I found that the biochemists were uninterested in bits and bytes, yet the computer geeks were wholly ignorant of a peptide bond. It is quite a mess, to be sure. For that matter, on what shelf should the poor librarian put this book? I suggest that it go on a shelf of its own, non-fiction, of course. Perhaps a whole new discipline is required: biomolecular informatics!

In the grand human taxonomy there are lumpers and there are splitters. I’m a big-time lumper because I tend to see the symmetry in things. I can see how one discipline is remarkably like the next, and their subject matter bleeds unavoidably into the other’s. I would not be very successful in big-time professional science today because I would be miserable at recognizing my natural boundaries and protecting my turf; and, after all, that
is what it is all about. I somewhat randomly choose a label to describe the fields in which I dabble, and I somewhat randomly choose biochemistry as the subject of this chapter. Our old friend, the dictionary, gives us this definition:

**biochemistry**

**bi·o·chem·is·try** [bī ə kēmə stree]

_noun_

1. **chemistry of living organisms**: the scientific study of the chemical substances, processes, and reactions that occur in living organisms

2. **chemical nature of organism or system**: the chemistry or composition of a particular organism or system

**-bi·o·chem·ist**, noun

*Microsoft® Encarta® Reference Library 2005. © 1993-2004 Microsoft Corporation. All rights reserved.*

But perhaps I should have chosen this one:

**molecular biology**

_noun_

**study of biology at molecular level**: the branch of biology concerned with the nature and function, at the molecular level, of biological phenomena, such as RNA and DNA, proteins, and other macromolecules

*Microsoft® Encarta® Reference Library 2005. © 1993-2004 Microsoft Corporation. All rights reserved.*

But why not this one:

**mineralogy**

**min·er·al·o·gy** [mìnnə róllajə]

(plural **min·er·al·o·gies**)

_noun_

1. **study of minerals**: the scientific study of minerals and how to classify, distinguish, and locate them

2. **Area’s mineral deposits**: a profile of an area’s mineral deposits

**-min·er·a·log·i·cal** [mìnnə róllajikəl], adjective

**-min·er·a·log·i·cal·ly**, adverb

**-min·er·a·log·ist** [mìnnə róllajist], noun

*Microsoft® Encarta® Reference Library 2005. © 1993-2004 Microsoft Corporation. All rights reserved.*

The plain truth is that the subject of Life will never fit into any neat compartment of human thought. Nobody has a patent on the field, and nobody has a monopoly on
ideas. There is nowhere to turn if we want to find “the truth” about Life. We must look inward.

Unfortunately, we need a framework for our thinking and a language to communicate and exchange ideas. I have found no single discipline that is adequate for either of these things. It is always the case where everybody assumes that the other guy knows what he is saying, yet the other guy rarely ever does. We have somehow achieved a malicious form of epistemic gridlock where separate disciplines can maintain boundaries built of self-sustaining contradictions between disciplines. One thing is certain, however, nobody has a monopoly on the truth, and nobody ever will. Yes, I include myself in that set, and I fully realize that not everything I say or believe is correct. However, I do find myself in a unique position of knowing just enough to realize that we need to start tearing down some of these dysfunctional boundaries and start building a new system of thought. We need to weed the garden of dysfunction language and ideas that seems to have folks spinning their wheels in their own little areas. We have already created some quixotic ideas and language here, and we will create quite a bit more, to be sure. But in the mean time, we need a basic foundation of molecules that we can use to put things forward for analysis and comparison. We can start with a conventional language of molecules, and biochemistry is a good place for that, but molecular biology might work just as well.

One could never hope to understand Life without knowing at least some basic biochemistry. However, the amount of biochemistry one needs to know is not that great, and in some ways, the less one knows about biochemical conventions the better off one might be. The proper foundation can be obtained from any introductory biochemistry or biology textbook, but don’t drink too much of that Kool Aide™. You see, there are some basic biochemical conventions that are flat out wrong. The professional biochemists of today perhaps need to spend just a little bit of time studying the basics of math, information science, and mineralogy. A touch of logic, linguistics, philosophy and art wouldn’t do them any harm either.

There are millions of biochemicals, and they are all distinct. Every molecule is distinct and therefore namable. We can break them down into rough sets and name the sets. There are sugars and lipids and nucleic acids and amino acids, and there are important combinations of all of these. Within each set, there are potentially millions of different variations. Life deals in the formation and relationships between sets, but Life has a strong tendency to bias the sets that it chooses. If Life did not do this, Life would not be Life. It absolutely must have a system for selecting molecules and intensifying its organizational scheme between the molecules it selects. The vast majority of the information contained within living molecular systems is contained within the organization schemes and not the molecules themselves. Therefore, we must know the molecules – to some extent – but it is more important that we know the organization.

Every molecule within biochemistry can be called a crystal. They all meet the criteria for being crystals. By normal crystal standards, they are all tiny crystals, to be sure. But more importantly, vast collections of these molecules can be called crystals as well. There is nothing that applies to our understanding of crystals that cannot also apply to our understanding of the organization of these complex collections of molecules. Granted, they do not conform to the rigidly simplified logic of simple symmetry groups, like common crystals do, but they must be built of these groups, and the logic of their
behavior must be somehow built up from this as well. There is no law of biochemistry that says that biomolecules do not need to follow the same rules that all molecules must follow. The addition of metabolic pathways does not exempt biomolecules from following the same basic rules in space and time. We just don’t happen to know what those rules are or how these special complex molecules happen to be following them.

The most striking features of life can be seen in two basic molecular sets and the relationships between them. These are the set of nucleotides and the set of proteins. Nucleotides include DNA and RNA, and there are many forms and functions of both. Proteins are the most active and diverse set of biomolecules. They are composed of chains of amino acids that are linked together by peptide bonds. There are many possible nucleotides and therefore many possible combinations of them. There are many possible amino acids and therefore many possible combinations that can form chains. However, Life has chosen to greatly bias its preferred sets of nucleotides and amino acids. For all intents and purposes, there are only four nucleotides in DNA and twenty amino acids in proteins. The relationship between these sets is poorly understood, but we will delve into that a bit later. Suffice it to say that Life has biased these sets for a reason, and the relationships between them contains a tremendous amount of information within Life’s grand organizational scheme.

Biochemistry is mostly about pathways, and pathways are complex, interesting and important in Life. However, too much emphasis has been placed on pathways, I think, and this emphasis has perhaps distracted us from recognizing more general functions in Life. Schrödinger recognized the need for and essential role played by biochemical pathways, especially with regard to metabolism, yet he clearly said that these things could not explain Life to him in any satisfactory way:

“That the exchange of material should be the essential thing is absurd. (Exchange of energy) … needless to say, taken literally, this is just as absurd.”

“The essential thing in metabolism is that the organism succeeds in freeing itself from all the entropy it cannot help producing while alive.”
IX. What Are Boundaries?

One of the key features of Life is that it is able to distinguish self from non-self. For instance, you are able to determine where you end and the rest of the world begins, for the most part. You have a fabulous collection of cells called the immune system that also does this for you. The immune system acts on a body-wide level, as well as cellular and molecular levels. If the immune system detects an invader at any level, it acts quickly and aggressively to remove it. Most countries have clearly defined borders, and they have elaborate defense systems to protect those geographic boundaries. They also have monetary and economic policies that protect their national systems of commerce. There are time zones as well. Life is like this on all levels, and there are many different kinds of boundaries.

When a sprinter runs a race he is bounded in time and space. The earth itself forms a vertical boundary consisting of gravity and the surface of the earth. He is also bounded horizontally by lanes, a starting line and a finish line. The sprinter is bounded in time by the starter’s pistol and a stopwatch that measures the amount of time it takes to reach the finish. The sprinter is also bounded by the rules of the sport and any particular race.

When we talked about ordinary crystals we said that they were essentially unbounded except for the symmetry groups that define their structures. Symmetry provides the bulk of the rules for how any particular crystal might form. Of course, any particular crystal can be bounded by several other factors. Organic crystals, however, have many more boundaries. When a protein is formed it is translated from a string of nucleotides. This string, like a sprinter, has many spatial boundaries, most of them predicated on the nature of sequential polymers. There is a start molecule and a stop molecule. There is a surface on which this translation must be performed. There is a physical space in which it takes place and this space will impart physical constraints on its formation. There are countless other molecules in the translation environment that are also specifically build to provide spatial guidance for the protein. There are molecules that transport the nascent protein to the right spaces at the right time. In addition, there are many time boundaries to the formation of any single protein and all proteins collectively. There are start signals to begin making a particular kind of protein and there are signals to stop making that protein. There is a beginning, middle and end time to making any individual protein molecule, and the difference between the two can effect the final translation.

Time boundaries are essential to Life. There are natural “cycles” for almost all molecular events. There is a distinct periodicity to the time it takes for life forms to be conceived and die. Within an organism there are countless other cycles or biorhythms. The same is true of cells, cellular processes, molecules and molecular processes. Cells are born and they divided and eventually die. An unchecked cycle quickly becomes dysfunctional, like cancer, or like many metabolic diseases. Regulation of time is an important part of time, and regulation requires time boundaries.

Numerical sets require boundaries. We identify sets by their boundaries and this is true of molecular sets as well. The set of nucleotides in DNA has tight boundaries relative to the set of all possible nucleotides. Likewise, the set of amino acids in protein
is tightly bounded as well. Transfer RNA, a molecular set essential to protein synthesis, is also bounded, but these boundaries are quite a bit more variable between life forms. Just as consistent molecular set boundaries represent a form of information, the variability of these boundaries can also represent information. Generally speaking, the change in one molecular set boundary is related to the boundaries of other sets. A change in the boundaries of transfer RNA signify a change in the boundaries of the proteins they create.

Perhaps the most significant boundary in Life is the boundary between the inside of a cell and the outside of a cell. This is called a cell wall or a cell membrane. Many cells also have intracellular membranes that also form important boundaries for differentiated cellular compartments. All Life is cellular, so a good bit of Life’s functions can be found in cell membranes. They first and foremost distinguish self-from non-self. Cell walls are much like boundaries between nations. Sometimes they are impermeable walls, but more often than not they are selective barriers that regulate commerce in and out of a cell. Essential ingredients must be transported into a cell and unwanted items must be transported out. Sometimes the exchange is passive and sometimes it is active and requires energy coupled to complex mechanisms. In virtually all cases it is a highly selective process, and by being selective it is generating information and informative organization for the cell. In this way, cell membranes can be seen to consist of complex codes that function to keep the cell organized and alive.

Boundaries can be spatial and they can be temporal, but many other kinds of boundaries exist in Life as well. There are chemical boundaries, like pH. There are heat boundaries within cells and processes must be finely maintained. There are electrical boundaries. There are pressure boundaries and diffusion boundaries. Life is full of boundaries of virtually every conceivable parameter. It is respect for and utilization of precise boundaries that defines Life itself. There is one more kind of boundary without which Life cannot exist. Life is full of symmetry boundaries, just like ordinary crystals.

Symmetry is not limitless. In fact, symmetry is just the opposite; it is the placing of specific limits. Symmetry places limits that can be limitlessly apply and thereby it becomes the placing of the most efficient limits. Because symmetry represents efficient limits, it is frequently used by the universe and Life. Symmetry can be found everywhere in natural processes, like ordinary crystals, and it is equally prominent in the boundaries of life. If we start with the notion of all possible points in space, we can limit this set to only those points that can be found on the surface of a sphere. If we limit this set further to only those points that are identical in every way to all others, we are now limited to a very small set of points: {0, 1, 2, 3, 4, 6, 8, 12, 20}. This also limits the relationships between points that can also be perfectly symmetrical. However, these points can be extended in space in unlimited fashion, and that is why they are so frequently chosen.

The symmetry boundaries in Life are many and significant. They not only dictate spatial parameters, but many other parameters as well. Sexual reproduction, for instance is a highly symmetrical process. Genomes are highly symmetrical structures, and they are built to symmetrically transform into new genomes. The symmetry of any genome allows it to become many different genomes through symmetrical transformations and combinations, but these processes are not unlimited. The combinatorial limits of genome formation general define species. These symmetries are mostly embodied in symmetries of sequence.
Proteins, like genomes, are highly symmetrical structures. Indeed, the symmetries of proteins are symmetrical with the symmetries of genomes. Genomes make proteins and proteins make genomes. Just as a genome can transform and recombine so too can proteins. It is this intense symmetry in proteins that lead to its formation and also lead to the efficient search for new proteins. The number of proteins that can be made is not unbounded. It is bounded by many things and one of those things is symmetry. All proteins are bounded by the sequence symmetry of codons. Codons can be transformed in many ways, but the number of possible transformations is bounded by symmetry. Of course, Life wanted the most effective and efficient boundaries for its important sequences, so it chose triplets in codons because this provides the maximum amount of sequence symmetry with the fewest required parts. From this life was able to bound its other important molecular sets, like nucleotides in DNA and RNA, and amino acids in protein. When a single protein comes into existence, the number of new proteins that can be derived from this first protein is bounded by the symmetry in the number of ways that this protein can be transformed. Life has found the most efficient boundaries for this.

When we contemplate Life we must first contemplate boundaries. Natural boundaries in the universe tend to become boundaries in Life. Earth is bounded by earth, and this is the first natural boundary of Life. Life is not only limited to earths near surface, it is naturally bounded by earth’s shape, structure, mass and therefore its gravity. Certain forms are off limits on earth, while others are guaranteed. An insect can only be so big and of a certain form. Insects are prisoners of gravity and diffusion. So too are all other forms of Life in various fashions. Ordinary crystals are bounded by just thirty-two symmetry groups. Life has clearly broken those bounds by its clever use of combinations, but Life is still bounded in many ways, and symmetry is one of the important ones. Life combines molecules in time and space, and Life combines symmetries in time and space as well.
X. What is a Code?

We have addressed this question to some degree, but we will now address it much further. The easiest way to recognize a code is to first think in terms of sets. Elements within sets can have logical relationships, elements between sets, and whole sets can be logically related as well. A logical relationship between one thing and another is called a function. A code is a function, but I prefer to think of a code as a collection of functions or multiple applications of a single function. When we spell out the steps of a code we have created an algorithm. An algorithm is a set of steps that reliably produces a result. Any particular articulation of an algorithm can be called a language. All languages are codes that perform algorithms. All codes are functions and all functions operate on sets. There are many possible functions. There are more codes than functions, and more languages than codes because every function is a code and every code is a language. There are many possible languages for any code or function. Not every language can perform every code or function. But functions combine into codes and codes combine into languages.

Codes and languages are collective phenomena. In other words, there must be collections of things before there can be codes. The code itself does not exist within the sets that are logically related. We cannot locate codes within the sets of things that are translated between. For instance, consider Morse code. There is a set of dots, dashes and spaces. There also is a set of alphanumeric symbols. Morse code is the logical relationship between the two sets. The elements of the first set have rules of combination that allow the set to expand to match the second set. The logical relationship between these sets is Morse code, but where is this code located? The code itself is not located in the symbol sets but it is located in various algorithms for performing Morse code. We generally use English to teach the algorithm, but we need not because Morse code is itself a logical language that could be taught without using English.

Codes operate on sets but their function is to generate more sets from the logical relationships between sets. For instance, Morse code starts with a set of symbols, but then it makes combinations of those symbols to generate a larger set of symbols. This set can then be related to another set. These sets are then used to make much larger sets of symbol sequences that inherit the logic of the smaller sets. Information is contained within these sets, but the value of this information is dependent on logical relationships and the context of the codes in general. The symbols can be related to human languages, and of course this language will logically relate those sets to other sets. If the codes disappear then the sets, their logical relationships and the information disappear with them. Therefore, codes are information and they lead to the existence of still more information. Codes generate sets of sets that lead to evermore sets. Information naturally expands in the process, but before there can be information and information expansion their must be functions, codes, and a practical way to apply them in ways that we can call language.

When we turn our attention to codes in nature, we can see that this same logic applies. The universe has fundamental codes sewn into its fabric. We call these natural laws but they are really just codes of universal behavior. When these codes operate on all of space through long periods of time, many sets are created. These sets become well
defined and the logic between them becomes well defined as well. The sets naturally combine, and more sets are logically created, with them comes more codes and more information, which leads to an accelerating expansion of information. The universe will select the most efficient and effective sets and the most efficient logic between them. Symmetrical things are the most efficient things and so we find symmetrical sets with symmetrical relationships everywhere in the universe.

When we turn our attention to Life, we know that it is built from natural sets and natural codes. It inherits its sets, information and logic from the universe. Since the starting sets are so symmetrical, and since symmetry is the most efficient way to produce more of anything, we should not be surprised to find so much symmetry in the sets and codes of Life. We can broadly outline the process behind the formation of Life in the universe by logically nesting its codes of operation. I will outline three broad codes here. The primary code operates in the universe and logically leads to the existence of life. An intermediate level code observes the boundary of a single life form. Other codes observe the boundary of a single molecule. This last kind of code can exist in numerous forms within Life, and it is usually referred to as a molecular pathway, but the most prominent code in this category is the one that leads to the formation of a single protein - and it has been mis-named the genetic code.

We will begin examining nature’s codes by looking at the name ‘the genetic code’. This is an important name, yet its current usage is causing more confusion than any benefit it provides. We must, therefore, ask ourselves if this is at all a good way to describe the “code” symbolized by the codon data table, which is neither truly “genetic” nor really even a code. It lacks structure, meaning and much of the important detail required to understand this code. Is it our intent to describe a code that spells out the genesis of life, or even the genesis of a single life form with this one dimension of information contained in a single spreadsheet? This, I should think, is clearly not the case. This term was coined to describe the translation of a gene. Genes are strongly associated with proteins, but there isn’t even a good, consensus definition of a gene today, and it is no wonder. ‘Gene’ was a term originally proposed to describe a subunit of information, and an entire collection of genes makes up, in part, an organism’s genetic information. It seems wholly inappropriate then to say that the information expressed by a single gene is ‘the genetic information’ when it is merely sub-genetic or infra-genetic information.

Should we then call this code of protein synthesis the sub-genetic or infra-genetic code? Perhaps, but I certainly hope not. It is obvious that the process currently described by the term ‘the genetic code’ is merely a process of protein synthesis, and protein synthesis in isolation is not exactly synonymous with total heredity or true organic genesis. The only genesis directly involved here is the genesis of proteins, one at a time, so the code that guides it must therefore be a protein code of some sort. It is entirely proper to refer to it as the protein code, but historic precedent will surely be a nettlesome problem here. The name is already in so many brains and textbooks that it will take a good bit of time and discipline to weed out its buggy usage. Perhaps a compromise could be reached if we merely shortened the classic name to ‘the gene code’. Or better still, perhaps we could borrow a trick from a famous entertainer and call it the code formally known as genetic.
The code for making proteins not only contributes to but is derived from a much larger context. There are many different codes of life, thousands of them, and they must all be in equilibrium with each other because they are at bottom codes of organic equilibrium. So we must begin a discussion of the protein code from the viewpoint of a truly genetic code as it might be properly understood. On this score there are two potentially good candidates. The first good candidate for the official title of the genetic code is the code, or process, or algorithm, by which a single life form is generated. The developmental timeline of a single life form is known as ontogeny, or in the early stages it is a code of embryogenesis. This code somehow guides a single cell in its task of creating another cell, or in its task of producing a complex population of descendant cells, or more narrowly, in producing multicellular complexity.

This code of cellular operation and proliferation is rightly a candidate for the term ‘the genetic code’ and it will be a far more complex code than any code of protein genesis. After all, a code for making proteins is but a single subroutine for producing the a large number of the many agents needed to make any living cell. Using the title of the genetic code in this broader context will have the added benefit of moving it closer to an entire genome, which is already becoming a familiar usage. But we still should not confuse a genome with the genetic code, because they are two distinctly different things. DNA is not a code. DNA represents data storage, much like the coins that we considered earlier.

The second candidate for a general code that might rightly be called the genetic code is one that describes the recipe for the sublime pattern of all life on this planet. This might truly be considered a code of organic genesis, one that has somehow structured simple, random, inorganic, insentient matter into the vibrant complexity of all living things, the code behind the phylogeny of life. We already have an inkling of this code, since Darwin provided it in his description of natural selection. It is a deceptively simple algorithm where astronomical populations of individual life forms are generated in astronomical numbers and stunning variety. These complex populations of life forms constantly interact in complex networks, perpetually sorted by fecundity, or sorted on their ability to survive and somehow produce descendents. The steps always lead to descendant populations that are distinct from the originals. Each new population is sorted and searched again, ad-infinitum. Life on earth inevitably and perpetually evolves as a result of this basic code.

The steps for generating descent with modification are contained in an algorithm of natural selection that always generates a complex equilibrium. “Darwin’s law” (if we can call it that) is that life will always seek and find its point of equilibrium, both within itself and within the physical environment in which it exists. That equilibrium point is constantly changing, but time’s arrow is forever moving it in a direction marked by accumulation of design, order and precision, and so it is a direction of evermore complex and efficient systems. Life evolves as a complex system for evolving more systems that can generate more complexity. It consistently rewards the strategies that can fundamentally “take it to the next level.”

The equilibrium of life within its environment must also always be placed in the broader context of equilibrium within the universe. Life cannot disobey any general laws of the universe, so it should in no way be seen as an exception to any general laws. Life involves countless random events, yes, but it is a big mistake to ever think of the
processes of life as mere accidents. Life evolves precisely those complex systems that anticipate the randomness of large numbers of events. Life actually requires astronomical numbers of “random” events everywhere and at all times, without which not even a single protein could ever be made. All molecules at base are decidedly random, to be sure, but non-random complex behavior emerges through time from the processes of life. That is the negentropy essence of life. It is how life gets molecules to do what we ordinarily would not expect them to do. The codes of life are harnessers of random noise, converting the noise into logical signals. The codes of life are the specific adaptations that best anticipate the dynamic nature of random events and turn them to advantage. Life depends on thermodynamic forces but it is not controlled by them - it is controlling of them. Life has subjugated the entropy of thermodynamics and turned it into perpetually accelerating negentropy.

Of course, Darwin did not say those exact words, so I am putting words into his mouth when I speak of Darwin’s law, but that is the gist of his words. Perhaps that is how he might want his nineteenth century words translated into twenty-first century words of high-power physics and computational sciences.

The evolution of life is not a directionless process, as some so strenuously insist. Darwin made this point crystal clear. The word ‘progress’ has been wrongly verboten, merely rendered politically incorrect, but the arrow of progress in the evolution of life is clearly in the direction of more and more complexity, as noted by Darwin. Unfortunately, he did not have all the facts that we now have available. Starting with the simplicity of a lifeless planet and progressing to today’s earth, we can witness firsthand the breathtaking complexity of the patterns that this simple algorithm has generated. Starting with the relative simplicity of a single cell we can witness the miraculous complexity of a child. We can also witness the complexity of the molecular and cultural processes behind the many patterns of macroevolution on a large scale. On a smaller scale, starting with the relative simplicity of DNA we can witness the miraculous complexity of proteins. The pattern is the same everywhere and at all levels; life is all about pointing time’s arrow toward negentropy. Darwin’s recipe for life is a blind algorithm, but one with a guaranteed result of evermore complexity. The arrow of progress is clear, and it points toward the emergence of ever-increasing complexity and negentropy in the evolution of life.

The central method of producing this complexity is clear as well. It is an iterative process, or a method of recursion. Natural selection is a looping process where the output from one loop is fed back as the input for the next loop. It is an algorithm like those used in mathematics to produce fractals. It is no surprise then that we find so many complex patterns of life that are also fractals, because life is indeed a complex tapestry not a simple grid. The system is seeking constant novelty, a dynamic system of constant renewal, and the information within the system is translated by methods that unpack, embellish and expand. This method for producing complex organic systems is a great example of what is meant by the term ‘epigenesis’. In fact, the term literally means evolution in some dictionaries, as contrasted with involution. Epigenesis means to develop or grow complexity by repeated application of a method or process.

Unfortunately, epigenesis has many meanings in many quarters, including the meaning that describes the process of embryogenesis - for which it is entirely apt. But this narrow usage will once again cross contextual boundaries. Interestingly, the term has
‘Epigenesis’ or ‘epigene’ are terms used to describe a mineral process that occurs at the surface of the earth where the content or structure of a mineral or rock is changed. This, to me, is an irresistible term that is begging to be applied to the code of all organic life on this planet. It would, in fact, be etymological malpractice to fail to apply this word to the evolution of life on earth. No hardcore evolutionist would resist this temptation, so the broad code of life simply must be officially called the epigenetic code.

Having now somewhat forcibly pried this term ‘epigenetic’ away from the clutches of embryogenesis, we are free to apply the term ‘genetic code’ to the methods normally reserved for embryogenesis and ontogeny. Of course we must stipulate that these ontological processes too are clearly epigenetic in nature. All of life’s codes are executed by complex, evolute patterns of information unpacking. They inherit this property from the epigenetic code of life. But going still further, the code for making proteins is epigenetic as well. Note that the information in DNA is acted upon to form a string of mRNA, which is acted upon to form a tRNA string, which is acted upon to form a polypeptide, which is acted upon to form a protein, and this stepwise series of information unfolding, unpacking - string evolution - must be performed on many different proteins before the masses of proteins in a cell can be efficiently produced via protein polymerases, regulators, facilitators and chaperones. So the entire process of protein synthesis really does wrap back upon itself in epigenetic fashion. Proteins are the net result of a truly “evolute” pattern of translating the many informative molecular strings that have their core “information” somehow merely stored in DNA. It is a sequential but complex pattern - not a simple or linear pattern.

We can see now why it is so hard to define a gene. What is a gene? For that matter, what is the code formerly known as genetic? The problem that the founding fathers of our classic paradigms of biochemistry unwittingly ran into is that they went looking for life’s code in DNA and there isn’t really a “code” in there. There are many codes involved here, but DNA is data not a code. They started their search by aggressively simplifying the system so that they could take even the first steps toward studying it. When they started they were just sure that DNA is the code of life. After all, that’s what they keep saying, but they had only a rudimentary notion of any code, let alone a code of life. It turns out to be quaint yet entirely counterproductive to call DNA the code of life. Unfortunately, these pioneers set up their definitions of things as if data files where application codes, and they clearly are not. The codes for building life, life forms, and proteins are carried by the molecules themselves, by the molecular machinery that executes the codes not by the physical storage medium that holds the necessary data. Codes are collective phenomena and the logic is embedded in the machine not the data. Although the data is indeed part of the machine and does provide an essential physical framework for the machine itself, the code is a real entity that is independent of any single molecular structure. But when it comes to the “simple” act of making even a single protein, there are many nested codes to consider because it is a truly evolute process of molecular translation.

If we start with a single gene, we might simplistically say that a gene is the minimum amount of molecular information required to impact on the protein population in a living cell. This should at least include the data for a single protein, plus all of the supporting algorithms that make the protein come about in the cell. We must recognize
that a single gene is a tremendously complex thing, and ‘the gene code’ should be expected to be extremely complex as well. Drilling down, we should refer to the protein code as the algorithm that living cells use to actually make a single protein from the data of a single protein. This is a tremendously complex code as well, because there is so much machinery and so many options involved in making even a single protein. It will require an entire molecular system that depends on a huge array of molecular components. The actual algorithm will be complex, for it is charged with the task of getting lots of unruly molecules to behave in a consistent fashion in space through time. There is a lot of design work embedded in life’s winning strategies for protein making machinery and protein making algorithms. We should expect these algorithms to be sublime, just as their physical products and general results are sublime.

Finally, at the very bottom of the hierarchy of codes for making protein we can name the polypeptide code. The polypeptide code is that algorithm that generates the primary structure of a protein. Comically, we are still one level in the hierarchy of life’s codes above the code formerly known as genetic, because the codon table does not even spell out the polypeptide code, it merely gives us the primary sequence of amino acids but tells us nothing of the information contained in their bonds. Where is this information accounted for and how are the signals structured within this “linear” language? Another dimension is somewhere required in this one-dimensional language just to arrive at a primary sequence of amino acids.

So, we know that primary sequence has no chance whatsoever of determining a “protein,” a protein population, or the steps for making anything requiring any larger amounts of information, let alone the genesis of a single life form. Primary sequence and genetic code are two ideas that should never be put together in the same sentence let alone the same basic definition. This is why ‘the genetic code’ has been one of the greatest scientific misnomers of all time. That name fell well short of the mark in terms of accurately describing the real-world phenomenon for which it has proudly stood for the better part of one half century. Fortunately, we can clearly recognize the mistake now and perhaps do something about it.

I reject as false the premise that DNA in any useful way represents a genetic code. I therefore also reject as false the premise that a single string of DNA can represent anything of a code for making protein, and it especially cannot represent a code that functions like the code formerly known as genetic is supposed to function. The classic paradigm of the genetic code axiomatically accepts both of these premises which I just so easily rejected. The codon table, which has stood for too long as “the genetic code,” is itself nothing but a simple data set, and it does not represent a code as it should be properly understood. Even as data it is merely a small portion of the relevant data, and it has been empirically shown to be inadequate in understanding the true nature of protein synthesis. After all, “silent mutations” cause changes in protein structure. “Synonymous codons” are not even synonymous. However, because of the intense, pervasive and prolonged indoctrination behind this simple scientific icon, it is going to take some people quite a bit of time to get their minds around this basic idea. Perhaps some never will. It is difficult to accept the fact that we are making such a simple mistake of basic perception, but it will be still more difficult to accept that most people will absolutely insist that we keep making it. Change is painful but no less inevitable here.
Now, at least, we have here a more global perspective of codes in general, and the nested nature of life’s many broad codes, as well as how they inherit their properties from the whole. It is exactly analogous to object oriented programming techniques used for complex computer simulations. The master loop is the epigenetic code of natural selection. This creates sub-programs that create and maintain the genetic codes of individual organisms. Within the program of every genetic code there must be the essential sub-programs for assembly of protein populations. If we consider a single protein, we find the gene code using the protein code using the polypeptide code to complete the task of making a single protein from the data stored in a molecule of DNA. The code formerly known as genetic was so narrowly and rigidly defined that it doesn’t even merit inclusion on the current list, that’s how far off the mark the classic paradigm is, and that’s how much it is perverting our thoughts and blocking our ability to communicate important, complex thoughts. Granted, it is an ignoble demise for a cherished scientific icon as “the code of life,” but it is now quite well past time for a new paradigm.
XI. What is the Genetic Code?

The genetic code is a scientific icon. It is simple. People can easily see it and quickly recognize it. It is so simple, in fact, that to understand the genetic code today we don’t need to know much about chemistry, biology or even evolution, let alone math, language theory, complexity, computer science or quantum physics. It’s just remarkably easy to see and master. We have accepted a model for this living molecular code that is central and quite important in science, and so if ever there was a scientific icon, the genetic code truly is a scientific icon. However, there is now a widening gap between the general idea about “a language of molecules” and our accepted notion of the genetic code. As we probe ever deeper into the complex molecular secrets of life, this cherished scientific icon clearly fails us as a useful model to illustrate, understand and further explain the complex molecular reality that is operating through time and throughout the biosphere.

A big part of the problem today is that general awareness of the genetic code was heralded by great fanfare, treated as an unearthing of not just any code but as the unveiling of the code of life. Obviously, investigation of this natural wonder generated a lot of attention and interest. The same scientists and molecular code busters who painstakingly captured and enshrined this code in millions of textbooks and billions of brains also hinted that it would eventually give man control over his biological destiny. Thus, the great promise of this secret molecular code of life was made so tantalizing that it completely captured our collective imagination. There was a tremendous desire to have a simple model of this code that anyone could easily grasp, and so now we have one. Unfortunately, this model is more aggressively marketed as scientific truth than it is analyzed, debated and genuinely understood and described. What’s more, this simple paradigm for molecular behavior has gone on to become a gatekeeper for other thoughts and broader ideas about life on this planet. So the genetic code now represents a bona fide first principle of molecular information and genetic translation. However, this stunningly simple first principle must also somehow be awkwardly positioned in the center of a much bigger and more complex picture of life. Sadly, it doesn’t fit well there and it never will.

To be sure, things have not gone according to Hoyle with respect to investigations in molecular biology, and so the genetic code today is quietly and curiously disappearing from its high visibility position on center stage. Despite this, it is still touted as an unassailably great scientific achievement. Still, it doesn’t fit well into the current picture of complex life that is taking hold nearly everywhere, nor does it advance any global paradigms in evolutionary biology. Despite its iconic status, the genetic code - as it is described today - is actually useful for very little in a practical sense or in a larger theoretical framework. Truth be told, it does not even allow us to make proteins from scratch, which is ostensibly its most basic and essential role in the grand scheme of living systems. It is hard to go still further then and label this as any sort of general pattern in nature from which basic ideas can be learned and broader tenets of life can be expanded upon. This uneasy pairing between scientific hyperbole and scientific reality is a total shame, because in the proper light the genetic code should be a fabulous touchstone of thought in biology. It absolutely should be a centerpiece for our broader understanding.
of life on this planet, but only if it is properly analyzed and more fully understood. In the crystal clear and utterly logical reality of the natural world, this code is nothing short of a molecular miracle, and so it ought to be seen as an awe inspiring thing. We should be lauding the thing itself, yet it is increasingly ignored today as a ripe field of scientific investigation and wonder. The simple view of the genetic code, a view that I have named the classic paradigm, has become entirely dogmatic. It teaches us very little about basic molecular information, and what it does teach us is for the most part demonstrably wrong.

The name ‘the genetic code’ is the official title that is used to describe a specific molecular phenomenon in biology. But as a visual icon this “code” is simply data that is shown in textbooks as a standard four-by-sixteen grid. In other words, the genetic code is nothing but a spreadsheet, or what is known in computer parlance as a data look-up table. It is also commonly described as a substitution cipher. Within the classic paradigm, the genetic code is viewed simply as a table of “codons” and amino acids. The table tells us that for any combination of three nucleotides, such as are found in DNA and RNA, there is assigned one and only one of twenty corresponding amino acids in the “standard” code. Therefore, it is a purely “linear” relationship between nucleotides and amino acids, and in truth it can and should be fully illustrated on a single straight line. It would be visually inconvenient to present the code in a single line, to be sure, so the data is commonly broken up and condensed into a standard table.

Fig. 1. The Standard Codon Table of the Genetic Code

<table>
<thead>
<tr>
<th></th>
<th>U</th>
<th>C</th>
<th>A</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>Phe</td>
<td>Ser</td>
<td>Tyr</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Ser</td>
<td>Tyr</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Ser</td>
<td>STOP</td>
<td>STOP</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Ser</td>
<td>STOP</td>
<td>Trp</td>
</tr>
<tr>
<td>C</td>
<td>Leu</td>
<td>Pro</td>
<td>His</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Pro</td>
<td>His</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Pro</td>
<td>Gln</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Pro</td>
<td>Gln</td>
<td>Arg</td>
</tr>
<tr>
<td>A</td>
<td>Ile</td>
<td>Thr</td>
<td>Asn</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>Thr</td>
<td>Asn</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>Thr</td>
<td>Lys</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Met</td>
<td>Thr</td>
<td>Lys</td>
<td>Arg</td>
</tr>
<tr>
<td>G</td>
<td>Val</td>
<td>Ala</td>
<td>Asp</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Ala</td>
<td>Asp</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Ala</td>
<td>Glu</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Ala</td>
<td>Glu</td>
<td>Gly</td>
</tr>
</tbody>
</table>
bottom nothing but a collection of data that relates molecules in one line - nucleotides - to a molecule in another line - amino acids. Nothing could be simpler. And so it is true that the code itself is a single line of data that is merely cut up and repackaged into a table that is labeled The Genetic Code. This table dogmatically tells any chemist that “the genetic code” is all about composition and has very little to do with structure. But think about it: This is the molecular phenomenon that was at first considered the code of life. Could this code ever adequately be viewed and understood as a single line of data? The origin and operation of such a thing seemingly defies all logic. In truth, there is very little logic behind this view of things.

This particular “code” was auspiciously given the regal title of the genetic code. It is a grand title, to be sure, but it does seem a bit out of sorts with the thing it names. What’s worse, this grand title is undeniably seductive, so the name quite naturally tempts many uses. Upon brief reflection we find that it is clearly a much too grandiose moniker that is truly at odds with its dreary reality, so there is growing confusion about what the term actually stands for today. Officially there is no confusion, and it should be quite clear what a working molecular biologist means when he formally uses this term today. The genetic code is a codon table, nothing more, yet it is still frequently confused with DNA or the actual sequences of nucleotides in a living cell. Referring to DNA in this way is a more fashionable bootleg usage for the colloquial term ‘the genetic code’. It is a common and understandable confusion of terms between genes and the genetic code. It blurs the line between information and translation. And so the traditional view of genes is that they too are somehow made only of nucleotide sequences. However, ‘Genome’ is the appropriate official term for entire sequences of nucleotides within cells. But even the traditional view of genes has now been called seriously into question. It is a dirty little scientific secret that there is actually no official and widely accepted definition for a gene, nor has there ever been one. For that matter, there is no basic definition of life itself. Confusion, rather than an exception, is the guiding rule in biology.

We build models to sort through the confusion, but sometimes the models we build end up merely adding to it. It should not be surprising then that without basic definitions for genes and life that we find a tremendous amount of confusion when it comes to the code of life.

The plain and simple fact is that the name of this “code” is an obvious misnomer. Even within the formal confines of the classic paradigm, the genetic code is not truly a code but merely data used by some other thing that might properly be called a code. How did we get ourselves into such a bind? It appears that the basic confusion arose primarily from a logical inference that was made as the result of observed molecular behavior half a century ago. However, that observation was demonstrably inadequate, so all of the important inferences made from it are also flawed. The observation involved enzyme activity, the inference was that every protein can assume only one shape, and the universally accepted conclusion was that there can be only one-dimension of molecular information in this code of genetic translation. The accepted model of this code thus became founded on the essential premise that it is a simple, one-dimensional language of otherwise complex molecules. The observations were inadequate, the inference untenable and the conclusion, therefore, is clearly false. This is not hard to prove, but it has proven rather difficult to change. Folks love their heroes and their icons.
A genome is DNA, the double helix, the immortal coil, chromosomes, chromatin, the colored substance of nuclei, the genetic material of life, its idealized structure frequently referred to as “the greatest scientific discovery of the twentieth century;” and so some prefer to talk about genomes as the genetic code. As we shall see, this is not an entirely bad idea, but it is still a harmful confusion of terms, and they are important terms that should be – need to be - quite clear. Genomes and genetic codes are decidedly different things. They are apples and oranges, and they must be described as distinctly different things in reality. Granted, both things could legitimately be called codes in some loose sense, but only because the word ‘code’ has so many common uses. However, we absolutely must have a way to quickly and clearly make the distinction between genomes and genetic codes when we talk about these important concepts at work in nature.

The confusion of genomes and genetic codes today stems in part from a justified disenchantment with the codon table, but it also flows as a natural consequence of flawed perceptions and consequently a horrible imprecision and confusion of basic naming throughout the field of molecular biology. Many of the quaint but misleading names emerged from an exciting scientific era. It was a bustling cacophony of partial comprehension, which is to be expected from the initial stages of any brand new field of scientific investigation. The situation was understandably fraught with ignorance and confusion, but the names that our beloved pioneers haphazardly chose for central things only add to our problems in thinking about these things now. Further adding to the growing confusion is the fact that we now live in a time when the entire human genome, as well as many others, has been miraculously spelled out for storage in computers. How exciting! Indeed, it is now more common to see and hear the term ‘genetic code’ in the lay media in reference to the human genome than it is to hear anyone actually ever refer to the real thing officially known as the genetic code. This is partly the fickle finger of fate behind any popular fad, no doubt, but we should not be too quick to discount the fact that the classic paradigm is badly flawed, and perhaps common sense is merely trumping the dogmatic practice of experts in the field today. However, neither of these two things, genomes or the codon table, actually represents something that can properly be understood as a real molecular code if a code is to be properly understood in this context.

DNA is the double helix or the immortal coil that makes up genomes. It is the unique molecule, or the enduring storehouse of molecular information that records the nitty-gritty information, or a big part of the molecular essence of each individual that is passed along in some form from generation to generation. This buff icon of a double helix is, quite naturally, much more likely to capture our imagination today. It is far more seductive than a stale, flat, colorless, old look-up table. But even the image of DNA has been grossly oversold in the grand scheme of things, and the classic paradigm of the genetic code is the single most important factor leading this false but aggressive sales pitch for DNA.

A genome is much like the hard drive of a computer. It is a physical mechanism for storage of persistent molecular information. It is one form of molecular memory in a cell, just like a hard drive is physical memory for binary bits of digital information. On the other hand, the genetic code should be seen more as the logic of a central processor or CPU of a computer. It is a molecular code, a specific protocol of logic, a crystal recipe, a set of molecular events that define molecular behavior. Ultimately, the genetic code is an
algorithm or a set of steps used by living cells to process information from one molecular form to another. It is a molecular language used by insentient molecules to magically perform molecular translations. Dogma has directly led us into profound confusion on this important point, and the future promises nothing but still more confusion coming from the ideas that make up the classic paradigm. However, for the time being I will use the term ‘genetic code’ to mean the CPU as illustrated by the codon table, and ‘genome’ will mean the hard drive as represented by the double helix.

The genetic code works in all cells making a truly remarkable molecule called a polypeptide. Amino acids are building block molecules that are strung together in polypeptides by the genetic code. Amino acids are like letters that are strung together to make words, phrases and ultimately sentences. Eventually, the sentences work together to write the books of living cells and multicellular life forms. Amino acids are linked one to another by peptide bonds to make long chains of amino acids, or more accurately polypeptides. Poly-peptide simply means many peptide bonds between many amino acids. In living cells, these delicate, skinny, highly flexible polypeptide chains are somehow reliably folded into robust, precise, three-dimensional molecules of great durability, number, complexity and variety. Folded polypeptides have a truly remarkable ability to maintain their delicate shapes in space through time. As a result, they reliably and efficiently perform cellular functions for which they are so well suited that it is hard to think that their design was not sweated over for eons by an army of highly intelligent engineers. The intricacy and apparent purposefulness of these tiny molecular machines is truly breathtaking. One must be left in equal parts awe for the molecules and for the language and intricate molecular system that builds them.

As a group, the many polypeptides within a living cell magically form a complex colony of highly interactive molecules we know as proteins. All proteins are polypeptides, and all proteins are built by the systematic logic of the genetic code. Proteins are also frequently called enzymes, but because they perform so many different functions within any cell they are not just enzymes. Proteins perform myriad tasks within living systems, and as beautiful and elegant as DNA is, it cannot begin to approach the complex and majestic world of proteins. DNA is surprisingly simple but proteins are stunningly complex. DNA is impressive, to be sure, but proteins are sublime. So, the genetic code is involved at bottom in a system that somehow makes proteins from the molecular information stored in DNA. That is the genetic code. It somehow translates DNA into protein. As a code, then, it must be some sort of language that insentient atoms or unthinking molecules use to speak the words of proteins. It is a molecular language that cells use to speak protein from the information contained in DNA. It is a molecular language that reliably dictates the behavior of tiny, mindless molecules on a grand scale of space and time. The language consistently, quickly and efficiently builds, without a doubt, the most breathtaking physical machinery on the planet. It simply must be a miracle of molecular engineering to be able to do this, and so it should be seen first and foremost as an icon of Mother Nature’s coding brilliance. It should not be a dullard of a scientific icon. It is no kluge but a beautifully crafted language of molecules.

However, as sexy and intriguing as this description may sound, the classic paradigm has managed to completely flatten and totally pervert the genetic code as a scientific icon. So, it is not hard to see why, in its classic, drearily unimaginative form,
the genetic code has become wholly unattractive to scholars and poets today. But the experts of today have merely been deluded into thinking they know the genetic code. They do not. They emphatically do not. They have simply drunk a heavy dose of the dogmatic Kool Aide™, which has caused them to stop being curious and amazed about this miraculous molecular code. They have axiomatically accepted the simple and comforting fiction that is the classic paradigm, and they have now moved on to other things. They have missed, and they are still missing quite a bit of the fun, I assure you. This is quite tragic that the code of life has over the years come to be entirely represented as a simple, one-dimensional spreadsheet called a codon table. The standard codon table is merely a partially condensed shortcut notation for showing the linear relationship between nucleotides in ordered groups of three, called codons, and the twenty standard amino acids that make up proteins. This is but a tiny subset of the more robust translation of molecular information that is the genetic code. Unfortunately, besides a few minor details, this subset is pretty much all anyone needs to know about the genetic code today because any more details would be gilding the lily. This “code” can be used to decipher the composition of amino acids in a protein, but that is all. It has never been used to actually make a protein, or even to successfully demonstrate how a protein is actually made - a situation I find most curious. The classic paradigm has never been scientifically “proven” to any degree whatsoever. It has, however, been completely accepted as fact.

It seems odd to me, as a novice computer programmer as well as an imperfect practitioner of a particular human language called English, that I can be instructed in “knowing” the molecular code of life from a simple spreadsheet, when even a basic line of computer code is never quite this simple for me to understand. I also have yet to see a self-contained map or model of English, yet I would expect a language of molecules to somehow rival it in basic complexity and structure. If I can actually see the code of life in a spreadsheet, how then should I expect to be able to “see” English? Obviously, I would expect a more subtle and complex information handling system lying behind a fundamental language of self-organized sublime molecular structure. Where and how on earth did it ever arise? I can find nowhere in the classic paradigm an explanation of how and why this “code” of a spreadsheet somehow came into being as a spontaneously effective molecular language. I merely find a simple, flawed definition and a few simple instructions of how I should imagine a cell uses them. It seems to me a shockingly hollow and entirely inadequate view of things for such a cherished scientific icon.

There are several obvious problems with the flat view of the genetic code that today are leading it quietly off center stage, but these problems have nothing to do with the genetic code itself, the real molecular code in nature. The problems have everything to do with our flawed perception of the genetic code today and the language we use to describe it. The big problem now is in how we perceive this code and therefore how we think about it, model it, illustrate it and talk about it. What’s more, the problem only gets worse when these flawed ideas inevitably spread to how we talk about other molecular phenomenon in the grand scheme of life. Our perception of the genetic code is clearly and strongly influencing our thoughts and words elsewhere. This simply must change, and if you bear with me here it can change. We will attempt to break the dark spell of the classic paradigm that has our thinking about life on this planet pointed in the wrong direction.
So, it is sadly true that within the classic paradigm we view the genetic code today as a linear cipher represented by a codon table and nothing more. That is the best we can do with the model we have. Although the basic data, to the extent we actually know the data, is virtually unchanged over the past five decades, the paradigm that accompanies the genetic code is gradually fading into obscurity, and it is no wonder. The classic paradigm has been annotated, modified and generally whittled away since the time of its hopeful and showy appearance on the public stage. Its explanatory power has greatly diminished but its influence remains. It is of little use in advancing our understanding of the actual data, and observations continue to accumulate that do not reconcile neatly with the unlikely expectations created by the dogmatic simplicity of the classic paradigm. The result is that the once shining icon of the genetic code has lost much of its luster, and now it is being unceremoniously shoved aside. It is disappearing from its throne as the code of life. Good riddance to the old views, but let’s certainly be sure to replace them with something befitting of a cherished scientific icon. As we move forward, we will restore most of its proud luster, and add useful and beautiful embellishments to its structure.

It is most ironic that all this should be happening at a time while computers and codes have taken a place of such cultural prominence in modern society, because the genetic code is perhaps one of if not the oldest complex and elegant code on earth. It is a true language or code of molecules that somehow wrote itself, and so computer analogies are unavoidable here. It is a masterpiece of Mother Nature’s programming precision and efficiency, yet it is universally not seen in these terms today. It is commonly described as a “frozen accident” and any hint of shortcomings or flaws in it are attributed to this necessary icy encasement. It seems from this that we must accept that it is what it is because it is what it is. Not a very satisfying description. Yet these are exactly the descriptions one finds in virtually all of the creation myths of biochemistry today. But false contexts and false premises have grossly distorted our perception of reality. In reality, the genetic code is cleverly structured by nature and it is cleverly executed within a cell. It is not an accident of any sort because there is a fundamental rhyme and reason behind its existence and its ability to continuously hone its effectiveness. It is, in fact, still evolving right before our very eyes according to the same basic logic, physical constraints and selection pressure that gave rise to it in the first place. The sublime elegance, complex efficiency, versatility and broad utility of this code should rightly be the envy of any hotshot computer jock pounding out code today. As a code it really should be an icon. It certainly should be gaining not losing prominence in the computer culture today.

It is difficult to appreciate any romantic view of the genetic code from within the context given to us by the classic paradigm. It will take some effort, but if life is to be seen as a complex, self-organized, adaptive system of molecular information, then the genetic code should be seen as a centerpiece, or as some kind of “machine language” of that system. Its structure and function should serve as a template for other molecular processes. The higher codes of life should not merely employ it as a simple, linear subroutine; they should also reflect its fundamental structure as an actual code.

Paradoxically, the opposite has occurred. The genetic code in all its perceived simplistic glory has become a first principle of biology, and therefore our entire view of molecular information and life’s complex processes has been jaded by it. The molecular sequence has been crowned lord and king of all molecular information, and apparently all
else is merely entropy. DNA is now the undisputed king of molecular information, no less than the command and control center of life, the center of rotation in the biosphere. This is, I firmly believe, a false view of reality, but the rest of our complex paradigms are now tasked with escaping the intellectual gravity of this simple yet false king.

The stakes are high. A flawed model leads us astray in important areas of thought; whereas, a more adequate model of this code could take us much farther than we’ve been able to previously go. The key issue is no less important than an understanding of the first principles of life itself. What are the basic forces in nature that can help us distinguish between a simple carbon atom in diamond versus the complex behavior of a carbon atom in protein? What game or system has the universe put in place that allows insentient molecules to collectively become living systems and sentient beings? How does the genetic code fit into that general system? How can our basic understanding of the genetic code advance our understanding of these deeper, less penetrable questions?

The basic view afforded us by the classic paradigm is that life at bottom is a matter of molecular sequence. In this view a sequence of molecules is synonymous with its composition as dictated by the exact sequence of its constituent molecules. The actual structure of even the most complex molecules is wholly determined by its sequential composition. All fundamental truths about complex molecular structure are therefore derived from their sequence. The essential languages that build these structures must also reduce to languages of sequence. The genetic code is therefore a linear language of molecular sequence. The unpleasant truth is that even in common protein chemistry there are countless cases where the same sequence leads to many structures and many sequences can lead to remarkably similar structures. Sequence and structure are nowhere near synonymous, yet the subset – composition – has somehow come to conceptually dominate the superset – structure. The ramifications of this view are immeasurable and as it turns out highly detrimental to our deeper understanding of real first principles. Molecular information is made simple and cheap not complex and precious. Accumulation and exact translation of precious molecular information is ignored. Molecular languages are coarsened and cartooned. Their origin, operation and engineering remain hidden and the teleology of the entire system becomes moot. Important components of the system are tossed out, wrapped in the thin paper of thin data and still thinner ideas.

No life on this planet now exists without protein. No protein yet exists on this planet without the translation machinery and some form of code on which it operates, the code we call the genetic code. This is why it is called the code of life. But no life exists without a robust collection of elements found in the periodic table of elements. This table is not called the table of life, and rightly so. The codon table was at first believed to be universal and unchanging across life on our planet today – truly exciting if true - but that has been thoroughly proved not to be the case. There are actually many known cases of variation, and this is an important fact that is routinely discounted. The degree of variation between flavors of the genetic code has been lost in its simplification and flattening. The codon table is not universal and it does not reflect any universal truth of this code. It merely illustrates a subset or component of the code that remains remarkably similar across all life forms that use it. It is a result of extreme molecular bias, a common tactic employed by Life in setting up its intricate molecular systems. Beyond this, there
certainly are first principles of this language, but they are not an embodiment of the table of codons. However, there must be more fundamental truths that lie beneath this table. This must be true for it to ever arise in the first place, and this also must be true to account for the remarkable similarity observed in codon assignments across the biosphere.

How is the universe constructed to give rise to this natural phenomenon? What are the first principles of molecular behavior that allow this pattern? What are the forces that give rise to life and drive its tireless march toward higher molecular complexity? What is a molecular language in the first place and how could such a miraculously complex thing ever arise? These are the difficult and important questions that must be asked. To find the answers, I believe, we will need to deconstruct the basic tools of human thought and reassemble them into a more coherent model. Simple geometry is one such tool, but it has been completely ignored by the classic paradigm. Basic solid geometry appears to be one of the few tools available to serve as a basic truth on which molecules can begin building a complex language, a language so complex that it “lives.” All molecules must obey rules of geometry and so we should expect to find these rules somehow embedded in any language used by its molecules. Although definitive answers are now darkly hidden from human eyes, basic human instinct should lead us to examine the myriad clues that are everywhere, clues that tell us that the language of proteins is fundamentally a structural language and it is somehow anchored in the basic solid geometry of the universe. Life does not merely build molecular compositions in a simple way it builds complex molecular structures in a complex way.

Despite persistent claims confidently made to the contrary, the true genetic code, the actual language used in nature by insentient molecules to make protein, the complete translation of molecular information from one distinct form to another, that genetic code is still largely unknown today. More telling, we are still without a single piece of empiric evidence that the classic paradigm actually works as a scientific model for protein synthesis. It has been completely accepted as fact but never actually proven as reality. Certainly, something this cherished and important in science deserves to be debated, explored, properly understood, and marveled in this, a new era of the algorithm. But we must leave our prejudices and preconceptions at the door as we enter this exciting new era of molecular information. The field is still wide open and so should be our minds.

To understand the genetic code from the standpoint of the classic paradigm we must first examine the rules of its presumed behavior. We must examine the algorithm postulated by the classic paradigm. This algorithm is founded on three simple premises and they are all false. Each one is a proposition of absolute simplification with respect to translation of molecular information. The three premises are called co-linearity, the central dogma, and the thermodynamic hypothesis of protein folding. Since all three premises are integral to the classic paradigm, I will do my best to fairly describe each in somewhat more detail. We must first understand what they mean before we can understand why they are totally wrong.

The principle of co-linearity means that genetic information is contained in a “linear” molecule of DNA or mRNA, and there is a direct correspondence between this information and the “linear” molecule of a protein. Co-linearity spells out the idea that sequential molecular units of a polymer define the molecular information being translated
during protein synthesis. It is a process of virtually all composition and virtually no molecular structure. The concept of co-linearity is no doubt a useful expedience toward understanding some translation mechanisms, but it is also an absolute simplification with respect to the nature of genetic information.

The second premise, the central dogma, is a truly unfortunate term used to describe the direction or “flow” of co-linear genetic information during the process of protein synthesis. It specifically says that genetic information flows from DNA to RNA during ‘transcription’ and then flows from RNA to protein during ‘translation’. Protein synthesis is a clear-cut, two-step directional procedure of first transcription and then translation. There is then postulated to be an entirely separate process called protein folding where every translated sequence of amino acids finds, of its own accord, its inevitable final shape. The central dogma too has its uses for pedagogic expedience, but it is also a statement of absolute simplification with respect to genetic information.

The third premise is called the thermodynamic hypothesis of protein folding, and it is most aptly named. This is the lynchpinch of the classic paradigm. It defines our view of the genetic code and forges the general shape it has in our minds today. Without the thermodynamic hypothesis, the classic paradigm would surely be much different. But, the thermodynamic hypothesis of protein folding is actually not a single hypothesis; it is two completely separate hypotheses. The first is that thermodynamic molecular forces completely dictate the mysterious process of protein folding after translation. In other words, the random collisions between protein and water molecules will determine the persistent shape of a protein after the genetic code has strung the amino acids together. This is a logical idea, since thermodynamic forces play such a large role in all molecular processes. Life is an aqueous system, and so it is natural that the random motion of water molecules will strongly influence the ultimate shape of proteins. However, the second part of the hypothesis takes the idea much further. The second part of the theory is that for every sequence of amino acids there is a corresponding shape – a state of lowest Gibbs free energy, as they like to say - that should in all cases be observed within a cell. This is sometimes called the single target hypothesis of protein folding.

The single target hypothesis means that when a protein starts on its path of folding, from any form - but especially from the form in which it emerges from translation - the protein will unfailingly find the single most thermodynamically stable target at the end of the process of protein folding. This should logically mean that there must be in all cases a single most stable target that the protein can somehow consistently find. In this context, protein folding is a matter of pure statistical probability. It is a process of entropy. It is order from disorder. How else could it be? Another way to put it in chemical terms is that a protein, under physiologic conditions (a huge and necessary open-ended qualifier for the theory that we will discuss a bit later) has but a single conformational isomer. In other words, in a living cell a protein can fold only one way. That is simply an unavoidable molecular procedure forced upon a living cell in nature, according to the theory. The single target hypothesis is a statement of pure belief in the primacy of entropy. It is entropy and entropy alone that determines the shape of a protein within the classic paradigm. This is a most curious belief in light of the clear warnings issued by Shrodinger less than a decade before this firm belief became so completely accepted and so pervasive.
The single target hypothesis is the grandaddy of statements of absolute simplification regarding genetic information. It simplifies the realm of molecular translation to a single dimension, the dimension of amino acid sequence. This is specifically why the genetic code is described as one-dimensional because the only dimension of molecular information is defined by the order of nucleotides which is translated by the genetic code into the order of amino acids. It is the meaning of the familiar axiom that “primary sequence determines tertiary structure.” Who hasn’t heard that one? This axiom has defiantly stood ever since the thermodynamic hypothesis was proposed and it has stood firmly in the total absence of empiric evidence, I might add. It is a first principle of molecular information in living systems today. It is the reason that the classic paradigm views a protein as a sequence of amino acids when more logically a protein should always be seen as a sequence of peptide bonds. It allowed the total abandonment of molecular structure for the false promise of simplicity seen in molecular composition. Furthermore, it is the sole reason the genetic code is called redundant or degenerate. There are quite obviously more codons than there are amino acids. Each amino acid in the genetic code is assigned on average three codons, and since amino acid sequence is the only dimension of information translated by the genetic code, it is a one-dimensional, redundant, degenerate code.

The single target hypothesis is a doozy of a hypothesis, and in part it earned its author, Christian Anfinsen, a Nobel prize in 1972. This is the equivalent of a Supreme Court decision, stare decisis, it is now settled law and no longer a debatable idea. Perhaps Anfinsen deserved this prize, but the plain fact is that he never proved the single target hypothesis – he barely even tried to prove it. He later admitted his extreme experimental limitations. In reality, he was in no position to ever prove this doozy of a universal simplifier because he simply did not have the tools available to construct such a proof. My advice to all budding heretics in this field: when the experts come for your throat, and they will, ask to see their empiric evidence for this doozy of a hypothesis. I would be surprised if they can cite even one tangible piece of evidence to justify the universal belief in the entropy model of the genetic code. The insistence on an unquestioning belief in a one-dimensional molecular code is nothing but scientific smoke and mirrors, and the rest is just a house of cards. Thus, the unbridled anger toward anyone who raises questions about this central belief.

The main attraction to this theory at the time of its acceptance was not empiric proof but the promise of utter simplicity. It was a bright green light for the classic paradigm to proceed full speed ahead on the path of simplicity. It was a provocative suggestion to all scientists in the field that all of their work could and should now proceed within a very narrow focus on but a single feature of molecular information. Anfinsen proved the first part of the thermodynamic hypothesis of protein folding to everyone’s satisfaction, the part about the obvious influence of thermodynamic forces in protein folding, but the second part of the theory today is still awaiting prospective, empiric confirmation. So, it is merely axiomatic today, but it is unlikely to ever be proven, because it is unlikely to be true. It is logically untenable, and in fact it is not hard to find plentiful evidence to disprove the single target hypothesis, yet the classic paradigm continues to insist that we cling tightly to this premise. Why?

These three central premises work in concert to present the simplest possible and a very specific picture of the processes of protein synthesis. Together they unequivocally direct
us to view translation of proteins as a simple, unidirectional, stepwise process of sequencing amino acids and only of sequencing amino acids. The rest is just a grand and mystical match game of finding the one shape that any sequence must become. The sequence is king and in this case the sequence of amino acids is all that matters. DNA is the command molecule that dictates the information content in the world of proteins. Entropy is the force of nature that allows DNA its unchallenged reign. We are inevitably led to the conclusion that the composition and therefore the shape of every protein are essentially rigidly and singularly predetermined in the molecular universe, and predetermined in each case simply by a single sequence of nucleotides. In the world of biomolecular information today, DNA is the ruler and king, and the primacy of molecular sequences cannot be questioned.

Co-linearity seems the most innocuous of the three, so simple and apparently correct that only a fool could object to it. But co-linearity is in fact a grossly improper descriptive term, inaccurate as a general concept, and rotten to the extent that it is blocking a view of the real code of protein synthesis. It might be seen as “right” in such an extremely limited, almost useless sense, yet it is known to be false in so many important ways. It is nothing more than bad poetry in an area that calls for scientific precision. The term ‘co-linearity’ stands for the concept that the genetic code translates, in discrete stepwise fashion, a “line” of information in DNA all the way to a “line” of information in protein. It tells us that the genetic code is linear in two different ways. First, the molecules involved are made of long chains of individual subunits that can whimsically be thought of as lines. Second, a single dimension of molecular information is responsible for protein synthesis, and this one dimension is entirely defined by the identity of each molecule in each chain. Molecular information is therefore not only in a line; it is a line. The premise of co-linearity is false on many levels, but most objectionable is the casual use of the term ‘linearity’. Words must have meaning, and so there are real consequences to sloppy application of names. The word ‘linear’ is cheap in the sense that it usually costs nothing to understand yet it carries so much meaning, but in this case it carries far too much meaning. It quickly becomes an expensive word in this case because virtually none of the meanings of the word linear apply to protein synthesis, and its usage requires extensive apologies for what the term does not mean. Linear connotes and denotes several things. With respect to genomes, linear simply means not circular, and we can accept it for that limited use here, but that is all. In the colloquial, linear means a thing that forms a line, and when looking at a diagram of DNA and polypeptides we might see why that is a very tempting term indeed. However, in a formal, technical sense, linear means a one to one mapping of variables, or a plot on a graph that forms a line. But no use of the word linear is appropriate for describing protein synthesis or the molecular codes that perform it. The word ‘sequential’ is more apt, but even co-sequential is a poor way to describe the molecular processes of protein synthesis if they are ever to be properly understood.

A chemical polymer, such as RNA or protein, is usually thought of as a sequence of sub-units. Perhaps in the most simplistic view they can be referred to as a line of subunits, but this is an extreme disservice to the sequence and to the sub-units. Each sub-unit is a real-world assembly of atoms, and each of these assemblies forms a demonstrable shape. The composition of a molecule is merely a subset of its structure, but the opposite is
never true. Furthermore, the bonds between these amino acids have demonstrable shapes as well. It is wrong to imply that a polymer such as DNA or protein may be considered as anything less than a sequence of shapes. It in no way serves our purpose to conceive of this real-world assemblage of shapes as anything even approximating a line. The specific shape of each of these molecular structures at all times will determine its behavior and define the actual “information” of the molecule. So the polymer itself must be considered, at all times, a complex shape, regardless of its specific conformation at any given moment. The word linear is not only sloppy it is misleading and detrimental to proper thought here.

These polymers should properly be seen as a sequence of three-dimensional bonds, not as a sequence of dimensionless points defined by the molecular name of each point. The molecular information contained in a polymer is more completely defined by its sequence of bonds. For instance, any amino acid sequence only gives us a small percentage of the information in a polypeptide because there is a tremendous amount of information in the peptide bonds between each amino acid. Conversely, and quite logically, bond information always includes the identity of the sub-units that are bonding partners, so the sequence of bonds is still inclusive of the sequence of sub-units. We can always derive the sequence of amino acids logically from only the sequence of peptide bonds but never vice versa. An intellectually honest view of this simple concept must lead anyone to concede this seemingly trivial point. It is only seemingly trivial, though, not actually trivial; it is all-important to some of the very basic flaws in the classic paradigm. Only in the most extreme case of natural simplicity will this ever be a distinction without a difference, but nature did not give us that one case from a nearly infinite number of possible cases. It cannot, however, be proven logically false, and all evidence points to there actually being a very big difference. After all, this is ostensibly a paradigm about molecular information, but it is foolhardy in the classic view to ignore large obvious chunks of molecular information in the paradigm. The defenders now must continue to shrug them off as unimportant. However, a protein should at all times rightly and formally be conceptualized as at bottom a sequence of peptide bonds; not merely as a sequence of amino acids. The classic paradigm insists that we completely accept a logical absurdity.

Although any sequence is technically a “linear” set of elements arranged in a particular order, the concept of line and sequence part ways formally and dramatically in ways that are vitally important here. Sequence denotes an order in space and time, whereas a line does not. Length and distance between two points on a line are rigidly defined by a formula for the line. Unlike a line it is the number of elements in a sequence that determines its length, and most important to molecular sequences the “distance” between two points can vary in time and space depending on the actual medium in which it resides. Sequences have a direction or progression in time and space, whereas lines formally do not. A line with direction is called a vector or a ray. Anyone who confuses the concept of a sequence with the concept of a line will struggle mightily to interpret the mountain of facts now known about protein synthesis at every level.

In the real world, it is not all-important merely being next to things, like in a line on a graph, it is more important being before and after things in time, above or below them in space, or simply being nearby, accessible in both time and space. This is an essential concept with respect to molecular information in general, and specifically with respect to
the real world translation of molecular information in protein synthesis. The kind of sequence referred to by co-linearity is strictly a sequence of molecular identities - their order - but it says nothing of their overall physical arrangement in time and space. All of the actual molecules involved in the genetic code perform their tasks based on their order and their physical arrangements in both time and space. They are temporally and spatially active sequences of molecules, not lines. The word ‘string’ is borrowed from computer parlance, and it is, I think, slightly more useful and far less misleading when we wish to refer to a character sequence of molecules in physical space. At least physical strings of twine are a more apt visual. I am not opposed to referring to DNA, RNA and proteins as molecular strings. It is far less objectionable than molecular lines. Many of the fascinating codes of molecular information are keyed to events in time, or they require the physical juxtaposition of various structures in space. The sequential order of nucleotide elements in DNA plays a major role in these processes, to be sure, but not merely from the simplistic, one-dimensional, classic standpoint of a linear molecule. The real genius of “linear” DNA is not just that it records an ordering of molecules; it is that it serves as a physical platform for creating spatial and temporal sequences of processes, spatial and temporal arrangements of molecular groups, and spatial and temporal coordination of molecular events. All of the complex and intricate machinery built up around DNA takes advantage of this unique property of this intriguing molecule. The vast majority of real molecular information in complex, self-organized, adaptive systems emerges in time and space at ever-higher levels of complexity within this system. This is a valid first principle of molecular information, but this thought is entirely subjugated by the simplistic premise of co-linearity, which tells us that molecular order is the primary thing.

There is a more specific technical flaw with the term ‘linear’ as it gets applied to the genetic code. In the formal sense a linear relationship is a one-to-one mapping of things, like X = Y. With respect to the genetic code, this premise fails on two points. First, the actual mechanism of codon pairing - the molecular process inside the tiny black box - is now quite clear. Codons do not pair with amino acids; they pair with anti-codons. Codon “wobble” adds a fifth nucleotide to the mix which brought with it more anti-codomons than codons; therefore, in no practical way can we make a one-to-one mapping of codons and anti-codons. There is a one to many mapping (read as non-linear) and therefore codons carry with them a form of ambiguity with respect to their actual information content. Defenders of the classic paradigm can’t accept this as a blindingly stupid mistake in their paradigm, but it is just simple math. These basic numeric idiosyncrasies cannot be ignored because there is a mountain of them and they all point in a direction opposite of the classic paradigm. Second, the premise of co-linearity carries with it the inference that the genetic code merely assigns one sequence of amino acids to any sequence of nucleotides. It is a simple, linear relationship in this way and so the codon table is a simple, linear set of data. Perhaps one could technically take this position, but in fact the genetic code must logically also simultaneously assign a sequence of amino acids to every sequence of nucleotides. This also means that it must assign one to every conceivable common transformation of any sequence of nucleotides. This becomes a big issue once one realizes that there are many “easy” and common transformations of any nucleotide sequence, and so we should expect them to occur quite frequently in a living cell.
Logically, these transformations become a factor in the fitness of any living cell. We shall see in a bit that the genetic code has taken all of these transformations into account in arriving at the canonical assignment pattern. So codon assignments should not be seen as a simple single pattern but actually as a complex mosaic of many patterns. Therefore, the codon data itself is not a simple and linear pattern but a complex and non-linear pattern. Codon assignments seen in a proper context are, therefore, an example of a one-to-many relationship just as codons are to anti-codons, so the term linear is misleading even with respect to the codon table. This topic will be addressed in more detail later in a discussion of codon assignments.

Another subtle breech of the concept of linear can be found in the variability of importance in particular elements of any molecular string. Sequences of DNA, RNA and polypeptides can contain some individual elements that are “important” with respect to global meaning and behavior during translation, while other elements are seemingly not so important. Some characters in these strings have the ability to impact transcription and translation mechanisms and results in a big way, both in terms of what downstream strings will become, and in terms of what strings “are” or how strings behave in time and space. In other words, some character replacements or substitutions will change the shape of the molecule, and some will change the timing or fidelity of events. Some substitutions will change the actual length, composition and order of current and future strings. On the other hand, some characters or extended stretches of characters seem to have no impact at all on any of these things. They seemingly can be replaced with impunity. They might merely be physical placeholders that appear to have no impact on downstream strings other than perhaps changing the length of intermediates, but the impact on behavior in time, space and the importance of their order might be negligible. This property of characters having variable importance within a single sequence should be thought of as a non-linear property not a linear property.

Another quite significant problem with the premise of co-linearity is that it greatly distracts us from the complex multilevel process of building a protein in the first place. The information for building a protein is never entirely contained within a single, contiguous sequence of nucleotides, and it is not linearly configured in any real or practical sense as we are classically instructed to think. But actually think about it for just one second – use your own brain for a moment - do we really want to in any way call protein synthesis linear? Have we learned nothing about non-linear dynamics and complex information systems in the past several decades? Are we going to ignore all of the wonderful data collected in molecular biology and continue to pretend that this is somehow a linear process? Do we want the term ‘linear’ just lying around in plain sight of and co-mingling with descriptions of protein synthesis when young, innocent and impressionable schoolchildren might pick it up and begin playing with it? The information for building a protein is packed into many fragments of information and it is unpacked in a multi-step or iterative process of decoding. Many proteins are needed to make a protein at every single step in the process. It is a general concept that ironically might be rightly called epigenetic and non-linear.

Transcription and translation are the steps that are specifically supposed to be co-linear, but collectively they require many molecular codes of behavior, and these separate codes are embedded in some way within the molecular sequences involved. These additional codes must share the text with the “genetic code” of amino acid sequence during the
entire process. There are promoters, regulators, splicers and folders to name a few of the “codes” in all of these sequences, and all of the “meanings” of all of these codes must work in concert if a protein is ever to be made properly. It is naïve in the extreme if we are going to pretend that this will be a one-dimensional, linear process in any way shape or form. The molecular strings that execute the steps of transcription and translation do their work in both time and space, and rarely if ever is there a clear-cut case of molecular order in the sequence as the sole controlling factor of any step. There are many forms of molecular information that must all share the same sets of molecular strings, so to believe that there is a one-dimensional, linear process to be found here is to believe in the Easter Bunny. It is a comforting fiction with absolutely no basis in reality.

Regulators of transcription and translation are themselves not linear as well, and they are always entirely dependent on proteins. Plus, the transcripts that are made, called messenger RNA (mRNA), contain leaders, trailers and un-translated middle pieces that contribute vital information to the final proteins they encode. The rate of these processes can and do determine their outputs, and the shapes of these transcripts can provide additional meaning as well. The mRNA are frequently cut and spliced together based on codes within them that interact with environmental factors tasked with interpreting these splicing codes. The outcome of editing and splicing can vary greatly, making one protein or another, or hybrids thereof. There is no one-to-one relationship in this process, and “co” anything from DNA to mRNA is a tremendous stretch of the imagination. There is no “co-linearity” for even a single mRNA that can become many different proteins! Likewise, there is no “co” to the process of translation. “Co” and “linear” are not good words, not words that pop into my mind anyway, when I study what’s going on with this fabulously complex molecular machinery.

The complexity only intensifies when the transcript is actually translated. Large populations of monstrously large molecules – monstrously large when compared to codons and amino acids anyway – called tRNA. They are the foot soldiers of translation. The ‘t’ stands for ‘transfer RNA’ but the word ‘translation RNA’ is more appropriate, because translation is a term more in line with what these molecules actually do. When viewed from the perspective of the classic paradigm, these tRNA molecules are seen as mute vehicles of transport for amino acids. They are nothing but silent “adapters” to fit an amino acid into its codon socket. This is a quaint notion but a farce with respect to accuracy. In reality, tRNA are intriguing players in the game with important roles to play in the process. They aren’t mute; they’ve got something to say about the situation. The molecular information they provide is in large part a function of a tiny portion of their string order, called an anti-codon, but it is clearly not the entirety of their molecular information. The active and diverse populations of tRNA molecules translate mRNA in a way governed by population dynamics within a cell. The tRNA must interact with mRNA, rRNA, many proteins, and with other tRNA as well. Taken together, these interactions are complex and governed by the rules of complex self-organized, self-assembly systems. This behavior can be expected to fall into the non-linear category of molecular behavior on all levels.

Because tRNA have been mostly overlooked and ignored in “the genetic code” the operation and impact of tRNA on translation is unknown today, not much data, but it is sure to manifest itself in the temporal spatial sequences of polypeptides, as well as the emergent information contained in overall composition profile of entire protein colonies.
So the role of tRNA certainly cannot be completely ignored in this process, but tRNA are completely ignored by the classic paradigm. DNA is transcribed to mRNA and mRNA is translated into protein. Where is the tRNA? The ultimate irony in this classic jabberwocky is that if one wanted to actually “locate the genetic code” in a single physical molecule within the living cell, one would have to look at the cell’s entire population of tRNA and observe their complex interaction with all of the other molecules in their environment. Without any sense of irony whatsoever the classic paradigm instructs us to ignore tRNA completely and behold the beautiful code of life residing in DNA, the darling of molecular biology.

So the premise of co-linearity is false virtually everywhere we look. ‘Linear’ is certainly not a word that comes to my mind when I think about the complex operation and output of complex populations of insentient molecules. As a descriptive term ‘co-linear’ is nothing but grossly misleading with respect to protein synthesis. We might be tempted to replace it with the term co-sequential, but even this is misleading in too many ways. The best we can probably do is to describe the molecular processes that comprise protein synthesis as a complex, non-linear sequence of events. The “code” moves through time in complex fashion from one distinct molecular string to the next based on sequences of time, space and molecular order. These strings and procedures create emergent forms of molecular information from emergent forms of molecular information in a non-linear fashion. When a protein is made it is only in retrospect that we can isolate limited aspects of it and say that a series of molecular strings was created in the process. More accurate is to say that a series of molecular events occurred in time and space. But these strings themselves are not linear in their formation or their existence. They are not lines and they do not function as lines.

Moving on to the second of the three key premises, the central dogma, we can see that it should have been a stillbirth, if in nothing but name. But, more than etymologically challenged, the central dogma is completely flawed as an idea. It then, unfortunately, serves as the primary basis for more tremendously flawed ideas about genetic translation. It is bad form to name anything in science a dogma, let alone a freshly minted “central” thing. Dogma is the antithesis of science, or it is in theory anyway. The word dogma means a belief based on faith, and it usually connotes some moral shading to the belief. Francis Crick is the coiner of the central dogma and later admitted to coining the term in ignorance. What I suppose he meant was the central axiom, which is a thing taken as true but unproven. Axiom is an accurate word and a whole lot more tolerable within science when describing the thought being conveyed by the central dogma. However, it has now become a central premise of the classic paradigm. A premise is a proposition that forms the basis of an argument, or a premise is taken as evidence of a conclusion. Regardless of what he meant, it is yet another obvious example of sloppy naming within the classic paradigm, and a prime example of the reasoning that still underpins the defenders of the paradigm. Unfortunately, rather than merely quaint, the Mad Hatter language of the classic paradigm is a continued burden toward understanding molecular information. Rather than interpret what is supposedly meant by this central axiom of genetic decoding, we’ll let the man speak for himself. Here’s what Dr. Crick had to say about the central dogma:
The central dogma states that once “information” has passed into protein it cannot get out again. The transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein, may be possible, but transfer from protein to protein, or from protein to nucleic acid, is impossible. Information means here the precise determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein.”

This, of course, is absurd. Even from the viewpoint of the unnecessarily limited way in which it defines molecular information, it is comically flawed. Perhaps it is nothing but a simplistic attempt to define molecular information, but one is hard pressed to find any truth to it at all in nature even in this regard. Empiric evidence proves it false on all counts. That leaves us in the untenable position of either defining it in a uselessly restricted fashion - for historically romantic reasons only - or rationalizing its many obvious flaws and continuing to use it anyway, which seems to be the chosen path in molecular biology today. We might try to interpret this lovable axiom of science to mean that time does not reverse itself during protein translation, but then we must rename it “The Central Triviality.”

With all we know today, if someone naively offered this concept as an axiom now, they would be laughed at and ignored, perhaps even boxed around the ears. The two basic steps of protein synthesis, transcription and translation, have proven to be a complex web, or a torrent of molecular information flooding the process from all angles and moving in all directions, as one might expect from the complex activities of insentient molecules. The “information” is introduced in many forms: in the form of editors, support structures, co-factors, regulators and feedback loops. More important is the fact that proteins participate in their own creation in numerous ways, imparting molecular information on the process at every level, and this includes the actual string “information” of both nucleic acid and amino acid sequence order. Not only is the rate and timing of transcription and translation relevant, yet being defiantly ignored by this dogma, but spatial information is obviously entering the process at several points and from several unanticipated angles as well. There is clearly more than one form of molecular information participating in “the genetic code” but this “dogma” is declaring its willful ignorance of this important “extra” molecular information. In reality, it is far from a simple, smooth, lateral transfer of one-dimensional information from DNA to protein. The total package of information that winds up in a finished protein can properly be seen to come from many sources, several directions and much iteration.

More concretely, the central dogma directs us to conceive of protein synthesis as a simple two-step process: transcription and translation, and translation is painted as a single, homogeneous step from mRNA to amino acids. But, in fact, even within the simplest interpretation of these events there are at least two demonstrable steps of molecular string translation during “translation.” First, information is translated from mRNA to tRNA; then, information is translated from tRNA to polypeptides. Compressing translation to a single step would perhaps be an acceptable linguistic shortcut if - and only if - it is proven that no additional information is generated or involved in the two distinct steps from mRNA to polypeptide. Even so, at least a courtesy should be paid to formally and individually naming the steps as ‘translation one’ and ‘translation two’ instead of hiding them as is done within the central dogma. But, in fact, the speed of translation, called translation kinetics, has been conclusively shown to be variable during these two steps,
and regardless of the actual consequences, this is at least one demonstrable level of molecular information that must be considered in the translation process as a whole, if only for intellectual honesty, completeness, and academic rigor.

In truth, the central dogma is merely a restatement of the single target hypothesis of protein folding. Once it was agreed upon that protein synthesis would be seen as an isolated act of amino acid sequencing, the central dogma declared that it would be seen as a simple two-step process. The central dogma reformulates, validates and amplifies the single target hypothesis of protein folding, but it adds little new or useful to the discussion. If the single target hypothesis is false, and it demonstrably is, then we must re-evaluate the crux of the central dogma because if more than amino acid sequence is impacting on protein synthesis we must make efforts to identify the origin, operation and “flow” of this additional molecular information. We cannot merely declare by fiat what something will be and then pretend that, against all evidence, it actually is that way. That is religion not science.

The biggest problem with the central dogma, besides being utterly false, is that it assumes and reiterates a brutally simplistic definition of genetic information. It actually stands as the definition of molecular information in the classic paradigm of the genetic code. It is a false definition, so there is no chance whatsoever that the classic paradigm can be anything but fundamentally flawed. After all, genetic information is at base molecular information, so it is not proper to exclude known forms and sources of molecular information in defining and then discussing the methods of genetic translation.

The painful, unspeakable, yet plain truth is that we have yet to identify and classify all forms of molecular information involved in the broad process of protein synthesis - far from it - so we certainly should not feel compelled to make any absolute statements of simplicity about them. But, more unfortunate is the fact that not only have we accepted this grotesque act of simplification we’ve used it as license to go much farther. And so the profound implications of this famous and authoritative declaration of the central dogma have been deeply internalized by broader paradigms in biology. The central dogma now represents a first principle of genetic information. We are directed by it to view the simple, fundamental hierarchy and “flow” of molecular information as follows:

\[
\begin{align*}
\text{DNA} \\
\uparrow \\
\text{DNA} \longrightarrow \text{RNA} \longrightarrow \text{Protein}
\end{align*}
\]

Fig. 2. The Central Dogma

This is the classic graphic statement of the central dogma unabashedly telling us that, for lack of a better word, DNA makes DNA, and DNA makes RNA, and DNA makes protein. DNA makes protein in the sense that it supplies all the information that a protein needs to become a protein. It all starts with DNA and the molecular information it contains. The molecular sequence of DNA is the molecular information of the biosphere, the power that DNA holds and the sword it wields over all other molecules. DNA is master and commander of the biosphere when it comes to molecular information. It is the command and control molecule of life. The classic paradigm of the genetic code makes this point very clear to us through the central dogma.
This dogma that is so central to molecular biology is a very intriguing proposition worthy of some serious consideration. It quite obviously has profound implications for our paradigm of genetic information in general and genetic translation specifically. It is not merely quaint and innocuous. But only if there were some evidence of its truth would it be worth taking the premise of the central dogma very seriously at all. Clearly, it turns out to be utterly false at all levels, with more than enough good, hard, empiric evidence of its fallaciousness that we can be absolutely 100% certain that it is in fact false. What’s worse, in the grand scheme of things it actually gets the molecular roles exactly reversed: it is not only tragically incomplete it is completely backwards. In truth, all of these different molecules, DNA, RNA and protein, participate in the process of translation, and all of them contribute important “information” to the process. Even if we accept the ridiculous definition of information in the dogma itself, it is false. But DNA actually “makes” nothing. Protein makes DNA, and protein makes RNA, and protein makes protein. It does so by using DNA, exploiting its unique molecular properties, but it does so by adding every conceivable form of molecular information at every step of the process. So a more nuanced and enlightened map of genetic “information flow” should rightly look like this:

![Diagram of protein, DNA, and RNA relationships](image)

**Fig. 3. The Central Dogma Corrected**

This map puts protein in its rightful place at the center of the biomolecular universe. In the world of molecular information, protein is the master and DNA is its slave. DNA is a kept molecule. Sure, at bottom there is always a chicken-and-egg type of question, it is the Escher drawing of two hands drawing each other, but there can be no doubt today exactly which molecule is in control of molecular information - and it is protein! Proteins turn things on and turn things off. They decide what will be done, when it will be done, where it will be done, how it will be done, how long and how fast it will be done, in what proportions and under what circumstances it is done. Proteins provide the rules for storage and the rules for access to information, while DNA merely provides the actual physical storage of information. DNA is molecular memory that provides some key physical properties of the global configuration of that physical information. DNA is a scaffold and a template for the elaborate realization of molecular information processing in space through time. It is not a code. Repeat that: DNA is not a code. If you ever find yourself confused while hacking through the thicket of gibberish in molecular biology, get your bearings by repeating that sentence: DNA is not a code. Granted, some of the important molecular information is stored in DNA, but DNA contains no protocol or algorithm for its own translation. Still more molecular information emerges in space through time at the direction of protein. It is quite true that some additional nucleotides play intermediate roles in the form of mRNA, rRNA, and tRNA, but they do this by mimicking the role of proteins, and they never do so except in
full concert with proteins. Quite simply, proteins are the boss, the king, master and commander of the biosphere.

As an analogy to help get our minds around this startling, no doubt offensive to some, inversion of thinking we can again imagine an ant colony. The ant colony is a complex, distinct entity that consists of large numbers of individual ant agents that interact with each other. The colony itself is an emergent phenomenon that has a timeline, a birth, growth, senescence and death, yet the individual ants don’t span the entire timeline. There is no individual leader who commands the trajectory of the colony from a master plan or blueprint. The colony itself emerges from the first principles of complexity and only as a function of time. This self-organized system is based on complexity and adapted for the purpose of producing yet more complexity. The life of the colony requires performance of many tasks and coordination of many factors and variables, yet no individual ant is in control. The queen of an ant colony is no leader in any real sense, and she issues no global orders to guide the overall colony. She is kept, harvested for the production of more ants. She is merely an efficient mechanism for ant production, a veritable ant factory in the grand scheme of things. Similarly, an ant colony can subjugate, cultivate and harvest other life forms, like fungus, or even other ants! The ant agents build physical shelters for the colony, and fabulous fungal gardens with nary a drawing, no foresight, nor any general clue as to what’s happening.

This is precisely the view we should adopt of living cells. They are protein colonies. They are complex protein computers of molecular information, and DNA serves merely as their primary storage device for processing complex molecular information. DNA serves at the pleasure of proteins. DNA is a kept molecule. Proteins use DNA to make more proteins, not just in terms of absolute numbers but also in terms of novel protein forms and combinations thereof. It is a complex entity from which more complex molecular information emerges through time. DNA is stored, manipulated, cultivated and harvested by protein. This self-organized system is based on complexity and adapted for the purpose of producing yet more complexity. The physical form of a DNA molecule serves as a structural template of the system, but its specific form is entirely determined by complex interactions within the protein colony. The entire drama plays itself out in time, space and molecular sequence. But protein holds the reigns even when it comes to deciding the novelty of future nucleic acid sequences. Protein insists on producing novelty in future DNA sequences. The intricate sequences now stored in DNA texts were primarily written by the complex interaction of proteins with proteins, as well as the interactions between proteins and nucleotide sequences.

Replication is an important part of life, but we rarely see it anywhere trudging down the long path of life. Protein has taken great pains to ensure that DNA can be replicated, but the goal in life is not replication. The end zone, according to Darwin, is descent with modification. If replication were the goal then cloning would be the norm, but sexual reproduction is the big winner when it comes to production of new life forms and future generations of entire populations. When you have a child you do not reproduce yourself you produce an entirely novel human. You are also comprised of seventy-five trillion cells that all started from a single cell. You can rightly be called a complex colony of cells, but the first cell did not reproduce itself seventy-five trillion times. Your cells have not been replicated because all of your cells are different. It is also statistically very unlikely that all of the DNA molecules in even a small percentage of your cells is the
exact same. But more important than differential replication of DNA is the fact that the basic molecular information in virtually all of your cells is totally different. That is because DNA does not even come close to defining the totality of molecular information in a living cell. You have cells in your body whose sole function is to create brand new sequences of DNA, containing only half the DNA in your original cell, and composed almost entirely of novel chromosomes compared to your first cell.

You, as a complex multicellular organism, do not send “your” cells to make a child, you create entirely new cells that are unique from your original cell, and this new cell must then combine with a wholly independent life form, preferably of the same species, before you “reproduce.” Does this sound like a system that’s trying to make exact replicas? For all the machinery that protein has at its disposal to replicate old DNA, it has twice that much machinery, strategies, and protocols for ensuring novelty in future DNA, as well as the novelty in the expression of its own DNA. Protein is not lazy and it has no intention of ever standing still in its relentless search of the biosphere. It has no qualms about boldly manipulating DNA to do its bidding in this task.

DNA is not just operated on by proteins it is coated with them to produce secondary codes in time, space and sequence. New codes emerge from the platforms of other codes. Proteins “read” DNA, and proteins “read” other proteins. These emergent secondary codes dictate the protocols of replication, recombination, storage, packing, unpacking, transcription, translation and even protein folding. Higher levels of molecular information emerge from this, and they determine what the colony will look like in terms of type, number, concentration, location, access, shape and behavior of the molecular agent population through time. This is the proper broad view of genetic information in the context of molecular information in all of its forms. It is not defined by any single form of molecular information or at any single level of complexity. It is never simple or linear, but complex and non-linear in its overall effect. Molecular strings are not absolutely controlling but utterly controlled, even to the extent of determining what the future order of them will be. The codon table should not be viewed as reflecting a single, simple, linear process in one direction but a complex, non-liner language with arrows pointing in both directions. DNA is a kept molecule, kept as an excellent physical means of storage, by which information can be efficiently stored, extracted, expanded, and further manipulated. Molecular information emerges in space through time from complexity not simplicity.

The biosphere clearly does not revolve around the relative simplicity of DNA; it revolves around the absolute complexity of protein populations. DNA is a sublime tool by which this system can be animated in time and space, both in the short and long terms. But if we dogmatically insist on using the central dogma as our guide, we will spend considerable time, effort and frustration in breaking the mental inertia needed to redirect our thoughts from the false direction in which dogma is clearly pointing us. The genetic code should not be defined as a protocol for DNA to make protein. But, unfortunately, the premise that the genetic code is a protocol for DNA to make protein is a false premise that is made in one form or another throughout the entire classic paradigm. The correct premise should be that there is a code that protein uses to use DNA to make more protein. Granted, RNA is an effective intermediate in the process, but it generally mimics protein, and at the same time it is totally dependent on protein to fulfill its role. The central dogma represents an intense focus on a single tree at the expense of a view of the entire
The thermodynamic hypothesis of protein folding is the final premise of the three fundamental axioms of simplicity propping-up the edifice of the fading classic paradigm. It is the worst of the three, I think, both in the way it was eagerly accepted and installed as the centerpiece of the paradigm, and in the impact that it has broadly had on our thinking. I obviously have no intention of hiding my contempt for this premise, despite any accolades it may have received or how cherished it apparently still is. It set the tone early for absolute simplification of genetic information. It was the license needed to proceed boldly down the path of simplicity. We now struggle upstream against the strong current of utter simplicity, and as we do, we will find the single target hypothesis in its headwaters.

The thermodynamic hypothesis of protein folding, as we saw earlier, is in fact two big hypotheses rolled into one. The first part - that thermodynamic forces play a dominant role in protein folding - cannot be in dispute. However, this is self-evident: thermodynamic forces will always play a significant role in the behavior of insentient molecules. But the second part - that all polypeptides have a single target of folding - has been convincingly disproved. In retrospect, the single target hypothesis appears to be a radical simplification with no chance of being correct. It was more a wish than an idea, and those many experts who still believe in it are merely engaged in wishful thinking. It is absurd on its face and strains credulity that a huge, complex molecule, such as a lengthy polypeptide, can find but a single, stable conformation in all instances. The statistical odds of this being true approach zero. More amazing than the firm belief in this axiom is the fact that nobody has apparently seen fit to even bother to prove this all-important hypothesis, yet disproof seems plentiful. It’s everywhere we look.

Dr. Anfinsen set about investigating the single target hypothesis by taking a small, folded protein and disturbing its environment. He then noted that the protein stopped functioning. When Anfinsen returned the protein to its undisturbed environment he noted that it started functioning again. This is a very interesting observation, it really is, but it is a long way from proving the single target hypothesis of protein folding. It is backward logic. To prove it with forward logic we would need to start with an unfolded protein and observe its many possible paths to a single, final, folded form. Of course this has never been done. It would be technically difficult to do now, but it was technically impossible to do then. Anfinsen merely proved that proteins are robust in their persistent shapes, as no doubt they are, but he did not prove that their shapes are absolutely predetermined entirely by their sequence. He is to be congratulated for his work, as he was, but perhaps we have over-congratulated him just a bit.

The single target hypothesis is today, as it always has been, an axiom of faith. No wonder, since the task at hand is to prove a negative, to prove that multiple stable conformations for any protein do not exist in the wild. It is analogous to proving the hypothesis that no blue cows exist. When proving a negative such as this it is not adequate to merely glance out the back window and note that no blue cows are out there. One must conduct a thorough, prospective search of the countryside for blue cows, or even for the potential existence of blue cows. And in the absence of any blue cows one
might still want to hold open the possibility that they exist but have yet to be found. Certainly when a blue cow, or in this case an entire herd of them, wanders in the back door one should sit up and take note.

The process of protein folding – under physiologic conditions – involves several different kinds of proteins. There are things, such as post-translation editing proteins, that play key roles. But more significantly there are folding “helper” proteins known as chaperones. These chaperone proteins certainly are odd ducks if in fact there is but a single way to thermodynamically fold a protein. What assistance could possibly be required in such a process? But folding clearly goes awry in the absence of chaperones and other editing or regulatory protein. Further disproof of the mythical single target – a big blue ox actually – is found in the protein form known as a prion. A prion is classically an oxymoronic definition of a “misshapen” protein. How can a thing with only one shape possibly be misshapen? Fittingly for this metaphor, it is a prion that is responsible for mad cow disease.

So we now conclusively know that proteins can assume more than a single, stable conformation under physiologic conditions. It matters very little how they got “misshapen.” The fact is, in the case of a prion they somehow consistently find at least a second, stable form. We have no way of knowing at this time just how many shapes any given protein might take, but we do know that there can be more than one, and that is all that matters. From this we should be eager to note that the entire paradigm of rigidly pre-determined protein shapes is thrust into crisis. It is no longer possible to axiomatically grant the single target hypothesis and all of the simple ideas and language it spawned. The defenders of the classic paradigm just cannot accept this, because if they do, their paradigm vanishes into the mist.

To counteract this kind of hard empiric evidence, the defenders of the paradigm like to invoke the qualifier, “physiologic conditions.” This has morphed into an undefined, ad hoc, self-protective, contingency clause that seems to have no utility other than to serve as ideological protection for the classic paradigm. It is yet another Easter Bunny. It is a garbage bin term used to dump all the trash that offends the classic paradigm - out of sight out of mind. I have even heard “physiologic conditions” invoked in the sense that “physiologic conditions” means anything that happens in a cell that is not related to the primary sequence of a protein. Isn’t this type of reasoning just a bit childish? Surely this definition, or argument, if you can call it that, cannot be considered “scientific” in any meaningful way. Certainly if we are going to use the single target axiom and its protective “physiologic conditions” companion as a basis of thought we must rigidly define it in a more useful way.

To begin moving in a more productive direction I propose a thought experiment. Let us suppose that we are able to construct a machine that will reliably produce sequences of well-defined peptide bonds however we wish, and it will use the standard twenty amino acids in the genetic code. Let us suppose that this machine, call it the peptomatic, operates in what reasonably can be called physiologic conditions, but it operates in the absence of the other molecular machinery found in a cell. How would this machine operate, and what would be its output? I would gladly assume that when all other variables are held constant, the peptomatic would reliably produce the same protein at all times. The polypeptides will be thermodynamically folded and thermodynamically stable in this folded form. But even this is a dangerous assumption because it is possible that
early production of a single protein could influence later production of the same protein. In other words, it is logically possible that the presence of individual proteins could provide “information” to their cloned brothers and impact on their own shape as they are formed, and we can no longer take for granted that they do not. We might also suspect that some polypeptides might just randomly fold “wrong” because they are complex and it is a potentially chaotic process.

Even if we clear this relatively trivial hurdle, there are still very many basic variables that must be considered. For instance, we can imagine that the peptomatic is able to put the amino acids together in their standard sequence but from end to start. They are in proper bond order and direction but merely out of sequence with respect to time. Would these proteins still be the same shape? I doubt that anyone believes that they would, but protestors quickly shout, “those are clearly not physiologic conditions!” Well, we didn’t change the primary sequence, except in time, but, okay, so now we must apparently add some sort of time element to our definition of physiologic conditions. But how important is the time element? Could we merely change the rates at which each bond is made and then produce different results? What if all the bonds are made lightning fast versus glacially slow? What if we alternate individual bonds in time or whole sections of bonds? How does this change our expectations, and where do we pin this on physiologic conditions?

What about the precise shape of each bond as it is made? If we could set a dial on the machine to produce only cis-peptide bonds would the proteins still look the same? What if one in two, or one in ten, or one in two hundred bonds are cis bonds, would that change the output? Could we knit one pearl two and knit two pearl one and get the same protein in both instances? I think everyone would agree that it would change the output. What if every bond twisted slightly to the left, or to the right, or alternated from one to the other, or in long stretches that shifted from right to left, would the proteins all still ultimately become the same shape in the end? These elements of shape, in this case put into the string at the time of bond formation and not some mystical time later, would certainly not rigidly define the ultimate shape of the protein, but they would serve as pointers in space for the entire structure to interact within itself and assume larger structures within the assumed background of thermodynamic activity. It would be a non-linear process, but not one that anyone should assume has but a single possible outcome.

Once we get done fiddling with these basic parameters of peptide bonds - time and space - we can start to imagine that we could then add additional elements or cofactors into the machine environment. We could add a huge number of different proteins that would interact with the polypeptides in real time as they are formed by the peptomatic. This would surely put “information” into the proteins in many forms, and it might actually change the profiles of the entire protein population that the machine creates. No single protein should ever be considered entirely in isolation. It is only within the broad context of an entire complex population of all proteins can the “meaning” or the information content of any protein be considered. The additional proteins in our “physiologic conditions” would begin folding the proteins with the creation of the first bonds; not after all of the bonds are made. These extra, added proteins could mold, twist and even edit the length and quality of the proteins as the peptomatic produces them. In other words, primary sequence doesn’t really even determine primary sequence, let alone tertiary structure. This has been empirically demonstrated in living cells. Surely these added
proteins would change what we perceive to be the output of the machine. Is this what we mean by physiologic conditions? That’s what I mean by molecular information of the complex, emergent, non-linear variety.

To my knowledge, such a machine does not currently exist, but in theory it could. Still, speculation about the peptomatic is far more enlightening than the current forms of speculation of pure semantics about the proper meaning of “physiologic conditions.” Life has assuredly built this type of machine, and man has now undoubtedly found a way to partially hijack it into doing his bidding. We use living machines to “make” proteins today, because we cannot do it without them. Our peptomatic would not make proteins like those we get from living cells. We also must now realize that we use the output of living molecular machinery to study proteins. But merely studying the output of protein synthesis is not the same as studying the mechanisms and rules of protein synthesis, and so we are probably not able to draw any absolute conclusions about those rules. It is an understanding of just those rules for making protein that we seek, not by simply defining them but by observing, quantifying and perhaps even controlling them.

At least now we should feel free to begin to imagine that protein synthesis is not really a physically rigid, linear process analogous to a ball rolling down an inclined plane. Perhaps protein folding is a complex, even a chaotic yet non-random process more analogous to the formation of weather patterns, and perhaps there are a handful of butterflies flapping their wings – under physiologic conditions. It seems more likely than not that protein synthesis and folding is in fact a complex and chaotic process with several potential outcomes rather than a simplistic linear process with but a single possible outcome.

More significantly, all of the basic questions that were mooted for decades by the single target hypothesis must now be asked. How many targets of folding are there for any given polypeptide? One? Two? Ten? Ten thousand? A billion? If more than one target of protein folding can exist under physiologic conditions, by what mechanism is each target either found or suppressed? What strategies are available to life to build or dismantle the mechanisms of hitting various targets? What cost and what benefit do the different strategies carry? What complex, self-organized, adaptive molecular language will emerge from the strategies that are chosen? Quite simply, the lid on Pandora’s black box of protein synthesis must come off because Schrödinger’s cat is out of the bag. (How’s that for a wildly mixed metaphor.)

The information content of a protein cannot and should never be collapsed to its sequence of amino acids, but that is the classic paradigm. The axiom that primary sequence determines tertiary structure must be rejected. The primary sequence of amino acids is clearly not the sole determinant of its tertiary structure, but it is, no question about it, a huge part of how a protein will eventually fold. The axiom should properly read: primary sequence strongly influences tertiary structure. The information content of a protein should at all times be seen as its sequence of peptide bonds, and our view of the molecular language of proteins should start but not end there.

The single target hypothesis was adopted virtually without empiric data, and no definitive study, that I can find, exists today to demonstrate its veracity. It is a classic case of the dog that didn’t bark. It was a position taken to simplify and therefore accelerate investigation, which it did. But in the absence of empiric proof it should now be rejected. Unfortunately, rather than rejection, today it is still used as a starting point for teaching
and thinking about genetic translation. I’ve been told a thousand times that it’s been proven “a thousand times,” but I’ve yet to see a single study. I agree that proteins are robust thermodynamic warriors, but not nearly as hardy as we like to think. Either a study must be performed, or the hypothesis must be formally rejected - and let’s move forward.

In truth, the single target hypothesis was a simple choice between entropy and negentropy. If one chooses entropy in this process, then the single target hypothesis makes perfect sense. The information stored in DNA is in its most expanded form, and the only thing the genetic code can do is shepherd it through to its inevitable thermodynamic conclusion. However, if one chooses negentropy as the guiding principle of protein synthesis, then one can expect complex yet tight control of molecular information and behavior through time, and perhaps even information expansion somewhere along the way. Entropy is simple and negentropy is complex in this context. The classic paradigm is founded on a belief in entropy as the guiding force of molecular information, and the single target hypothesis is a grand thumb of the nose to Schrödinger and his prescient observations about life and how we should expect to find it behaving.

Today the single target hypothesis has formed a simple and useless tautology in thinking: since a protein can only fold one way the genetic code is one-dimensional, and since the genetic code is one-dimensional a protein can only fold in one way. All of the basic questions that were never properly asked must now be asked. How many dimensions of information are translated by the genetic code? What are they? What is the proper relationship between the dimensions of information in the genetic code? What are the results of translating these various dimensions with respect to protein synthesis and the overall functioning of a living cell? What, precisely, is the genetic code?

The three fundamental premises behind the classic paradigm of the genetic code have conspired to give us a false context for thinking and talking about gene translation. They represent three mutually supporting statements of simplicity with respect to molecular information. They represent a circular definition that is not based in reality but yet all of them are sustained by the ideological support they give to each other. What’s more, the additional language and ideas that these three erroneous yet fundamental premises in the classic paradigm have spawned must be examined and revised. It is simply not acceptable to speak of a “non-synonymous synonymous codon,” a term most assuredly contributed to this party by the Mad Hatter himself. If we think of theories as mental factories of language and ideas, this factory must be closed down, stripped out and retooled completely.

The natural process of protein synthesis that is complex and multi-dimensional can no longer be described as simple, linear and one-dimensional. The process is not black and white, flat and lifeless; it is intricately shaped, complex and operating in Technicolor, and so the machine language of these fabulous molecular machines must be as well. The paradigm must come alive again. The problems with the classic paradigm can be traced to many sources, but the origin of its fundamental problems begins with the single target hypothesis of protein folding, the total embrace of entropy within the paradigm, and the extension of this idea to its two derived companions, co-linearity and the central dogma. The shocker that most people don’t quite seem to realize is that these three premises don’t just describe and support the genetic code in the classic paradigm; they are the
The genetic code in the classic paradigm. These premises define the essential steps that make it so that a flat codon table can stand alone as “the genetic code” in our hearts, minds and textbooks. The three premises spell out the protein recipe within a cell, and they simply use the codon table as a list of ingredients. But the codon table is not a code! It is merely data to be used by a code. It is a look-up table within a code. The actual code is the simple molecular algorithm described by these three premises.

The defenders of the paradigm apparently don’t even understand the basic structure of a code, or recognize a need for this, and so they can’t accept and mentally process these statements of simple fact. They say, “well, of course, I know it’s not a code, but it is the genetic code, and everybody knows that. Why are you making such a fuss about semantics?” And so they have all ended up speaking gibberish from the bottom of a rabbit hole. The problem they have is that the classic paradigm does nothing but consistently confuse data with algorithms. A code in this specific context is an algorithm, a set of steps for reliably doing something, a method of decoding. This is not semantics; this is reality. The codon table is not a set of steps but a set of data used by the steps. The actual steps of the genetic code are not spelled out by the codon table but essentially by the three premises of the classic paradigm. There is absolutely nothing wrong with the data. But the steps that we take as true in the classic paradigm of the genetic code are all three demonstrably false, yet the paradigm marches proudly on with nary a scratch. It is time for a new paradigm.

Besides the total disintegration of the three pillars beneath the classic paradigm, what else went wrong? Here again, there are many things that have changed, so there is much to consider. The classic paradigm is metaphorically linked to computers, and rightly so. Unfortunately, the referent for the metaphor, information theory and digital technology, has radically evolved, yet the classic paradigm has not. It is now a hollow metaphor with nothing but an anemic referent from the distant past of digital technology.

Computers have changed dramatically over the past half century, and our perception of them is no-longer dominated by thoughts of monolithic chunks of iron loaded with spools of linear tape and spitting out streams of blinking lights. Programming a computer is no longer a rigid, top-down, sequential process of logic. It is a decentralized process of nonlinear logic that is determined by the multifaceted properties of coded objects, the inheritance of those objects, and the effects of multiple iterations of logic when they are applied to the central core of a problem. Furthermore, a single computer can no longer stand as the icon of computer technology. In today’s highly networked world only a nuanced understanding of networks and how they operate will serve as the proper referent for a suitable metaphor of genetic information. The earliest digital computers, it turns out, were a grossly premature referent for the classic paradigm of the genetic code. They are digital dinosaurs. All computers have been upgraded since then, and now it’s time to upgrade the computer metaphors of biology and the codes of life.
Along with the revolution in digital technology has come a raft of powerful new tools, both physical and intellectual. Our powers of resolution when it comes to directly observing things have been magnified many-fold. Not only can we see smaller and smaller objects and processes with more and more precision, we can observe them in a more natural state, without total disruption of the thing or process under observation. This, of course, produces more and more data, for which we now have fabulous relational databases in which to store and access this data. The precise sequences of entire genomes are now commonplace and the intricate details of protein shapes are plain to see as well. The molecular machinery of transcription and translation is remarkably detailed, even when displayed in the most basic biology texts. The texts themselves are now digitized and accessible virtually anywhere in the world at the pain of only a few keystrokes. Discussion of the information within the texts can be conducted at light speed with virtually anybody anywhere on the planet. Clearly, information and ideas are forming and evolving at an ever-accelerating pace.

The new intellectual tools of this era are no less robust and no less impressive than the new physical tools. The guiding philosophy of better science through simplification is not only passé it is rejected. The related concepts of complexity, emergence and chaos now hold sway. It is not enough to isolate a phenomenon and study its parts. The whole is no longer perceived as the sum of the parts – more is not more. More is different. There are movements afoot to describe all universal processes within the framework of information theory, complexity and the emergence of all things. Certainly the pinnacle of earthly complexity - life and all of its intricate processes - cannot be exempted from this trend. Certainly the popular icon of life’s codes, the genetic code, cannot be exempted either.

Perhaps the latest and greatest intellectual tool is a rediscovery of sorts, one where a full appreciation of Darwin and his proposed mechanism of evolution have risen to a new intellectual prominence. It is not so much a primacy of evolution; it is a primacy of algorithms. Since all computer programs are algorithms, it is no great surprise then that folks today have a different perspective on the power and utility of algorithms, and natural selection is now seen as a basic algorithm for making more complex algorithms. We are witness to a new modern synthesis of computers and biology, and it is mind boggling that so much form, function and complexity can be at bottom the result of algorithms. However, it is no-longer appropriate to merely discuss the effects of natural selection, it is now necessary to include the contributions to natural selection made by all of life’s intricate, interrelated processes, and this most certainly should include the genetic code. It should now be clear that this broad context of self-organized complexity is virtually ignored by the classic paradigm. It is time for a new paradigm.
The genetic code is a code of protein synthesis. It should rightly be called the protein code. The protein code represents the logical relationship between nucleotides and proteins. It is a relationship based on a symmetrical structure between these sets, and it is a structural code that logically relates molecular structures to each other. It is poorly understood today, and will go so far as to say it is misunderstood. Here is a definition that can preserve most of past definitions, but still allow us to include the new things that we know are true:

The protein code is an incompletely understood natural phenomenon that once was called the genetic code. It is a molecular system of logic that relates nucleotides to proteins. Proteins are produced from information stored in DNA, and novel proteins are anticipated by symmetrical relationships within this system. These two functions are achieved through a complex series of relationships between four distinct sets of molecular sequences: DNA, mRNA, tRNA and protein.

In my opinion, this definition is adequate, but it probably does not go quite far enough toward dispelling the false beliefs and disinformation that currently dominates all discussions of this important code. Biomolecules are tiny crystals and this is a crystal code for making more crystals. The majority of the information is dependent on the code itself to represent any information value at all. The code does not exist in any single location or molecule, it is a collective phenomenon existing only in a complex collection of molecules in time and space

The basic structure of the crystals of life is formed by the tetrahedral system of water and carbon, but many more complex forms are built up within this basic tetrahedral matrix. A tetrahedron – much more than a cube - is an excellent choice for building complex, dynamic shapes. Carbon is the foundation of life as a crystal and it proves itself facile at forming more complex, basic organic shapes. For instance, the carbon atoms in nucleotides that make up the DNA double helix create a larger structure that is actually a remarkable model for a sequence of dodecahedrons. The double helix mimics a string of, or a concatenation of dodecahedrons. It is made of roughly ten dual pentagonal molecules facing opposite directions per complete rotation, so it takes very little imagination to see a string of perfect solids within the immortal coil itself. Of course, we already knew that tetrahedrons could be used to make a dodecahedron, but this is a complex horse of a different color. The dodecahedron is a platonic solid that trumps the tetrahedron in geometric complexity, yet at heart it is merely a complex composite of many tetrahedrons. The vertices of a single tetrahedron can be matched one hundred and twenty different ways within the vertices of a single dodecahedron – lots of crystal choices. If a simple tetrahedron is seen as spatial “information” then this same information is “compressed” to quite a degree within a more geometrically complex dodecahedron. One dodecahedron can literally stand for 120 tetrahedrons, so the dodecahedron is rich with tetrahedral information. It is an embarrassment of spatial riches, really. What’s more, there is a natural, platonic system of translation built into the relationship between these
two shapes, as demonstrated by Code World, and so the dodecahedron would appear to be an excellent choice of shape for crystal information encryption, storage and translation. In fact, there is seemingly no better choice of a shape than a dodecahedron to store and use tetrahedral building information. As we will see shortly, nature has coincidentally used the dodecahedron to embed the logic of sequence symmetry into the translation machinery as well. The character sets of this molecular language are curiously built, both literally and figuratively, around many logical facets of a dodecahedron. Plato was right, as far as life is concerned, the dodecahedron is the cosmos. It is seemingly everywhere and all things to the logic structure of living molecules.

Likewise, the tetrahedron is a seemingly good choice of basic molecular shape within the dynamic matrix of water for generating larger dynamic crystal morphologies. We might go so far as to say that the tetrahedron is the perfect choice for the basic building block in an organic crystal erector set for both of the basic crystal functions of life, search and build. We are therefore quite happy to find the tetrahedral structure of carbon being passed on to the higher-level tetrahedral structure of amino acids, which then get passed along to the higher-level complex structures of proteins. This being the case, we simply must ask: what about the protein code itself? Has nature actually used a platonic spatial language as the logical basis of this code? Are there, or were there units of shape and or geometric logic in the machine language of the primitive crystal code? We simply do not yet know, but we shouldn’t hesitate to ask. It sure looks like a safe bet that there are, but nobody seems very interested today in actually trying to find out. The question has heretofore lain hidden by the oppressive blanket of the classic paradigm.

The protein code then appears to generally be a relationship between two broad but distinct sets of crystals, one set is fundamentally dodecahedral and the other set is fundamentally tetrahedral. These two basic crystal types are co-dependent on each other for their existence in every respect. Their relationship is complex not simple, and their raison d’être seems to entail the ever-present goal of making more crystals, not just in terms of absolute numbers, but more so in terms of exploring different varieties of individual crystals and different complex sets of inter-related crystals. So the relationship between these two crystal types should be expected to reflect an inherent need and an inherent ability to efficiently generate novel crystal forms – and so it does.

We should rightly perceive that the polypeptide code was founded not on a single shape but on the complex interaction of two shapes, the dodecahedron and the tetrahedron. The code itself does not exist only within the molecules of either shape but within the complex rules of interaction between the two basic shapes of these molecules. The exact relationship is still a mystery, but it is somehow achieved through a series of relationships between distinct sets of sequential macromolecules, or crystal strings. The relationship involves molecular identities, as well as complex molecular timing, quantities, and exquisitely specific shapes. There are four obvious sets of strings within the protein code: DNA, mRNA, tRNA and protein. Each set has a relationship to the other sets; in other words, there are codes of behavior within and between each set. We do not presently know all the details of these relationships, primarily because we do not know all of the types or meanings of the molecular information that is passed back and forth between these sets. We do know that the net result of the overall relationship is a high
degree of productivity and endless novelty within both crystal types. In other words, the relationship seems to work really well for its apparent purpose. The protein code is a crystal code used by crystal computers. I think we can safely call this a crystal code because if this is not a code or a language spoken between crystals, then the term ‘code’ can have no meaning whatsoever. Crystal information is being efficiently created, stored and communicated by this crystal information system. It has been all too tempting to extract a tiny subset of the code – the codon table – and argue that, “of course it is not really a code but merely a substitution cipher.” This misses the point entirely. It is no code at all, substitution cipher or otherwise. It is merely data used by an actual code. The codon table is not the protein code; it is an extreme simplification of a more complex, more powerful language entirely. The table is merely our way of storing the data, but the code is the precise set of steps used to extract the data. Furthermore, the codon table is a miserable way even for humans to represent this data in the first place. There are better ways. So, it is a huge mistake to confuse the codon table with a more robust molecular language. It is, unfortunately, a very common, an almost compulsory and universal mistake made today, and one that the classic paradigm insists that we keep making.

The protein code is a true language of crystals, a crystal code. It is an ancient, sublime system of complex crystal generation, a self-organized, adaptive system of complex molecular auto-assembly. What’s more, it is an example of molecular negentropy not entropy. It appears that the classic paradigm not only fails to answer Schrödinger but utterly fails to even understand his primary question. They have given him the answer he said he would not accept. But as a code, the protein code cleverly combines the dual function of an operating system and a search engine. The data is not the logic, but the pattern of some of its logic can actually be seen in this particular data! This clever code builds intricate crystal forms, and it continually finds new ones within the vast space of all possible crystal forms. It forever sifts the randomness of the universe for nuggets of crystal utility. Crystal information is stored in a deceptively simple dodecahedral format, but it is somehow translated into a larger, more dynamic and complex tetrahedral world. This is perhaps an example of steganography, which is a cryptographic trick where the actual message is obscured by its physical format.

We know that this crystal code is at bottom a language used between atoms and molecules, so at some level it is a language of shapes and atomic structures. It is first and foremost a language of molecular structure not molecular composition. We also know that ultimately the product of translation is nothing but a shape. We simply do not know at this time how the language works or why it works in the way it does because we do not know all of the factors that contribute to the translation of these remarkable crystal strings. We do not know which forms of string information can and can’t participate in the intricate crystal dance and therefore which ones do and don’t play a significant role in this vital language.

We flat-out do not presently know the crystal recipe for making a protein; otherwise, we’d be doing it – making proteins that is – and we clearly are not making proteins. We have not invented a peptomatic that has been put through its paces and empirically demonstrated to work on basic principles of molecular information. Remarkably, an expert once told me that we had invented and are actually selling this amazing device – we clever, clever humans. It was his way to quickly convince me that I was clearly
wrong, but he later told me “never mind.” It is logical to think that if the classic paradigm is correct that we should actually have such a device, but we do not. If we have in fact invented this miraculous peptonatic, it is a well kept secret. We should be on national TV demonstrating our mastery over nature by designing and building new proteins with Microsoft Protein™ or some such thing, but we aren’t. When I try to point this out, I am constantly reminded that every scientific theory has its little exceptions to the general rule. Well, this one seems to be 100% exception and no useful rule. I challenge anyone to find anywhere any piece of empirical evidence that supports the classic paradigm. There is not a single documented case that I am aware of that shows the classic paradigm in action. We are today without a single example of de novo protein synthesis, or even reliable folding prediction. We are no doubt quite good at letting a living cell make a protein for us, but that’s not exactly the same as de novo synthesis. Studying these proteins and relating them to DNA sequences does not crack the code of primary sequence folding – I’m oh so sorry to say. Just because a cell can do it does not mean that we know how the cell does it. There should be a limit to our patience, but to date there appears to be no limit to our patience with the classic paradigm and its many obvious failures. Therefore, the question is still a wide open one: what precisely is the code that life uses to make proteins in a living cell?
XII. What are the broad patterns of life?

They are broad and they are patterned.

Symmetry.
Information
Complexity
Catalysis
Structure

Note that the classic response to Schrödinger’s recognition of the obvious negentropy in life is to simply say that the earth is not a “closed system” and it therefore receives “energy” from the sun. This hardly qualifies as a convincing explanation. It is actually a horrible explanation when you think about it, one that completely ignores what Schrödinger actually said. It is just a fancy way of saying that the answer can’t really be found here on earth, and the whole system cannot be understood unless it is placed into a more proper broad context. So apparently the proper context for life on earth is the rest of the universe, or at least the sun, which is okay in as far as it goes. But it is a self-evident concept in the pagan view of things, so much a part of a globally equilibrating universe headed in the general direction of more self-organization. However, in the classic scheme it seems to be a non-answer, a backwards answer actually, or at least it is such a bad almost-answer that it could never really even be wrong. It barely shoves aside the larger question of how any undifferentiated, disorganized system, open or closed, large or small, could begin a journey of self-organization of such unimaginable complexity and scale as we can so clearly witness in the pattern of life on earth.

The earth has accumulated a phenomenal amount of negentropy through time not entropy. It must come from somewhere and there must be a fundamental explanation for why it is here. Regardless of how much energy the earth receives from the sun, how does that energy find a way to seemingly disobey the law of increasing entropy in the larger pattern of things? It seems as though the classic pattern should expect increasing entropy with increasing time and pattern size, yet the trend on earth is clearly the opposite. However, in the pagan universe the grand pattern of life on earth is not seen as an exception to a larger rule; not something to be explained away but something through time to be expected. Life is not just a reflection of limited properties of bizarre crystals on earth but a fundamental property of the universe itself.

The issue at hand here is negentropy not entropy, but the sun is a clear source of entropy. We cannot therefore propose the sun as our primary source of negentropy. I am sure that the explanation that the sun provides energy and therefore the earth is an “open
system” never would have satisfied Schrödinger. It is exactly the explanation he said that he didn’t want to hear. He was well-aware of the sun, I’m sure, but he was looking for a more detailed explanation of the source of the phenomenal amount of negentropy that is displayed in living systems. He made a clear distinction between energy, material exchange and negentropy, which we can view by splicing together just a few sentences from his book:

“That the exchange of material should be the essential thing is absurd. (Exchange of energy) … needless to say, taken literally, this is just as absurd.”

“The essential thing in metabolism is that the organism succeeds in freeing itself from all the entropy it cannot help producing while alive.”

The question that Schrödinger was asking is not even slightly answered by the “open system” explanation of life with respect to the second law. He knew that life was a system that as a brand new quantum physicist he would not recognize, but still he rejected the notion that it would actually fall outside the realm of physics. He knew that his understanding of physics must, therefore, be incomplete, because he could not use his view of physics even in a cursory way to explain life on this planet. He knew that his understanding as a physicist was perhaps backward, or at least inadequate to deal with the bizarre reality of life as he then understood it:

“How can the events in space and time which take place within the spatial boundary of a living organism be accounted for by physics and chemistry?”

“To put it crudely, an engineer, familiar with heat engines only, will, after inspecting the construction of an electric motor, be prepared to find it working along principles which he does not yet understand.”

“The orderliness encountered in the unfolding of life springs from a different source. It appears that there are two different ‘mechanisms’ by which orderly events can be produced: the ‘statistical mechanism’ which produces ‘order from disorder’ and the new one, producing ‘order from order’.”
“We must therefore not be discouraged by the difficulty of interpreting life by the ordinary laws of physics. For that is just what is to be expected from the knowledge we have gained of the structure of living matter. We must be prepared to find a new type of physical law prevailing in it. Or are we to term it a non-physical, not to say a super-physical, law?

No. I do not think that. For the new principle that is involved is a genuinely physical one: it is, in my opinion, nothing else than the principle of quantum theory over again.”

Schrödinger knew about the laws of thermodynamics, the workings of quantum mechanics and the existence of a power source from the sun, but he still felt that there was something missing in the picture. He had no doubt that global chemical energy exchange equations would ultimately be shown to add up properly, but he literally said that this was not the answer he wanted and that life actually “feeds” on negative entropy, “attracting a stream of negative entropy upon itself.” He was saying that our understanding of the random heat motion of molecules in the universe was not enough to explain the logic of living systems. Merely noting that the sun provides the living system on earth with a limitless supply of energy would not have swayed him in the least, I don’t think. In fact, he would surely see the sun as a source of entropy not negentropy. The sun increases the heat motion of molecules on earth, so they become more random not less. In my worldview, the sun is simply a power cord for earth’s crystallizing computer. If we plug a power cord into an undifferentiated black box of silicon we do not expect the box to become more organized and differentiated, to ever spontaneously grow a CPU. We simply expect it to get warmer, to become a less differentiated black box, and the longer we wait the less differentiated we should expect it to get. No, the sun does not drive the complex logic of life; it merely charges the battery packs that living cells need to perform their essential logic functions. The sun provides nary a hint about the logic structures that will eventually self-assemble within earth’s complex crystal computer.

Schrödinger also saw a global pattern that he tried to connect in the dots between bio-molecules, human culture and human consciousness. He was plotting a path in his thinking from the bottom to the top of the pyramid and trying to find the common thread or the driving force. He wrote essays on consciousness that equated consciousness with novel environmental stimuli. Something that is old hat can be moved to the subconscious, he argued, but when we encounter new stimuli we are forced to deal with it and practice it until it no longer requires our conscious attention. He drew an analogy
between this and organic molecular ontogeny, and I’m sure he would not have objected to the idea that this same analogy could extend to the beginning of molecular phylogeny. He was essentially saying that molecules could somehow “learn” just like we do, and somehow they store that learned behavior. To him, molecules, embryos and human minds must display a pattern of learning that somehow must be stored in the substance of their living matter. The patterns must then somehow overlap.

How crazy is this idea? Let’s see if we can find out. Imagine a photon as a billiard ball at the surface of the sun. Imagine this photon leaving the sun at time A and traveling eight minutes in a straight line to the surface of your eyeball at time B. Now imagine the earth at the time of its formation, again time A, and then imagine the earth as it is now at time B, covered with a thick coat of tenaciously complex life. The earth traveled in a straight line through the universe from A to B just like the photon. Now try, if you somehow can, to imagine that they are the same thing. In other words, replace the history of the photon in your mind with the history of the earth. Further imagine that the history of the entire universe is again the same thing. All of these things are physical patterns with a logical path on a definite timeline. What kind of universe would this have to be where all of these things might somehow be different versions of fundamentally the exact same thing? How could we imagine that this universe is built in just such a way that the eight minute path and pattern of a simple tiny photon can be explained by the same basic set of rules as the four or five billion year path and pattern of the gigantic earth? If the laws of the universe must somehow account for both, what could these laws possibly be, where will we find them, and what kind of universe could contain them?

The answer is remarkably simple and obvious. The question and the answer are one and the same. Everything that you just imagined took place merely in your brain, which just translated this specific, bizarre illusion, as directed by me, from the printed page of my universe into yours. The “universe” is indeed an elaborate illusion and the only thing that we can “know” about it is a product of whatever models, paradigms or worldviews that our brains can create and show to us. The entire universe, as far as you are concerned, exists between your ears. The “universe” is whatever simulation your universe, in the form of a human brain, can muster up and show to you. However, now we can start the same process of questions all over again and wonder about what this new universal simulation of a universe is actually made of. Of course, I am still lightly tethered to reality and I firmly believe that there is a “real” universe “out there,” but you can never actually “know” anything about it. This is not just useless philosophy this is a useful physical fact of logical reality. Perhaps one can merely take a position similar to Descartes, “cogito ergo sum,” and quit there, but this is not exactly an explanation of how it actually works. It has always rung a wee bit hollow to me. Should we quit and go
home with that philosophical zinger? However, in today’s computer saturated culture we can easily translate Descartes as follows: We can only “know” that the universe is able to build a machine that can simulate the universe. This is good in as far as it goes, but I want to know how that actually happens because I “believe” that it actually did happen. The absolute knowledge that we think is real and useful. There logically must be a real context for that knowledge, and this new knowledge is also real and useful. How is the universe constructed, in what particular way is it built that it can build a machine capable of creating a useful and fairly reliable simulation of itself? Common sense suggests that the universe and any simulation of it must share some fundamental properties of logic and construction. How could this be even remotely possible?

The obvious place for us to now look for these answers is in other machines capable of simulations, and so computers are the most obvious place to start. In this light, perhaps, computers as a whole should be seen as fundamental discoveries of universal behavior and not merely human inventions of convenience. Perhaps humans have discovered the fundamental laws of the universe, in some rudimentary form, and computers somehow model them. Perhaps somehow the fundamental laws of computing reflect the fundamental laws of the universe. This is truly the deepest part of the rabbit hole, but once you go there, it is hard to see how there is really anyplace else left to go. In my mind there seems to be no close second. If not here, where else might we expect to find an explanation for how the universe is constructed? Geez, Louise, are we really completely stuck in our own brains? Is there some way to use this knowledge that our brains actually simulate the universe to help get us logically unstuck?

Look at your computer as it functions beautifully in front of your very eyes, and ask yourself what you see. Whatever it is, it is an illusion, I can assure you. Your computer is nothing but cleverly crafted human artifact to provide you with illusions of complex behavior - because at bottom nothing is happening inside your computer but a happy shuffle of just two digits – zero and one. To your computer there is no “reality” because there is nothing except a zero and a one. Everything else is merely an elaborate illusion created for you by your computer and the clever engineers that built it. If you have ever wondered how one thing can literally become another thing, then study your computer closely because it is a case were a set of zeros and ones can become anything we want them to become. When a set of atoms becomes a rock or a tree is the universe performing essentially the same kind of logical process as a computer?

The movie Toy Story is the first full length feature motion picture to employ digital animation entirely. It is essentially a computer simulation of trillions of “photons” logically behaving in an imaginary digital universe. It is awe inspiring for its degree of
realism. At bottom it was achieved by human imagination combined with the precise logic of a purely informative universe. Why should we imagine that this excellent simulation of photons is not somehow constructed on fundamentally the same kind of system as the real universe of photons? What part of the real universe could not be simulated in a similar way?

Try this thought experiment. Imagine that a clever physicist will come up with a “perfect” computer simulation of the universe tomorrow. It’s not so terribly hard to imagine today. Suppose this fabulous future simulation is capable of simulating virtually any natural phenomenon. Heck, just try to imagine any idea, formula, behavior or “natural law” that you could not today crudely simulate on, or somehow program into a computer. Imagine any known “law” or “idea” that we cannot busily get into a computer code somehow, right now. In this light, it is not at all hard to imagine that all of our computer models will only continue to get more and more accurate at an accelerating pace, so where will it all end? Ask yourself: in this, the world’s best possible future computer simulation of the universe, will we need to have one set of rules that leads to the pattern of a photon and a completely different set of rules that leads to the pattern of the earth, or will they share the same basic set of rules? More obvious, how will the rules for the earth work if not somehow based on Darwin’s clever observations? So what about the photon? Could the same observations somehow explain it, too? Now ask yourself what you expect to find at bottom in any simulation of the universe, brain-based or computer. It will surely be pure information and cold, hard logic, and probably in a form that somehow can be cleverly reduced to nothing but zeros and ones. That is the nature of information. Given this deflating, deterministic reality, how can we imagine a universe that at bottom is not fundamentally an informative universe? How could we ever imagine that the total pattern of information in this universe isn’t increasing through time? As Schrödinger said, it will be quantum mechanics all over again. How else could it be?

I suppose it takes a certain kind of bizarre person, but the conclusions are logically unavoidable in my mind. It is obvious. No? We simulate the universe with our brains and our brains are somehow an inevitable product of the universe. Our brains have then somehow gone on to build “machines” that are merely prosthetics for our brains and they oddly can also somehow simulate the universe. Did I miss a day in Sunday school? Was this predicted long ago in the Bible or anywhere else in science, philosophy or even in fiction for that matter? There must be some fundamental logic and basic construction to the whole system of the universe that can account for this bizarre fractal reality of a universe creating a simulation of the universe creating a simulation of the universe.
As humans we do not merely imagine the magic of computer technology. It is quite real, and it is quite obviously an accelerating reality. We have had computers for what, fifty years? From a universal perspective this is nothing. Where did these magical things start and where are they now? How could anything that remarkable and distinct within the overall pattern appear that quickly and advance that rapidly in a system otherwise measuring small steps of “progress” in units of tens of thousands of years? What can computers not now do? Why could they not do anything that we or they are somehow able to imagine them doing? Furthermore, I don’t care a wit specifically how particles collide, I do care how we can “know” that they will collide, and how we can “know” they will collide in a predictable way. That is truly bizarre. How is the universe constructed so that this “knowledge” is even possible, let alone necessary? Think about it. After roughly 14 billion short years of evolution, certain parts of the universe are anticipating the future behavior of other parts of the universe. If this is not a self-organizing pattern what is?

In my book, and this is my book, life needs to be seen as the product of some fundamental force in the universe that is acting on the universe through time to produce not just the self-organized patterns of life but the self-organized patterns of everything. This book, however, is not all about the universe, philosophy, quantum physics, or even chemistry. This book is about a new paradigm of the genetic code. My paradigm of genetic information says that negentropy is the measuring stick for all of life, and it is divided into units of complexity. I am also placing that paradigm in a larger context by proposing a truly hardcore worldview based on the idea that we are the product of an informative universe perpetually running complex codes that lead to more complexity. This is, I believe, the proper context for contemplating the codes of life. Anything less falls comically short in my imagination. If you can’t explain the appearance and obvious progression of the general pattern at this basic level, then something is lacking in the explanation. Furthermore, it should now be crystal clear that I believe that any paradigm will always reflect the belief systems, individual desires and worldviews that went into its construction, no matter how “scientific” it purports to be. I don’t give a wit for science if it can’t explain the proposed ideas in full. If I can’t understand the logic of an explanation, I simply don’t believe the explanation. I’m not stupid.

No scientific paradigm today can be expected to explain everything, but it should be internally and externally consistent, at the very least, and we certainly should never become wedded to the paradigm at the expense of repeated and contradictory observations. No paradigm should force us to accept ridiculous explanations of clearly observable phenomena. The classic paradigm of the genetic code is very bad for all of these reasons. It is defined in such a way that its scope is brutally limited, and it asks no
questions but merely begs questions to not be asked. ‘Please don’t go there’ is all it says today. Ostensibly, it is a theory about a very basic process of life, but it concerns itself very little with the broader processes of life. There can be no attempt to integrate it into important questions of evolution and living systems of molecular information. It reduces to silly conclusions like, ‘life will always waste molecular information if thermodynamics is involved,’ and ‘life sometimes gets stuck with bad solutions if the problem is old, basic and extremely important.’ The classic paradigm in this light is not an explanation of these things but a basic definition, and an embarrassingly poor definition at that. The absurdity of this is nowhere more apparent than the feeble attempts that were made after, the fact, to fit the paradigm into a broader concept of evolution. It is an historical fact that after the paradigm was made from whole cloth, then we were told by some of those same people: ‘It must have been a single, isolated, random accident, and it is now frozen in the amber of time, so it can have no discernible meaning in today’s broader scheme of things. It therefore quite obviously came from another planet. Please be a love and pass the pastries. Next question?’ Sure, I buy that one.

Fortunately, nothing today is exempt from the sharp blade of digital dissection. Every economist, physicist and biologist must submit his ideas to the cold, harsh, septic reality of digital logic. The logic of any idea in any area will eventually be laid bare by the “inhuman” logic of computers. Sure, the mathematician today must teach the computer how to think, but no human can now begin to match the heavy digital lifting that can be done by the computer. The computer has no trouble handling the complexity of crushing amounts of information that can be endlessly layered and complexly interrelated. The results of this mechanistic processing have now only begun to teach the mathematician how to think. This is the new reality of our times, and there is no place for anarchists in the future of this picture. If you cannot defend your position with a computer and to a computer and within a broader information paradigm, then you obviously have no viable position. Nothing is immune anymore. This is the attitude that we must adopt in biology, because biology is made up of pure codes and molecular information on some level. This is the only viable attitude we can take in attacking the understanding of the molecular codes of living systems. This is now the most logical starting point of any debate or discussion.

The genetic code is not merely a topic for chemistry and computers, although they are vitally important. The genetic code is essentially a topic that is all about Darwin and what he has to contribute to this biochemical paradigm of complex molecular information. If we totally exclude Darwin, we are bound to end up with a very limited paradigm, and that’s clearly what we now have with the classic paradigm. The reigns were apparently handed over to a small group of fun-loving chemists half a century ago,
and they literally held parties to discuss the paradigm as they created it. Quite naturally, they simply attacked the problem basically as chemists. That’s what they were. Unfortunately, nobody thought to invite Darwin to the party, so the paradigm completely ignored what he had to say about how we should expect to find this code behaving in nature. There weren’t really any hotshot computer jocks around to consult; consequently, the information paradigm is quite anemic within this classic version of the code of life. Darwin is no longer around to engage the debate, but his ideas in the form of his printed words are still around, going stronger than ever. We should listen carefully to what he had to say.

People wonder how I can remain optimistic and cheerful in the face of such a sterile, deterministic worldview. Why am I not profoundly depressed by the belief that I am somehow a computer simulation? Well, I know that I must be something, so what’s so bad about that? Rather than oppressively existential, I find this view uplifting and reassuring. It says that no matter how deterministic the universe may actually be at bottom, it is still at top a fundamentally unknowable universe and therefore up for grabs. At our level, there is obviously a huge place for “free will” in this universe. We are granted the great joy of being independent agents in this grand and exciting game, and our actions are our own. I am tickled to be invited to the game. Our actions have clear and significant consequences, and we can impact our worlds quite profoundly. Rather than an impersonal, meaningless universe it is a highly personal, highly logical universe where there is always a rhyme and reason, despite the fact that we may never actually know what it is.

Imagine that you walk out of your front door in your mind’s eye and you see yourself sitting beneath a majestic oak tree. You are reading this book and holding a big, beautiful diamond. The tree is pungent and flowing, and the diamond is crisp, beveled and explosively brilliant. What are the fundamental differences and similarities between the things in this scene? The diamond is a simple collection of atoms. On a tiny scale it is pure carbon arranged into carbon tetrahedrons the same way everywhere within the diamond. It had to form under a specific set of circumstances that included the atomic composition of its environment, as well as unusually high heat and pressure. Its timeline of formation started with the few atoms of carbon that first anticipated the final structure of diamond. The diamond structure itself now records the mindless choices made by each carbon atom in the overall structure. But because the structure is so simple and repetitive, the diamond contains very little molecular information, no matter how big the diamond may be. Diamond is just an inherently uninformative molecular structure. There is not a tremendously large amount of molecular information recorded in a
diamond structure, so the entire pattern can be compressed into a simple model of diamond.

How is the tree any different? It includes carbon in tetrahedrons but it includes many other kinds of atoms as well, all of which are every bit as mindless as the carbon atoms in diamond. How are the choices made by these molecules any different than the choices made in non-living compounds? The composition and the structure are much more complex but the tree also needed a unique set of circumstances and a specific timeline of formation. The environment and the timeline that led to the formation of the tree are much larger and much more complex than that of any diamond. The entire environment and timeline of formation for the tree goes back billions of years and extends over the surface of the earth. The tree itself records the molecular choices made in arriving at the specific molecular structure of this particular tree, and since there are so many choices to make, the information content of the tree is remarkably large. Unlike the diamond, the tree can also lead to new and more abundant molecular information by making novel molecular forms in the future. The information content of the tree is not only large it comes from a historically broad base and it surely will propagate into massive amounts of diverse molecular information in the population of its progeny.

What about you, the reader? This is the most complex and interesting thing in the scene, certainly, but how different really are you from the diamond and the tree? Start answering this question by thinking for a moment about your actual contemplation of the scene. It is a simulation of the universe that was anticipated by me, the writer. I encrypted this specific contemplation and stored it in the characters of this book. This too represents stored molecular information, and like the tree it has the ability to make new and more abundant information in the future of your thoughts. Your thoughts here clearly represent organized behavior in the universe. Molecules are organizing as a result of your thoughts. This is truly stunning. One portion of the universe, my brain, anticipates the future organization of another portion of the universe, your brain. There is no apparent end to where this kind of molecular behavior might lead through time, but the bottom line is that the universe can and is organizing itself. What other basic explanation is there?

Like the tree and the diamond, the book required a specific environment and set of circumstances in which to form. The book itself is a physical record of choices made, they are complex, and it is a tremendous amount of information. The history of the entire system that led to the book is embedded in many ways within the book itself. English is just one example of its embedded history regarding its environment of formation, but for another example, consider that I used Microsoft Word to write this book. Only a few
decades ago, Microsoft the company did not exist, let alone their product that I used to write this book. The company itself is a product of the imagination of Bill Gates and Paul Allen, merely two humans. The environment and the circumstances had to be just right for Microsoft to form, but the company itself started in the minds of just two people. It was a collection of ideas represented by the molecular patterns in a couple of human brains. Bill Gates and Paul Allen and their unique environment were the seed crystal for Microsoft to eventually form. They anticipated and communicated an idea of the future and the idea became a reality with all of the resulting massive human behavior and productivity, all of the attendant complex organization it brought to the future universe. They are yet another example of limited parts of the universe anticipating and organizing much larger parts of a more complex future universe. Microsoft the company and all of their many products and services represent a huge amount of new information in our universe. If this is not a fundamentally informative and self-organizing universe, then how do we account for the complex things within it, things that are becoming more complex and adding new information to the overall universe in the form of trees, people, thoughts, languages, books and companies?

Perhaps a broader question is, how do things become other things? Sections of the universe become protons, electrons and neutrons. These then become other things like hydrogen and oxygen, which become water. Water becomes oceans, and oceans become ecosystems. Ecosystems become cultures and cultures become books. Thoughts become English and English becomes thoughts. Thoughts become more thoughts, and thoughts become things like inventions, languages and companies. Human thought and invention combined with the simple foundation of two binary digits now can become anything your computer can make them into. Molecules become cells and cells become people and people make cells to become more people. DNA becomes protein. How do things become other things? How is the universe constructed so that this basic pattern is so prevalent? Which of these things that become other things can be considered alive and how do we formally make the distinction between living and non-living?

A virus is an organic particle. It is constructed of a protein coat, usually a repeating pattern of one or a handful of different proteins that assemble themselves symmetrically according to simple rules of solid geometry. The protein shell contains a genome and perhaps a few enzymes. A virus cannot replicate itself without the help of a fully functioning cell, but with the help of a cell it replicates with high fecundity. It hijacks the machinery of a living cell for the purpose of making copies of itself. Viruses tend to mutate rapidly and so they have the capacity to generate diverse populations of progeny. Is a virus a living thing? Most people say no; however, it is a purely philosophical question. I say yes, a virus is a living thing. Certainly, a virus cannot
continue its path of existence without the unique environment provided it by a cell, which is obviously a living thing, but the cell too cannot continue its path of existence without the unique environment provided it by earth and the benefit of the historical path it has taken over billions of years. Viruses are integral parts of living systems. They share many but not all of the basic properties and components of things that are clearly alive. They are self-contained entities that merely happen to require a unique environment for their existence. But a virus bridges the conceptual gap between inorganic crystals and organic crystals quite well.

If we are to admit viruses to the club of things deemed to be living, what of the other things mentioned above? What about languages, thoughts, cultures, businesses and computer systems; are they living things too? I think, by similar reasoning, that they are. Clearly, they require a foundation of organic structure and a robust population of human hosts for their existence, but how could they not be construed as living manifestations of complex molecular behavior on the surface of planet earth? Without life there are no viruses and viruses follow the long term pattern of living things. Likewise, without life these other higher-level things would not exist and they too follow the broader long term pattern of living things. The virus blurs the lower boundary of what is living and the more complex things blur the upper boundary. The trick, I think, is to view life as not a thing but a process or a system. While some things can easily be considered living apart from the system, like you or me or a tree, we cannot exist independent of the larger system. The virus, however, demonstrates essential elements of the molecular behaviors that are a part of fundamental living codes.

The universe must contain certain elements that can lead to living systems. The universe must have the ability to create and store information. It must be able to combine discrete elements of information to form more complex information. It must have mechanisms that allow a broad exploration of variations on complex information. It must have a basic mechanism to evaluate “useful” forms of information, and natural selection is that mechanism. It is a simple matter of keeping a record of forms of information that not only can persist but can expand to greater levels of complex information. There is one other fundamental element the universe must have to do all of these things, it must have fundamental rules of position and interaction. Geometry is the platform on which these rules are built. Things in the universe must always obey the rules of geometry, so the interaction of things must always obey the rules of interacting geometries. These are the bottom floor first principles that underpin everything else in the universe, and I believe they underpin the basic construction of all molecules participating in living systems. The fingerprints of these rules should somehow be found in the machine language of living codes.
A virus is an excellent example of the bridge between the living and non-living systems on our planet. A virus is a relatively simple auto-assembly of existing molecules according to the geometric first principles of auto-assembly. It also contains both nucleotides and proteins and is entirely dependent on the genetic code for its existence. But what about the proteins that are essential to virus assembly, are they also an example of auto-assembly? Absolutely! Therefore, since proteins arise from a still more primitive auto-assembly process, we should rightly expect their formation to also be predicated on geometric first principles of the universe. They do not represent a single geometric form but the interaction of at least two geometric forms, the dodecahedron and the tetrahedron. The language of proteins is a language of interactive geometries and ultimately it will be decoded by us through an enhanced appreciation of this first principle of the universe. Before any language could exist the universe must contain basic relationships for complex molecular interaction. These relationships must logically be based on the relationships between forms of solid geometry. The geometry of the information storage – DNA’s double helix - is essential to the system, but so too are the geometries of more complex molecules. There also is an important geometry to the language of translation between DNA and the more complex molecules that arise from DNA’s basic molecular structure.

Now, return in your mind’s eye to the majestic oak tree, prick your finger and place a single, small, perfect drop of deep red blood on the surface of the diamond. This drop contains trillions and trillions of hemoglobin molecules, and each one is an atomic arrangement of mostly carbon and water in tetrahedrons. Each hemoglobin molecule represents a tremendous amount of molecular information that must come from a specific environment and a certain timeline. The whole timeline goes back billions of years, although each individual molecule was formed within your body within a matter of seconds. This scene must be explained by the genetic code, but the classic paradigm has no chance of doing so. Place a Code World in your hands, and we can finally get started thinking about first principles of molecular behavior.

This will not be a simple language of molecular composition but a complex language of complex molecular structure. We simply do not know what nature is doing when she executes this code or on what level she is doing it. We do not yet know where, why or how she found this particular system. But we do know that the complex foundation of the system is laid down in time and space by the physical properties of insentient molecules, and the output of this system merely provides the context for more complex higher order codes and systems. We rise through the hierarchy into populations of molecules that make populations of cells that make populations of organs and...
populations of life forms until we reach the highest level represented by the entire biosphere. The language will still be represented in this highest level by the structure and strategies of its basic logic. The starting point is complex and the ending point is complex because it is a system built upon complexity by complexity for the purpose of generating more complexity.

Today, the issue has been framed by the premise of entropy. This is what the thermodynamic hypothesis is; it is a bold statement in our belief in the global entropy of protein synthesis. Of all possible systems it is the minimally informative system. It means that the system of translation cannot add molecular information to the initial information that is stored in DNA destined to be translated by the system, but it can and does erase huge amounts of information from it. The information starts out maximally uncompressed and the system maximally compresses it during translation. It erases information during translation. However, it is an accepted engineering principle that even the erasure of information from a system requires the input of energy to achieve reliable erasure. It will require physical mechanisms and strategies to perform this reliable erasure if proteins are to be reliably made in this way. Even if this particular molecular information system is doing this, why would it ever want to do this? What benefit is derived from the loss of information during translation? How could such a maximally inefficient system ever out-compete any system that is in any way not this wasteful? Why would a negentropy system of translation not be more competitive? Certainly, error correction in translation is a proven winner at some level, but at how high a price? Is error correction a goal worthy of exclusion of all other functions? I think not. There are plenty of ways to put the many dimensions of molecular information to good use.

Even if the system was miraculously able to achieve a single possible kind of peptide bond at the time of peptide bond formation, this is not necessarily an entropy system. When we take account of higher levels of molecular information within the system as a whole, we may find that single bonds will preserve higher levels of molecular information in the long run. This accounting will be completely dependent on a calculation of all probabilities of all choices being made throughout the system, a complex calculation to be sure. For instance, translation of a single string will in all probability lead to many translations of other closely related strings. These related strings are highly likely, even perhaps certain at some time to also be translated. If these additional translations turn out to add useful information to the system, then the net effect will be a rise in total information of the system by virtue of the system somehow finding the original strand of information. It is like a lay-away plan of information management. If a single kind of original bond can lend support to just this circumstance, then a “waste”
of initial string information *might* make some sense. We do not have the data to make this determination, but my intuition tells me that the system is much more elegant and efficient than this at this basic level. This opportunity for leveraging string information clearly does exist within the system and probably is utilized to some degree. It is quite conceivably a good information management strategy, but it would still have to be a negentropy strategy in general; otherwise, it never would have won this intense competition.

These are precisely the kinds of discussions and explanations that are absolutely required by any model of this system. However, it is exactly this kind of thinking that is obviated by the one-dimensional definition of the classic model. Once we acknowledge the possibility, nay the certainty of additional dimensions of molecular information, then an explanation of how and why the system could or would be limited to only one-dimension becomes more critical. The explanation that other dimensions are not possible has proven false, so the classic model should become immediately untenable by virtue of it being so illogical and improbable. So, even a discussion of this model virtually disproves the model and requires mountains of empiric evidence that currently nowhere exists to convince us to continue to believe in its veracity. This is why defenders of the model are so obstinate about not even being willing to discuss their model. It is a useless and indefensible belief system, but it is nonetheless the universally accepted belief system. There is strength in numbers.

Despite the fact that we do not know the exact logic of the molecular language that we have comically been calling the genetic code, we can still speculate about the broad types of molecular information that emerge at higher levels of living systems. Ultimately, the broad context for the information is the living system itself, and it is recursive, or self-referential at every single level. A more primary context for the information is provided by the molecular components, and they are provided by the context of atomic patterns in space through time. All of the molecular information has a strong time vector component. Without both the long and short term time vectors the information becomes completely meaningless. In the end, the code is a sequence of events not a sequence of molecules. The events are played out by a huge set of molecular components. Granted, many of the time vector properties are derived from the spatial vectors inherent in the physical quantities and specific ordering of molecules. This is the inherent genius of basing the system on sequential polymers, as was done with DNA by protein. The genius of the system has nothing to do with being a linear system. A truly linear system would seem to be a remarkably stupid system.
Higher order molecular information emerges in the form of shapes, charges, and the actual atomic constituency of molecules. Still higher order information exists in the concentrations of molecules, or the ratios of one form to another within a population of molecules, much like the pH of a solution. Another thing to consider: the physical location of molecules, like in or out of the cell, or on one side or the other of a sub-cellular membrane, this is another form of emergent information. It is information that can be quantified and accounted for. This kind of information could be called molecular accessibility of molecular information. For instance, a DNA-reading protein has a completely different “meaning” if it is in the nucleus and has access to DNA versus in the cytoplasm and does not. Likewise, a “gene” might be physically present but packed tightly in protein coils and therefore not accessible. An mRNA molecule might hold codons that are physically inaccessible to translation because they combine with each other. A protein might hold a string of water-loving amino acids, but if they are wrapped deep into the interior of the structure they are not accessible to water, and so they have a different meaning than when they are on the surface of a protein’s structure and therefore accessible to water. Heck, the information of one primary sequence apparently can somehow be influenced by access to itself in a different form – a prion. How shall we classify this kind of molecular information? All of the levels of components, all of the complex events and all of the dimensions of molecular information go into the structure and function of the molecular language that we hope to understand. It is a complex language, to be sure, but it is beautifully efficient in its overall operation.

The highest forms of emergent information manifest as behavior patterns of organs and organisms. The ultimate bit of emergent information in life is a binary property of the entire complex system that could be named ‘alive’. When the alive-bit is true the system is functioning, but when the alive-bit is false the system is dead, no longer actively processing and adding to complex molecular information in a sustainable fashion. This is when negentropy of this limited section of space in time gives up and entropy seemingly takes over. On this note we can now comfortably conclude that the alive-bit on the classic paradigm is clearly set to false.

The nesting of master codes and sub-codes that life has developed is a surefire recipe for intricate, complex and ingeniously sublime patterns of organic evolution. And so the pattern of increasing complexity is unmistakable over the surface of earth and back throughout the history of time on earth. If the starting date or time of first iteration for earth goes back 4.5 billion years, then the first fossil evidence of life on earth appears remarkably quickly, sometime between 3.5 and 4.0 billion years ago. Life seems to have
been chomping at the bit and raring to go. I think that trying to draw a line in the sand to
decide that life started at the line but did not exist before it is a fool’s errand and a false
line. Similarly, I think that searching for a last common ancestor at any point in the
process does not put into proper light the actual nature of this complex process. I think
we must view life as a pervasive web of inter-related complex processes that started with
the start of the earth, or perhaps with the start of the universe. I think too much
simplification and too many simple lines have been drawn in this grand picture already,
and we certainly don’t want to needlessly repeat those mistakes here.

More remarkable than the rapid fossil appearance of life is the relatively long
period of time that it takes eukaryotic or nucleated cells to make their first appearance.
They seem to have waited about 2 billion years to show up. Why so long? I would
venture a wild guess that many complex codes tried and failed their ascendancy to
today’s throne. Consider the unimaginably large number of ways that even the small data
set we find in nature today could have been arranged, and then consider the unimaginably
large number of sets of data with we could have possibly found. At the very beginning
we were dealing with an “all possible” scenario of molecules and codes. The most
successful molecular and cellular codes would be the ones best able to equilibrate all of
the competing forces of change being applied to them at any given time. The winning
codes were the ones that formed the best platform for more complex systems, the ones
that best harnessed the equilibrium of many dynamic forces. The protein code, like all
molecular codes, must be at heart a code of molecular equilibration.

Many systems and many codon matrices were rejected before the current system
was generally found and developed to near molecular engineering perfection. This is the
platform that could eventually lead to the kind of high level, multi-dimensional
functioning that we see today. After all, construction of the eukaryotic cell is a
monstrous achievement in complexity, but it merely set the context for higher orders of
multicellular complexity. It surely must have been a real grind to come up with it, so
billions of years is not an unreasonable time of expectation. Anything that could merely
be called a “cell” is nothing on the scale of complexity when compared to a modern
eukaryotic cell. Rather than freezing out Darwin by trivial necessity in explaining the
existence of today’s code, we should invite him to make a guess as to how we arrived at
this particular point in molecular translation systems. He most assuredly would tell us
that the code or codes selected two billion years ago were merely among the ones that
could eventually lead to the codes we have today. All others must be eliminated
sometime between then and now. This is not the same as saying that we can have only
one code now; quite the opposite, it is to say that the code we have now must not only be
able to evolve but it must also be able to somehow facilitate future evolution. This
always has been and always will be the basic winning strategy in the grand game of life. So although there could have been many successful codes then, the successful codes now are the ones that could lay the groundwork for future complexity, the ones that would incorporate the basic principles of the epigenetic code into their logical structure. Anything else is a loser today. Yesterday was different from today, but yesterday laid the groundwork for today. Tomorrow will be different still.

But from there, from the delayed but seemingly sudden appearance of complex eukaryotic cells, the process merely accelerates in the progression of complexity, heralding the appearance of multicellular organisms less than a billion years later. Then all hell breaks loose and the complexity of whole body morphologies takes off. Man eventually makes his fashionably late appearance, bringing an intriguing new level of complexity: bringing advanced language, culture and science with him. This last level of complexity – complex thought – is virtually new on the scale of evolutionary complexity, yet look at how robust it became almost instantly on a geologic time scale. It is a new form of complexity in the biosphere, one that rapidly obtains and retains new characteristics, one that clearly separates itself from the rest of the biosphere. It is a brand of directed change and its pace is Lamarckian not Darwinian; therefore, the curve of increased complexity within our complex thought as a species is exponential not geometric. The digital revolution has now punched the pedal to the medal and accelerated us to the next level.

All of the advancing complexity of the biosphere is the product of systems of logic or systems of information management. It is a process of having choices, making choices and recording choices. This is how our brains work. Or brains are systems of information management. For instance, should I stand and fight the tiger or should I run away? Should I turn left or right? Should I tell her that her new dress is pretty or should I be honest? The choices we make affect the molecules that are us. Individual molecules have choices too, but their choices are not conscious in the same way. Molecular choices are defined by their physical circumstances. Life always has countless choices and Mother Nature must make those choices in life according to the logic spelled out by Darwin. So, the “consciousness” of the broad pattern of life might be seen in the act of making those choices, as Schrödinger pointed out. The consciousness of life is always at the point of decision, the process of constant renewal that brings each generation of life into being. Each life form is a new hypothesis of life, and the false hypotheses are rejected while the ones that ring true merely serve to generate the next round of new hypotheses. The entire biosphere is constantly “thinking” about new ways to do more. There is a clear direction and logic to life, a method to the madness of Mother Nature.
The direction of life points toward the accumulation of information, negentropy, choices made, choices recorded, in an evermore complex pattern leading to evermore choices and complexity. Along this path man has clearly taken the lead. Man is at the top of the pyramid, achieved relative dominion over the rest of life if not exactly dominion over Mother Nature. This is politically incorrect but no less valid from a point of logical analysis. Man has begun manipulating the basic information systems of life in countless ways, but where this will all lead is still anyone’s guess. One outcome is clear, however, man has greatly accelerated the rate of negentropy on this planet with his culture, languages and now his digital systems of information. The general process is clearly accelerating today and it is a human foot on the pedal.

A broad analogy between the development of life and computers as it relates to negentropy and complexity is unavoidable. One need only look at the many new genome maps we have today. In general, the simplest organisms have the smallest genomes, and the most complex organisms have the longest genomes. More important, the maps of the simplest genomes and their distribution of information resemble the hard drives of early computers, but the maps of the most complex genomes resemble the hard drives of today’s more complex computers. Files and applications are not only more fragmented and scattered about; they are more widely shared and widely interrelated.

The same kinds of analogies can be drawn between the biosphere and modern information networks. Life and information networks are both logic protocols for large numbers of independent agents that become self-organizing entities in equilibrium, but the technology is quite primitive in today’s computer systems when compared to life. However, the same patterns of evolution have occurred in the computer revolution as in the evolution of life on this planet. The same general kinds of patterns of information storage and handling that have led to the increasing complexity and diversity of computer systems also lie behind the strategies that life used to give us this complex and grand view of life that we have today. Our human simulations of the universe - our brains - are now contemplating simulations of the digital universe that are “living” and have digital “brains.” We have yet to determine the “vital” difference between bits and brains, but information management is at the heart of both. I suspect it is only a matter of time.

The biology-computer analogy might seem overly harsh to pure naturalists, and it does require some not so subtle tricks of imagination. We must juxtapose the lightning speed of computer loops with the glacial pace of organic generations. Computers deal in nanoseconds, but life evolves in the timeframe of generations of populations. Although individual molecules operate with an incredible speed on a molecular scale, over the
scope of an entire cell the pace of change is remarkably slow. On the scale of populations of whole organisms, evolutionary change is often marked by units of time in 10,000 year increments. This merely reflects the magnitude and overall complexity of life’s pattern in space through time. Fortunately, at the highest level of bio-complexity the human mind no longer labors under this glacial pace of organic change. The brain’s neuron, like the computer’s electronic circuit, operates on a scale of nanoseconds, so thoughts can develop quickly. The brain also has a facile storage capacity – memory – allowing useful changes to be retained. The evolution of thoughts and culture – memes - flies forward many-thousand fold faster than does the natural evolution of genetic material. Computers have clearly accelerated the pace of change.

But any seeming fundamental superiority of digital computers over biological systems is merely an illusion today. In reality, the molecular systems of information processing are still far superior to any artificial system that man has yet created. And although the digital computer of today is obviously superior to a 1950’s computer as a metaphorical referent, it is still an anemic referent, at best, for the computational power of complex biologic systems. Digital systems are still completely dependent on logic operations of one-dimensional strings. These operations must be forced through narrow pipes called processors, and the sequence of events must be rigidly defined in time; conversely, biological systems do not observe these absolutely rigid restrictions in space or time. They never rest in performance of widely distributed events, and the many molecular multiprocessors form a complex arrangement in space through time. Digital systems choke on recursion but life seems to relish hardcore recursion. So, we apparently have barely begun to glimpse this brand of complex multi-layered logic that life uses, but it’s sure to continue to surprise us with its elegance and efficiency as we probe deeper into its secrets. The codes of life are sure to be far more complex than the codes of today’s computers.

The main goal of biological computers seems to be an execution of a broad, endless search. It seems to be searching for stable patterns of complexity that can serve as the basis for ever-more complex patterns that will have the ability to persist in space through time. The net result is an accumulation of information and complexity – negentropy. The prized achievements in this relentless search seem to be the patterns of behavior, both molecular and organism-wide, that can accelerate the search. We should then expect to see search enhancing patterns in the genetic and protein codes, and so we do, happily.

Within this broad context of biology, complexity and computers, the classic paradigm of the code formerly known as genetic is standing firmly in the way of our
advance. It is the antithesis of a complex code, yet it is a fossil that is strangely frozen in the amber of our culture. We must break open the amber and re-animate this fossil. We must bring it back into the realm of the living by incorporating entirely new computer metaphors and transplanting a new language of complexity into its heart. It is an exciting prospect, since we are so familiar and facile with digital technology today. This knowledge and technical familiarity was not available to the founding fathers of the classic paradigm all those decades ago when the amber encased this cherished paradigm. But now we can study the path of technology in a more advanced computer revolution, and draw analogies large and small with biology. Perhaps more exciting is the prospect that we might also do the converse and study biology with a new eye for detail, a new feel for life’s subtle strategies and patterns, and then perhaps use this knowledge to further advance our digital technology.
XIII. How could life begin?

Not the way they say it does.
XIV. What is life?

Life is a crystal.

Raw.

Embracing a computer metaphor of molecular bio-information is all well and good, but it begs yet another more basic question: what is life? This is the question Schrödinger asked himself in 1943, and the answer he came up with is not exactly clear, or not as clear as it seems in the retelling. Schrödinger wanted to describe life in terms of known physical laws, and life surprised him by appearing to him as a process of thermodynamic hill climbing which he called negentropy. However, he does not come right out and say, “life is negentropy.” He did not know how to answer the question because life made him realize that there were gaps in his understanding of physics. Schrödinger also contributed the phrase ‘aperiodic crystal’, but he meant for it to apply only to the hereditary substance or chromatin. He does not say that life is chromatin and therefore life is an aperiodic crystal, which was wise, because it is not. He does, unfortunately, come out and say that the genetic “code script” will be found in chromatin, and we know now that this is false. It was, however, very insightful of him to observe that life would need a physical medium in which to generate and store molecular information, and also that the general trend of this process would go opposite the second law of thermodynamics as he understood it. But he did not have enough specific information at that time to give us an appropriate answer to his own question: what is life?

Schrödinger was particularly brilliant in that he started with the most basic questions about life. He looked at the question through the eyes of a child, albeit a quite precocious child. For instance, is life a solid, liquid or gas? This seemingly simple question turns out to be remarkably difficult, even now, because in fact life combines the physical forms and some features of all three, solid, liquid and gas, yet it is not clearly any of the above. His cutting logic was to note that a hallmark of life involves its persistence of form through time, in the sense that an individual retains its form during its lifetime, more or less, and some feature or essence of that form is passed along to its descendants. So he sensed that there is some persistent physical form to an otherwise abstract concept – molecular information. What could it be if not a solid? Specifically, he reasoned, the type of solid it must be is some type of orderly arrangement of atoms. ‘Crystal’ is generally the name for a solid with an orderly arrangement of atoms. However, he reasoned that a perfectly ordered arrangement of atoms would have no information value. For instance, one salt crystal is essentially like the next, it has no algorithmic complexity, so salt contains very little molecular choice, which means that salt contains shockingly little molecular information. So, if life is to be a crystal, it is a crystal that is fundamentally unlike the types of crystals with which we are casually familiar, the inanimate crystals like salt and diamond. The special crystal-type of life must have some disorder, or non-repetitive nature to its crystalline pattern. Ordered atoms form periodic crystals, so he described the crystal behind heredity as an aperiodic crystal. He then said that this aperiodic crystal would somehow record a code-script for building the organism – a genetic code. He guessed that it should involve a process of
isomerization, or a physical switching of molecular form, which is hyper-technically correct, but it is not generally the way we think about it now. He did not guess that the genetic information feeding the code-script would be based on the information content of sequential polymers as was later demonstrated with DNA and protein. Nobody at that time could have possibly imagined the bizarre reality that was later discovered. Consider if you can, it was a truly remarkable situation. A founding father of quantum physics realized that the usually random behavior of unthinking individual molecules could ever be so orderly, purposeful and efficient on an atomic level. The situation went against everything that he thought he knew about physics and chemistry, and he said as much. Schrödinger was rightfully dumbfounded, because from a molecular standpoint, it’s just like magic.

Of course, we have the great advantage to know today that Schrödinger was speculating on the nature of DNA, and that sequences of nucleotides are to a large degree, but not entirely, aperiodic. Periodic means regularly repeating, and on a different level, though, DNA is a crystal that proves to be remarkably periodic, in that it has the ability, under the guidance of protein, to repeat its entire large and complex pattern with uncanny speed, economy of space, and impressive fidelity. DNA molecules have a high fecundity, they like to repeat, or in Richard Dawkin’s terms, they are aggressive replicators at heart. So, even the simple act of replication must be tightly controlled within the broad scheme of life. The number of copies of a particular molecule is a vital piece of molecular information to a cell or to an entire organism, and in the case of DNA it means the difference between life and death.

Consider for the moment that you are somehow the product, in no small part, of the molecular information stored in the pattern of your DNA. This pattern of your DNA sequence represents a quantity of information, or a data file that when plugged into the genetic code will produce you. DNA is a crystal. This specific crystal pattern, to a very large extent, exists within the nucleus of each of your cells now and in most all past cells that have been you. (In this limited sense you really are just a number, sorry to say.) The pattern of this molecular sequence is comprised of a relatively stochastic sequence of some three billion nucleotides; however, you are comprised of some seventy-five trillion cells, all in a complex process of constant renewal, and this is a much bigger number than the number of nucleotides in your DNA sequence. So the complex crystal pattern of your DNA is phenomenally periodic in the sense that it has been and hopefully is being faithfully repeated many, many more than seventy-five trillion times. The sequence of nucleotides is relatively aperiodic, but the complex crystal itself is stunningly repetitive, or periodic from an information standpoint. One of DNA’s many charms is the fact that it can be periodically repeated in its totality, when protein so chooses.

In the exact same sense, the other crystal forms that contribute to your being, proteins like hemoglobin, for instance, are also constantly replicated at incredible rates with exquisite precision (hopefully). This could be described as a complex, but remarkably periodic pattern of hemoglobin. Although, the individual hemoglobin molecules do not arrange themselves in a globally ordered pattern within a cell, like salt or diamond, they themselves are repeated with uncanny precision. So we might think of it as an amorphous arrangement of complex yet highly periodic molecules. Add to this the fact that the hemoglobin that you are producing now is not the exact same form of hemoglobin that you produced as a fetus (hemoglobin patterns shift from fetal to adult
form after birth) yet your hemoglobin now is remarkably similar to the hemoglobin of a chimpanzee. Worse still, aquatic mammals, like whales and seals, have brands of hemoglobin that surely put yours to shame. There is not only a code that produces a single hemoglobin molecule, but a broader code, a hemoglobin code perhaps, that produces the time-path of hemoglobin within individuals, both phylogeny and ontogeny, species, and within the entire pattern of life on this planet. The pattern of a hemoglobin molecule spans time and species, so now we can begin to see there is an essential element of relative time, timing, space and number to the overall pattern of bio-molecules no matter where we look. We can never freeze, flatten or simplify the pattern completely and still pretend to be studying the real pattern.

An aperiodic sequence within a crystal polymer is therefore certainly a defining feature, but it is clearly not the defining feature of life. On one scale these crystals are aperiodic, but on most other scales they are not. If life is to be defined by its aperiodicity, or by its purely informative nature, then the term ‘informative crystal’ is perhaps more appropriate. It captures both the key elements of aperiodicity and periodicity within the larger scheme. In this sense the terms ‘negentropy’ and ‘aperiodic crystal’ convey virtually the same message, which is that life is a physical system of information storage, management and processing. Molecular information, its creation, storage, processing and replication are very important aspects of life in general, but there are two features of this process when viewing life as a crystal that are still more unusual and more defining than this.

We still need, however, to address head-on the basic question of whether or not life actually is a crystal. I believe it should be seen as some kind of crystal system, but consider that approximately seventy percent of most cells are composed not of solids but of liquid water. Liquids are defined by the dynamic, amorphous nature of their molecular structure; they are constantly changing form. So if life is generally amorphous and seventy percent composed of liquid water, is it right to refer to life as a crystal? Perhaps it is a liquid crystal. This should not be too objectionable to very many savvy readers, because it combines the precision of organic crystals repeating themselves with the generally amorphous nature of liquids.

Certainly, water and its unique physical properties are essential ingredients of life in the liquid not the crystal form, but life itself seems to take the overall form of a solid. To reconcile, we might note that in many ways life can be seen as a process that exists right on the boundary, or at the phase transition between liquid and solid. This is a familiar and exciting boundary in many academic disciplines, such as chemistry and physics. Chaos, complexity and life are interrelated dynamic processes that exist right on the edge. Compared to salt, which dissolves in water, life is more like a complex carbon crystal that depends on a dynamic liquid matrix to stay at this exciting boundary, or on the edge between liquid and solid. So life is a crystal, but it is always a quite dynamic type of crystal system in every sense of the word.

Life depends on crystal motion – thermodynamics – to execute many of its molecular codes. Merely changing the temperature up or down in a cell will adversely effect the operation of these codes and the actual content of the molecular information. The codes themselves are calibrated for a certain statistical amount of motion, but they generally govern this motion and are not simply at the mercy of it. Sunlight is another example of the necessary element of atomic dynamism in life. The system requires the reliable
source of energy from the sun to perpetually and reliably excite the basic subatomic, atomic and molecular structures into action within the complex, dynamic system of life. Life is based primarily on carbon and water. Both of these atomic structures like to form molecular tetrahedrons. Water molecules - individual molecules of H2O - form dipoles that jostle around in the liquid form to create a dynamic but persistent tetrahedral matrix. Even as a liquid, water has notably crystal-like properties in its structure, and it is the unique properties of water that make life possible. Remember, ice is just crystal water, but ice floats whereas most crystals sink in their own liquid solution because most crystal structures are much smaller and therefore denser than their liquid counterparts. But it is the dynamic nature of liquid water that drives the processes of life at its base. Carbon has a valence of four, and so it likes to covalently bond with four partners in a tetrahedral arrangement. This obvious crystal part of life, carbon, must play its role within a dynamic matrix of water molecules.

Unlike most crystal forming systems, organic crystals have the unique ability to change their physical form through time. They are not static morphologies, like salt that dissolves in water; they are always dynamic morphologies that only exist in the dynamic matrix of water. Organic crystals also have the unusual and highly useful ability to briefly associate with, and then separate from other organic crystals. A single organic crystal also has many stereoisomers, or many distinct shapes, and it has the curious ability to quickly explore and re-explore several of these throughout its existence. This inherent flexibility of shape gives many organic crystals their utility within life’s processes. More importantly, there are so many potential organic crystal morphologies that it is virtually impossible that every single one of them could ever be explored within the lifetime of the entire universe. It is a highly versatile system of molecular combinatorics. Yet life has acquired the dynamic ability to methodically and systematically explore a large number of these potential shapes. These various forms of dynamic crystal behavior, taken all together, are emblematic of life. A hallmark of life as a crystal is its ability to be a dynamic crystal, not just a single dynamic form but a nearly infinite population of many interrelated dynamic forms.

From this foundational view of life as a large, dynamic population of informative crystals, we can see clearly that complexity is the most defining feature of life as a crystal. It is not a single crystal, or even a population of single crystal forms. It is at all levels a complex interaction of large populations of many crystal forms, on many levels, played out on the backdrop of time. Life emerges from the complex and dynamic interaction of many crystal agents. It is a crystal system, a complex system that combines crystal machinery with complex crystal logic. It is not a simple system of crystals but a complex system of complex crystals. There are codes of interaction that not only take into account crystal identity and form, but the complex interaction of many crystals that play out on the necessary backdrop of time. Life, therefore, cannot be viewed merely as a snapshot but a movie. Energy must be supplied somewhere in the system, but energy is not a substitute for complexity and increasing self-organized inter-relationships. Energy is entropy that is converted by the system through time into negentropy. Therefore, it involves a language of crystals that has an important diachronic facet to its meaning, and the relationships of the whole are never independent of the time element on any scale. It is a crystal system that has a complexity of both time and space. Complexity should be considered the overriding hallmark of life as a crystal at all levels.
We still cannot truly answer the question, what is life? But we now have several general descriptors. Life is a complex, dynamic, informative crystal. Because of its extreme complexity, a single crystal, or even a single crystal system cannot define life. It is a complex interplay of many crystal systems. It is the interplay that makes it complex, and it is the interplay that allows it to be dynamic. The system is not static; it cannot tolerate stasis, it requires constant novelty, perpetual renewal leading to advancing complexity, and so it is actually driven by change. Our description should then include a nod to this relentless change: Life is a complex, dynamic, informative crystal system that constantly makes new crystal combinations throughout time.

More subtle, but more significant, we must recognize that life is not a thing but a system or process. It is based on time and embedded in time. It is not a single event but a complex sequence of events dependent on the profound interdependence of its fundamental components in time and space. We must consider the entire pattern of life on earth back through the entire history of earth. It is a gigantic and ancient pattern of crystal growth that has created a complex crystal information system. We cannot freeze time, or take a snapshot of a pattern, and then declare, “Here is life!” In contrast, we can easily paint a still life of a mineral or inorganic crystal and understand its basic structure or its informative essence. Crystal in the colloquial sense means a simple, periodic and adynamic arrangement of atoms. The codes of molecular behavior in inorganic crystal formation are quite simple. It is a still life, but a still life of life is the opposite of life – it is death. Conversely, the codes of organic crystal formation are tremendously complex because they are generators of complexity.

If life on earth is in any way to be seen as a crystal, then the pattern of its growth must extend over vast expanses of time and space. Any part of this pattern can be isolated from it for only brief periods of time and still maintain the general trend of the pattern. It started billions of years ago in the watery environment of vast oceans, and it progressed to fill those oceans. It then advanced onto land and brought the essential components of its environment with it. It has now tenaciously extended to radically different environments and so it blankets the earth. As it has extended outward in space and onward in time it has also moved upward at an accelerating pace on a hierarchy of increased complexity. The general growth patterns of this remarkable crystal system are unmistakable, yet the specifics of its molecular behavior still lay hidden to our eyes. The codes of living crystals are embedded in time and dependent on the specific crystal sets by which they are executed. The codes themselves are evolving with the total pattern, but there must be some first principles of molecular behavior to anchor them to the physical laws of our universe. The laws of the universe breath life into the codes of molecular behavior, and it is these codes that breath life into the molecules of earth.