

# Rafiki

## At The Edge

by Mark White

## Foreword

You are walking a midnight street under a moonless sky. A streetlight reveals a man searching a vacant lot. You offer to help, and the man accepts.

“I lost my keys,” he explains.

“Where was the last time you had them?”

He hesitates... a second at most, before pointing across the street.

“So why are you looking here?”

“Because the light is better.”

## Patterns

The universe is a pattern. It is very big and very old, making it extremely complex. It is truly incomprehensible to any human, but it and everything in it are patterns. To understand the universe, or any part of it, we must appreciate its patterns and the methods it uses to create them. Pattern recognition is therefore a valuable human talent, and it is in fact the root of human intelligence. On this score God and science are partners. We focus here mainly on science and its handmaiden, mathematics. Science attacks a problem through simplification, which is the best - perhaps the only tool available. Science must first simplify a system in order to work an otherwise incomprehensible problem. It is no surprise then when science is found guilty of oversimplification.

Numbers are the language of science, and numbers themselves are a gross simplification. Mathematics concerns the nature of numbers, and more importantly the relationships between them. Therefore, mathematics is also limited by simplicity in what it can teach us, but numbers and mathematics are the best tools we have in searching for patterns. There is something intrinsically human in the spiritual rapture of seeing “the numbers add up” as the old one chooses to reveal a frayed corner of a secret to us. It is essential then that we augment these tools in any way possible if we hope to create a more robust view of the patterns around us. This augmentation process is called building a model. Models are scaffoldings that hold the metaphors we use to expand and explain our understanding of patterns. Models are paradigms, or perceptions, or conceptualizations of the patterns we are able to tease from observations of the universe. Seeing is believing, and models are the eyes through which our brains can see that which cannot be seen.

Numbers are indispensable in this quest because the universe has the curious property of appearing quantized. At every level the universe behaves, contrary to our intuition, as if it can be made discontinuous or segmented. It can be broken up into discrete units of time, space and energy, and these units can be counted and arranged. This is most curious but most convenient, because a quantized thing is an informative thing. It reeks of dualism and determinism, but calls longingly for context. We have yet to find the level - big or small - where this property seems to fail us. In fact, the further we dive into the problem, the more examples we find, and the more convinced we become that this is indeed a quantized universe. It is certainly an informative universe. The debate of universal quantification might never be resolved, but the patterns made by the various quanta can still hold our collective fancy.

Numbers are the essential language of these patterns, but numbers are always represented by symbols. Symbols aren't things, they are representations of things, and they are therefore changeable. We could use the symbol “3” to represent the concept of three things, or we could just as easily change the symbol for the same concept to “Γ”. For that matter we could replace this symbol with a color or a musical note; it all depends on the pattern, the model attempting to find the pattern, and the person attempting to understand the model. It is the concept of “three” that interests us, not so much the symbols representing them. More confusing is the fact that the symbol “3” is merely a printed or spoken symbol that serves to trigger a less tangible symbol for the concept of three in our brains. But how sure can we be of the concept of “three” and its relationship to anything “out there” in the space that we know as reality?

Numbers come in only two basic types, those that represent something, and those that represent the absence of something. We have zero and everything else. From this perspective zero was a remarkably late arrival to the drama of numbers. Humans got along fine without zero for millennia. Now “0” is an indispensable concept and symbol for the elegant models we build of patterns in the universe. Our base ten, or decimal system requires ten symbols: “0”, and

nine others. From this humble platform we can combine and relate things, producing patterns that are unimaginable to all but a few gifted pattern alchemists. How sure are we that we understand zero?



Life is a pattern. When I say this I mean it on two levels. You have a life (I know because you are reading this fine writing) but you are part of a larger pattern of Life on this planet, or Life in this universe. There is a big difference between the pattern that is Life as a universal process, and the pattern of life that is the individual instance that we call you, or me. I will denote this difference throughout by the use of capitol "L" for Life the process, and lower case "l" for life the individual instance. This book shines light on Life the process, and we will examine it in relation to our understanding of patterns of the universe. Likewise, thought is a pattern, and we will use Life as an example of how we can better, more effectively pattern our thoughts about the universe.

Models are lights that illuminate our search for patterns. There are different sources of light, but as the above joke implies, some lights are better than others. It is very common that scientists select the wrong light and end up searching the wrong areas, sometimes for long periods of time. It is not that these searches come up empty, but the keys are never found under the wrong lights. Usually, maps are drawn during the searches that can then show where the keys might actually be found. In this way science is not a linear march as is commonly thought, but a jagged saltation forward. It is Brownian motion with a perceived goal; unfortunately, the goal cannot be known. The beauty of science and the models it produces is in the art of explanation, not prediction, and frequently beautiful patterns of Life blossom for us only after we have constructed some radically new models. Many of these models are simple and straight forward, but others are elegant and complex, and they are necessarily heavily metaphorical. Most constructions have a marked bias; they have a foundation that is a general model of the universe, or what we typically call a worldview.

The worldview on which this book is founded has already been stated: The universe is a complex pattern. This is too general to be entirely useful so I

will do a few things to it. I will name it and I will expand it. Rafiki is the name I chose for my worldview. I was recently asked why I chose the name Rafiki, and I couldn't think of a good answer, other than, why not? I didn't want to burden it with my name, and I didn't want to give it a long, pretentious, unpronounceable, pseudoscientific name that would end up as initials anyway. The world has enough initials already, so I chose Rafiki. The Rafiki model is a casual collection of ideas about patterns and how they behave, or fit together. Rafiki is a room light that we will try to focus through a lens and light up the pattern of Life. This book tries to explain Life and how it is better understood within the context of the Rafiki model. Without Rafiki, the model of Life can lay there flat and colorless.

Rafiki contains general rules of thumb to help develop a global perspective about patterns of space and time. Of course we already have a model to teach us perspective on space and time – it is called relativity. Rafiki is a relativity of sorts, but it is far less formal. Rafiki is more about the relativity produced by networks, about how all things in networks serve as the context of all other things. It is about how new things can emerge from networks of completely different things, and how these new things themselves become networks that serve as the next unexpected level of emergence. Rafiki is about how patterns, the really interesting ones at least, are really just networks of things, and how the universe is just a nested hierarchy of these interdependent networks. If this over-generalization seems scientifically implausible to you, I suggest that you read *The Emergence of Everything* by Harold Morowitz, or *A New Kind of Science* by Stephen Wolfram.

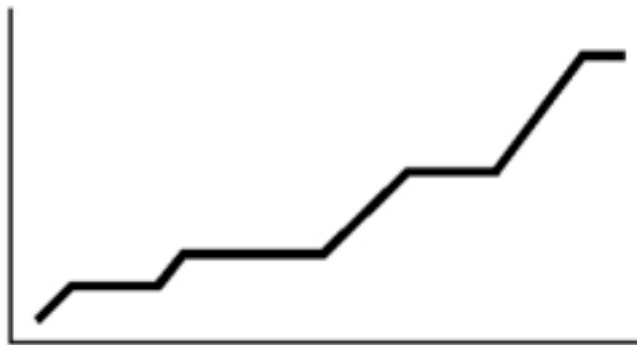
The first Rafiki rule is that all patterns are ultimately patterns of space, and all patterns must change. This is where time comes in. Let's say we freeze a moment in time, and we examine space during that frozen moment. If we notice a property, or wave, or energy distribution in space at that moment of time, the only certainty we can have is that when we let go of time, say for one quanta, or one tick on the universal clock, every individual element of the pattern will change in some way. In this model, the only constant is change, and doesn't that seem to be the case with our lives? The Rafiki model relegates time to the role of counter in a series of universal patterns. Granted, it is brutal and crass to simplify things to this extreme, essentially reducing the universe to one big, deterministic cellular automaton, but at least it's a place to start. Time is the necessary consequence of the first rule of patterns, which is that they must change.

The second rule of patterns is that they will equilibrate. This answers the question begged by the first rule: How will patterns change? Now we beg a new question: What does it mean to equilibrate? That's a whopper, because Equilibration can mean many things, in many settings, as we shall see. One of the tricks to recognizing patterns is to guess what an equilibration might look like, or where it might be, and then go there with a flashlight and look for it. In general, it means that every portion of the universe is copasetic at every moment of time. Equilibration means that things are as they should be at all times and

that they are following the rules. The universe and everything in it must follow a small set of immutable laws. There are no special rules and there are no special rule breakers. At any level, at any time, the laws of nature are always followed, leading to equilibration at the very largest and the very smallest levels. Constant change in a state of equilibrium seems paradoxical, but it is not. Perhaps it would be more understandable if we just said, “it’s God’s will, and only God can know the answer”.

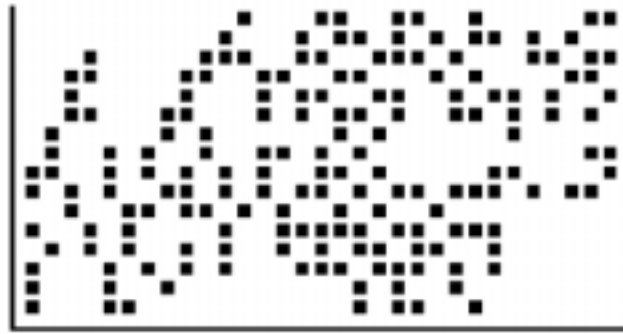
Patterns will fluctuate. Constant change in a state of equilibrium will lead to a fluctuating pattern. An undifferentiated universe would display no pattern at all and would be uninformative. Therefore, the universe must have uniquely identifiable components or elements, and patterns will be an intricate dance of these elements. Patterns constantly seek a sweet spot, only to pass through, never allowed to rest because of some pesky local perturbation, or some global trending force. But they will return again and again from differing angles, never straying far or long from their ideal. This of course will lead to patterns in time of other patterns, which gets rather confusing. This can be seen as a hierarchical series of patterns. In fact, most of the patterns we humans can identify in the universe are incredibly complex aggregations of other patterns. They are networks that spread their dendrites above and below through the levels of the hierarchy. It is rare when we can find a complex pattern that can be perfectly excised and painted with simplified numbers. When we can, we like to think of this pattern as linear. For instance, the pattern of mass being influenced by force demonstrates acceleration. Given any mass (M) under common conditions, an acceleration (A) will be produced by a force (F) with such regularity that we describe it by arranging the symbols as follows:  $F=MA$ .

This is what we call a linear relationship, because any one input can be expected to produce precisely one output. We can use a simple numerical tool to demonstrate the concept of linear - we call it a graph.



One-to-one relationships plot nicely onto a line, and this is why we like to call them linear. There is only one dimension of influence that needs to be

explained by the graph. One-to-many relationships plot poorly into a line, and we can unimaginatively call them non-linear. Frequently, we will try to find lines in non-linear patterns, primarily by using statistical tools, but often there simply are no lines to be found. This does not stop us, many times to our detriment, from drawing the lines anyway. It is our evermore-ambitious quest for enlightenment through the power of simplicity. This is where we've historically tended to find the most light. However, sometimes we encounter a pattern that we can all agree is not linear, and should never be made as such. For instance, the following graph in no way appears linear, but many still insist that we find a single line there.



Rule #3: Patterns accumulate. This is the primary mechanism of fluctuating patterns, and it is here that we find the most interesting properties of the universe. Patterns emerge from the universe, and more patterns emerge from those patterns, etcetera. They are fractals, networks of patterns becoming nested like so many Russian dolls. However, an accumulation of patterns does not create a bigger version of the previous pattern, it forms a different pattern altogether. This is the fundamental principle of emergence, a difficult to define and therefore little understood concept, but intuition tells me that it will be the foundation of the next great scientific revolution. Incidentally, intuition is an emergent phenomenon of the mind. Complexity is another concept that defies definition, but like art or pornography we know it when we see it. Complexity and emergence embrace each other so tightly that it is impossible to separate cause from effect. Complexity emerges from our universe at the boundaries of patterns, ones that aggregate order and randomness. The ordered part of the pattern is not complex, but neither is the disordered part. The edge between randomness and order is the most complex, the most interesting and the most useful part of any pattern. Life fits this description, and in fact it is a property of the edge - Life exists only at the edge.

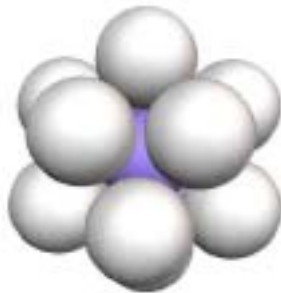
Many of the biggest missteps in human thought occur because of a failure to recognize the properties of accumulating patterns. These misguided models begin by asking us to “imagine a spherical cow”, or “consider a perfectly rational agent”, and they end with a conclusion about the state of our universe less than a tiny fraction of a second after the big bang. Let there be light. Alas, to a hammer all the world is a nail, and to me all is emergence and complexity. I will therefore pound the theme that more is not more; more is different.

This all sounds so horribly deterministic, as if the universe is a large state function device, a digital computer. Perhaps it is, but I am comfortable knowing that we will never know. The equilibration function has a decidedly historical input, and therefore it is subject to wildly chaotic behavior. The precise, detailed history, and therefore the input can never be known. The boundary conditions of equilibration are unknowable, and therefore the state of the universe cannot be determined by anything other than the universe, or something capable of containing an exact replica. More importantly, the nature of equilibrium itself is changing, and the direction of change is clearly toward more complexity. Emergences are accumulating and complexity is the result of emergence. Let’s start with an example that is purely abstract, and move toward something tangible.

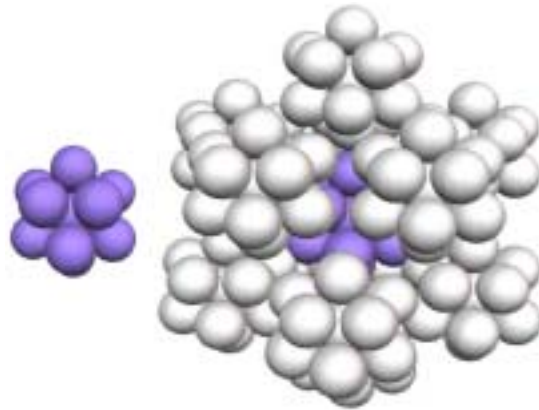
Please consider the following pattern.



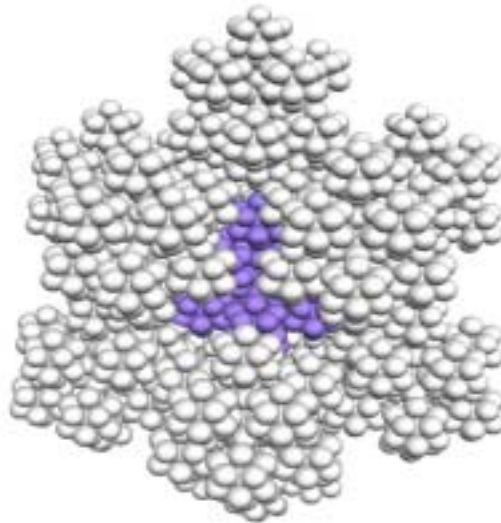
Not very interesting, but we can surround it with copies of itself, and generate a larger aggregate pattern.



The new pattern is clearly different from the original, but our new aggregate pattern can be considered as the base for a new level as follows.



Of course we can continue to repeat this procedure indefinitely, and we will start to see completely new patterns emerge from the old.



The purpose of this simplified geometric model is to help visualize an aggregation of patterns, and to point out that they must be considered relative to scale. How many iterations of this simple pattern might we perform before something of enough complexity might emerge where we could look at the pattern it produced and say, “there I am”? How many iterations before a universe could recognize its own pattern? Where could we store such a pattern and how could we view it? The missing element in all of these questions is the same; it is context. All patterns have a need for context, and when we shift from one scale to another we will unavoidably produce uncertainty regarding the pattern. To jump from one scale to another is not to examine the same thing on a different scale, it is to examine something completely different. Since time defines patterns in space, we will need a tool that scales time relative to space. Fortunately, with a little poetic license we can create one.



$$\Delta d_{\min} / \Delta t_{\min} = \Delta d_{\max} / \Delta t_{\max}$$

This is a minor twist on our old friend called special relativity. We are not going to use this for physics per se; we will use it as a universal tool for scaling time and space in our search for patterns. When we pattern our thoughts to conform to the very small in space, we must change our view of time accordingly. When we think about patterns in time we must likewise scale our reference in space. This gives us a nifty metric for how much relative scaling of either is required by our models. Time, space and proportion become the universal trinity of patterns.

Whenever we jump from one scale of time or space to another, we necessarily introduce a concept known as uncertainty, which is a fundamental reality of patterns in the universe - no way around it. We cannot contemplate patterns on the order of one second in time, and be certain about patterns on an atomic scale in space. A second to us is roughly a million years to an atom, per our scaling device. Not much happens to us from one normal second to the next, but an atom can perform many tricks during that same period. The pattern of the atom has fluctuated countless times in a fraction of a second, and the net result is an absolute trick to us. The trick that the atom is found to have actually performed during that second will be the one that is most likely to be found by us when we look under a particular set of circumstances. Unfortunately, the set of circumstances is itself a pattern requiring an unknowable context. Our certainty about what we find will be a function of the scale we use to find it, and how we look. To mistake this actual finding of the atom as the only possible trick, or the only actual trick performed during that portion of time is to make a mistake of scale.

Similarly, the pattern of Life in the universe has formed over a vast amount of time, and we should seek to locate it in the very large, not the very small. To trace life backward in time to a cell, or worse a molecule, is to make another mistake of scale. The above ratio is a nifty tool to help calibrate our thoughts for any scale, but we must first be aware of the scale on which we are searching.

# THE ORIGIN OF LIFE

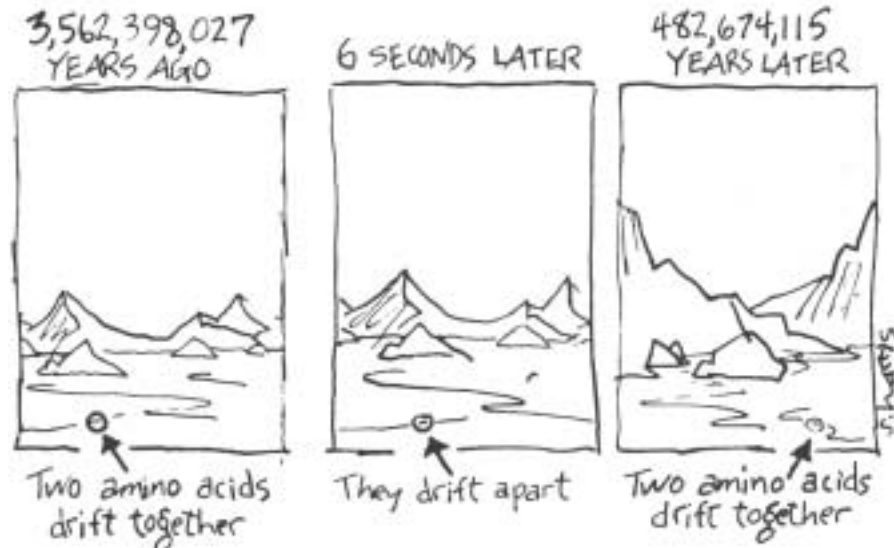


FIGURE 7.1

Uncertainty is relative. Our level of certainty manifests as a spectrum across the patterns in the universe. It scales according to time and space. At one end of the spectrum is absolute certainty, or total order, and at the other end is absolute uncertainty, or total randomness. The Rafiki model says that we might find absolute certainty only at the absolute smallest and the absolute largest scales of the universe. All other certainty is relative to the scales of the patterns observed and the patterns of observation.

The behavior of a pattern is defined by its scale. A diamond billiard ball's behavior is dictated by the behavior of countless, smaller carbons. Likewise, the carbon pattern is locked by the countless smaller patterns of its parts. Carbon is a nested network of sub-atomic particles, and diamond is a nested network of carbon. We can imagine a system of diamond billiard balls that will behave in an orderly fashion, much like the system of carbon that makes up each ball, or the system of particles that make up each carbon. Staying at any one scale produces relatively predictable behavior on that scale, but jumping scales produces relative uncertainty.

Consider the simplest known network of sub-atomic particles - the hydrogen atom. Hydrogen is at the base of the atomic scale, and we can generally understand its behavior by understanding its parts. An oxygen atom is on the same atomic scale, but it is constructed from a more complex nesting of the same parts. Hydrogen has one proton where Oxygen has sixteen. We combine the two atoms in a ratio of two hydrogen to one oxygen, and produce an

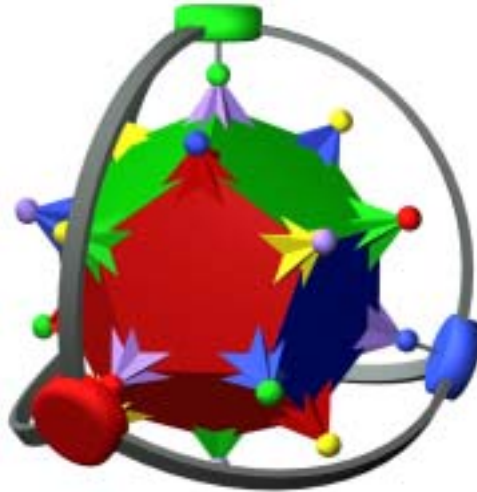
ocean of molecules that behave on a new scale - H<sub>2</sub>O. A molecule such as H<sub>2</sub>O might be expected to behave as an atom with 18 protons, but we all know that nothing could be further from the truth. Our oceans are made of water, not Argon, and their behavior could never be predicted by adding the behavior of Hydrogen and Oxygen. Why? Because patterns have scale, and we must respect the scale of patterns to understand their behavior. To do otherwise is to commit a grievous error of scale, similar to thinking that oceans should behave as Argon.

Defining scales, and therefore defining patterns is a very messy business, as we shall see in our saltation forward. We will start with a general road sign for our universal exploration of patterns.

**Universe**  
**String**  
**Quark**  
**Sub-Atomic Particle**  
**Atom**  
**Molecule**  
**Macromolecule**  
**Life**  
**Thought**  
**Culture**  
**Universe**

I have no absolute way to determine a valid metric for the divisions of this road sign. It's just a rough gauge that I like to use as I search for patterns. We hope that in our search mathematics will create patterns in numbers that will model patterns in the universe, and thereby we might attain a small measure of hard fought enlightenment. The numerical patterns we contrive must respect the dynamic nature of the universe. They must perform magic along an axis of time and thereby develop a rhythm. More importantly they must respect a scale, and as we jump from one scale to another in our searches we must shift our pattern senses. Mostly we move down in scales, and as Richard Feynman wrote, there's plenty of room at the bottom.

The Rafiki belief in patterns is nothing more than a simple belief in logic. If a process in the universe is operating on some logic, then that process should leave telltale fingerprints in the patterns of its behavior. Natural laws should create natural patterns, but we frequently miss this evidence in our searches from lack of perspective, usually a failure to recognize the relevant scale of the questions. Our preferred scale for this journey is Life. Let's begin.



## A Code World View of Life

All life on this planet is based on a genetic code. It is a system that somehow defines the construction of living things by directing the processes of replication and protein synthesis. In 1953 James Watson and Francis Crick described a double helix as the structure of a huge molecule called DNA, which was known to reside in the cell nucleus and store the secrets of the genetic code. Excitement grew, and by 1960 leaders in science were predicting that nature would be laid bare within a year, creating justifiable fears. If man actually controlled the genetic code, what would happen to life on earth? Salvadore Dali seemed to anticipate man's dominion over nature and its relationship to a higher truth, as shown in his painting *The Temptation of Saint Anthony*.



*The Temptation of Saint Anthony*, 1946  
Dalí is about to leave the earth in order to reach the heavenly spheres. The dimension mediating between heaven and earth is embodied in the elephants with their spindly legs. They introduce the theme of levitation that was to be fully developed soon after in his "mystical-corporeal" paintings.

The predictions and accompanying fears proved unfounded, however, since the code wasn't completely "broken" for another ten years. Entirely synthetic life has yet to be created, and today, despite tremendous strides in genetic engineering, there is a general disaffection with the code. It appears that the code alone was not enough to allow man dominion over nature. The full glory of protein synthesis remains a mystery, so we have now moved "past the code" and on to proteins themselves. According to conventional thinking, the genetic code is so simple and buttoned down that its logical foundation appears remarkably trivial. Instead, today's glamour boy is the protein – the idol to proteomics. It is the study of proteins and their many eccentric habits of folding that dominates the search. Proteins are so devilishly complex that "Breaking the protein" makes "breaking the code" look like child's play. Fortunately, we have a tremendous amount of technology to help with the task as compared to 1960, and some of the greatest scientific minds are focused on a solution.

Surprise!  
The genetic code *is* child's play.  
Enter the child.

A funny thing happened on our way to dominion: somebody... everybody forgot to "break" the *other* half of the code. A central premise of this book is that the genetic code is far different from our conventional view of it. This book attempts to illustrate this "obvious" fact, and the implications of having missed it. We also intend to swing a machete in the general direction of any sacred cow that ambles into view.

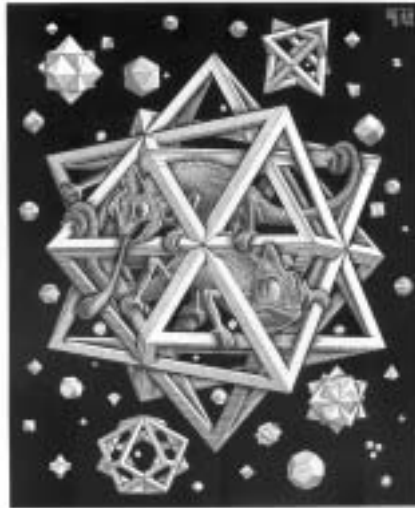
That's how children are – childish.

This is a story, a mystery adventure and a thriller involving molecules, ideas and people. We want to include all of the relevant characters in this mystery, so unfortunately we cannot avoid describing the molecules. I will attempt to make the nagging molecular aspects of this drama as palatable as possible, but no gain is without some pain, so hang in there. Mitigating the pain will be some visual pleasure in the form of art.

Semiotics is the study of meaning, specifically non-verbal meaning, as opposed to semantics which deals entirely with verbal meaning. Some concepts just cannot yield to words, especially for some people. Semiotics of molecules is a huge player in this mystery. Fortunately, we have art. Art is semiotics, and therefore we cannot possibly ignore it here.

Since molecules cannot read, and you are reading, you are not a molecule. (N'est pas?) You must begin to think like a molecule, so we will use art to bridge the gap between human thought and the thinking of a molecule. There is plenty of splendid material from many fabulous artists that we shall draw upon in making parallels, but I favor a few. Salvadore Dali, the surrealist we have already encountered, and M. C. Escher, the graphic geometer both have a powerful patch into the semiotic bandwidth of our strange and beautiful universe.

Another artist, Michael Teague, is a lesser-known contemporary of mine who has the ability to jar the senses, transporting the viewer into an alternate universe where things aren't as they might appear. In this altered state we find the ability to recognize the pieces and reassemble them into a less familiar, yet more comprehensible whole. Plus he's just plain entertaining. There will be others along the way; the common thread is the element of surrealism of patterns.



M, C. Escher  
Stars

## Amino Acids

We must begin our tale with some yucky school-like basics about chemistry. I promise to be brief, and recess will be here before you know it.

Proteins are large molecules – macromolecules - long chains of building blocks called amino acids. All organisms use the same twenty amino acids for the construction of proteins. They are strung together by connections called peptide bonds, so we can refer to them as polypeptides. There are many more than twenty amino acids available in the universe, but these are the twenty we always find in proteins. These twenty amino acids are therefore called the *common* or *standard* amino acids, all of which are a variety known as  $\alpha$ -amino acids. An  $\alpha$ -amino acid is a relatively simple organic molecule with a central carbon atom, the  $\alpha$ -carbon, to which four additional atoms or molecules are attached:

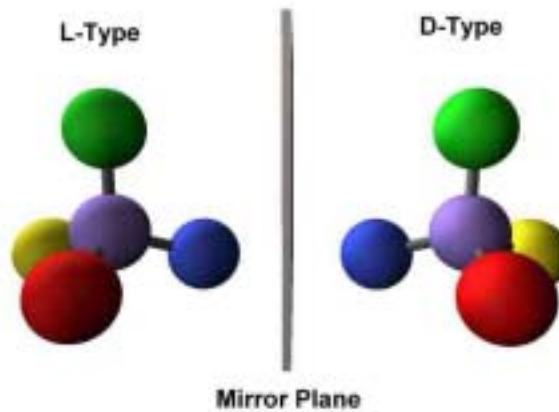
- amino group ( $\text{NH}_3$ )
- carboxyl group ( $\text{COOH}$ )
- hydrogen atom ( $\text{H}$ )
- functional group ( $\text{R}$ ).

It is the R-group that varies from one amino acid to the next, and therefore defines every standard amino acid from its brothers. There is considerable variation in the properties of the R-groups, creating variation in the properties of the standard amino acids, and a good deal of sibling rivalry, I'm sure.

Our chemistry here is kept brutally simple. Pictures and colors will be the standard mode of communication, and these pictures will generally treat atoms as balls and molecules as shapes. Since all standard amino acids are molecules with the same basic structure, the same basic shape will represent them all - a tetrahedron.

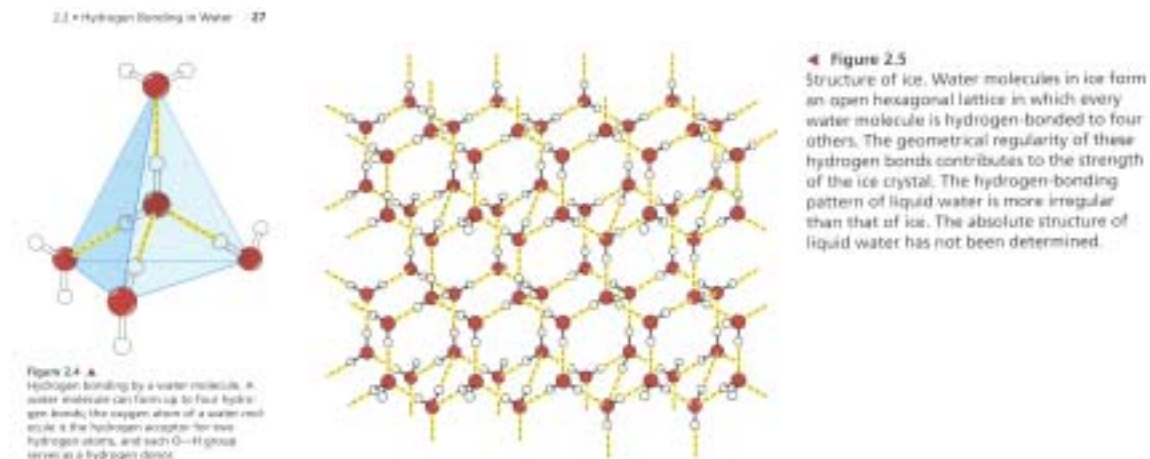


The  $\alpha$ -carbon lies at the center of four vertices in the tetrahedron. In nineteen of the twenty standard amino acids (the exception being glycine) all four vertices are different from each other. Therefore, there are two possible forms of these amino acids. This is because there are two possible ways to label the points of a tetrahedron: **Green-Red-Yellow-Blue**; and **Green-Red-Blue-Yellow**. All other ways can be rotated into one of these two. The two forms of each standard amino acid are mirror images of each other, and both exist happily in the universe. By convention, one form is called D (dexter, right) and the other is called L (laevus, left) but all of the standard amino acids are L-amino acids. This is most curious, don't you think?



## Water

Organic chemistry is the study of carbon-based compounds. Biochemistry is the study of carbon and water. Water accounts for 60 – 80% of all living things; therefore, it is fair to say that biochemistry is the study of complex carbon-based molecules arranging themselves in a water matrix. Carbon, as we have seen, is a tetrahedron. Water molecules also form a tetrahedral structure. Therefore, at the simplest level biochemistry involves the arrangement of tetrahedrons within a tetrahedral matrix. (I told you this would be simple.)



There are so many remarkable chemical properties of water that descriptions of it could fill oceans. Suffice it to say that of all the chemical ingredients of life, water is the key ingredient.

## Nucleic Acids

Nucleic acids are the grand conductor of biochemistry. They somehow orchestrate the assembly of the complex carbon compounds within a water environment. They are the heart of the genetic code. Two highly similar types of nucleic acids are important to life: DNA and RNA. Together they direct the selection and positioning of tetrahedrons within the tetrahedral matrix. DNA and RNA comprise a three dimensional system of storing and using information required in building complex organic molecules - proteins. Each link in the protein chain - each amino acid - is specified by a sequence of three nucleic acids. These nucleic acid triplets – tri-nucleotides - are called codons.

Just as proteins are chains of building blocks, DNA and RNA are chains of nucleic acids. Amino acids and nucleic acids both form sequential macromolecules, which is exceptionally useful to Life. Similar to the standard set of amino acids, there is a set of nucleic acids in DNA, comprised of four distinct building blocks: Adenine (A), Guanine (G), Cytidine (C), and Thymidine (T). RNA has the same four as DNA with the one glaring exception that Thymidine (T) is replaced with Uracil (U). These blocks are called bases, and they represent the letters of the genetic alphabet. Information is somehow passed to amino acids



by first grouping three bases together to form a codon. Therefore the genetic language is seen to have four “letters” and sixty-four “words”. A genetic numbering system within these parameters can represent 64 values.

Since there are only 20 standard amino acids and there are 64 possible words to describe them, the genetic language has considerable redundancy. This means that most amino acids in the common set have more than one genetic word to describe them, but no amino acid has more than six. The following is a table that lists all 20 standard amino acids and their codon assignments.

**Table of codon-amino acid assignments**

Amino Acid	Codons					
Isoleucine			AUU	AUC	AUA	
Phenylalanine			UUU	UUC		
Valine			GUU	GUC	GUA	GUG
Leucine	UUA	UUG	CUU	CUC	CUA	CUG
Methionine						AUG
Tryptophan						UGG
Alanine			GCU	GCC	GCA	GCG
Glycine			GGU	GGC	GGA	GGG
Cysteine			UGU	UGC		
Tyrosine			UAU	UAC		
Proline			CCU	CCC	CCA	CCG
Threonine			ACU	ACC	ACA	ACG
Serine	AGU	AGC	UCU	UCC	UCA	UCG
Histidine			CAU	CAC		
Glutamate					GAA	GAG
Asparagine			AAU	AAC		
Glutamine					CAA	CAG
Aspartate			GAU	GAC		
Lysine					AAA	AAG
Arginine	AGA	AGG	CGU	CGC	CGA	CGG
STOP	UGA				UAA	UAG

This table, in its various forms, has today come to represent the entire logic of the genetic code. I arranged this table on a somewhat eccentric scheme, one that will become less confusing as we progress. The important thing to note, however, is that we can arrange this table any-ol-way we like. There is no “correct” way to arrange and display this data according to our conventional view of the genetic code, and many different ways are in use. Since we can’t say for sure where nature got the data to begin with, and we believe there is no absolute meaning in its arrangement, we are free to view the organization of this data as arbitrary. This is strongly related to the paradigm that the genetic code is “one-dimensional”, which means that it contains only one dimension of information.



communicated? If one accepts conventional wisdom, the answers are, “not much, and with simple linear correlations.” These answers are incorrect, and the insistence that we cherish them as we have for so long has led to a truly comical view of the genetic code. More comical is the defense of it, as history will record. Despite a mountain of data to refute it, the genetic code is seen as “linear” in the sense that it is one-dimensional. The whole of science is in the trance of a more than forty-year post-hypnotic suggestion, causing obvious anomalies of the paradigm to go unnoticed. This is most unfortunate, so it is our job to correct it. We will start with some basic questions.

- What is the origin of the language, or how did Life get started?
- With so many  $\alpha$ -amino acids to choose from, and room for 64 in the code, why does the standard set only contain 20?
- What is the logic behind the arrangement of nucleotides, codons and amino acids?
- Since the mirrors of  $\alpha$ -amino acids (L and D) are equally stable and exist in equal proportions within the abiotic areas of the universe, why are all of the standard amino acids in the L form?
- In such a beautifully rapid, accurate and efficient information system, why is there such an ugly redundancy?
- With few exceptions, the above system appears to be used in all species and presumably back through time. Given the ravages of evolution - changing properties of organisms rapidly and constantly - one might expect some branching into competing dialects of the genetic language. At least the redundancy of the language should be subject to widespread change, since it has no absolute meaning. How could this exact system exhibit such dogged durability across time and throughout species?
- We now know the shape of DNA – a double helix – and we know the functional significance of this shape, but this is only genetic storage. After all, it is a complex 3D information system, and shape imparts structure, function and meaning. What is the fundamental shape and meaning of the genetic code when it performs its magical role during protein synthesis?

## Answers

To pick up a good biochemistry text today one might imagine that either there are answers to these questions, or the questions are unimportant, unworthy of real answers. To wit:

Why only 20? “The fact that all living organisms use the same standard amino acids in protein synthesis is evidence that all species on Earth are descended from a common ancestor.”

Why all L-amino acids? “Like modern organisms, the last common ancestor (LCA) must have used L-amino acids and not D-amino acids.”

And this is from an otherwise excellent book!

That’s it? That’s the best we can do? At least say, “we don’t know and we don’t care.” We mustn’t pretend to know, or imply it’s unimportant that we don’t know. These aren’t answers; they are fables. They are known as “just so” stories. They equate to, “they are because they are, and they must need to be because they are.” Not knowing these answers is a very unsettling concept for a lot of very intelligent people, creating more than a small component of denial.

Furthermore, there is another anomaly, a gaping hole so to speak in the same texts. Leaf through them and what do you see? Information, lots of it and presented beautifully. They illustrate an abundance of knowledge representing some of the greatest achievement of human thought and investigation. The trend is toward shape, fit, three-dimensions. There is a tip-of-the-cap to the idea that the meaning of the things lies in their shapes and their space occupying attributes. Some even provide 3D glasses, and most offer links to animated web sites to facilitate the spatial effects. The double helix is celebrated and dissected. Proteins are unfolded, folded and fit together. Electron prowling domains are drawn and speculated upon. Yet at the point where the rubber hits the road, where the genetic code performs its magic, the descriptions revert to 1930’s flatness and they are presented in living black and white.

“Toto, I have a feeling that we’re not in Kansas anymore.”

Dorothy  
The Wizard of Oz

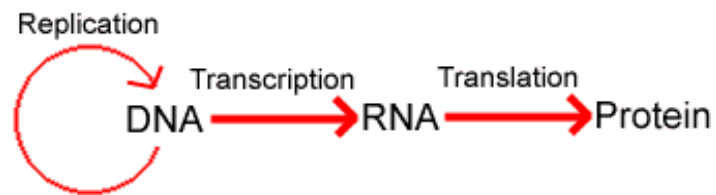
How can it be that the double helix has this wonderful relationship between form and function, yet the nucleic acids have no form-function relationship during protein synthesis? Discovering the double helix and proclaiming that the “shape” of the genetic code has been found is analogous to describing the hard drive of this computer and saying that the CPU is not important. Certainly the form-function of our genetic information storage is important, but what about the genetic processor? It is likely that it has a distinct shape as well, and the shape of the processor is somehow logically related to genetic information storage and retrieval. What is historically responsible for this omission and the resulting confusion?

A prime suspect is a culprit called *the central dogma*. One must have big, brass-like cajones to name a scientific theory “the central dogma”. The whole point of science is to shun dogma in favor of inquiry and investigation. The

dogma moniker is a gauntlet, an a priori challenge to science that says, “attack me, I dare ya”. It must be something so obvious that it is unassailable: right? Like the central dogma of astronomy where the earth is flat and the sun orbits the earth. It would take a real fool to approach the church on this one. Somebody’s got to play the fool, and I can think of no better candidate than yours truly.

Copped from the Internet:

The central dogma of genetics is essentially a "roadmap" of **how information within the DNA is transferred to protein**, the building blocks of your body. Here is a diagram of this process, "the central dogma of genetics":



The foundation of the central dogma is that the information is “co-linear”. In other words, there is a line of information in DNA that is communicated, somehow, to a line of results in proteins. This is taken to mean that the code is one-dimensional. There is believed to be only one dimension of information passed from line to line - the one dimension being the identity of amino acids or links in the protein chain. Due to faith in co-linearity, and due to the nascent digital information industry in the 1950’s, the code itself came to be seen as linear. As I’ve already stated - **this is a big mistake**. There is nothing really linear about the code, unless you believe that a swarm of bees is in some way linear. Nature ignores lines; it’s all about shapes.

The existence and translation of information in matter is not mystical. It is a nitty-gritty process of quantizing and selecting possibilities from a defined set of possibilities. The mysticism lies in the process by which the universe methodically bootstraps information in an evermore-complex cascade of emergence. The correct term, I believe, is sequential. I will grant that the genetic code is to an extent co-sequential, but I will not concede that it is co-linear, because these illusions of linear paradigms are clouding our eyes and our brains. The linear indoctrination process is intense; I know, because I’ve been through it. However, we can loosen the reigns on our senses and find some sense in the madness. With the help of some recent discoveries, some clear, rational thinking, and some bodacious art, we can see the order in the chaos.



M.C. Escher  
Order and Chaos

In the middle of a table of periodic elements sits logic. As with any pattern there is an organizing force, something that drives the formation of complexity and order. The laws of nature are in play, and the patterns are there for us to see, if only we have the light and courage to look.

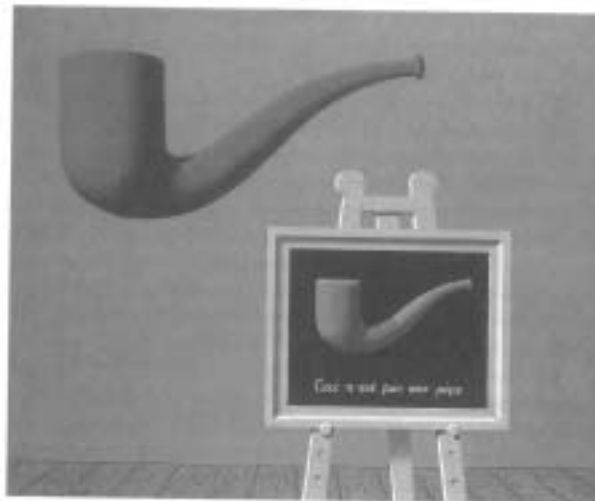
## Language

The codes of language and reason – as well as all other codes – are *constituted, instituted, constructed* and, therefore, presuppose something that is not coded, order is always implicated in something that can never be ordered, order is always incomplete and codes are inevitably partial.

**Mark C. Taylor**  
**The Moment of Complexity**

There is an antecedent to the genetic code, but what is it? Without one there could never be a code, let alone a mechanism to build one. The code must be built by and around a logic structure – it must be a manifestation of that logic. Just as Life extracts oxygen from air it extracts logic from the code.

Language informs thinking. We know this intuitively, but in the pursuit of a genetic language this is a much bigger issue and a huge potential problem. All languages are complex, and the genetic code is no exception. In this book we are investigating a flaw in the conventional view of this language. Specifically, we suspect that the conventional view is overly simplistic and therefore incomplete in how it informs us. We are using a completely different language, or languages, to describe **and** solve this mystery. This issue is an intractable problem for any profound inquiry into the processes of Life or thought, but especially so with respect to the genetic code. All languages, especially natural languages, inevitably create ambiguities and misperceptions. Our thinking can just as easily become *misinformed* by language, and it usually is. Unfortunately, there are few options. Sometimes our best forays into epistemology come in the form of graphic images.



© 1964 by René Magritte, Two Mysteries (© 2003 E. Hazzard, Museum of Modern Art, New York)

Rene Magritte  
Two Mysteries

“What we have here is failure to communicate.”

Captain  
Cool Hand Luke

In the case of the genetic code, we are contemplating a molecular system, representing it in another medium - human language - and then using this representation as a functional artifact, or “the genetic code”. This is a complex hierarchy of symbols precluding any one level from fully encompassing any other level. Despite this, the system can be coherent, and a global intelligence observing the operation of the system can find the logic in the whole. Dali excellently provides a test of our ability to appreciate this concept: *Venus with Drawers*, a painting of a famous statue of a human - used pragmatically for storage. So which is it, paint – stone – human – fame - dresser?

TOP RIGHT  
*Venus de Milo with Drawers*, 1936



## Answers Revisited

Intuition predicts that the answers to our previous questions about the genetic code will share a common foundation. That foundation should reflect something sublime, fundamental and unchanging in the fabric of our universe. In the parlance of a linguist, we expect a synchronic language, as opposed to a diachronic language. Since DNA is a three-dimensional information system, and shape is all-important, there should be answers to these questions somewhere in a study of geometry.



It is helpful to take a pragmatic approach initially, one that is teleological. We can frame the questions in terms of the “intent” of the genetic code. Life is a system of building organisms through the block-by-block construction of complex molecules. The blocks are amino acids and the instructions for the addition of each block is contained in nucleic acids. How could it do this - what strategy might it take? Does Life intend to build a particular structure? If so, how does it know and execute its intent? Surely there is not a master blueprint for the working whole of any organism, but just as surely there is some form of blueprint on some level. There must be some logic to the construction methods Life employs; otherwise, the building process itself would be random and unrepeatable. Life, we know, is the contrary. It is precise and repetitive. Trillions of cells in a single organism can consistently produce identically well-defined complex molecules. The element of chance must be small, and control over the process approaches absolute.

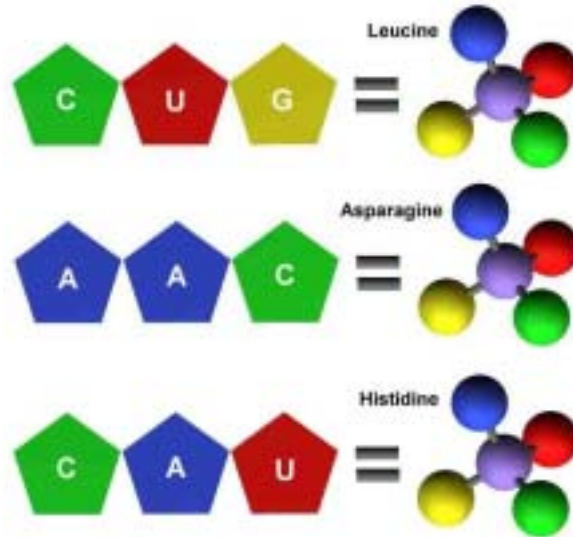
It seems that from a three-dimensional structural viewpoint the “meaning” of each block should be contained in the instructions themselves. Since the consecutively added blocks are all tetrahedrons, and they are arranged in a tetrahedral workspace, it seems that the instructions should reflect something useful about the arrangement of tetrahedrons in a symmetric world.

Taking the purely teleological position of designing such a system, we start with what we know nature has actually done, as if this is what it wanted or intended to do. The genetic language has only four letters: U, G, C, A. These letters are permuted in sixty-four unique triplets, or words. If we somehow were able to identify sixty-four unique tetrahedrons in nature then our system could specify each one with a single unique word.

In reality, each letter of the genetic language is a nucleic acid, which is a five-carbon sugar attached to one of four nucleotide bases. We are now wading knee-deep in the Venus with Drawers phenomenon of languages. In merely defining the genetic language we are required to employ symbols from several levels of the linguistic hierarchy. Confusion is inevitable. We can help our cause if we simplify our usage of symbols and identify each base schematically with a colored pentagon as follows:



The five-carbon sugars are rings that link together to form long chains. Groups of three consecutive rings form a codon. Each codon defines an amino acid from the standard set. Of course there is a complex and elegant process of translation passing the information through various forms, but for now we will keep it simple.



These schematics - pentagons in groups of three that define an object - immediately suggest another naturally occurring phenomenon. They suggest a dodecahedron. Three pentagonal faces meet to define a point.



Gosh, Dr. White, that's really keen, and thanks for the geometry lesson, but what could this crude association between real things and idealized schematics possibly have to do with Life on this planet?

We *don't* know... yet, but hang in there. We are in the process of building a system of symbols that is capable of holding our language and concepts, and I promise you that it will profoundly inform our thinking. Our system is about shapes, so we will use shapes to build it. Shapes are called polyhedrons, and we are going to use this system to speak about shapes so we will call the language polyhedrish. We'll first create a four-color dodecahedral system and see how she flies. Consider the following parallels:

- Nucleic acids are five-carbon rings, so we can equate this to the five-sided face of a dodecahedron.

- The genetic language has selected only four bases, so we can equate this to four possible colors on any face of a dodecahedron.
- Bases are grouped in threes to define an amino acid, so we can equate this to the faces of a dodecahedron grouped in threes to define a vertex.
- There are  $4^3$  or 64 possible permutations of triplet bases, which is the same as the maximum number of ways to define a vertex in a four-color dodecahedron.
- There are 20 amino acids in the “standard set” so we can equate this to the 20 vertices of a dodecahedron.
- Once a point has been identified as “special” three other points will potentially become “special” three-face groupings as well. In this way two points can completely define the dodecahedron. There is a set of only six dodecahedral configurations made in this way from the primary point, which is equal to the maximum redundancy for any amino acid in the genetic code. Ultimately, one point can be made to define the entire dodecahedron in six different ways.

If nothing else we can win bar bets with the following question: can you name two physical systems that incorporate the following set of numbers – (3,4,5,6,20,64)?

Yes, I can: our genetic code and a four-colored dodecahedron.

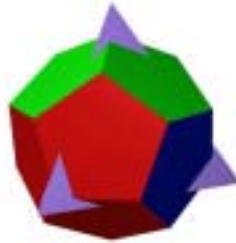
Whoa, Dude, that’s pretty salty!

This obviously proves nothing, but it is very cool, and we can see it as a **CLUE** for potential answers to the questions we are asking. More than that, it provides us with a solid foundation for a system of symbols to describe the components of the genetic code. At the very least, we can now have a visualization of the code that is brutally simple and imminently useful. However, like any good potential answer, it raises more questions. My first question - what is a dodecahedron anyway? Plato said that it was the cosmos, but what did he know?

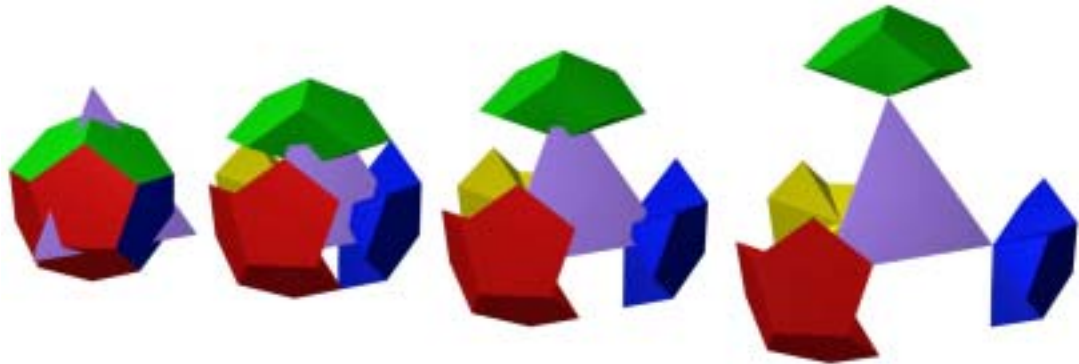
Consider this: a dodecahedron is a tetrahedron in drag.



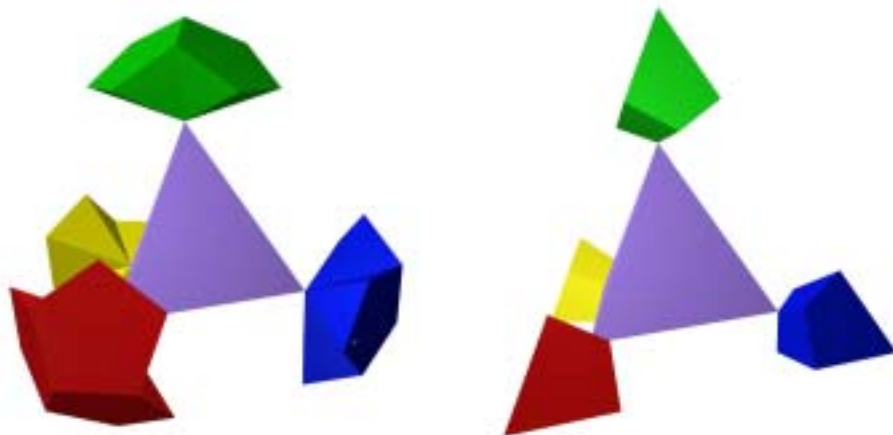
Starting with the above dodecahedron, it is easy to imagine a tetrahedron sharing the same center, and four vertices poking through the dodecahedral vertices.



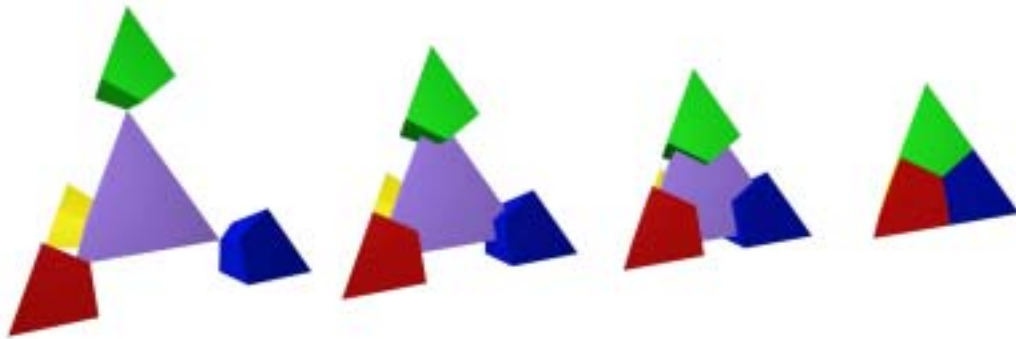
A more interesting way to look at this is as if the dodecahedron actually is a tetrahedron, one with fancy vertices. Each vertex is an identical collection of pieces that represent one fourth of the dodecahedron. This is easier to see if we explode the dodecahedron out to the points of the tetrahedron.



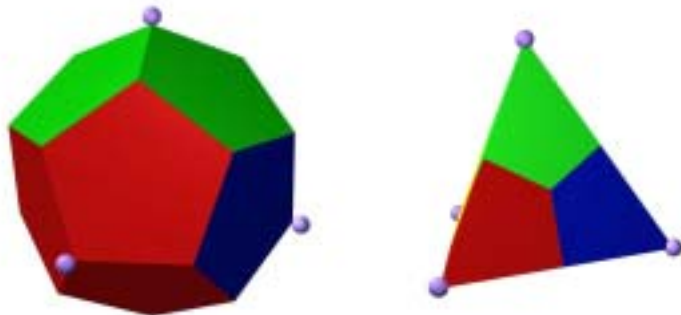
When we morph the four exploded dodecahedral fragments into four tetrahedral fragments we will see the following.



Sucking these back to the center of the tetrahedron will give us a new visual perspective on the dodecahedron.



We can add purple balls to the vertices of the tetrahedron to remind us that we are looking at a four-color dodecahedron melted around a purple tetrahedron.



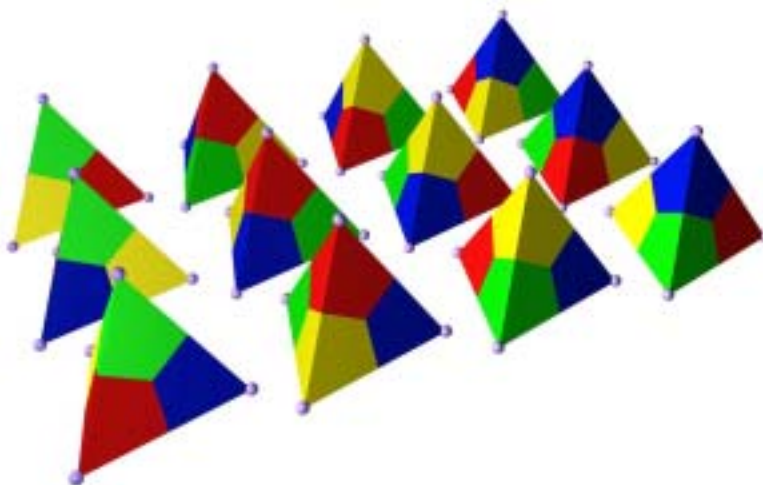
From here we can set things in motion. Every vertex has an axis through its opposite face. We can rotate the tetrahedron around any of these four axes in  $120^\circ$  increments. If we did this with our original purple tetrahedron, to our eye it wouldn't change, because it has no identifying markings. This is called symmetry.

**When an object can be transformed in some way but not fundamentally changed, then that object has an element of symmetry.**

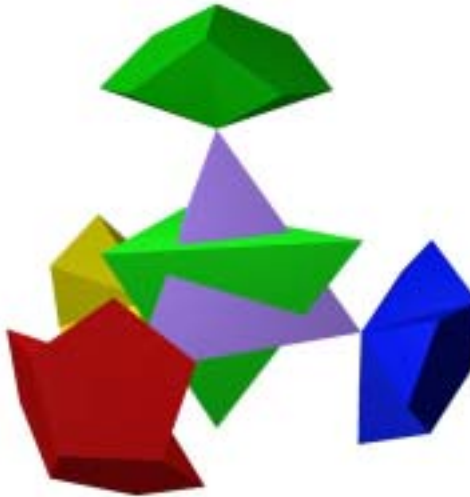
There are different kinds of symmetry, like rotation and mirroring already discussed, but we will only need rotation here. When we apply this rotational symmetry to our new, fancy purple tetrahedron we find the following.



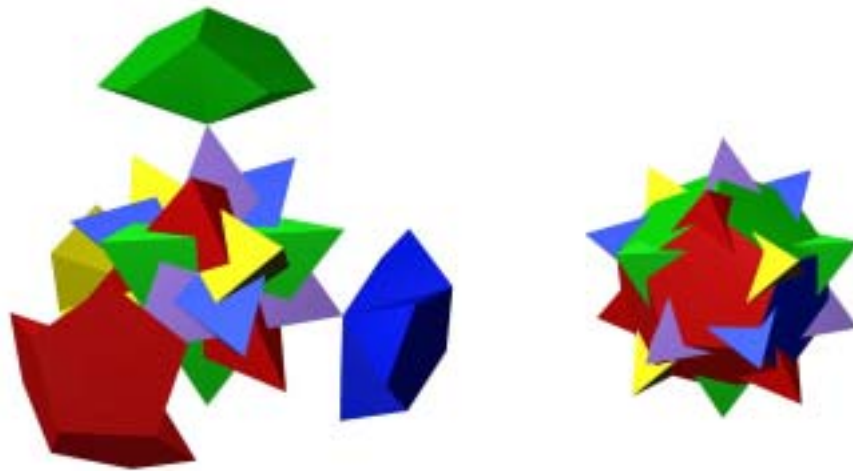
We see that by systematically utilizing all of the rotational symmetry in the above tetrahedron we can find twelve possible tetrahedrons occupying the exact same space as our original. In other words, there are twelve discrete combinatorial rotation symmetry options for any tetrahedron. A blueprint for building molecules by specifying one of these rotational options would be an inherently logical system. All that is required is a language that can communicate a specific choice from one molecule to another.



If we consider the mirror symmetry of the tetrahedron, we could produce a mirror twin of each of the above twelve, so then we would have 24 tetrahedrons potentially represented by just this one. As stated, we do not have those twelve mirrors available in this particular system. However, there is really nothing special about any of the four points of the dodecahedron that we have chosen to represent this particular tetrahedron, and we could have therefore just as easily chosen four completely different vertices. If we do so, we generate another tetrahedron, such as the following:



In addition to the original purple tetrahedron we have generated a second, green tetrahedron. The second one shares none of the vertices of the first. In fact, we could do this three more times without using any of the vertices more than once.

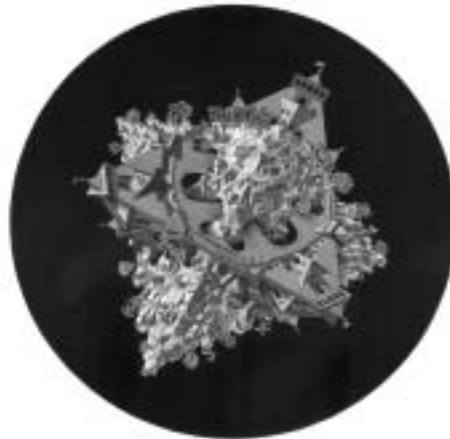


We now have five completely separate tetrahedrons located inside our original dodecahedron. Except for the colors we have assigned, all of these tetrahedrons are indistinguishable from the next. A tetrahedron is a tetrahedron, which means that each of these tetrahedrons can be rotated into twelve equivalent options as we did before, creating 60 distinct tetrahedrons. None of these five groups of twelve tetrahedrons share any points; therefore, none of their sixty rotational equivalents overlap either. The group is balanced, and as tetrahedrons they are all interchangeable. The only distinguishing properties are found in the colors we have arbitrarily assigned.

It is mighty convenient that five tetrahedrons with four vertices apiece can perfectly consume the twenty vertices of a single dodecahedron. But this is

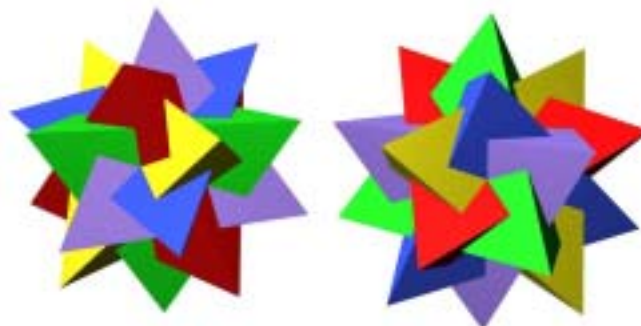


literally only half the story, because if we go back to the first step in the process, the one where we placed the first tetrahedron, we can see that we actually had two distinct choices. It is true that we could have started with any of the above five tetrahedrons and ended up with the same final formation, but each of the five has a non-equivalent twin called a dual tetrahedron, and the two of them together form a cube.



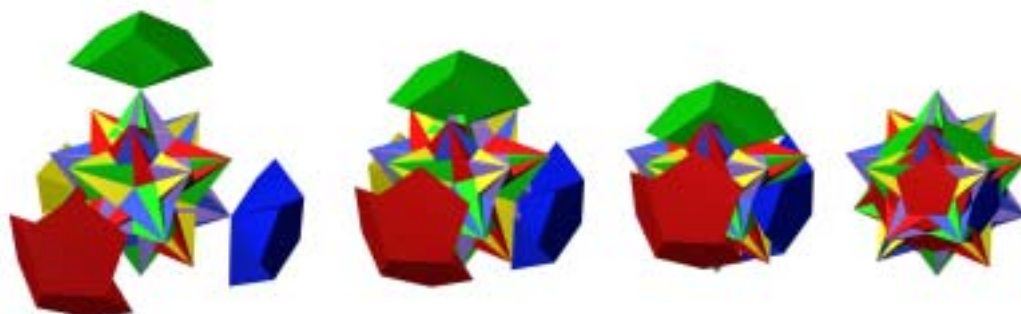
M.C.Escher  
Double Planetoid

Every one of the five original tetrahedrons has a dual, so we can repeat the above steps of adding one dual twin tetrahedron for each in the original set. We end up with a configuration of five tetrahedrons that is a dual twin of the original five.





Of course, each of these dual tetrahedrons has twelve rotational equivalents, so we are adding 60 new tetrahedrons to our original 60. We now find a total of 120 distinct tetrahedrons in the dodecahedron. When we combine these dual configurations together and suck the dodecahedron back around it, we see the following.



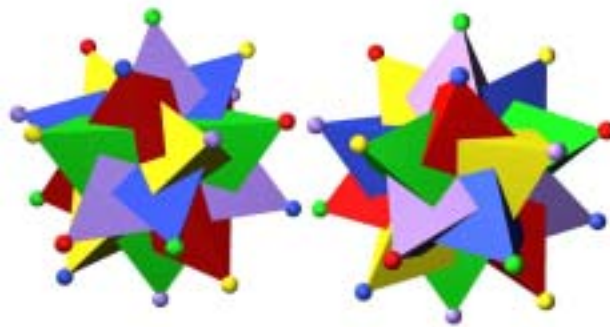
This suggests that either the points of a dodecahedron provide a logical way to group tetrahedrons, or a tetrahedron provides a logical way to group the points of a dodecahedron – or both. Regardless, I am sure we can find a use for this somewhere in our quest for a logical way to build a molecule. However, with all of the fun we've been having adding cool new tetrahedrons to our dodecahedron, we failed to realize that we have created a serious problem. Who can keep all this crap straight? I don't know about you, but this looks like a jumbled mess to me, so let's try to clean it up a bit, shall we. Start with the fact that all of the tetrahedrons have duals, and duals make cubes, so we can just as easily view the dodecahedron as five cubes.



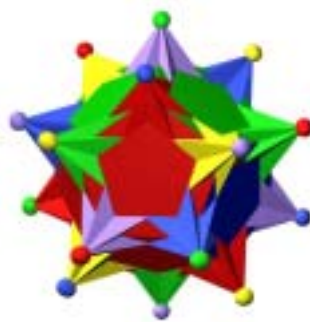
That didn't help much, because although there are only five cubes here, they are still too difficult for lowly humans to comfortably differentiate. Notice, however, that every vertex of the dodecahedron is a composite of exactly two cubes, and the faces form five-color stars. Perhaps we can use this fact to take advantage of the dual tetrahedron cubes. We can add colored balls to the first purple tetrahedron and replace its dual with our melted dodecahedron.



The two dual purple tetrahedrons are now more individually complex, but they are also more identifiable because all eight vertices are in some way marked by one of the other four colors in their global configuration. Now the dual configurations appear as follows.



It's hard to believe that this new complexity will mitigate our confusion about identifying tetrahedrons in a four-color dodecahedron, but in fact it simplifies the process in fabulous fashion.



We now have strong visual cues to identify the original five tetrahedrons and their five duals. More importantly, each of these ten tetrahedrons has a pattern that will allow us to identify its twelve rotational equivalents. We now have a picture of the dodecahedron that will allow us to identify 120 equivalent alternate representations of that dodecahedron using tetrahedrons. Where could we possibly find a need for such a thing?

## Data compression.

The dodecahedron is a natural and phenomenally good compressor of tetrahedrons. In virtually no more space than a single tetrahedron, a dodecahedron can be made to represent 120 unique tetrahedrons. The difference between one tetrahedron and the next is simply its spatial orientation. A simple language describing dodecahedrons allows us to easily and powerfully talk shapes with tetrahedrons.

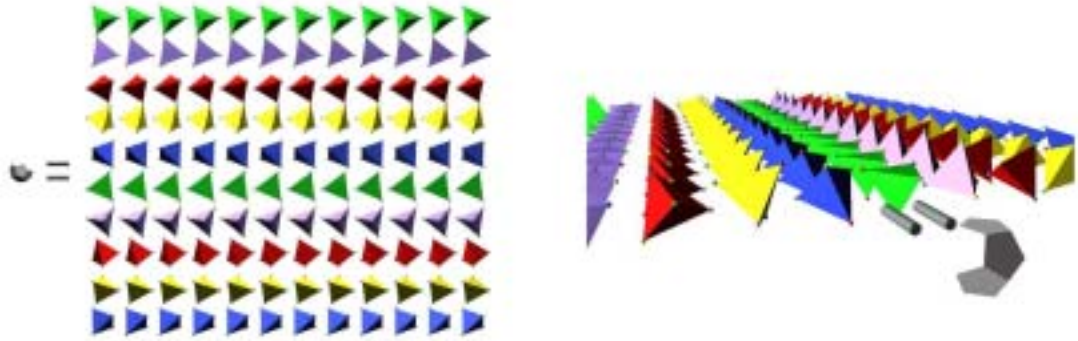
Look at it this way (you might regret this). Let's say we go to the store and buy 120 tetrahedral dice and one dodecahedral die. We paint each tetrahedron in one of ten patterns, and then, based on its pattern, we physically orient it in one of twelve ways. Rather than hold it there indefinitely like an idiot, put each one on a tiny little stand so it will be preserved for posterity. We started with 120 identical dice and industriously created a collection of 120 unique, identifiable objects.

Now, invite all of our friends over to marvel at our organizational and painting skills. They will feign interest and patronize our enthusiasm, perhaps recommend a fine Lilly product - Zyprexa, maybe high doses of Prozac. Remain undaunted. We carry our tetrahedrons with us everywhere; discuss them incessantly - *even* try to *build* things with them. They are everything to us, but we soon tire of the burden. These damn tetrahedrons are taking up so much space. If only we had a simple way to describe a single one out of the many.

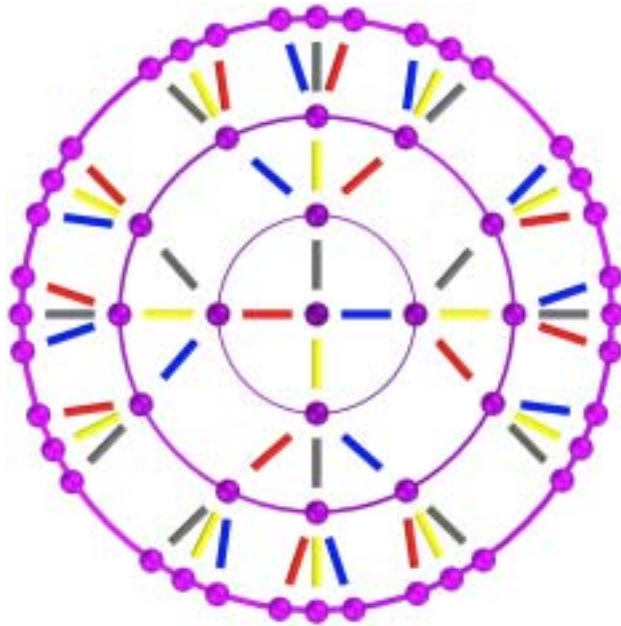
But wait! Suddenly it hits us like a nightmare - the dodecahedron!

If we had a simple language of the dodecahedron we could replace all 120 tetrahedrons in this pesky fanny pack with a single dodecahedron in our pocket. Perhaps a series of colors on the faces would indicate vertices. Knowing the language, we could quickly and accurately orient the dodecahedron according to a prescribed code. Then the dodecahedron would come to literally "mean" the individual tetrahedron and all of its exact angles, just waiting to explode from the dodecahedron.

I am hurt that you even wonder whether I am nutty enough to have attempted the above. Well... I am nutty enough, but at \$0.40 apiece, dice are too expensive. How about some pictures?



But this is only the beginning of all the fabulous tetrahedral information stored inside a single dodecahedron. Remember that we found five non-overlapping tetrahedrons originally, and we added five additional tetrahedrons as duals. We now have forty tetrahedral points (10 X 4) but only twenty points in the dodecahedron. This means that each tetrahedral vertex is linked to one other tetrahedron from the dual set. Therefore, every tetrahedron is linked to the four dual tetrahedrons in its five-tetrahedron configuration. These linkages form a network of tetrahedrons. We can walk this network, one link to the next, and get from any of the 120 tetrahedrons to any other in six links or fewer.



Above is a map of just three of the six possible steps in our walk from any given starting place (each sphere represents a tetrahedron, and each color represents a step). A new “best” map must be created with each step. All maps should contain at least six rings, not just the three we have space for here. This is an exceptionally complex relationship between these two shapes, but it is the only way to precisely detail the relationships and distances between tetrahedral

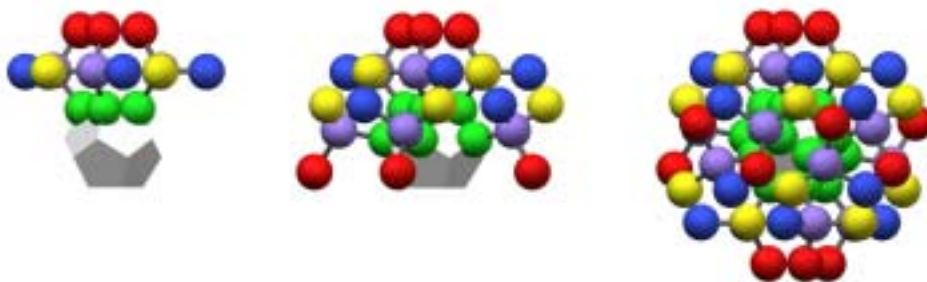
walks in a dodecahedron. There is so much complexity here that it is hard to imagine any regularity to this relationship at all. But all of this complexity could be handled by a permutation set of four colors grouped in threes. Where might we find a use for this? The information needed to communicate these relationships in the language of a dodecahedron is tiny compared to the complexity of shapes that can be created by it. It is a system of coherent logic, and any construction blueprint based on it will inherit that logic.

## Another Angle

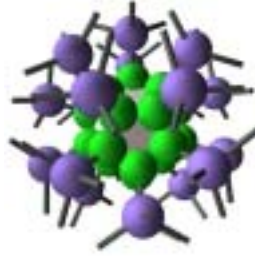
We approached the dodecahedron from the inside, but an equally viable approach is from the outside. The parallels that we drew earlier between dodecahedrons and the genetic code suggest to me that we try to schematically assign a tetrahedron to each of the twenty points of a dodecahedron. There are a tremendous number of ways I can conceive of doing this. Rather than search for a “correct” way I will do an experiment and assign them in a simplistic, stylized way to see what happens.



No meaning is implied by the colors on the tetrahedron other than to illustrate its chiral orientation. In this case the green point was arbitrarily selected to form the axis of connection to the point of the dodecahedron. Proceeding in the same fashion we will fill the dodecahedral points with tetrahedrons.



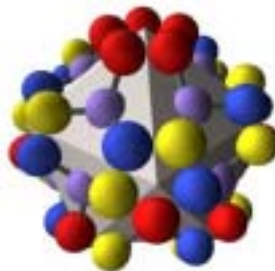
Very nice, but what is it? It is a cool collection of shapes, Dude. That's what it is. We can describe it as a dodecahedron with tetrahedral vertices - that is what we set out to build. We could also view it as three concentric dodecahedrons, one for each level of the attached tetrahedrons. This is a little easier to see if we strip away the outermost layer:



The inner dodecahedron defined by green spheres should be totally expected. After all, the first step in the process was to put a green sphere at every vertex of a dodecahedron. The second layer made by the purple centers of each tetrahedron is also pretty much what you'd expect. The outermost layer is less intuitive. It is not clear to me what to expect from the properties and behavior of this layer. For me, it is helpful to take advantage of the duality between a dodecahedron and an icosahedron, converting the purple layer into an icosahedron:



Now, instead of vertices, the purple centers of the tetrahedrons represent the centers of faces, and the relationships between the shapes becomes obvious. If we add back the outer layer we can get a better understanding of the relationships between layers:



Forgetting for one second that we are actually trying to idealize something with physical meaning (the genetic code) we can examine the number of ways to run this experiment and create a different pattern. Consider first that we could have two types of tetrahedrons: an L and a D-type. Either type could fit into its spot in the above pattern in one of twelve different color configurations; therefore, each tetrahedron presents us with 24 options (2 X 12). This assumes, of course, that the purple centers are fixed and the colored vertices are interchangeable. If



we consider the above pattern as an icosahedral frame holding tetrahedrons at their centers, we can imagine one of 24 orientations for each of the 20 tetrahedrons. There must be  $24^{20}$  possible patterns. Thank goodness I didn't spend any time trying to find the "correct" pattern. It might have taken all day!

However, in the real world of biochemistry there are no idealized tetrahedrons, each vertex is not interchangeable with the next. Not all vertices work and play well with others. There *are* limits, and choices must be made. One choice that we could easily make, if we were molecules, is to eliminate the mirror symmetry - only consider tetrahedrons of one type. Let's pick one. I arbitrarily pick the L-type. Having done so, we have a much better chance of fitting our non-idealized biochemical tetrahedrons into, and inter-changing them within the idealized pattern. Albeit, our options for pattern formation are trimmed in half, but we started with such a huge number there is still plenty to go around.

There are two important points to these geometric exercises. First, regular solids - like tetrahedrons and dodecahedrons - love to pop up in situations where space gets filled by stuff. Second, very simple geometric models develop a capacity for staggering complexity very quickly. The actual number of possibilities explodes within a short combinatorial walk, and "number of possibilities" combined with the ability to specify one from many means that information is created. Languages are required if this information is to be communicated. These shapes supply a huge information capacity, as we will see later.

The first series of diagrams demonstrated that dodecahedrons carry 120 tetrahedrons internally. The second series shows that a dodecahedron or icosahedron can arrange a staggeringly large number of tetrahedrons externally. Together, the two concepts point out that from simple shapes complex patterns quickly develop. A simple language that can juxtapose one shape with another would be very powerful in generating a diversity of shapes. This diversity will be hierarchical in nature, as Escher aptly suggests:



M.C. Escher  
Concentric Rinds

A diverse pattern generation system welcomes Escher's concept of concentricity. We are now in the realm of our third rule of patterns, which is that patterns will aggregate. As we saw above, a dodecahedron can surround itself with a layer of tetrahedrons and generate a huge capacity for pattern generation. But this is only the first concentric layer, and each tetrahedron, we have learned, can be thought of as a dodecahedron, and the process can begin anew twenty more times. The complexity explodes from a single dodecahedron, but there is a simple language of shapes that could handle it.

Polyhedrism is the name we have given this language of shapes. Later we will delve more carefully into information theory to see if the infrastructure of polyhedrism is truly appropriate to carry the information we speculate is contained in the genetic code, but for now let's assume that it is a viable mechanism. Perhaps Escher had polyhedrism in mind when he created the following:



Gravitation  
M. C. Escher

Emergency dispatch received a frantic call from a man out hunting. The man said that his friend had been accidentally shot, and the man feared that his friend was dead.

Dispatch calmly instructed the man, "first, make sure he's dead."

There was a brief pause, and then:

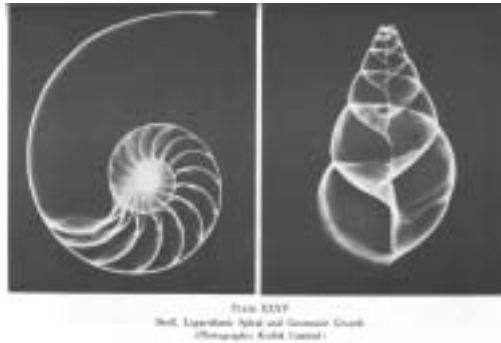
**BANG!**

"OK, now what?"



## Space Filling

We still don't have any useful answers; however, we might suspect that there is something here. It seems quite possible that these geometric coincidences have a universal impetus. To advance we must develop some hard-core abstractions regarding the process of Life.



Erwin Schrödinger was an Austrian physicist who shared a 1933 Nobel Prize for new formulations of the atomic theory. Around 1940 he trained his big brain on biology and famously asked the question what is life? Naturally he attacked the problem from the point of view of an atomic physicist, and he made the observation that life is an aperiodic crystal. Most crystals are periodic, which means that their structure repeats very regularly. Table salt, for instance, has a cubic repeating structure. Every atom of sodium and every atom of chloride have a cube in mind when they decide to participate in a salt crystal. Each atom in the structure follows a logic that is crystal clear, and every atom could answer questions about precisely who and where all of its neighbors are. Conversely, an aperiodic crystal must have a structure that does not repeat regularly. It is without a periodic structure. One could only imagine what an atom has in mind when it participates in an aperiodic crystal, and God only knows the logic it uses.

I happen to share Schrödinger's view in this context, but I would add a caveat. We must specifically identify the feature of Life that we believe is aperiodic. The conventional wisdom holds that the aperiodicity of Life is contained in the random appearing nature of the sequence of nucleotide identities in DNA. I disagree with this as the appropriate context for the aperiodicity of "a living crystal". The crystalline structure of DNA is brutally periodic, that's one of its many charms. The nucleotides in DNA might be irregular from one base to the next, but this is seemingly trivial in light of the fact that these nucleotide sequences, when taken as a whole from one cell to the next, are completely periodic. In other words, every cell is virtually identical in its sequence. Humans, for instance, have many more cells in their bodies than nucleotides in their DNA. It is more appropriate to conclude that the sequence of bases is actually completely periodic from one cell to the next. It is a point not worth arguing about, however, because there is a much better candidate for the

source of the aperiodic construction of crystals – their blueprints. A linear genetic mechanism, however, does not have the capacity for this type of behavior.

In contrast to salt, the structure of glass is random. The silicon atoms in your windowpane are a random, jumbled mess. Whereas sodium and chloride have a crackerjack drawing in their heads of the entire salt crystal, silicon has no clue about its surroundings in glass. It seems unlikely that glass is the best candidate for the molecular logic on which all life is based, because precision replication is out of the question and consistent protein synthesis is difficult to imagine. There is no inherent logic to the structure of glass. So if DNA is completely periodic, and proteins are synthesized with total precision, where is the aperiodicity in Life? The answer is in the translation from DNA to protein; it is in the genetic code. The operation of the code is completely regular. The output is precise but consistently irregular, so the logic of the code must contain the aperiodic component of the crystals of Life.

Schrödinger and the rest of the physics mafia have focused on two aspects of Life: energy and reproduction. I too am interested in both, but my focus here will only be on reproduction. Can you reproduce something that you've never produced? This is literally a chicken and egg question (egg is always the correct answer). What does it mean to reproduce; is it the same as replication? I don't know the answers to these things – I will leave them to the professional philosophers. What Schrödinger brilliantly identifies, however, is that we must find the crystal algorithm of Life. Crystals have maps and maps are algorithms. There are two problems with this task. First, he proposes an aperiodic crystal, which implies an aperiodic map. This must be one fantastic map. Second, the algorithm must be recursive; it must somehow get itself started, and it then must reference itself. These are two big problems.

My peanut-sized brain was briefly trained in the fashion of a geologist, which explains a little of my bent. Not only are geologists taught to appreciate scales of time, and to a lesser degree scales of space, we are taught to appreciate natural beauty. In the good ol' days we actually went outdoors. We also couldn't help but be exposed to fossils and questions of ancient life, and origin questions of Life and earth. (We also were taught to appreciate beer, which plays a role in a lot of my thinking.) I was killer with the patterns and the symmetry of mineralogy, and since half of my brain probably never got formed, the other half tries to fill in where it can. The pattern and symmetry half is doing double, perhaps triple duty. It is no surprise then that I see patterns and symmetry in everything – all the world is a nail.

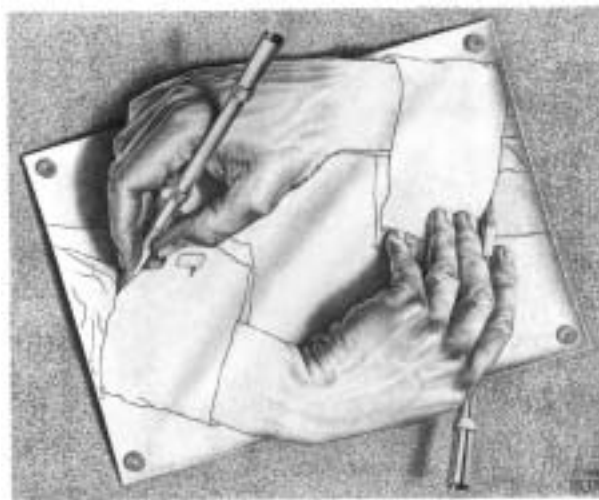
More abstraction is required, and the secret to a good abstraction is to make it as abstract as possible. We must strip away the shell, all extraneous higher order details, and try to peer down into the marrow, the most fundamental mechanisms of a system. An excellent example is a book written by Daniel Dennett called *Darwin's Dangerous Idea*. It is crammed with gold, and the take home message is that complex systems, such as Life, can be based on simple algorithms. Dennett describes how Life on earth can be seen as a type of sorting

algorithm defined by Darwin's description of natural selection. He does a fabulous job by describing, among other things, the library of Mendel. In a nutshell, the library of Mendel is a place to store books that represent all potential life forms. The information in the books is written with a fictitious set of genetic symbols given specific parameters. The key question is, by what metric do we interpret, value and sort this information? Darwin's brilliant contribution was to suggest a simple algorithm, a process we call natural selection.

But what is it in Life that is being sorted? In the library of Mendel, books can be sorted. Each one is identical on virtually every physical parameter, give-or-take some ink. Therefore, information contained in each book is the only basis for sorting. But what are the books "about" and can they "mean" anything? According to Dennett they are about every possible thing, every conceivable combination of ideas, events, symbolic patterns in every conceivable language. Some will have meaning and others will not. I believe, however, that every book in the library of Mendel has the exact same title:

### **How a book makes a book** by DNA

Dennett uses the library metaphor to make several excellent points; such as most of the books are total garbage (aren't they though). An excruciating few are pure gold, however, and they have perfect "meaning" for some purpose or another. The metaphor also brings to light the fantastic logistics problems presented by the library. The space and material requirements are unimaginable. The process by which the library came into being cannot even be addressed in the metaphor. In the library of Mendel the books themselves use matter to fill space. Parts make wholes, and wholes make parts. It is a mind-bending circularity; it is recursion.



M.C. Escher  
Two Hands

The real subject of every book concerns the “strategy” a collection of matter used to fill a portion of space. The system used to “publish” these books must be a space filling system. The symbols used must comprise some space filling language. However else the books can be evaluated or sorted, they must at least be sorted on a matter consuming, space filling metric. Despite the complexity of the results, the method of filling space must also be a simple, recursive algorithm.

In *Darwin’s Dangerous Idea* an abstraction is made of a library to illuminate the pattern behind all potential Life. I want to formalize this with a mathematical representation, or function, so we must symbolize some fairly complex concepts with compact notation. This is never pretty. The total pattern of organic matter can be symbolized as LIFE. We will call the function that created the pattern SORT, and we will say that it operates on something that we have yet to clearly define, call it “x”. We can write it as a symbolic formula as follows:

$$\text{SORT}( x ) = \text{LIFE}$$

The **SORT** function is a simple, mindless, recursive algorithm, described by Darwin, and operating in some way on the library of Mendel. But what is the Library of Mendel - what is x? I submit that x is itself a simple, recursive algorithm that functions to fill space. The substrate is space and the operation is FILL, so we can re-conceptualize the formula as follows:

$$\text{SORT}( \text{FILL}(\text{space}) ) = \text{LIFE}$$

A true, hard-core abstractionist would take it one step further and substitute variables to create an infinite recursion that could make any fractal blush.

$$\text{SORT}( \text{FILL}(\text{universe}) ) = \text{UNIVERSE}$$

A logician might view this as some tricked up lambda conversion. I view it as the most logical starting point for any analysis of the genetic code, because at the atomic level this is all there is. Also, this has bootstrapping potential, and it clearly is recursive. From this perspective we can begin to grasp the burden of any scientist. He must first carve an idealized section from this fractal, but no section can divorce itself from the bias of the whole. Even Einstein required an absolute vacuum, which can nowhere be found or created. All space in the universe is filled with something, if only gravity.

Through this function the universe creates an interwoven network of patterns. There are not patterns and gaps between patterns; there are only continuous interlocking patterns. A scientist must struggle to isolate any pattern, and even if he should somehow succeed his conclusions should expect to suffer from oversimplification. The nested function above calls for the logical synthesis of biology and physics - a system where Darwin iterates Newton in a quantum

universe. The details, one might imagine, are slightly more complex. We are merely interested in the logic that might spill from attacks at questions of Life from this direction. The universe is complex, and so is Life. We cannot shy away from the non-linear strangeness inherent in complexity if we hope to understand it.

You can't have everything. Where would you put it?

Steven Wright

The universe is a big place with a lot of stuff and nothing but time on its hands. The stuff now isn't where it was a minute ago, and it will be somewhere else a second from now. How does it know where to go? This is the question on everyone's mind in one form or another. The trick is to know the past so you can predict the future. Any rocket scientist worth his salt can tell you that. The only hope is in recognizing the patterns that nature makes. Regardless of the process, the result is that stuff goes places. A good way to keep track of it, or predict its behavior is to make understandable rules based on the patterns we are able to observe. The rules serve as instructions for predicting where stuff is and where it's going. Some of these rules are "common sense", intuitive, a priori, i.e. stuff falls down. Others are obscure, counter-intuitive and sometimes downright nutty (take a look at quantum mechanics). But the rules generally serve a purpose otherwise we discard them. This is loosely what we call science.

A good starting point in any search is to ask what things are available to create patterns in the first place. There are surprisingly precious few things available in the universe to create action in the form of patterns. There are forces, which generally fall into two categories, those that attract and those that repel – push or pull. These forces conspire to form shapes, and most forces cooperate to form shapes that display symmetry. Forces like to balance, so we should let them. Complex balancing forces will have a hard time arranging into a line, as far as I can surmise. They will in all cases prefer a shape. Our numbers or mathematics bias toward a line – a number line – and that is typically the first place we look. That is our brightest light. Is it any wonder then that we occasionally find lines surreptitiously? Many times the act of finding a line blinds us to the nature and origin of the apparent linearity. We are just so tickled with possession of the line.



Creepie Peeple Hula  
Michael Teague

If we ever hope to tease out the strategy that Life has adopted for filling space - the function behind the results - we must find the patterns in its behavior. A good place to start looking is with the map that we already have, our data, the map of codon and amino acid assignments.