Snowmen

Emergence is above all a product of coupled, context-dependent interactions. Technically these interactions, and the resulting systems, are *nonlinear*: The behavior of the overall system *cannot* be obtained by *summing* the behaviors of its constituent parts. We can no more truly understand strategies in a board game by compiling statistics of the movements of its pieces than we can understand the behavior of an ant colony in terms of averages. Under these conditions, the whole is indeed more than the sum of its parts. However, we *can* reduce the behavior of the whole to the lawful behavior of its parts, *if* we take the nonlinear interactions into account.

John H. Holland
Emergence, From chaos to Order

Men and snowmen have so little in common that it is ludicrous to compare them, but at the deepest sub-molecular levels there are interesting parallels. There is pedagogic value to building a snowman at this time (and you thought this was all work and no play).

For all intents and purposes a snowman is 100% water – H_2O. At most, a man is 70% water, but much of the rest of him is carbon. We will generalize from this that both are heeps of jumbled tetrahedrons. The two biggest differences between the heep on the left and the heep on the right are metabolism and self-replication; man has both and snowman has neither. There is no universally accepted definition of Life, but virtually any contender will include both metabolism and replication. The snowman is obviously not a living thing, but man meets anyone's definition of living.

Metabolism is a tricky, squishy subject, and I don't pretend to understand it in any profound sense. For me it boils down to energy capture, storage and usage. It is peripherally of interest to this discussion, but the primary nut here is replication. Replication is a much cleaner subject. Although there can be disagreement about referents and complexity of results, replication can ultimately be reduced to a numeric function, which is nice. At the end of the day we can all

settle on whether two things are the same or different, and then put a number on their sameness or difference, such as 100% the same, or 20% different, etcetera. This leads to graphs, which we have learned leads to patterns, and patterns lead to insight. See how easy it is to be really smart?

The take home message is that snowmen have no metabolism, all they do is melt, and they generally can't even do that without help from the sun. Men are metabolically linked to the sun as well, but in a far more complex way. Our good buddies, the plants, play the role of sun worshippers. They store the experience for us in molecular batteries, and we ultimately obtain the batteries for our cordless energy needs. This is an elegant and useful trick, and suffice it to say that we kick a snowman's ass on this front. Further comparison would be a gratuitous gilding of the lily.

It is the replication front that holds our interest in this discussion. Replication can be viewed from a molecular level as the lining up of atoms. Atoms line up with other atoms to form molecules, and molecules line up with other molecules to form macromolecules. From the viewpoint of the atom or the molecule, the process must look the same whether it is in the snowman or the man, so we need to pull back the lens and see the situation from a broader perspective. At what level of activity does it become apparent that the two situations are completely different? The most obvious departure is the difference in composition. The snowman contains only Hydrogen and Oxygen atoms as H_2O molecules. This does not necessarily have to be the case, however, because we can imagine dumping chemicals into water, running it through a snow machine, and clumping it into a snowman. We can pretend that this was how our snowman was made. He is now a dirty snowman (DSM) but a snowman none-the-less. Atom for atom he is identical in composition to man. I am not a chemist, but I suspect that DSM is really closer to Jello than snow. We could call him Jello-man, but this gets us away from the crisp visualization of water freezing into snowflakes. I will call him DSM and you can visualize this crystalline monster however you like.

Nobody is going to argue against the idea that Life transcends atoms and molecules. I am here arguing that there are natural laws that subtend that transcendence. The universe has fundamental properties in its operation upon which we can make a distinction between Life and not Life. Science and mythology are riddled with schema to explain the "vital force" within the universe. I do not think it is a single entity. I think it is a combination of properties, circumstances and time.

Our lens must pull back from the molecules so much more in a search for the differentiation of man from DSM. It will not be on the basis of composition, because they are identical, it will be pattern; the patterns will start to diverge. Therefore, the first distinction seen in the lens between life and non-life is a distinction of pattern. It is the pattern of how atoms and molecules chose to line up, or how they chose to fill space. But this distinction is not inherent in the atoms, nor can it be located in the molecules, so where does the distinction arise? It must be inherent in some component of the environment of the atoms and molecules. It must be a property of some environmental force, but what?

Change

The biggest difference between the living crystal and the non-living crystal is that the living crystal must deal constantly with change, and the non-living crystal does not. A salt crystal or a snowflake can exist indefinitely without change, but an organic molecule has a shelf life with a short expiration date. If we examine the process at the level of the atom, we don't see the overall strategic difference between the two, all we see are two crystals growing. But DSM hopes to peacefully exist in an unchanging environment in an unchanging way. Man on the other hand must anticipate a changing existence in a changing environment. The blueprint for the pattern must somehow reflect this process.

Crystallization is a process of transition from the chaos of a liquid to the order of a solid. The most intriguing portion of the process is the interface between liquid and solid. It is this interface between order and chaos that defines living things and makes man more interesting than Jello (most men). It is therefore prudent to recognize the hallmarks of regularity, and irregularity, and somehow find a way to combine the two. A system marked by total regularity is called salt, and one marked by total irregularity is called water. A system that embraces and combines the two is called Life. Life thrives at the interface between regularity and irregularity, this is the living crystal's growth strategy. It is a process marked by change, and therefore it must simultaneously embrace change and impose order.

This takes us to the level of the cell. Look at it. Do you see it? It is a squishy bag of liquid jello, but what's in the middle of it? It's another squishy bag of jello, but what's in that? It's a crystal.

No way.

Yes… way.

What kind of crystal?

It is a carbon-nitrogen-hydrogen-oxygen crystal; we call it DNA. In the middle of the chaos of liquid jello we find a rock. Better still, the rock is a touchstone for the organization of the chaos, it directs the ordering of the Jello molecules, and the Jello reciprocates by making more rock. The jello makes the rock and the rock makes the jello. Chaos from order, order from chaos. But where could this mutual reliance, this self-propagating cycle get any toe-hold in the chaotic storm? We must start by imagining the crystallization of carbon in water – dirty salt water. This scenario has been played out in many beakers across the world since Miller started the game in the 1950's. Everybody has their own spin on the situation, but there are several major patterns that emerge.

The process is literally like watching paint dry, if not less interesting, so it makes sense to inject some energy of one sort or another. Miller used electricity, but others have tried thermal, radiant, and mechanical forms. The bottom line is

that without some energy source, nothing of much interest forms in our lifetime. I would put a different spin on the role of energy in the system; I see it as a form of change. The energy applied to the system produces a variable of change that would not otherwise be a factor to the crystallizing sludge. If we imagine the abiotic earth, which these beakers are meant to model, whatever the role of the energy source, it is playing a part in a changing environment, whether it is delivered in the form of waves, tides, sun, chemical gradients, temperature fluctuation, or even electricity. The only thing the molecules can count on from any one energy source, or all of them combined, is change.

Great Drama

In our staged drama there are two major parts. The role of chaos will be played by water, and the role of order will be played by carbon. Carbon and water are forced into an uneasy relationship against the evil forces of change, a role we have already cast as energy. Carbon and water would like nothing more than to settle down to an eternal, fixed co-existence, but the ravages of change continually spoil their plans. Our heroes eventually catch on to the methods of change and begin to construct a framework for dealing with it. Soon it becomes an integral part of their plan. But change is not an end unto itself, it implies the existence of its opposite, which is constancy. The role of constancy is played by repetition. Life is a song with a beat, a repetition of cycles or events. Things happen over and over and over with monotonous regularity. Events accumulate and cycles develop. Numbers become very large, and without very large numbers there can be no Life. Small numbers cannot cope with order and chaos; they cannot generate a meaningful interface between the two. Repetition and change are the necessary two heads of the same snake, the raw fuel of evolution. Life is the molecular snake charmer.

The stage is being set for our molecular drama, but we have no scenery. This is one of the most embattled decisions of this human production. There are scenery suggestions from all involved, which range from preposterous to useless. One suggestion is that the drama begins in outer space, an idea labeled pan-spermia, a setting of the play starting on Mars, or Jupiter, or you name it; life begins, advances, and travels via rock to Earth. This is not a suggestion that I would label as preposterous, it is quite viable, but it is useless for our purposes. Panspermia has not set the stage, it merely displaces the stage to another planet. That would be useful if this planet were not old enough to contain the entire drama, or lacked some essential character, but I have not seen anything to suggest that this is the case, so why not stay here for the show?

The biggest clues today toward deciding the “best” setting of the opening act are water and salt. We in the audience have a particular preference for water, owing 70% of our existence to it, and our salt content, or salinity, is precisely that of an ocean. Mighty convenient don’t you think? This is my drama and I chose to set it in an ocean. Another advantage to using oceans for scenery is that they are so damn big, so there is ample opportunity to generate repetition. The coolest thing about oceans is that they have shorelines, lots of them, and

shorelines are good places for our cast to assemble with great plausibility. So the curtain will open on our drama with a beautiful vista of coastal bliss. We hone in on the shallows and find lots and lots of dirty salt water. This scene is anything but serene. There are countless episodes of sludge globules churning, tossing, heating, cooling, drying, wetting - what's a carbon atom got to do to get a little peace around here? Our cast is caught in the whippersnitch between water and land, an environment of constant change. Bazillions upon bazillions of strategies are tried by sludge to separate itself from water, searching for order from the chaos. It is a state of flux within flux. This is no time to talk metabolism or even replication, this is a time to talk order and disorder.

Spooks of the Eyestalk
Michael Teague

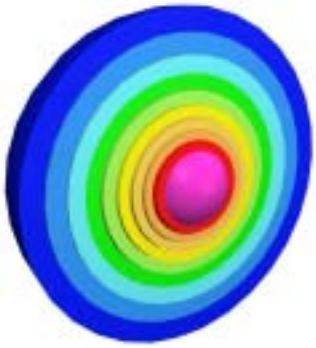
Carbon tires of the commotion, and seeks alliances with other carbon. Miller et al teach us that carbon finds easy alliances with other carbon in the form of amino acids. The trick of making amino acids is a relatively easy one, and it takes little time or elaborate stage setting for it to happen. The trick of tetrahedrons cooperating to form larger tetrahedrons is unremarkable. Our global coastline can easily support this plot twist. But the trick of forming nucleic acids is far less probable and far less secure. Nucleic acids are not born of the tetrahedron, they are worshippers of a more complex shape, the dodecahedron. They crystallize in the awkward symmetry of the dodecahedron and therefore require more complex support. There are fanciful accounts of cyanide "ponds" that freeze, spontaneously creating nucleic acids. Perhaps this is a common setting in an abiotic, volcanic earth, but I've never visited these places either in body or spirit. Furthermore, whatever body of water that is "freezing" we must imagine it happening billions of times, not just once as is currently in vogue. That is why I prefer to micromanage these scenes by segregating the action in our little sludge globules. If the freezing or crystallization occurs within pockets of

pockets of sludge. The dehydration process can partition itself in sections of sludge, and gradients are formed. The grand orchestration can split between the easy access of tetrahedrons and the more complex circumstances of dodecahedrons. With separation and gradients come interfaces and transitions. This is where the real drama rages, at the interfaces, the interface of chemicals, shapes and order.

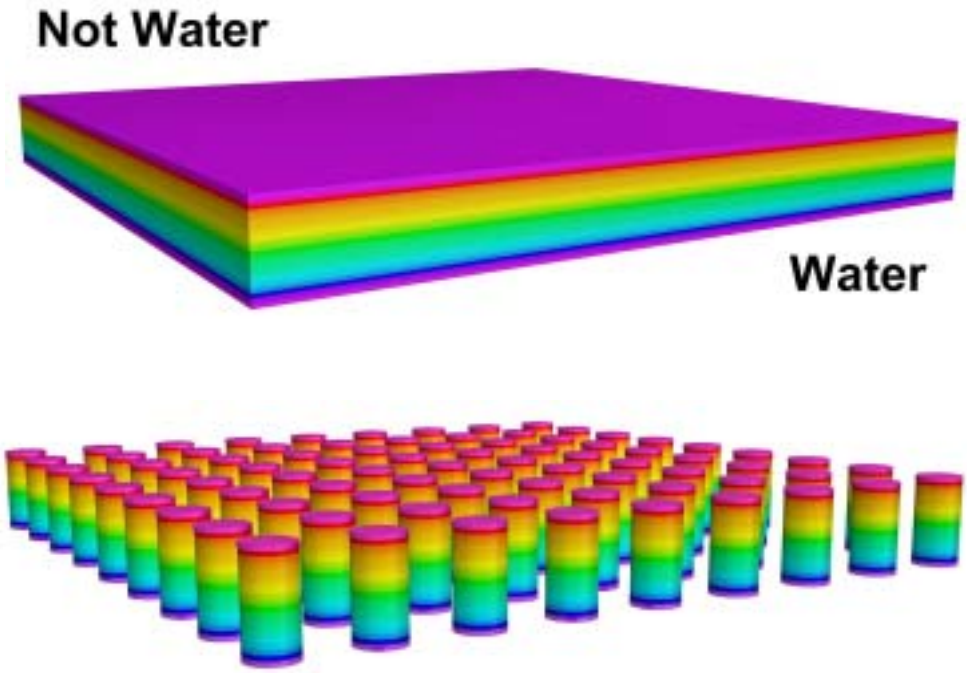
This drama is a process not an event. The process involves large amounts of repetition sprinkled with the essential spice of constant change. We mark the milestones in the process historically with the crowning of champions, symbolic kings of survival. But these inaugurations are chimera existing only in our mental models of complex, repetitious events, historical events with no special features during the times of their own existence. They carry the titles of “last common ancestor” and “origin of this” or “origin of that”. They are the pyric victories of posthumous coronations for detached, historical champions. The drama emerges not from imagining something improbable and isolated, but in imagining a process ubiquitous and inevitable.

Enough goth, let’s splash some color into our scene. Here’s a sludge globule.

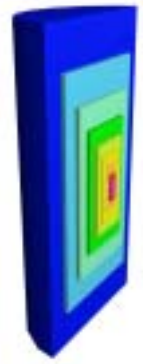
Pretty unimpressive, but we can imagine some simple sludge properties for this globule to set it up for some sludge tricks. If we consider that there are natural properties of sludge, such as its ability to get along with water, then sludge placed in a dynamic setting might entertain us. If sludge is not perfectly copasetic among water, then it will segregate itself into globules. If these globules are not homogenous sludge, then they will create gradients according to water affinity. The water-loving sludge will hang around the perimeter, and the water hating sludge will squirrel itself as far from water as it can.



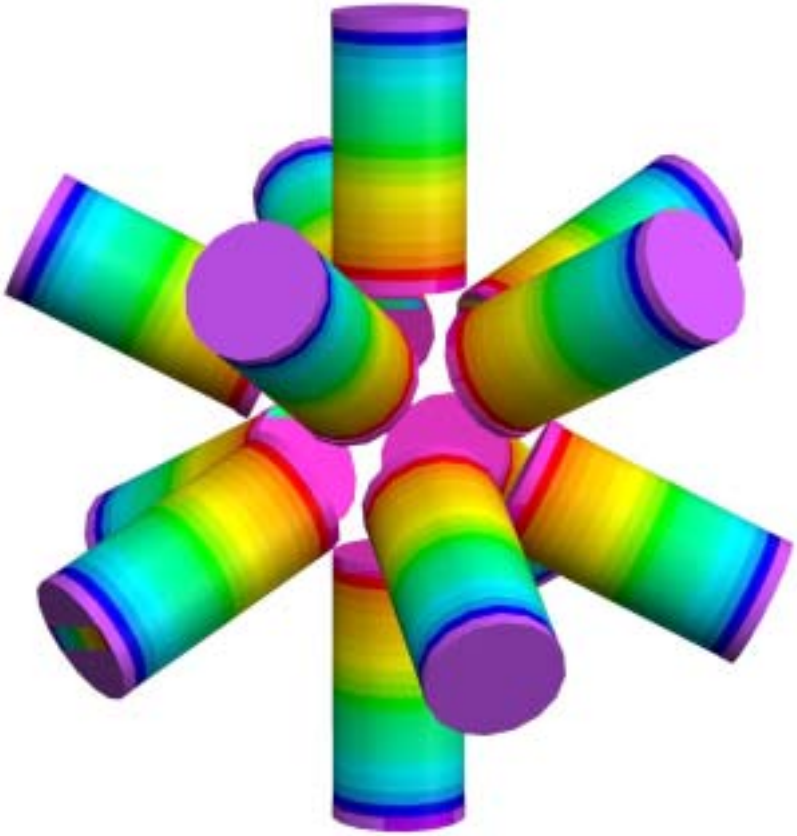
Of course no sludge is an island, so water will continue to be a part of sludge's life to some degree or another. It is only through dehydration, or the exclusion of water, that sludge can crystallize into a solid of its own. This process will follow a path as diverse as the number of sludge globules one can imagine. In general, the crystallization process will make sludge smaller, so we can imagine sludge columns forming within the water matrix. This is simpler to visualize if we start with a sludge layer on top of water.



Now we have a new sludge-water interface as the stratified sludge begins to crystallize, which is a happy circumstance because new gradients can form, and we love gradients.



The most concentrated water-hating sludge should be found at the core of the sludge crystals. This process is just as true with sludge globules as sludge layers. Perhaps it is a combination of the two, where sludge spicules form and aggregate into sludge clusters, much like lipid layers are known to do.



It is the core of these crystals and aggregations of crystals that I imagine environments for carbon crystallization much like a freezing pond. The specifics are beyond me by many orders of magnitude, but I generally sense that the

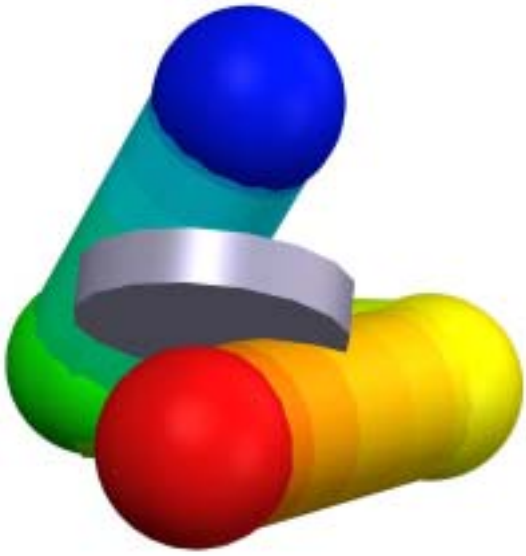
process of carbon jostling for crystalline position will be multileveled, complex, and interesting. Sludge at the core will be presented with different environments than sludge at the watered periphery, it will chose different strategies, and there will be an interface between strategies. We now have imagined a carbon sludge crystal that is stratified with a strategy predicated on water affinity. We can imagine this beast in bazillions of instances with bazillions of different strategies.

It is at this rudimentary level that the true dynamics of the drama can kick in. They invite new players called combination and aggregation. Complexity, network science, universality, chaos, emergence are all playgrounds for these new players. Simple things combine to form complex things. Things aggregate and accumulate, but these combinations do not become big versions of little things, they become different versions of new things. More is not more, more is different. In this new world of social mingling and sludge spicule accumulation it would be a useful trick for sludge to learn a measure of independence. If talented spicules figured out how to ball-up, they could travel from aggregate to aggregate and form new connections, a crystal network can evolve and complexity will emerge.

Buckminster Fuller teaches us the simplest way to ball up a spicule or rod; it's called a tetrahedron. This is the lowest energy folding strategy for a column, practically just two folds and a splice.

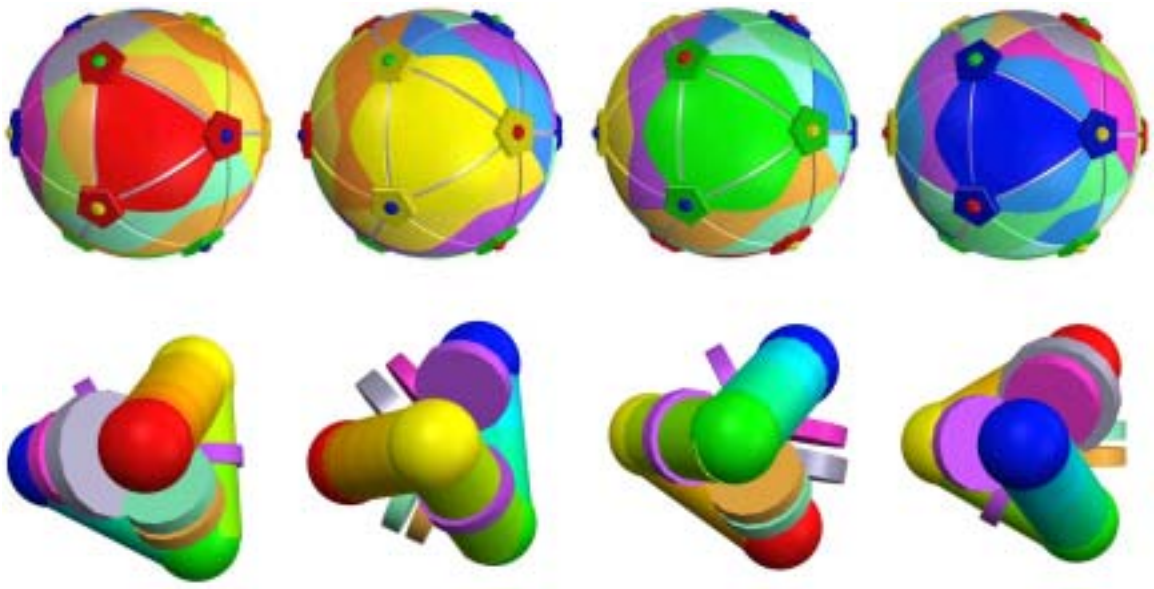
In the real world of sludge spicules the splice is a bitch, and the folds are no cake either. If these sludge regions actually wanted to get together they would have gotten together in the first place. They are apart because they prefer to be that way, so they will need special sludge chaperones, or glue to keep them

together. The region of most desperate conflict is the space between water loving and water hating, blue and red, so we will put a firewall in as a splice.



The other regions between the hydrophobic and hydrophilic poles of the folded structure will require creative arrangements to maintain a happy structure. We can represent the finessing of the balance at these junctions with some special sludge icons.

This is a quite fanciful representation of a mythical sludge spicule strategically built to form and travel in water. Such a beast could probably never exist on this planet, but it is starting to look like a pattern we've seen before.

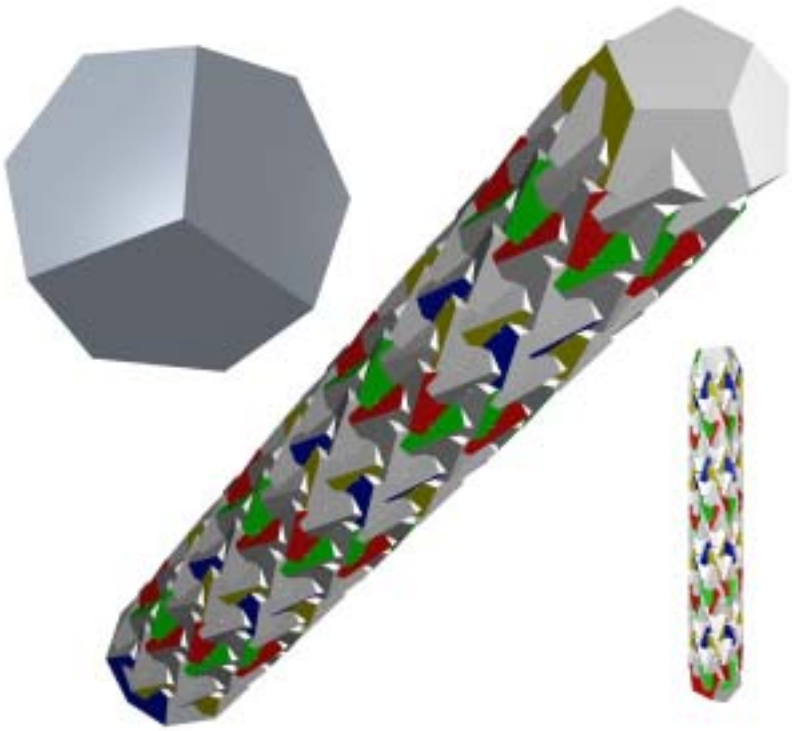


The spicule is mythical, but the map, or globe of the genetic code is not. The pattern demonstrated by that structure is self-generated. Even if you reject the notion that there is meaning embedded in the genetic code, you cannot deny the pedagogic value of studying this object. It is exponentially more powerful to comprehend codons on this globe compared to anything approaching linear. Assignments, properties, such as water affinities, similarities and differences of all parameters in the system can be illustrated relative to one another. It is the natural habitat of nature's data. These study tricks are founded on many of the principles we used to imagine the folded spicule. It is possibly a coincidence that these parallels are even possible, but the intellectual utility of this genetic globe is striking.

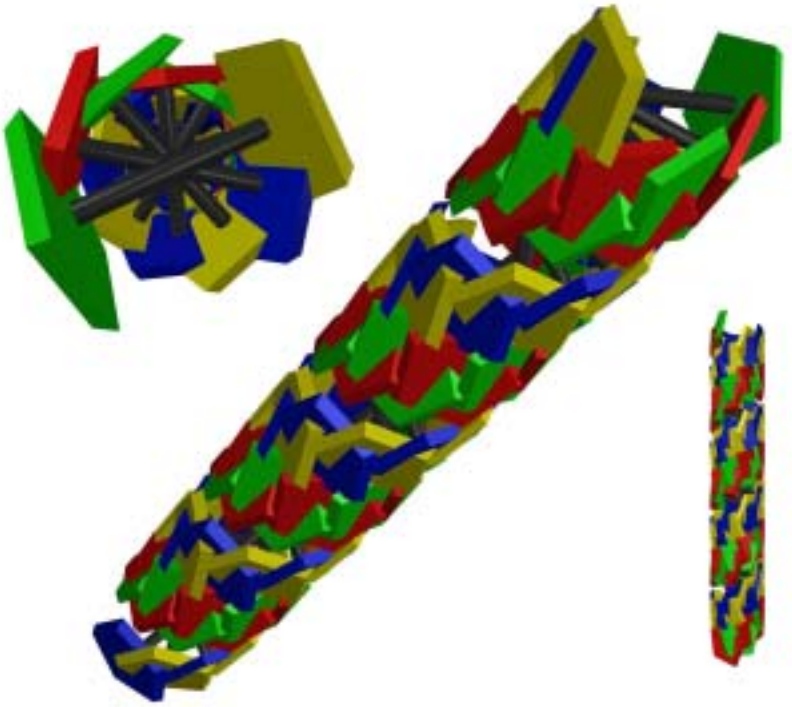
Just as our fanciful crystal spicules can aggregate, so too can crystal particles. These rough, imprecise seedlings will be polar, and they will cluster in water. This form of spontaneous aggregation or assembly is particularly interesting in light of the many modern homologs in biochemistry, viruses being a major example. In cases where aggregation of subunits is a major growth strategy dodecahedral and icosahedral symmetries are common.

As seed crystals, these large mythical beasts are now in a position to fold a diverse collection of smaller beasts - proteins - based on any and all platonic forms. Empiric evidence suggests that this is exactly how proteins fold. Is this the clue to a fundamental law of folding?

But are crystal spicules common, and what relationship would they have to dodecahedrons? Crystal spicules are fairly common phenomena, but dodecahedrons are not. Dodecahedrons are very rare in natural crystal systems because the symmetry is an awkward way to make an interconnected lattice. Consequently, inorganic crystals demonstrating the five-fold symmetries of dodecahedrons are unusual, but five-fold symmetry is the rule rather than the exception in living systems. What would a dodecahedral spicule look like?



Again, merely a fanciful illustration of a dodecahedral crystal spicule, intentionally designed for a demonstrative reason. I am not aware of any natural crystals that spontaneously form in this type of dodecahedral structure. I aligned dodecahedrons sequentially and proportioned them so that their complimentary faces can be colored and the cores removed:



I am not aware of any structure in nature like this, except…well…DNA, but that's not really a crystal, is it? This is the exact proportion of the double helix in DNA, which also demonstrates dodecahedral symmetry. It has a major and minor grove, complimentarity, the same pitch, the same number of faces per rotation. It is not meaningful in any sense other than as a fanciful parallel between DNA and a sequence of dodecahedrons. This is not a typically useful crystallization strategy for carbon, which will generally take one of two pure forms in nature: graphite and diamond. The precise form taken, and therefore the symmetry employed is a function of the environment in which it crystallizes.

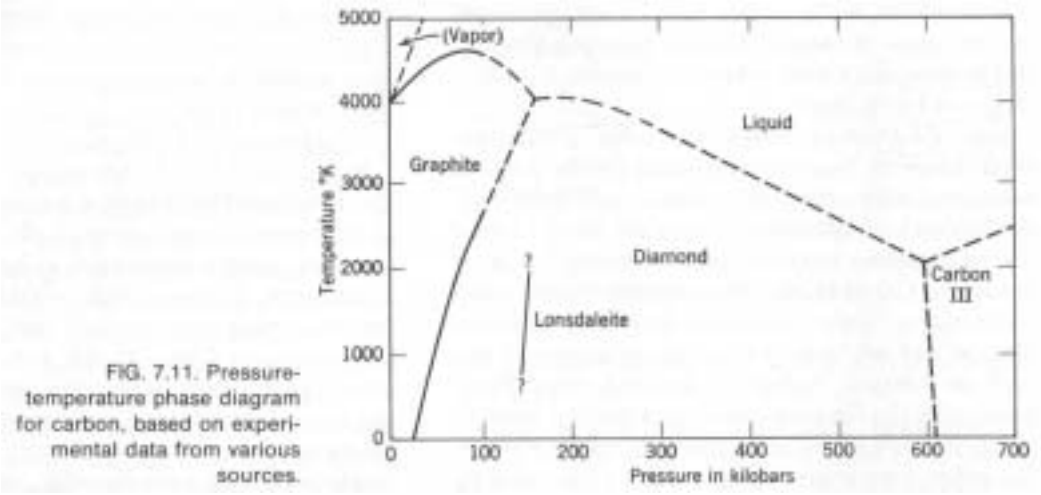
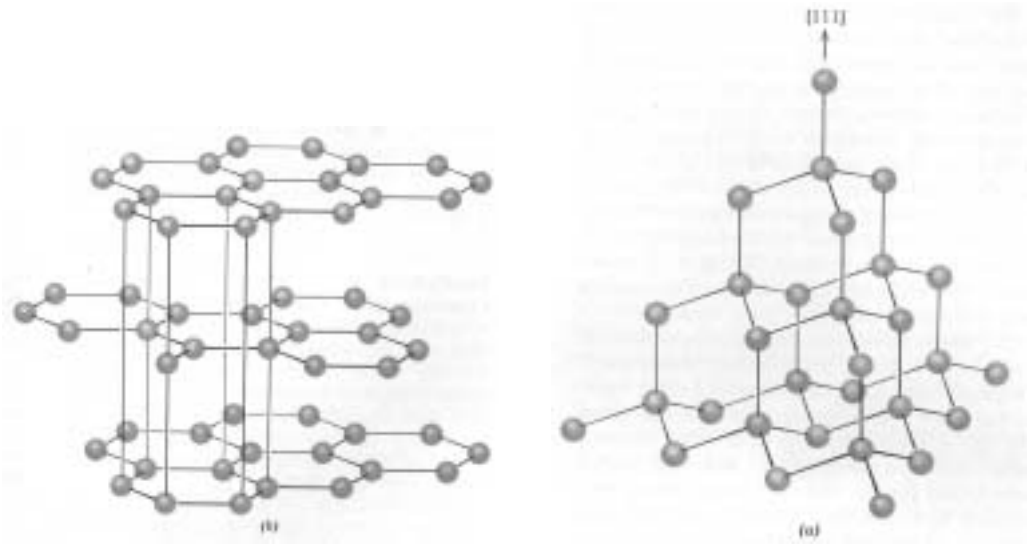


FIG. 7.11. Pressure-temperature phase diagram for carbon, based on experimental data from various sources.

In low pressure environments, such as at earth's surface, carbon will crystallize into the six-fold symmetry of graphite. At higher pressures carbon crystallizes into the four-fold symmetry of diamond.



Graphite Diamond

The differences in physical properties between these two structures are dramatic. Color, luster, cleavage and hardness are quite disparate, and they form the basis of widely differing commercial value. The environment produced the structure, and the structure produces the utility. An environment where regular crystallization is not an option will favor strategies that embrace irregularity. These are the type of strategies that demonstrate emergence. They are agent-based strategies, where simple agents are acting on simple laws in huge numbers. It is only through context and combination that more complex behavior begins to emerge from the system.

Grandma

The theory of Babbage accounts with great probability for the rise of ground in the vicinity of volcanos, and Herschel's theory accounts, perhaps, for the subsidence of deltas and other places where great accumulation of sediment occurs; and this latter theory has the additional advantage of accounting for metamorphism, and perhaps, also, for volcanic phenomena. But it is evident that some other and more general theory is necessary to account for those great inequalities of the earth's crust which form land and sea-bottom.

Joseph Le Conte
Elements of Geology – 1898

Life is a cascade of networked precursors. Flesh, bone, talent, ideas, culture, it all moves along with time, propagating from one network to the next. We like to see linear relationships in the process, but I doubt that the linearity is more than an illusion.

My last living Grandparent just died, Grandma Naomi. I learned a lot of things about her only after she died. She was a special lady on many counts, but she had horrendous handwriting; of course I knew that before she died. What I didn't know was how beautiful her husband's handwriting was, Grandpa's. She always wrote everything for the pair, but you could never read a damn word she wrote. Grandpa could write like nobody's business, but I never saw a thing he wrote until after they both died. My Dad got his mother's handwriting, which he didn't give to me, because I have no handwriting at all, so I print, or I use my computer. I can't really decide which of my four grandparents is to blame.

Naomi attended the university of Chicago before girls generally went off to college, and she took geology classes. My brother was a geology student as well, neither of us knew that she had preceded us in the study of geology. After she died, he found some of her projects and textbooks. Based on that material, she was clearly a better student and geologist than I ever was or ever will be. The above quote was from one of her books. It is remarkable for two reasons. First, they had no clue whatsoever in 1898 about geology in the context of plate techtonics. We now know that the crust of the earth is wandering around the

surface, but without that little tidbit of information there are a lot of facts that are absolute head-scratchers. For instance, an ancient shoreline can somehow be located at the peak of a contemporary mountain. They struggled valiantly to make sense of the data, but in the end they were debating the dance patterns of fairies on the head of a pin. Second, the statement is remarkable for the line:

> But it is evident that some other and more general theory is necessary to account for those great inequalities of the earth's crust which form land and sea-bottom.

They knew what they didn't know, and they were willing to withhold final judgment on the theory in lieu of better ideas to come. They did not vex their cultural descendants with dogmatic insistence on one theory or another. Apparently the excitement of bigger ideas outweighed the fear of no ideas at all. To them, the interpretation of the earth's crust without plate techtonics was like the interpretation of the genetic code without a dodecahedron. It shouldn't be done.

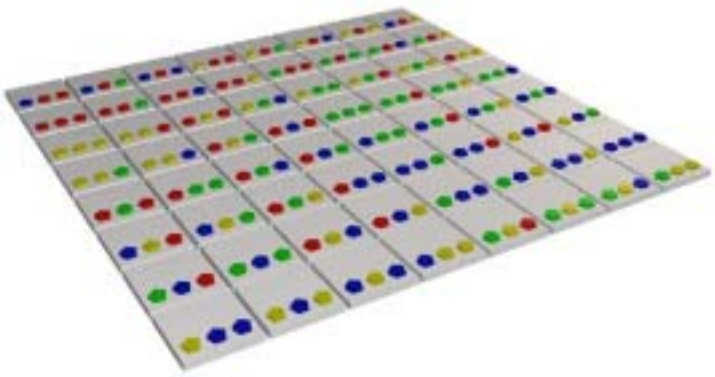
Differential settling was eventually replaced with a model of techtonics, but it was not an unreasonable guess, because heavy things sink. The really heavy stuff in the earth is toward the center, and it gets hotter from pressure as we descend to the core. Toward the periphery are left the lighter elements, the most interesting of which are carbon, oxygen, nitrogen, and of course hydrogen. The most interesting interaction of these elements at the surface of the earth - in the universe - is the interaction of hydrogen and oxygen to form water. The properties of water are oddball in innumerable ways, but without them Life is improbable. Seventy percent of the surface of the earth is water. The atmosphere is full of it. H_2O exists on the periphery of earth in many forms, gas, liquid, ice, and in that most interesting of all forms, the semi-solid carbon slushy form of Life. About seventy percent of Life is water. Life is like a continuous crystalline outer crust on the surface of the earth. What possible mechanism could get this crystal started all those billions of years ago?

Mechanism of Assignment

The linear genetic code is a system in which twenty amino acids are *arbitrarily* assigned to sixty-four codons - and that is all. This is the only dimension in a one-dimensional mechanism, accidentally matching one codon with one amino acid. The code arose by some unimaginably mystical mechanism about which we have little or no understanding. Regardless of the actual steps taken, they were taken only once. All traces of the mechanism must now be gone, because the process happened a long, long time ago, and couldn't continue acting on the genetic code in the same way today. It must have been an accident that got frozen, a frozen accident of monumental QWERTY proportions with which we are completely stuck. Unfortunately, we cannot precisely model and rerun the process to see what types of assignments are possible and perhaps even probable. Vast time and large numbers are required. Therefore, the genetic code is seen as the winner of some ancient, unknowable assignment lottery. In addition to linear, and arbitrary, the code is also degenerate.

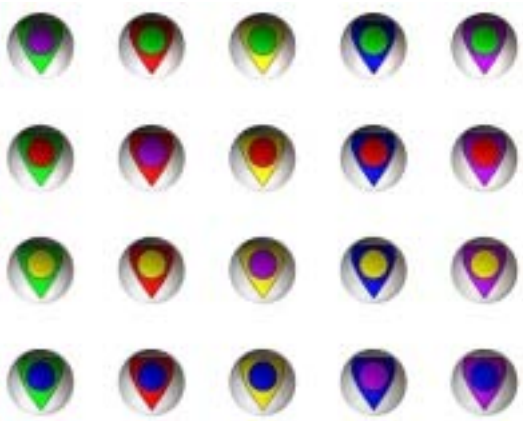
Call me sensitive, but it's offensive to think of something beautiful, dynamic, efficient and powerful as frozen, one-dimensional, arbitrary and degenerate. It seems that if we could somehow unlock the mystery of the lottery process, or assignment algorithm, we might gain insight into just what skill the components of the code possess, making them winners in what must be the contest of all contests. We will create theoretical assignment algorithms, looking for a possible optimization.

Despite its label, the linear genetic code is usually displayed in a two-dimensional grid. There is no stated interaction between the grid cells, or between the symbols of each grid identifier. In other words, one cell could be placed next to any other without affecting the actual code. If we use digits, 2 is 2, regardless of its position relative to other symbols, such as 421, versus 233. It is therefore easy to imagine a way to assign identities to a variety of things. Simply create a grid with every possible assignment, and begin filling the grid with objects to be assigned. It doesn't really matter what grid shape or symbol configuration we chose; at least we have no way at this time to evaluate how well one grid would serve us as opposed to any other.

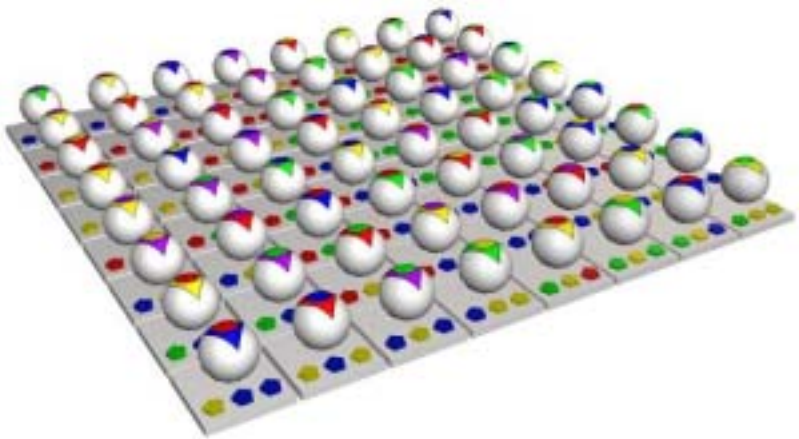


I call this an identity grid with a maximum carrying capacity of sixty-four. Any grid such as this might actually identify some number of objects less than it's maximum, which I call the optimized carrying capacity. For instance, the above grid will identify only twenty different amino acids, so we say that the optimized carrying capacity of this grid is twenty. The ratio of maximum to optimum carrying capacity is the degeneracy of the grid, which in this case is 3.2.

Starting from this simplified model, let's run a thought experiment and hold an assignment lottery. To play the part of assignment objects lets imagine twenty types of ping-pong balls.



On average, each type will appear 3.2 times. Now imagine that the identity grid has walls around every cell, and each cell perfectly holds one ball. Put all of the balls in a box and randomize them by shaking the box. Fun, fun, fun - drop all of the balls onto the grid so that each cell captures one ball. We might see a pattern as follows.



The result of this random lottery is highly similar to the genetic code. But there is also no reason to believe that any running of the experiment will produce a result any "better" than any other. This is why the genetic code is considered arbitrary; it doesn't necessarily mean anything that a ball occupies a cell. Any ball would do just as well in any cell. Our thought experiment by all appearances is a fair game of chance. No component need demonstrate any skill over any other component to "win" their cell. We could expect to hold a lottery tournament of these experiments over an eternity and never find a way to demonstrate accumulation of any meaningful properties. This is the one dimension of the linear code: assignment.

Of course this is the most simplified of all possible assignment algorithms. It is specifically designed to produce an identity grid with extraordinary similarities to the genetic code. But it is very easy to identify the source of these similarities – they are designed into the algorithm. Starting from these premises, we expect arbitrary results with no trace of optimization. To move away from this triviality, we will need to examine and revise our premises.

We have assumed in our running of the initial experiment that we already know some things about the final grid. Of course all of these assumptions came from our knowledge of the actual outcome of the lottery that we are theoretically modeling. We need to challenge some of the foundational assumptions, such as why our system is based on assigning words composed of 4 types of symbols in groups of three. This assumption is at the heart of the grid we have built. How do we know that there should be only 4 symbols? Why not 3, or why not 8? It is because the system is built on a 4^3 architecture that assures us that every result will have the same form. Also it appears we have assumed which 4 symbols will be used and in what proportions. Isn't this a part of the optimization process? Shouldn't the symbols and their arrangement also be drawn in a lottery just like the objects and the assignments? But wouldn't the symbol lottery and the word lottery and the object lottery somehow all depend on one another. There is a huge, ugly recursion in this process somewhere. Even if we were able to model these interdependencies, we still would have to build the model on some foundation of assumptions.

One assumption that is less easy to justify is that the optimized carrying capacity must equal 20. After all, it is hard to argue that anything predetermined, such as this, could somehow be the result of an optimization process. In fact, it is quite baffling why the optimized carrying capacity should turn out to be 20 at all. Why not 17, or 27, or 64, and why just those 20? We know with certainty that there were many other assignment objects available for the genetic code besides those 20. Even those had identical mirror twins, but for some reason only those 20 were drawn. Why not let the twins play the lottery too?

Perhaps there is some unanticipated affinity between the icons on the balls and the icons in the cells. Perhaps this mystical affinity is so small that we could never detect it in a hundred or a thousand or a million runs of the experiment, but over the course of a billion runs, and with the help of sensitive statistical tools, we might detect it. This imagined affinity is an assignment bias of some kind. During the course of a billion years we could run the experiment

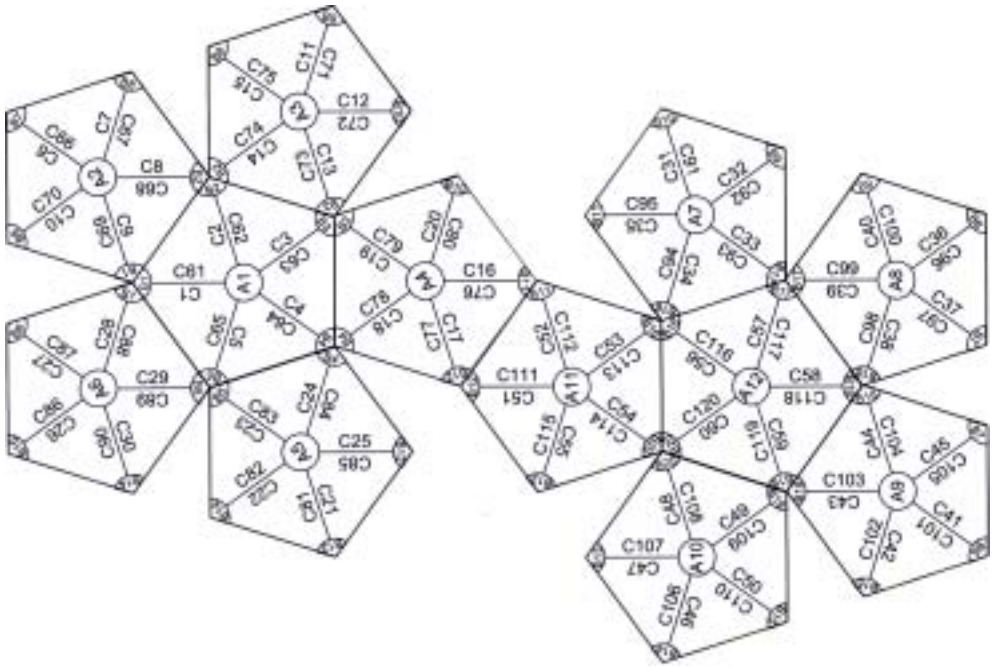
billions of billions of times and average the results. Real or imagined, at the end of that prolonged experiment, the winners among objects for filling particular cells might begin to be seen as optimized for those cells if such a bias existed.

We have, at this time, no reason to expect a resulting pattern that appears anything other than random. However, if any such assignment bias does exist, we would expect a pattern reflecting that bias to emerge, and the average of all possible cell results could in some sense be seen as optimized for that bias. The key here is that, bias or not, in the linear code there is no reason to expect a priori any one pattern or pattern type over any other. The accident could have occurred in any ol' way, and then it was inevitably frozen, by the law of QWERTY I suppose. Important to note however, with the addition of a bias in the system we are evolving from a fair game of chance to a fair game of skill.

Still, there is no reason, other than hope, to expect the results of this experiment to resemble the genetic code any more closely than would be caused by all similarities built into the algorithm. Furthermore, if we stopped any of our artificial practices of dropping balls we would expect the results to drift even further from the "ideal". If we are completely honest, there are a few more confessions we must make about this rigged game. A fair lottery would require a whole box of symbols dropped into many grids before the ping-pong balls are dropped, and hundreds of different ping-pong balls instead of just twenty. Finally, where did we get the apparatus to run this lottery in the first place – doesn't that have to develop before anything else? When all of these pre-ordained fixes get removed, all vestiges of the genetic code can be expected to disappear *completely*. There is no hope of even holding a lottery, and nothing in this thought experiment useful toward modeling the ancient mystical genetic assignment algorithm. That's because we are operating on little to no theory, or construct. We have an ad-hoc argument into which we are willing to place any and all empiric data. Regardless of our experimental findings, we have decided that they will support our theory, or lack thereof. To paraphrase Wolfgang Pauli, this theory has become so bad that it's not even wrong. This is truly a philosophical doctrine, not a scientific theory.

Rafiki has an actual theory that the universe will work toward balance, begging the questions that must be begged. What are we balancing, and what do we have to balance with – what is the apparatus? Rafiki knows that the universe has precious few choices, and molecules always chose shape, so we should too. We have been trying to model a multidimensional, interrelated network of components by sequestering them into arbitrary, unrelated grids. *We* might chose to perceive the genetic code as linear, arbitrary and degenerate, but I suspect reality at the molecular level is nothing of the kind. The first choice we must make is a choice of shape. Our apparatus must be three-dimensional, so we will chose a shape that matches our empiric numbers from the genetic data table. We have 3, 4, 5, 6, 20, and 64, so a dodecahedron might be an inspired choice. This is the exact type of reasoning, based on Chagraffs rule, that lead in 1953 to the assumption of base pairing and a double helix (not triple) in the

structure of DNA. If it's good enough for the double helix, it should be good enough for the genetic code.

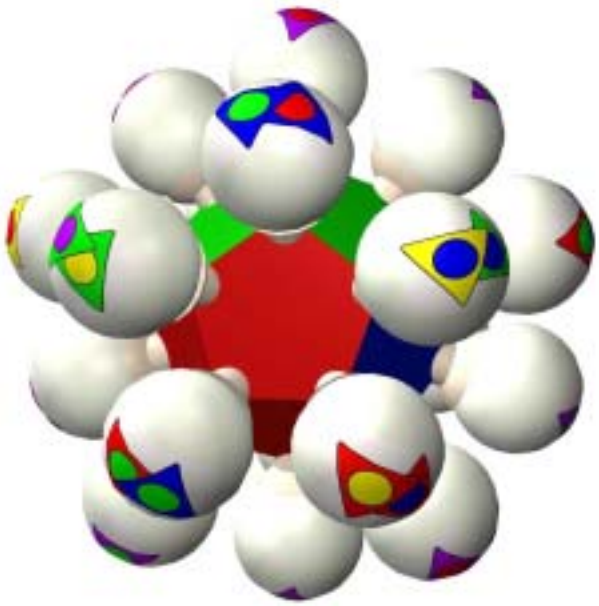


This is our identity network, and from this we can imagine a mechanism that will produce a network of identities. We can substitute all of the fabulous symbols and colors developed earlier. The dodecahedron as comprised of twelve interchangeable panels representing assignment symbols on the nucleotide side of the mechanism. At the junction of any three panels is a suction cup.

Instead of dropping balls onto a grid, we can envision a lottery process where a dodecahedron is tossed into a mass of ping-pong balls, a mélange we will call the sludge. After being tossed, the dodecahedron is shaken, not stirred, and then retrieved. Presumably, but not necessarily, upon retrieval the

dodecahedron will have a ball representing an assignment object stuck at each vertex. Right off the bat we can see how this lottery is different, and we can begin to guess at the implications. First of all, there is no absolute way to select nucleotide panels in building the dodecahedron, because we have assumed that it is substrate neutral. In fact, there is no implied optimum number or configuration of different symbols, only that no configuration can have more than twelve symbols. This is one conceivable way to include the symbol lottery in the process with the word lottery. We are similarly not constrained on the list of potential candidates for object assignments. Our sludge can include any and every possible option.

The big downside is that this model does not seem to hold out hope that we can run a single assignment event; instead it will require a lengthy, expanded, iterative lottery process, perhaps going on forever. Rather than speculate on what all of the parameters might need to be for a "successful" result, let's just start with a first iteration. We begin by selecting 12 panels; any twelve will do. Our box of ping-pong balls includes the standard 20, their mirror twins, and a gigantic host of other candidates. Since we have at hand a color system of icons capable of identifying hundreds of objects, let's imagine that each of the balls has one of these icons on it. The ball types can be equally distributed, and we can imagine an infinite supply of every type.



We must note that every ball is associated with three panels, but each panel is associated with five balls, and there doesn't appear to be a clear word to object relationship. In order to track the results of this and all future assignment events, we need to imagine a bookkeeping mechanism. We will use assignment cards. On each card we will record the results of each word-object assignment. Since this is not a one-to-one relationship, and since there is no way to dissect

out the potential bias of any one symbol from its five neighbors, or all 11 symbol companions for that matter, we will also record the sequence of all twelve symbols on the card. Each object can be identified by 6 permutations of its 3 associated symbols, so we need to make 6 cards for each ping-pong ball. The complete results of our first iteration will be reflected in a set of 120 cards (20 X 6 = 120). There is no reason to believe that this iteration is representative of an optimization except in a very restricted, local sense. First of all, the results were heavily dependant on our selection of the 12 symbols and their sequence. We arbitrarily chose this configuration out of a virtually infinite number of possible choices.

At this stage the process has decided nothing except the one thing we decided for it, which is shape. Therefore the shape of the results, at the end of the day, will necessarily in some way involve a dodecahedron. At step-one there are no nucleic acids, no amino acids, no codons, and no assignments, but the process can move onward from here.

Before we can start thinking about calling this an optimization, we are going to have to at least check all possible configurations. But since it is unlikely that a process existed to predict, force or duplicate any one configuration preferentially over any other at the time of the original assignment lottery, we must see the process as a configuration lottery as well. No single iteration can be taken as “optimized”. The “winner” will have to be some form of statistical winner. Each card is taken as a “vote” for some particular symbol, word, configuration, object and assignment. We will therefore have to run billions of iterations for each configuration. This has turned into a huge undertaking, reminiscent of the Florida recount.

Let’s imagine how this lottery might evolve. Pick a day to start – let’s say next Tuesday. Mark the day on your calendar with a big red X, and block out enough time to complete the process – say… a billion years should do it. Show up bright and early on start day and we’ll jump right in. We already have one toss down, one set of cards made, only a bazillion bazillion bazillion to go. Each toss starts with a different configuration and ends with a new set of cards. After we feel that we have created a sufficient sample, we take our total collection of cards over to our new, gleaming, ultra-efficient computer (donated by the good folks at Apple). The computer has a high-speed (very high speed) card reader with a hopper the size of Texas. The computer hums comfortingly, relentlessly gobbling cards, putting out little heat, requiring no maintenance. In a short period of time the vote counter outputs the results.

God only knows the format these results take. The important thing is that these results are used to guide continued iterations. (You didn’t think we were done, did you?) Results can accumulate and the system can begin to move away from its starting point of complete disorder. The system is being biased. Perhaps the vote counter tells us to abandon certain configurations because nothing sticks to them, or they don’t pick up a full load of objects. Perhaps the universal counter suggests that we need to change the recipe of the sludge, add a little aspartate, for instance, or lysine. Perhaps the computer tells us to start

trolling for objects with more than one network at a time. It is not as important to speculate on what the exact results are as what *type* of results we might look for.

The first conclusion we are likely to draw is that each assignment word will include 3 symbols. This is not the same as concluding that each word only needs 3 symbols. The two do not necessarily follow. Any and every assignment might be contingent on a part of, or the entire configuration in which it appears. The only way to know is to run the experiment and tabulate the results. Another conclusion that we can expect is that the number of symbols that "win" will somehow involve divisors of 12. Since 3-symbol words have already been decided upon, it is pretty simple to expect that sets of 4 symbols are going to be hot in the running, but we can't rule out the 2 and 6 families. (DNA is actually a 2^{nd} order 2-symbol set.)

The most interesting results we can expect involve the maximum and optimum carrying capacity of the network, and the degeneracy ratio. Now, new questions need to be asked about what constitutes a separate word and a unique object. Perhaps there are several varieties of previously perceived distinct objects. It would no longer seem foolish and trivial to consider 6 unique permutations of the same symbol. (This is where you, the reader, should be screaming and jumping up and down on the table, "you fool, you fool, you bloody fool! You act like these damn ping-pong balls are amino acids. Where are your tRNA?")

Oh yeah, thanks, I forgot… Calm down.

The selection process must somehow include tRNA and a second dimension. We know that in the first dimension codons match tRNA, and in the second dimension tRNA match amino acids. We know that one codon can match multiple tRNA, and one tRNA can match multiple codons. There are two obvious dimensions to this mechanism, and there are no linear relationships between the components. This is not news to anybody. Somebody please tell me why we continue to insist that this process is linear and one-dimensional, when no aspect of it matches that definition. This is mass hypnosis. It is blind faith in a philosophical doctrine. Hello, anybody awake yet? I'm going to count down from three and everybody will awaken from the spell: three, two, one. Good morning.

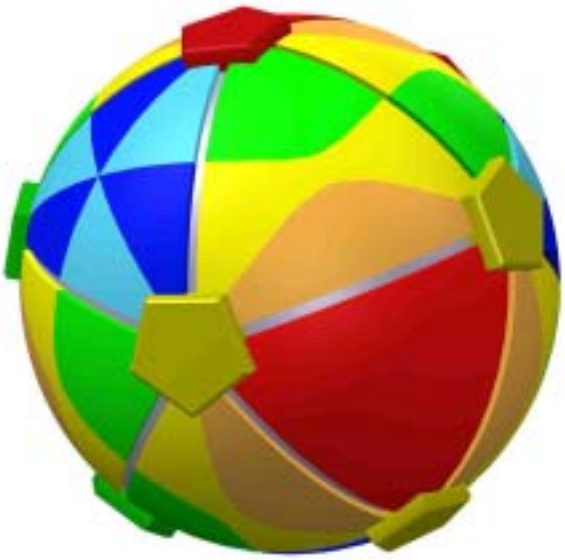
We must finally see that this mechanism includes tRNA on some level. It cannot be only about codons and amino acids. The selections come as sets, or kits for building. It is a symmetry kit of shapes and other properties. We might expect that amino acids and nucleic acids get selected simultaneously, not individually. We might also expect the ratio and balance of sets, four nucleic acids and twenty amino acids, must also be simultaneously selected. In this way these sets can be seen as optimized, perhaps so much so that they are relatively forced moves. This does not mean, however, that the interactions of these sets are forced, they are irregular enough that they cannot be forced. They remain dynamic and changeable; they are free to adapt. Storage and translation of

information must be addressed simultaneously. The interface, or genetic code can continue to evolve despite being forced - or allowed - to work within the constraints of these strongly forced moves.

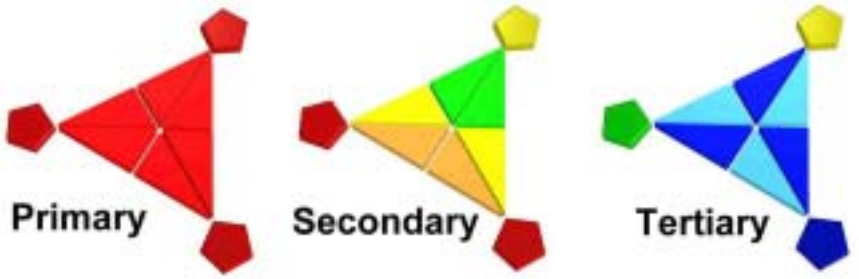
If we accept this model, we might also want to alter some of our other dogmatic thinking about the genetic code. Perhaps it's not a single, linear code, but a network of interrelated codes. We can then call into question the view that the code is meaningless and arbitrary; each word means something different in its own context, and it was optimized for that meaning within all contexts. It is a genome of massive numbers of overlapping sequences ***and*** the ability to translate them into proteins. Just by changing the ratio of nucleic acids in the genome, the code in that organism changes, which has in fact been observed in the data. Finally, there is a very good reason to suspect that it is not degenerate, at least not in the way currently proposed.

Good morning, Dr. Pangloss, how do you do?

We can now combine all of these seemingly disparate approaches to the distribution of assignments, beginning with the un-weighted table that we developed earlier. We know that the table creates three kinds of triplets and six kinds of codons

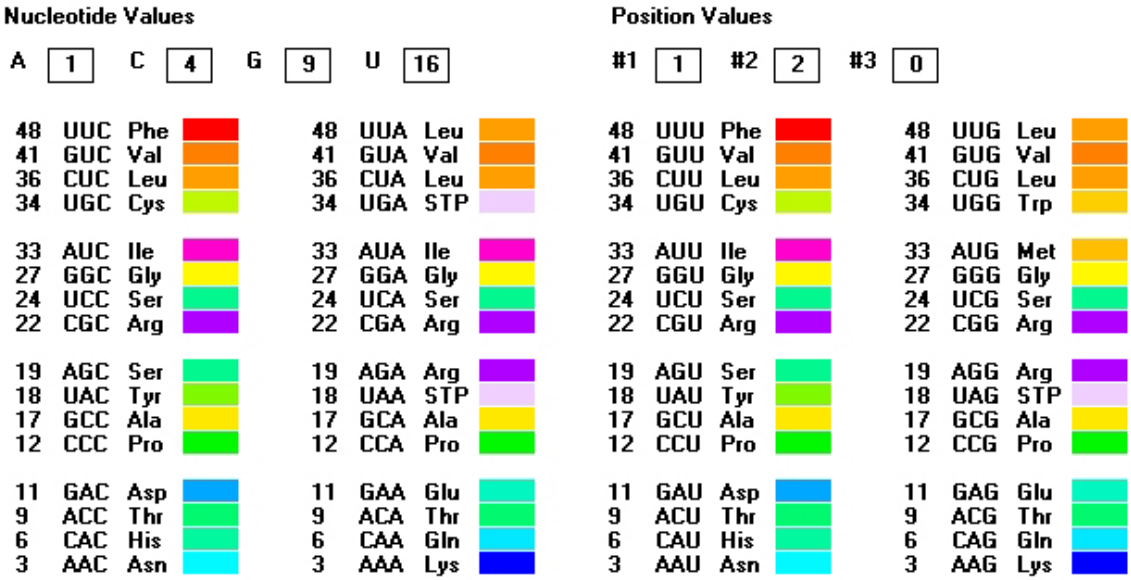


I	II1	II2	II3	III1	III2



Class	Type	Number
Primary **(I)**	1	4
Secondary **(II)**	1	12
Secondary **(II)**	2	12
Secondary **(II)**	3	12
Tertiary **(III)**	1	12
Tertiary **(III)**	2	12

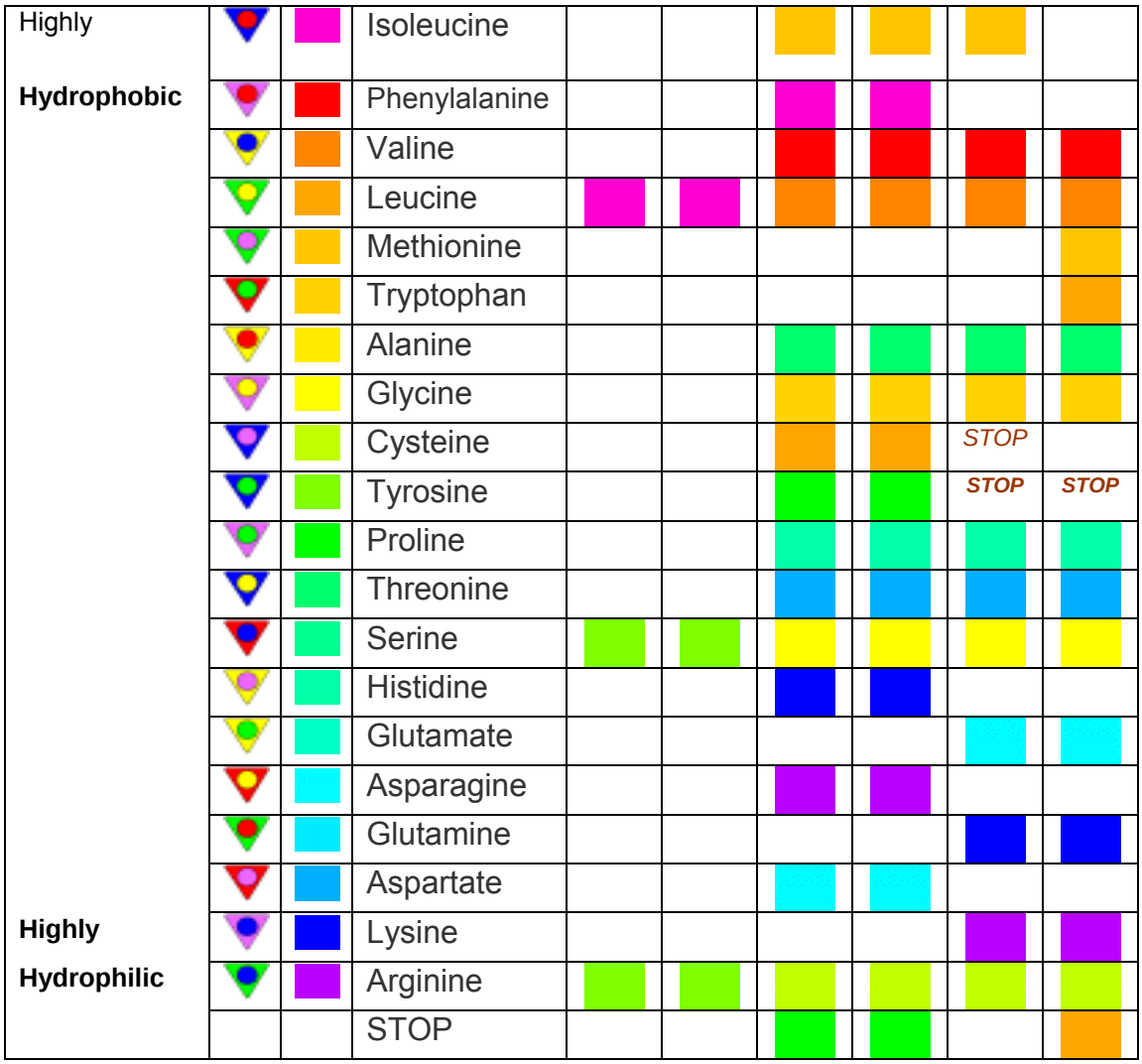
We are interested in biasing the unbiased data to see if we can demonstrate a correlation with the possible assignment bias inherent in the actual assignment mechanism. To do this we borrowed the formula from the textbook grid to create our own weighting of the assignment data, and we found a rainbow. The rainbow began with a systematic biasing of codons - one to sixty-four - and combined this bias with amino acid water affinity. We also used a similar formula to bias triplets with respect to nucleic acids “not chosen” in the codon, which seems kind of funny, but we are searching for a balance of forces. The Gamow test demonstrates that there is a correlation between the triplet energy and the codon energy. Now we can finally apply the weighting of codons and triplets to the un-weighted grid. We will do this by subtracting the triplet energy (energy hole) from the codon energy (energy present). This can rightly be viewed as creation of “wobble” groups, because mathematically it has the effect of dropping the last nucleic acid in the triplet and only weighting the other two.



This data table is essentially the subtraction of the triplet table from the rainbow codon table, and it seems to have the least of all possible patterns. However, it creates sixteen distinct four-codon energy categories that can be ordered as follows:

#		E	C1	C2	C3	C4
1		48				
2		41				
3		36				
4		34				
5		33				
6		27				
7		24				
8		22				
9		19				
10		18				
11		17				
12		12				
13		11				
14		9				
15		6				
16		3				

Each category is a collection of four codons that share nucleic acid identities, but interestingly they are created by different nucleic acids in position #1 and #2 of four different triplets. This is a distinction that only comes to light from the un-weighted twelve-symbol grid. All groups can be ordered based on the common energy of the group. When we plug the assignment data that nature has given us into this presentation format, we see a 100% correlation between assignments and these categories. No amino acid is assigned outside its category unless it has more than four codons.

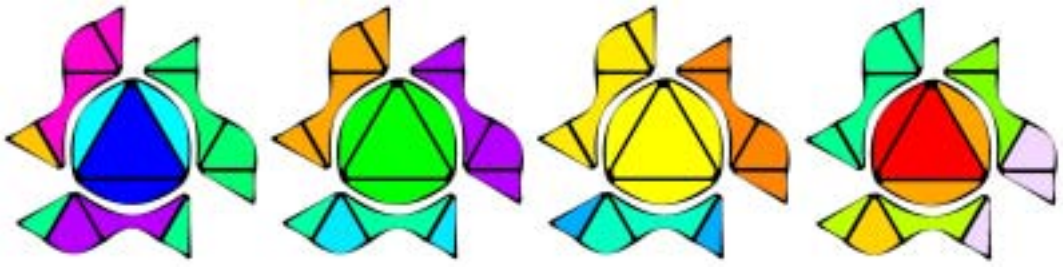


Most people will be tempted to attribute this perfect correlation to wobble, but this is a mistake. For wobble to be the force behind this perfect assignment pattern, it would have to be important enough to actually work. Remember, wobble was a factor proposed to explain and mitigate redundancy. It was a way to reduce tRNA populations below sixty-four, but there are more than sixty-four tRNAs in an organism. Either wobble doesn't work, or it can't work for this tRNA

reducing trick; either way it cannot be the force driving this pattern. It is likely that wobble does work as a serendipitous advantage of something else. Without wobble, an organism requires even more tRNA. The pattern here is related to the symmetry of a dodecahedron, and a weighting of nucleic acids on some metric within the dodecahedron. I suspect that water affinity is playing a role.

Many times energy, or force will obey an inverse square law. The energy of light and the force of gravity are two prime examples. In this case we were obligated to square our values of nucleic acids so that they would create distinct "energy" values. We can now demonstrate that the assignment of amino acids follows some rule that not only respects these values, but depends on the symmetry of a dodecahedron. It now seems implausible that the mystical, ancient assignment process was not following a scheme that turned on the symmetry of a dodecahedron, and somehow obeyed an inverse square law of molecular forces.

We can speculate on what these forces might be. If we orient the sixteen discrete energy categories within their natural habitat of a dodecahedron, we see their unmistakable pattern, and can begin to guess what they might be balancing.



There are at least two apparent forces balancing in the assignment data. The most obvious is water affinity, as there is a strong suggestion of a rainbow running from Methionine at the start to STOP at the end. This presentation of the data recalls a fanciful crystal that we created earlier.

The second force being balanced is perhaps less obvious. There is also a clear consistency to the overall pattern of assignment, but it is only apparent when viewed from the context of shape. Note the middle of the progression. The two middle clusters comprising the 32 codons in the center of the code demonstrate the strongest symmetry with respect to assignment consistency. They seem to support two amino acids, Glycine, the universal swivel, and proline, the universal latch. A code that placed importance on these properties of amino acids can be expected to build around them.

We can suspect that the code is a recursive process, which begs the question of what is being recursed. The linear model is based on the concept of recursion of points, generating a line – a number line. Rafiki sees the process as a recursion of shapes, generating a sequence of shapes. This recursion is of course itself recursive, and complex shapes will rapidly accumulate.

Constrain, Leverage, Build – Languages Revisted

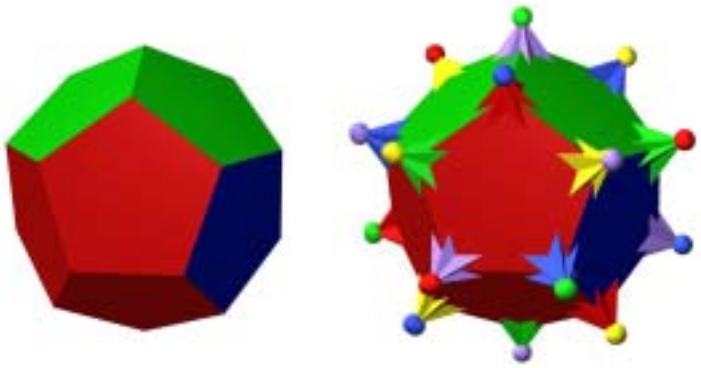
If we view a living system as a crystallization process, as crass as that sounds, the interface of crystal symmetries highlights a point of interest. Every living thing on this planet (that we all agree on as living) is based on carbon. Organic carbon organizes primarily around two key symmetries. The nucleotides take care of information storage and transfer, and the amino acids are primarily responsible for building stuff. Nucleotides favor the dodecahedron due to the five-ness of their ribose rings, and amino acids are tetrahedrons due to the four-ness of the α-carbon.

The dodecahedron plays the role of a seed crystal to the tetrahedron, despite the symmetry incongruities. Unlike ordinary seed crystals, however, tetrahedrons do not stay with dodecahedrons; they dock only temporarily. It is just long enough to link up with other tetrahedrons, and they're off again to frolic in the ocean of chaos. In this way the dodecahedrons teach the tetrahedrons a sequential alignment. The interaction of these two symmetries is not permanent, but it is repetitive. The exact same molecule can play the role of seed crystal countless times; it is a re-usable template, a teacher to millions. The system has the ability to learn, and Life here is at least a four-plus-billion year accumulation of learned crystal growth tricks. Life has the capacity to save a trick once learned and build upon it. Tricks can then be modified and combined. No single trick can amount to a hill of beans, but when taken in context, repeated bazillions of times, and timed just right, the whole network of tricks is truly magical. The resulting complexity of the system is staggering, but the fundamental rules at its base are not.

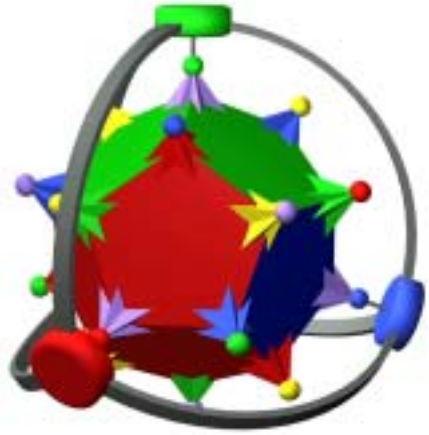
All of the clever tricks life has learned are the result of the interface between one set of molecules and another, which amounts to an interface of shapes, the forces they represent, and the information they possess.

M.C. Escher
Reptiles

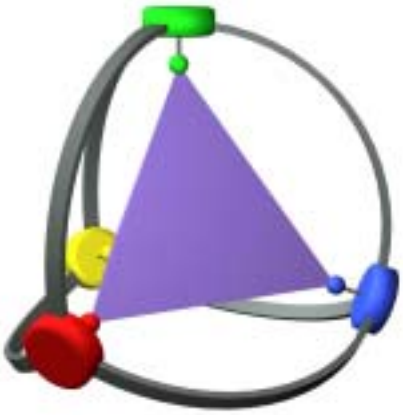
Of all the interfaces in the natural world - gas-liquid, hot-cold, order-chaos - the fundamental interface enabling Life is the interface between shapes. It is a basic set of shape-based relationships, logically and consistently producing irregularity and diversity. The shapes themselves are completely regular, but the interface between them produces output that appears irregular. If the system is to be dynamic and interesting, and indeed this is exactly what it is, it must escape simple regularity. It functions as if it were aperiodic. This should be unmanageable from a builder's perspective, especially for a builder so simple as a molecule. Fortunately, there is logic in the madness. The shapes themselves serve as holding platforms for the information about this aperiodicity, creating a form of language between them. This brings us back to polyhedrish – the language of shapes.



We return to our four-color dodecahedron, and the same dodecahedron with the dual tetrahedrons embedded inside. We now have enough background to ask how these structures might store information, perform self-replication, and direct the assembly of proteins. More interesting, we can ask what kind of language might get these jobs done, and what system could bootstrap itself into a position to do them. We know from our previous work that we can find 120 unique tetrahedrons inside this object, but keeping all of them straight is a real… mental chore. We will help ourselves in this task by imagining a place-keeper, or reader tetrahedron. We can surround our amalgamated shape with this reader, and label the vertices to track its position.



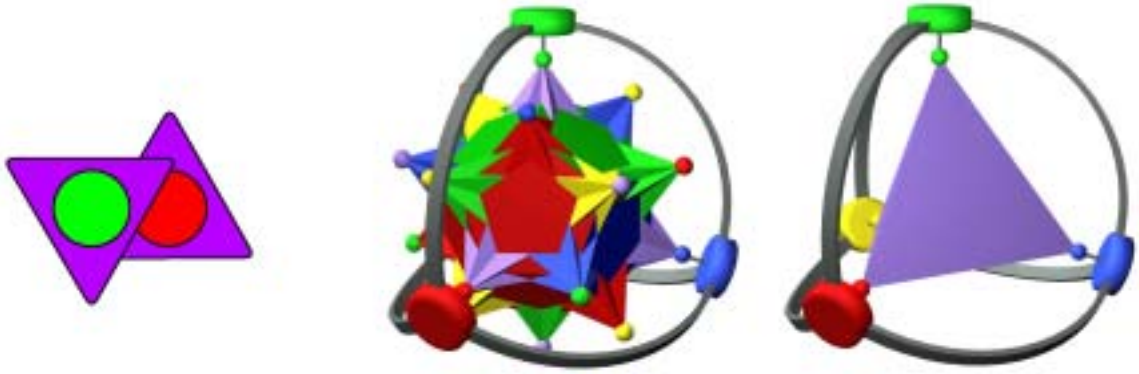
The tetrahedral reader, which we will call the glider, serves as a marker for the shared points of a given tetrahedron within the dodecahedron, which we will call the globe. For instance, the glider above is aligned with the solid purple tetrahedron within the globe as follows.



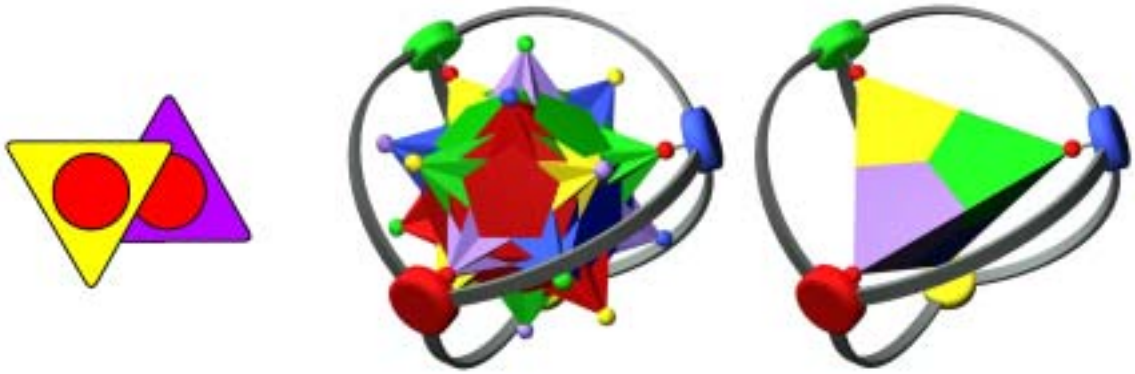
There are many possible languages that can direct the glider – and our thinking – to the exact tetrahedron within the dodecahedron. We have a huge number of possibilities for the structure of this language. We humans will understand any language differently than a molecule would understand that same language. Humans – surprisingly - require more information than

molecules to get a clear picture in this domain. Molecules require very little information to get the picture when it comes to shapes. However, Life is a very complex set of hierarchical information processes, and we would expect the language employed by Life to be at least as powerful and complex as the English language, if not many orders of magnitude more so. After all, molecular languages subtend spoken languages, and the sum of all living processes owes its existence to a long development of the “codes” of Life.

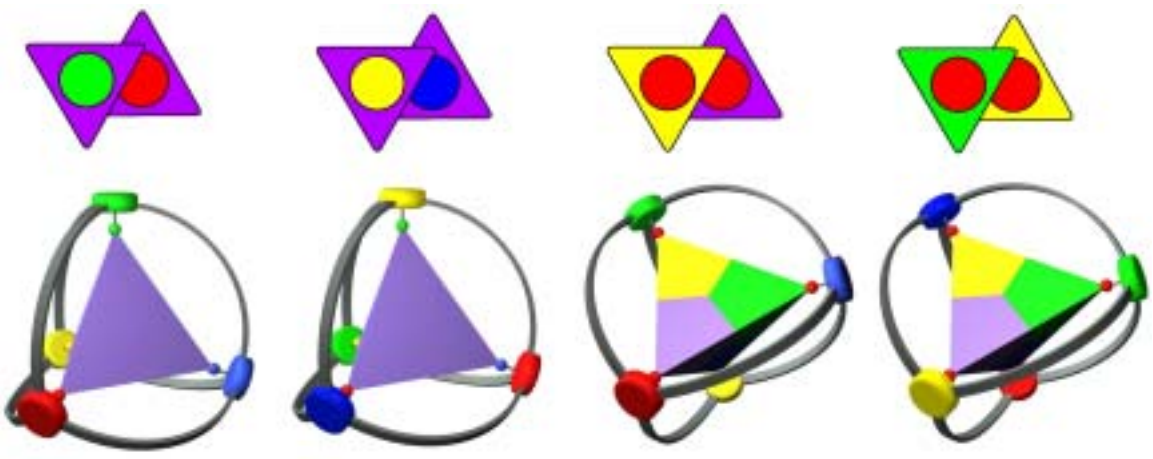
We will take advantage of the relationships of tetrahedrons within a dodecahedron to create clear symbols for humans to talk shape with each other – I call these symbols glyphs. Start by recognizing that the above tetrahedron only requires two points of specification, allowing conventions for our glider. The first point specified on the dodecahedron will be the green point on the glider, and the second point specified on the dodecahedron will be the red point on the glider. Since a tetrahedron and a colored ball identify every point on the dodecahedron, we can specify any dodecahedral point with a triangle and a circle. If on our glyph we put the first point in front of the second we can always identify first and second no matter what our orientation to the glyph. With this system we can identify tetrahedrons as follows.



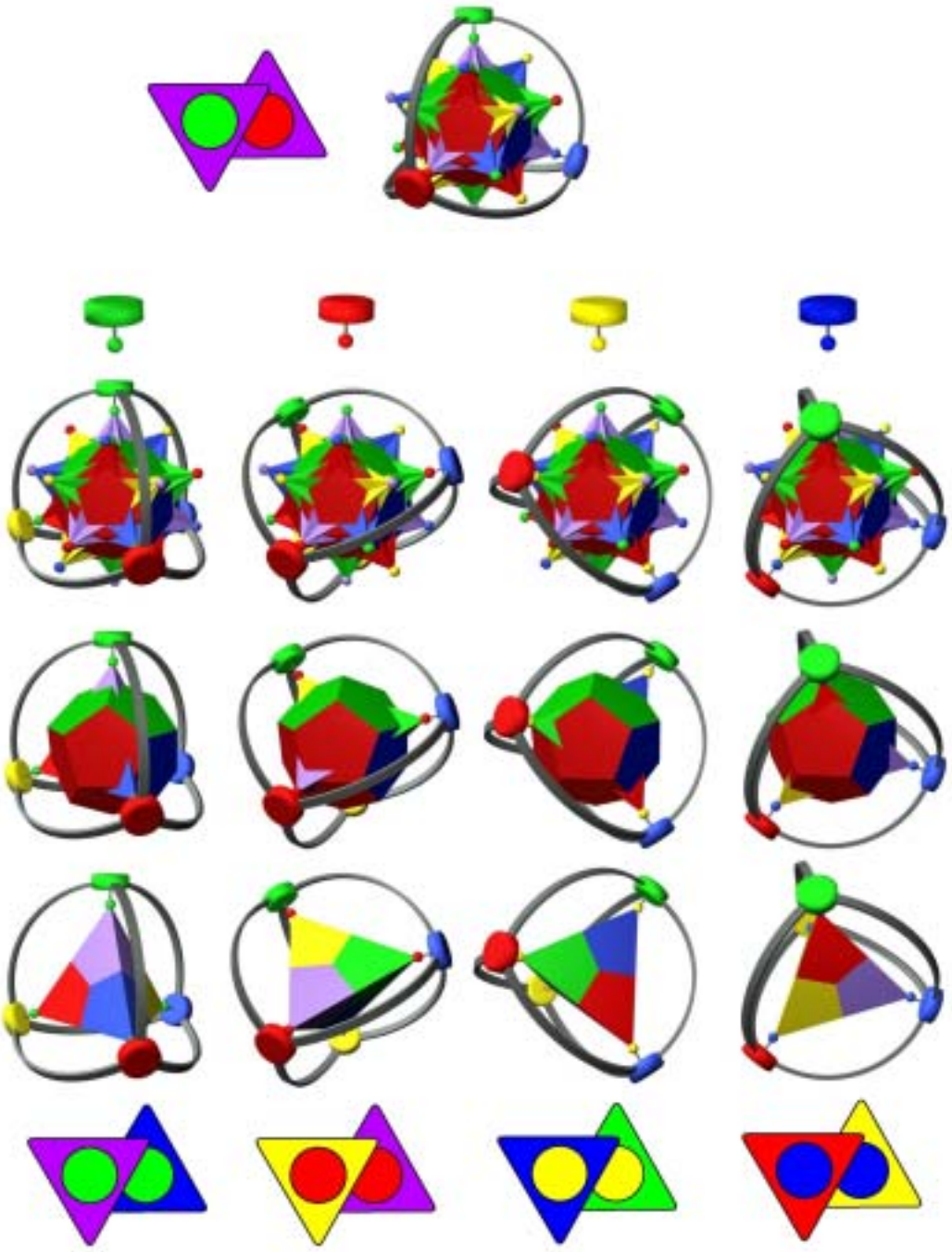
With this glyph we can easily identify the tetrahedral “meaning” by the visual information carried in the glyph. Another example of a glyph and its tetrahedron follows.



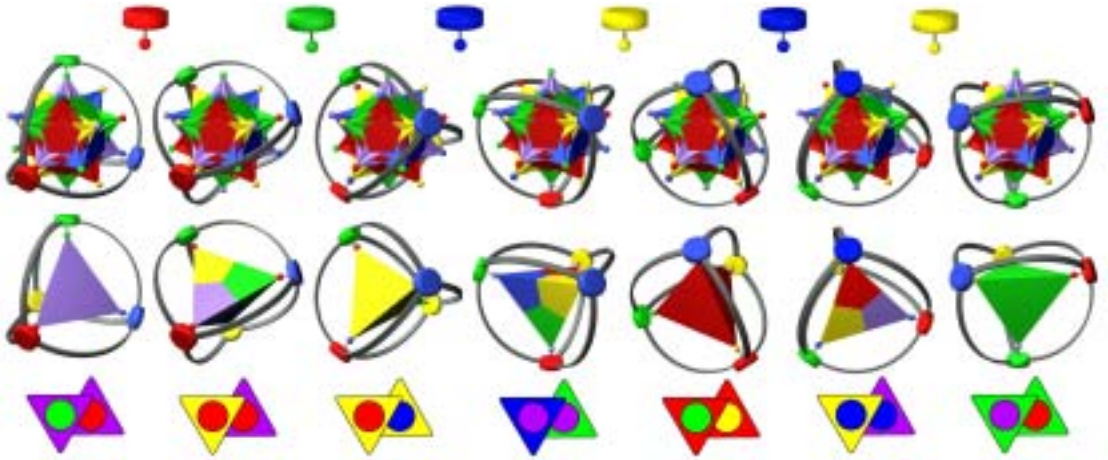
Not only are these two tetrahedrons identified by these glyphs, but the exact glider rotation out of the twelve possible rotations is identified as well. Following is an example of both tetrahedrons with two of their twelve rotations.



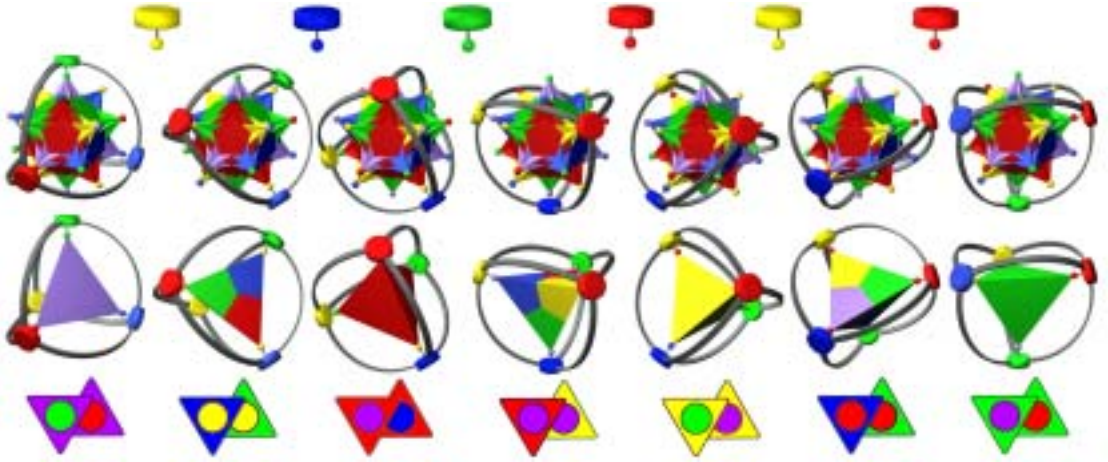
There are 120 tetrahedrons, so we need 120 glyphs to implement this language. This is almost seven bits of information. But we saw earlier when we tessellated the dodecahedron that we can leverage the symmetry relationships to reduce the information requirement considerably. Every tetrahedron shares each of its four points with four other tetrahedrons. To reference one of these linked tetrahedrons we merely specify the color of the glider aligned with the shared point, and rotate the other three in the only way they can be rotated without disrespecting the dodecahedron. Each tetrahedron can therefore support four different rotations to another linked tetrahedron as follows.



This provides us with a simple system or language for recording a “walk” through the linked tetrahedrons within a dodecahedron. Here is an example of a six-step walk.



Interestingly, here is another six-step walk that is completely different, but it starts and ends in the exact same place.

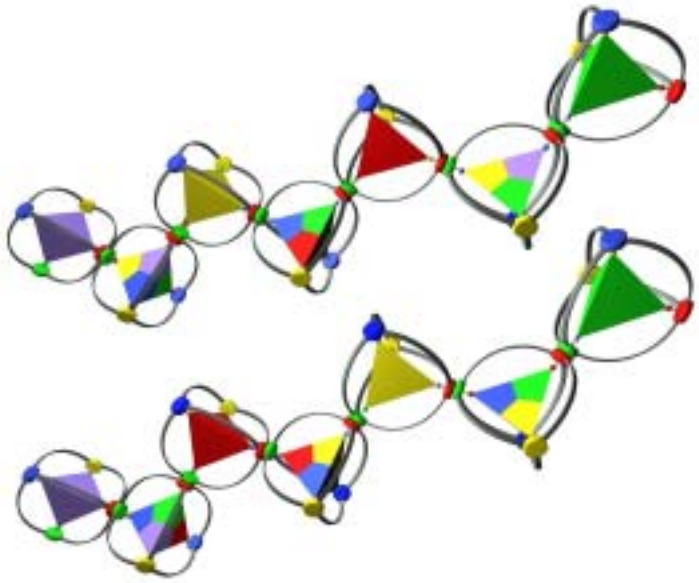


There are in fact 16 unique six-step walks that start and end in this place. Amazing! There are no “target” tetrahedrons that require more than six steps to reach, but any of them could have valid walks that require ***more*** than six steps without duplicating a position on the path. This means that these maps or paths or sequences are nonlinear; they are aperiodic.

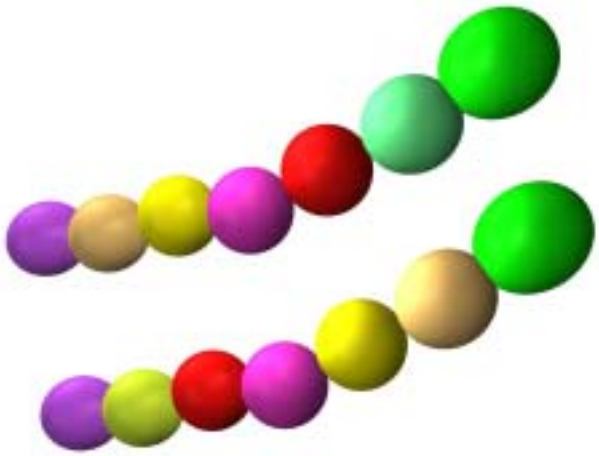
Two periodic shapes that want to team up to form aperiodic crystals would benefit from an aperiodic language and maps such as this. If the interface between them, and therefore the maps, were regular or periodic, then the resulting behavior would be regular crystal growth. Life would be frozen. Regular shapes interfacing in a regular way have no advantage in dealing with chaos, and certainly they have no way of embracing it. For life to thrive it needs a logic structure and a strategy that is not regular or FROZEN. It must embrace change and benefit from it.

Above is a subtle but exceptionally powerful demonstration of the missing element in the linear model. We can get a better feel for this with a side-by-side comparison. Let’s say that we somehow stumble upon a mechanism that joins

the above tetrahedrons consistently first point to second. What do we get? In the case of the Rafiki model we get the following.



Here we have started and ended with the same two tetrahedrons, used the same amount of information, but created two entirely different chains of seven tetrahedrons. We know the exact pattern and rotation of each tetrahedron in both chains because of its relationship to the first. In the case of the linear model, we do not have the advantage of describing actual tetrahedrons, so we could never be this specific. In the case of the linear model we get something like the following.



This isn't too awful, except for one thing: It took a helluva lot more information to produce the sequence in the linear model! Each strand in the first illustration required twelve bits of information to generate. The linear model requires 36 bits of information to do its "magic". Not very impressive is it? From

the linear model we are told nothing of the really important stuff, like which tetrahedron in which orientation. Linearity seems content to hemorrhage valuable opportunities to say something important. It requires three times more information input for less than one-sixth the information output. How could this system win any competition, let alone the most important, ceaseless, brutal competition in the history of this planet?

It couldn't.

The advantage enjoyed by the first system is unexpected – it is called constraint. Sometimes a constraint can be an advantage when it comes to information. This was the brilliant insight by Wolfgang Pauli when he posited his electron exclusion principle. Let me say that again. Sometimes a constraint can be an advantage. When one develops an information system, one should leverage constraints. If you begin programming computers using a PC Junior running DOS, you will learn this or you will die. If you know your constraints, you can use them and get more from less.

There are two questions that must be asked of an information system:

What are the constraints?
How can we use them?

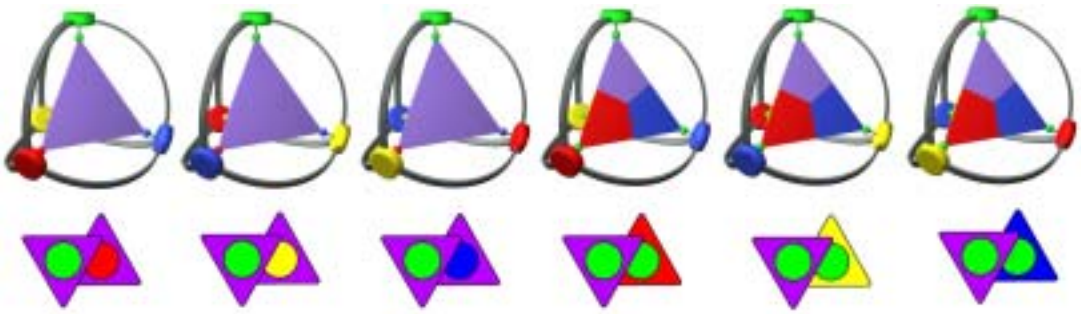
If we begin a study of an information system, like Life, we must ask the exact same questions. Life is constrained to a molecular system, and molecules are constrained to shapes. How can we use this? The above example illustrates this perfectly. Because we are constrained to tetrahedrons, a language that leverages tetrahedrons will win the day.

So, is *this* the language of the genetic code?

No.

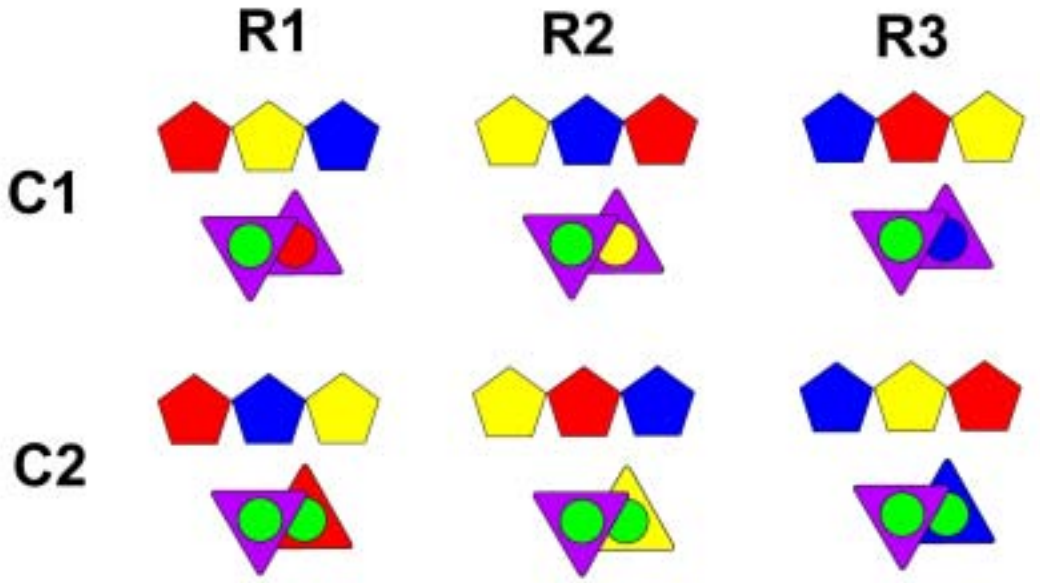
What… wait, I … what?

The above illustration is for humans; we are not molecules. We could never appreciate this point on a molecular level, but we can get close. We are required to visualize the information in order to understand it, and our visual information processing is more cumbersome. However, we can begin to guess at the constraints, and begin to guess at the best system to take advantage of these constraints. We must start with the difference between the dodecahedron and the tetrahedron, recognize their relationships, and imagine systems that would allow them to talk. The above illustration demonstrates that it is useful to first identify a point on the dodecahedron. From this initial point we are left with six options of vertices to assign as the second point.



We can easily see that there are two basic choices of tetrahedrons – purple major and green minor – and that each of these has three rotational choices as well. These are constraints of the system, and it would be nice if we could find a place to use them. We know that the peptide bond contains similar constraints in that it can take one of two major bond configurations – cis and trans – and each bond has three major stable rotational energies as demonstrated by the Ramachandran plot. Let's look at our dodecahedral information and see if it can easily assimilate these constraints.

We know that three faces can specify a point on a dodecahedron. If each face carries one of four colors then each point can be specified by six bits of information. We also know that there are six permutations of three faces, and each face will initiate exactly two permutations. We could easily match all of these constraints in a table.



Things get even more complex when you consider that molecules can cheat. We humans can observe the rule that four colors means two bits, but a molecule might not be this formal. Every molecule feels the influence of every molecule around it. When we take this into account, we must consider the context of every nucleotide as well as its identity. This might be the real value of the McNeil subscripts; they can turn two bits into four. Now there is at least a

reason to suspect that the system is capable of making distinctions between permutations of the same three nucleotides based on their context.

This requires that each codon must have a context. It must take into account the surrounding nucleotides in order to determine its own identity. With these new constraints molecules are capable of identifying any tetrahedron in a dodecahedron from a single point. They can leverage every drop of information and then some out of the system because of, not despite its constraints. There is plenty of evidence available today that demonstrates that this system is interested in leverage, not waste, and why shouldn't it be?

Polyhedrish appears to have a perfect set of constraints to serve as a platform for packing in the shape based information in the genetic code. When the system got started there was no possible mechanism to do anything else. Identity of a molecule would take a backseat to shape. The string of beads produced by the information in the linear model would require quite a bit of infrastructure to get going, whereas shapes are plug and play. It seems logical that identities of molecules would be far less important from the word go, and shapes would dominate the action. The information system capable of storing and repeatedly utilizing any crystal trick must be all about shape. With time, lots of time, and errors, lots of errors, the system would learn how to leverage its own constraints. My guess is that it is still learning. Some tricks will inevitably work in some environments better than others, and environments are constantly changing, so it is a good idea to keep options open, rather than frozen, to the extent possible. The functional imperative is change, not freeze.

Symbiosis is a powerful pattern in the history of Life. We could never have gotten this far without it, and it's natural to draw a parallel to these shapes. It's as if two shapes act symbiotically to get a job done that one could never do alone. There clearly is a non-linear logic at work beneath the genetic code, but we have somehow been blind to it. It is as if the symbols we see, and the meaning we infer are completely different from the universe in which the molecules operate. To go beyond our natural limitations, we must imagine creative ways to shed our anthropomorphic biases, and we must imagine bizarre new tools to help us do this.