Ron Maimon: Collected Works (Biology, Cognitive Science and General References)

Edited by Keith Y. Patarroyo

keith.patarroyo@gmail.com

October 30, 2021

Contents

Introduction

Ron Maimon is a theoretical biologist and physicist that was very active at the beginning of the 2010s in question-and-answer websites like the StackExchange family of sites and Quora. There he answered over 2000+ technical questions in topics such as physics, mathematics, biology, among others. Ron always emphasized strongly originality in both well-known and open problems, he kept this belief so firmly that at some point he claimed: "If nothing is original, everything is dead". I believe this element made him stand out among the average textbook or consensus-based answers.

I got to know Ron's ideas when pursuing my physics degree in the early 2010s, when browsing for information I was always pleasantly surprised with Ron's answers, they had an insight that was not found in the mainstream literature. Moreover, it was Ron's post about Stephen Wolfram NKS that convinced me that Stephen was not a total charlatan (as the media sometimes portrays him), but one could find deep insights in his content if one knows where to look. At the time I was mostly interested in Ron's insights on the sociology of science, the relation between private and public science, sadly I skipped the biology part due to ignorance of the topic.

It was only years later that I got interested in the topic of replicators and the origins of life when I came back to Ron's post about Stephen Wolfram NKS. It occurred to me that some interesting insights might be found since I remembered he touched on some of these topics in the post. There Ron describes a personal theory about the Origin of Life using ideas about computation, chemistry, and biology. For the past months, I've been reviewing the literature on the topic and I can mostly say that Ron's ideas are original, but to my knowledge unconfirmed. In this collected works my aim is to present all Ron's content available online related to the ideas of biology, computation, and the origins of life. I believe that at the very least they present a new perspective to the field of life origins and can be of value to any researcher on the field.

ArXiv

Ron's main ideas about biology are presented in the following *preprint:*

Maimon, Ron. "Computational Theory of Biological Function I." *arXiv preprint q-bio/0503028* (2005).

Abstract: This series presents an approach to mathematical biology which makes precise the function of biological molecules. Because biological systems compute, the theory is a general purpose computer language. I build a language for efficiently representing the function of protein-like molecules in a cell. The first paper only presents the kinematic part of the formalism, but this is already useful for representing large-scale protein networks. The full formalism allows us to investigate the properties of protein interaction models, ultimately yielding an estimate of the random-access memory of the proteins, a measure of their capacity for computation.

Quora & StackExchange

Theoretical Biology

1. - Is Stephen Wolfram's NKS, an attempt to explain the universe with cellular automata, in conflict with Bell's Theorem?

Questions: Is Stephen Wolfram's NKS, an attempt to explain the universe with cellular automata, in conflict with Bell's Theorem?

Stephen Wolfram's A New Kind of Science (NKS) hit the bookstores in 2002 with maximum hype. His thesis is that the laws of physics can be generated by various cellular automata--simple programs producing complexity. Occasionally (meaning rarely) I look at the NKS blog and look for any new applications. I see nothing I consider meaningful. Is anyone aware of any advances in any physics theory resulting from NKS? While CA are both interesting and fun (John Conway, Game of Life), as a theory of everything, I see problems. The generator rules are deterministic, and they are local in that each cell state depends on its immediate neighbors. So NKS is a local deterministic model of reality. Bell has shown that this cannot be. Can anyone conversant with CA comment?

While NKS came out with much hype, and with a lot of skepticism from scientists, the scientific ideas there are not completely trivial. I just think they are not foundational for the science of physics (at least not as we know it so far), rather they are foundational for the science of biology.

The main discovery made by Wolfram (although with an important confusion which I will explain below, and with an extremely significant precursor in Conway's game of life) is that a simple 1 dimensional cellular automata whose rules are chosen at random will have a finite not so small probability of being a full computer (in Wolfram's system, 2 of the 128 possibilities). The proof that the system he found, rule 110 in his terminology, is actually a full computer only came two decades later, thanks to the pioneering work of Cook (working under Wolfram). But it justifies his focus on the system as central to science, since before, it was often implicitly assumed that to get a certain amount of complexity, you had to put in complexity by hand. This result is also present in Conway's system, but Wolfram's work is somewhat complementary, because the information flows in 1d systems make it more difficult to imagine a full computer emerging. The fact that it does anyway (although, as Cook's construction shows, with horrible running times, because of the difficulty of shuttling information long distances) is surprising and notable.

This is not so important for physics, because any attempt to model physics with cellular automata will have to be grossly nonlocal in order to avoid Bell's theorem. This is not so implausible today, given gravitational holography, but Wolfram suggested that there would be a direct correspondence between local elementary particle paths and automata structures, and these ideas are flat out impossible, and were ruled out before he proposed them, by Bell's theorem. This means that the chapter of his book dealing with physics is completely wrong, and may be ignored.

But this work is important in a completely different way, it is the foundation of biology!

(EDIT: Chaitin's new book makes some brief comments about NKS which echo the main biological points below. I am not cribbing Chaitin, his book postdates this.)

Biology and Religion

The most puzzling aspect of the world we find ourselves in is that we are surrounded by complex computing devices not of our own design! Namely ourselves, other people, animals, plants, and bacteria. How did these computational structures get built, when we have to work pretty hard to make a computer? It seems that there is a puzzle here.

The puzzle has, in the past, been resolved by assuming some sort of magic put life on Earth, a supernatural agency. This idea is clearly at odds with the laws of nature as we understand them today, but it is important to keep in mind the superstitious answer, because elements of it are salvageable.

The superstitious answer is that God came down into the primordial soup, and mixed up the molecules to make life. The notion of God is not clearly defined in religious texts, where rigor is not the top priority. But I will try to give a positivistic definition below. I find that using this positivistic definition, which does not mention anything supernatural, I can translate the thoughts of religious people and make complete sense of what they are saying, where otherwise it just sounds like the ranting of delusional people suffering from severe brain-damage.

In order to discuss biology sensibly, I believe one must understand this religious point of view thoroughly, in a logical positivistic way, because it is important in biology to the same extent that it is completely unimportant in physics.

In a complex system, such as human social structures, we tend to observe patterns which cannot be attributed solely to the actions of individual people. For example, the protestant reformation seem to have happened all at once, within the span of a few decades in the early 16th century, where Church reformers were active and working for centuries before, with very little success. What made it happen? It wasn't just Luther and Calvin, it was also a network of businessmen and bankers, and disenchanted Catholics. The discovery of America was important in some way, as was the expulsion of Jews from England. To my mind, the most important was the 14th century edict forbidding usury by Catholics, which prevented the formation of banking. But it clearly wasn't one cause, nor was it the work of one person working alone.

When we see such complex phenomena, it is reasonable to attribute them to the working of a larger intelligence than the intelligence of any individual, and this is the intelligence of the collective. Just as a person is a collection of neurons, not any one of which is responsible for her intelligence, the society is a collection of individuals, no one of which is responsible for everything the society does or thinks collectively. The collective pattern is in many ways smarter than the individual--- it contains collective memories, in traditions and conventions, which inform individual action in complex ways.

The notion of of god (lower case g, like Zeus, or Mars) in ancient cultures is the name given to the entities formed from collective human actions. They are nebulous, but important, because the decision to go to war cannot be attributed to any one person, but to an entity, the god of war, formed from many individuals working together with the aim of forming a coherent collective which will lead the society to make that phase-transition of behavior which is going to war. Identifying a notion of a god, and explicitly setting people working for this god, makes them aware of the fact that they are working as parts of a machine, not solely as individual actors. Further, it can inspire them to act without direct orders from a King, or a priest, just through their own introspection, so as to best achieve the goal.

The notion of god was refined somewhere in India or Iran into the notion of God (upper case G), from which the Brahma cults and Abrahamic religion, and Zoroastianism emerged. This notion suggests that the conflict between gods is similar to the conflict between individuals, the gods also make collectives, and some win and some lose. In the end, there is a notion of a supreme God, the God which is the limit of the collective of whatever gods survive, defined as infinitely high up the god hierarchy, and demanding ethical actions.

This limiting conception of God was considered so important by the ancient thinkers, that they let all their other ideas die away in the medieval collapse, choosing to preserve only this through the middle ages.

But in addition to the practical notions of guiding behavior in collectives, the ancients also attributed all sorts of supernatural feats to God, including creating the universe and

hand-designing life. These ideas about God are out of place with the conception as a meta-property of a complex system, and are completely contradicted by modern scientific discoveries. They are superfluous to religion, and detrimental to it, because they make people expect miracles and divine intervention in ways that violate the laws of nature, and such things just never happen.

The notion of God, as far as I have been able to make sense of it, is essentially a limiting computational conception--- it is the limit as time goes to infinity of the behavior of a complex system where the computational entities combine and grow in power into ever larger units. The idea of the limit suggests that there will be a coherence between the units at all levels, so that in the infinite time limit, for example, all societies will agree on the ethical course of action in a given circumstance, and will agree on how to organize their economies, and structure their interpersonal relations. These predictions are surprising, considering the divergence in human behavior, and yet, history suggests that such a convergence is slowly happening.

This computational decidability in the evolving limit has a direct counterpart in the idea that as mathematical systems become more complex, by reflection, they decide all arithmetical theorems. This is not a theorem, but an observation. It is noted that as we go up the tower of set theoretic reflection principles, more and more arithmetic theorems are resolved, and there is no in-principle limitation that suggests that the theorems will not all be decided by strong enough reflection. This is Paul Cohen's "Article of faith" in mathematical practice, and I will accept it without reservation.

Further, the article of faith tells you that we already have a name for the mathematical idea of God, it can be identified with the concept of the Church Kleene ordinal, the limit of all countable computable ordinals. Any computable formal system is only able to approach this ordinal gradually, and this ordinal is infinitely rich. If you have a description of this ordinal, you have a reflection principle which should be powerful enough to decide all theorems of arithmetic, to decide what consequences of any axiomatic system will be.

Because this ordinal has all the theological attributes religious folks attribute to God, in relation to pure mathematics, I consider it a sort of heresy to assume that there are larger ordinals. In particular, any notion of the first uncountable ordinal, or inaccessible ordinals, are only meaningful once they are placed in a given axiomatic system, and then they should collapse in the appropriate countable model to be less than the Church Kleene ordinal. This is not technically precise, but it gets the main idea across (it is easy to collapse the ordinals to be countable, but it is not so easy to rearrange the scheme to make them less than Church Kleene, but this is because within any deductive system which is of a set-theoretic nature, you can give a name to the Church Kleene ordinal, and define this ordinal plus 1, etc. These technical considerations are not so significant for the philosophical ideas)

So the interpretation I will take for religious doctrine is that God is to be identified with the Church Kleene ordinal, no higher ordinal is to be interpreted as actually higher, and gods will be identified with human collectives acting together to form a unit greater than the individuals. The monotheistic law of complex systems will state that all gods converge to

the ideal represented by God over time, as they battle it out in a Darwinian struggle.

Automata and Darwin's experiment

When you have a cellular automata capable of universal computation, there is a strange phenomenon--- sub-parts of it are always in competition with itself. To explain this, one needs to look at Darwin's experiment, detailed in the Origin of Species.

Most of the Origin is theoretical, but Darwin did do one important experiment. He took a square plot of land, and carefully removed all visible living things from the soil. He uprooted all the plants, sifted to remove insects, and left the plot alone to see how it would be recolonized.

What he observed is that the plant species that recolonized the plot were first of the fast-growing unstable variety, that a whole bunch of weeds and bugs spread over the new area. Then, over time, other more hardy species slowly took over from the weeds, until, many months later, the plot was indistinguishable from the remaining land in the lot.

The purpose of the experiment was to see whether there is an actual struggle for resources in nature. Darwin hypothesized that if nature is in constant struggle, different elements, which are more successful but slowly replicating, will only win out after a time over elements which are less hardy, but whose strategy is quick colonization of new territory. His observations were consistent with the idea that the living things in any area are continually struggling for primacy, and that the limitation is the finite resources in any given plot of land.

This idea can be tested in computing cellular automata. By zeroing out a square patch in a 2d cellular automata which looks stable, one can see whether the remaining data colonizes the space in a way that is uniform, or in a gradually transforming way. I did this experiment using an 8-bit cellular automata (256 values) with random rules, and I found that in many cases, those cases which are complex, the colonization is in stages, much as in Darwin's plot of land. The stages are short-lived, perhaps reflecting the limited computation possible in a small region with 8-bit values. It would be interesting to repeat the experiment using arbitrarily large integers on each cell, which can be thought of as representing a complex polymer, which catalyzes transformations on its neighbors

But the inhomogeneous colonization suggests that once you have a computing cellular automata, there is a constant competition between parts of the automata, which make collective computations, for resources. In other words, that Darwin's struggle is begun.

To make this idea more precise, consider dividing a CA in two, by placing a wall between the left and right half, and not allowing the halves to interact. If the CA is truly computational and complex, the two halves will not come to a statistical equilibrium, but will have complex structures on either side which acquire new characteristics at random over time, as its subparts evolve.

If you now remove the wall, it is unlikely that the left half will have compatible

characteristics with the right half. They will not be able to mix. So in this case, the two halves must battle for domination, and whichever half wins will impose its characteristics on the other half filling the whole space with cells which are compatible with its characteristics. These characteristics include typical CA "animals" or structures which are qualitatively similar in their relations, particular configurations which are only stable in the environment of other structures around them. It is difficult to extract these characteristics from a running simulation, because you don't know a-priori what to look for, but I am confident that it can be done.

This type of thing implies that there is continuous competition in a CA which appears the moment it is first seeded, and continues as long as it is operating. In this environment, Darwinian selection and evolution are possible even without any explicit self-replicating structure. Any self-replication is of very high-level qualitative traits, not of low level bit structures.

Replication and evolution

This point of view is different from the most usual point of view regarding evolution (which is not the one originally proposed by Darwin). The usual point of view is the modern-synthesis evolution, which suggests that evolution proceeds by copying bit-strings in molecules, with errors, and that the result is that optimized bit strings are eventually selected.

This point of view is extremely poor in modeling actual biological evolution. First, nothing you are familiar with actually replicated itself. People have sex, bacteria share genes, and crossing over is complicated on non-genetic sequences, it is only a simple shuffling on genes.

Further, mutations seem to be produced by shadowy internal mechanisms directed by complex RNA networks in egg cells and in testicles. They are not random copying errors. To assume that the biological world is produced by a process of copying with error, coupled with selection is as silly as the following parable suggests:

Many years ago, there was only one book. It was a cookbook, with detailed instructions on how to make macaroni and cheese. The book was copied by scribes, who made an error here, an omitted passage there, and these books then competed for attention. Some recipes were improved by the errors, others became unreadable. Eventually, the books grew in length, with new passages produced by accidental duplicated copying, until today, behold! The library of Congress!

This story is ridiculous. But it is this ridiculous story that is currently sold as dogma in the biological sciences.

It is my view that any realistic theory of evolution must be closer to Darwin than to the modern synthesis. It must take into account that the process of mutation is authorly, it proceeds by complex RNA editing of DNA sequences. It must take into account the idea that sexual selection is primary, so that mate-selection is the dominant driving force of

evolution in sexual species. It must also take into account the idea that the competition begins well before replication, and requires nothing more than a computing CA.

There is support for this position from computer experiments on self-replicating evolution. In order to test natural selection, little chunks of code were allowed to replicated and self-modify in the 1970s-1980s, to see what the end result would be. The end result was that the programs modified themselves until they found the shortest fastest self-replicator, which then filled up the computer memory.

At the time, this was considered a positive sign, the programs had evolved. But the obvious stasis in the final state leads me to see this as death of a complex system. There is no further progress possible from the end state, without an external agent to kick things around. The result is not a complex system, but a system trapped in a stable equilibrium of parasitic fast replication. Far from being a model of life, it is a model of a self-replicating cancer killing all evolution.

CA properties: Wolfram's Annoying Error

Wolfram classified Cellular automata into four types:

1. homogeneous end state
2. Simple periodic structures, perhaps separated, with different periods
3. Self-similar ("Chaotic") structures
4. Complex structures

Type 1 are automata that die. These just have a single stable endpoint that you always reach. Type 2 have infinitely many endpoints, but they are as simple to describe as a classical integrable motion--- you just have cycles of certain types, and to specify the endpoint, you give a list of all the cycles, and where you are in the cycle, and this specifies the result of running the CA from a given initial condition. These first two types of automata obviously will not reproduce a general purpose computer.

Type 3 are those automata that lead to self-similar fractal structures, like the Sierpinski gasket. These are more complex, so that the end-state requires an actual computation to specify, and wolfram identifies these with classical chaotic motions. I think this identification is wrong, but this is what it is.

The Type 4 are the complex automata, where you have to actually run them in full to figure out what they do. I don't like the final category, so I will now give my personal classification.

1. homogeneous end state
2. simple periodic end states, perhaps separated with different periods
3. Self-similar or statistically self-similar fractal structures
4. random automata, chaotic stable endpoint, stat. mech.
5. Complex automata, biology.

The class 3 is expanded slightly, and class 4 is divided in two. There are random automata,

which act to produce a randomized collection of values which wander ergodically through the allowed value space, and class 5, those automata which produce true complex behavior, with a way to map a computer into them with a map of reasonable complexity, which can actually be described by a finite procedure.

Because Wolfram doesn't distinguish between 4 and 5, he lumps together automata that are purely random, thermalizing into a Boltzmann type chaotic equilibrium, like automata 25, together with truly complex automata like 110. The distinction between the two is all important, but perhaps out of a pig-headed inability to admit his earliest classification was incomplete, Wolfram refuses to make it.

I will make this distinction. Type 4 automata are the analogs of chaotic classical systems, randomizing their information into a strange attractor, defined by the allowed values of clumps of sites, and a probability distribution on these. Once you know which allowed clumps occur with what probability, you can generate a typical output with absolutely no work, using a random number generator. It will not be the actual output, since this is deterministic, but it will be indistinguishable from the actual output for all intents and purposes.

Type 4 automata are just as a-biological, just as dead, as types 1-3. CA 25 is not alive. I am 100% sure that I am not misinterpreting Wolfram, because I specifically asked him, in person, at a seminar, whether he believes there is a map between CA 25 and a computer. He answered that he believes there is, but that it is extraordinarily complicated, and random looking. I am sure it does not exist.

Type 5 automata are exemplified by 110. Those are the ones which have predictable structures with non-randomizing behavior. These can be used to encode full Turing computation. That these are not measure 0 is an important discovery--- it gives an explanation for the origin of life.

The existence of typically computing CA's means that life can emerge naturally as soon as a system that can store large amounts of information spontaneously has interactions which are capable of forming a computer. This happens with 110, but it also should happen with random proteins in a pre-biotic soup, because, here we are!

The evolution of life, as I believe it happens, is purely molecular for most of the early stages. The proteins compete and evolve, producing a more precise class which can survive, which eventually catalyze the formation of nucleic acids (among other things), and learn to store data for later retrieval in nucleic acids. The nucleic-acid protein complexes then compute more, and learn to store data in DNA, for permanent storage (since DNA is much more stable). Finally, they package all this up in cells, and you have modern life.

This is a just-so story, but it is important because at no point does it postulate a self-replicating molecular entity. Such entities are poisonous for the emergence of life (as the computer experiments show), and it is good that they do not exist, otherwise life would not be able to emerge.

Wolfram's Sociological Agenda

There is a separate reason for Wolfram's lack of success in penetrating the scientific world which has nothing to do with the quality of his ideas (which are really not that bad). Wolfram made the conscious decision to pursue his science using private money which he raised by producing closed-source software, Mathematica, for sale to universities. In this way, he was producing a model for science research financed by private capital, rather than state money. Because Mathematica is so successful, many saw his work as a model for a New Kind of Capitalist Science.

This idea was very current in the pro-capitalism climate of the 1980s, where state sponsored and funded things were looked down upon, because of the constraints on individual freedom which the modern state imposed. The Soviet Union was the extreme example, there all science was driven by state decisions, which stifled certain fields like genetics, based upon the ideological position of the government. In the U.S., science was taken over by the government and made into big-science in the 1950s, explicitly so as to compete with the Soviets, and many people felt hampered by the big-money, big-science system, which excluded promising research avenues from consideration.

The lack of freedom in the state run system led many individuals to oppose it, and one of the ingredients in this fight was private financing. This was obviously only available outside of communist controlled regions. Wolfram politically made the decision to pursue private financing for his research.

The result reflects all that is good and all that is bad about privately financed science. It is good, because it allows the individual with an idea to pursue it indefinitely, and no outside criticism can stop or kill the work. They can self-publish, without worrying about peer review dismissing their ideas before they have time to germinate.

It is bad in several other ways, which have been the focus of academic criticism

- Self-financing requires the individual to amass large amounts of wealth, leading to sycophantism in those around them which prevents them from hearing cogent criticism, so that mistakes go uncorrected.
- In private enterprise, one does not cite sources. One makes it seem that one did everything on one's own. This is not compatible with academic conventions for citations and respect for the history of a field. While Steve Jobs arguably can take credit for the work of his employees, it is difficult for Wolfram to justify taking credit for Cook's work, even if he paid his salary.
- The Citizen Kane effect: the isolating and corrosive power of money easily lead to megalomania and isolation, which leads one to dismiss the ideas of others. This unfortunately can be seen in Wolfram's blithe one sentence dismissal of the exceedingly important work of Post, Friedberg, and Munchnik on Turing degrees below the halting problem. He claims that all natural CA's either are equivalent to the halting problem, or else random, or else trivial. This is the principle of "computational equivalence". But this is a nontrivial statement, and requires more evidence than what is presented in NKS.

The problems of private research are entirely complementary to the problems of public research, and there is no reason to dismiss the one entirely in favor of the other. But NKS shows those flaws in spades, and this is particularly grating to relatively low-paid public researchers, who have worked equally hard on their ideas, if not harder, without the megaphone of money to shout them out to the world.

I think that the newest thing in NKS is the financing model--- the idea that one can do research privately and independently. Perhaps this is the model of the future, but considering the relative success of publicly funded science as compared to private science, even in the most extreme repressive case of the Soviet Union, I am not optimistic that this is the best way. It is likely that one will have to deal with the annoyances and suboptimal features of public funding for the indefinite future.

Perhaps with an appropriate internet structure, like stackexchange, some of the censorship and group-think of public science can be mitigated.

2. - How did life on Earth begin?

Question: How did life on Earth begin?

There is a reasonable answer to this question, which is a synthesis and extension of some ideas of Stewart Kauffman (1969), Stephen Wolfram (1981), and other stuff that is not in the literature. These ideas are essentially updating the cybernetic approach due to Turing, Von-Neumann, and Wiener, which was marginalized and suppressed in biology once DNA was discovered.

The reason the computational ideas were put on the back burner is because DNA has an obvious replication mechanism, and the molecule gave people a picture of the origin of life immediately--- a nucleic acid formed, and began to replicate, and then evolution proceeded.

This idea is seductive, but I believe it is completely incorrect, and many others who thought about this, including Francis Crick, eventually came to the same conclusion--- self replicating nucleic acids are not the likeliest candidate for the origin of life. Crick was mystified, and proposed half-jokingly that it was pan-spermia. I don't think this is a reasonable answer either, since it just pushes the question to the origin of pan-sperm.

* Computations in modern cells

The main characteristic that distinguishes living from nonliving systems is the ability to do Turing complete computation, in a finite approximation, with an essentially limitless

memory capacity, and a processing quantity per unit volume. Each cell has an enormous store of stable memory, dwarfing the best solid-state memory chip in bit-density, and the processing speeds are on the molecular scale.

An RNA read and write can be done at thousands of bases per second, with error correction, and you can stuff millions of these things in a cell volume. So the potential RAM of a cell is 10s of gigabytes, and the processing speeds for this data is the rate at which the data can be copied and transformed, which can happen at rates of megabytes per second. These are comparable to a modern artificial home computer.

To see that modern life is potentially Turing complete is not very hard--- you can easily engineer a Turing complete system using bio-molecules, and it is easy to see that the storage capacity of DNA and RNA is sufficient for running software of the kind you have on a modern laptop. Further, we have a system capable of computing in biological systems for sure--- our brain. The computer itself was originally defined to abstract out the information processing done by the brain of a mathematician. So biological systems can compute, and do compute.

But the processes that biologists recognize as happening in a cell are not always sufficient to produce a full computation. at least not one of a significant size. If all that happens in a cell is the central dogma, then DNA produces RNA and the RNA produces proteins, and then only the proteins are computing anything. The proteins compute with a random access memory which is determined by their potential different chemical bonds to each other, and this is only a few kilobytes of RAM at the most. It's still computing, but it's a very small computation, compared to the amount of frozen data stored in the DNA.

It is unreasonable that a system that has gigabytes of ROM should only have kilobytes of RAM, especially that the DNA has to get written and proofread in the process of evolution. I will argue later that the proper computational ideas demand that there are exceedingly complex RNA networks active in modern cells, which compute at the gigabyte/teraflop rate.

This idea that RNA networks are required and appear in modern cells in a way that can do gigabyte computations is implicit in recent work of John Mattick. It is experimentally more and more certain every year, as new functions for RNA are discovered. I take this idea for granted, as it is the only way I can see to make sense of the computational capacity of modern cells.

* Computations in non-living systems

If you start with a pre-biotic soup of molecules, it is very simple to make a naturally computing system. This became clear after Wolfram's work in 1981.

The basic idea of a cellular automaton is that it is a model for information transformation in a system which can store stable discrete data. An example is molecules, which store data in the pattern with which they are bound to one another. These molecular patterns are transformed by catalysis, using other molecules, and the result is that certain bit-patterns

rewrite other bit patterns in a rule basd way.

Bit rewrite rules were studied by Von-Neumann, Coway, and Wolfram, and it was discovered in each of these cases that a relatively simple system will produce full Turing computation. Von-Neumann had a many-state one-dimensional automaton, with relatively complicated rules, but it was proved Turing complete relatively easily. Conway used a two-dimensional automaton with very simple rules, and this was proved Turing complete in the 1990s (although it was pretty clear that it should be Turing complete in the 1970s too). Wolfram found a very simple nearest neighbor automaton which was proved Turing complete around 2000 by Cook, a Wolfram employee. The proofs are relatively difficult, because they require building a computer from the information transformations in the cellular automata, but the general program makes it clear that as long as an automaton has "complex behavior", which means that the system doesn't die out to a stable pattern, doesn't devolve to a simple fractal pattern, and doesn't wash out to completely random noise, as long as there are identifiable structures that persist long enough to impress their data on other structures, then you have Turing completeness.

This is not a theorem, it is a principle, and the principle was called the "Principle of Computational Equivalence" by Wolfram. It generally says that whenever an automaton looks complex, when it isn't trivial, then it is going to be Turing complete. I will accept this, because it is true in simple examples, and it is difficult to construct something complex which is intermediate in Turing degree of complexity, Friedman and Muchnik needed to work hard, starting from something that is already a universal computer.

So in order to make a computer, all you need is a system with information stored in molecules, with rewrite-rules in the form of allowed catalysis. The peptides produces on the early earth from the atmosphere, together with primordial hydrocarbons from the Earth's formation, can produce polymers at the interface of the primordial oil and water which have these properties, simply by joining peptides into polypeptides. This is the computing soup. I believe that a sufficiently large and sufficiently fast computing soup is necessary and sufficient to explain the origin of life, there is nothing more and nothing less required.

* Self and non-self

A computing soup is seeded random data, but the data doesn't stay random. It gets reworked depending on the local environment to acquire the characteristics of the molecules surrounding it. These characteristics build up progressively, because the system does not reach any sort of statistical steady state, and different regions of the large computation produce a different ecosystem of interacting molecules.

None of these molecules are self-replicating, but they are all self-replicating in a certain sense, in that they weed out and digest molecules which do not conform to the pattern that is compatible with the other molecules. This is a collective sort of replication.

Collective replication was proposed by Stuart Kauffman as an alternative to the

self-replicating molecule idea, back in 1969. The idea was that a collection of independent molecules can each catalyze part of each other, so that together they autocatalyze the whole set. Kauffman argued that such an autocatalytic set is inevitable given a large enough diversity of molecular species.

This is probably true, to a certain extent, but one must keep in mind that this is also true in a computing soup, and without a computing soup. But a simple autocatalytic set suffers in general from the same problem as other replicators--- getting stuck in a rut.

In order for evolution to proceed, it is not enough to be replicating, you need to make sure there is a path for further evolution into ever more complex systems. The simplest replicators have the property that all they do is replicate themselves, and then the only evolution is a quick minimum-finding where they find the quickest and stablest replicator.

An example of such a parasitic replicator is fire, fire metabolizes and reproduces itself, but it is incapable of evolution. Similarly, small self-replicating computer programs with noise are capable of filling up the computer memory with copies of themselves, but they don't evolve past this point.

The systems that are capable of further evolution are those that are not precisely replicaitng, but that are precisely computing.

The recognition of self and non-self by computing automata means that if you divide an automaton in two, and wait, the two halves do not mix together well after a while, because they acquired different characteristics. The result is that if you allow the two halves to touch, they will compete, and the best one at spreading will take oer the computing volume.

This produces Darwinian competition long before any precise replication. The Darwinian competition allows for selection of traits that are favorable to spread throughout the computing soup.

* Emergence of life

It is likely sufficient for life to have a computing soup of molecules, as these will then compete locally to make better and better synthesis systems, and eventually they will make compartments to localize the molecules into cells, long after developing nucleic acids, ribosomes, and all the other ideas we see in modern cells. The stable replicating DNA molecule, in this view, is the last to form. It evolves when there is a need to store RNA in a more permanent fasion.

This idea is proteins and hydrocarbons first, RNA and genetic code second, DNA and cells last. It is hard to test the later stages, but early stages can be tested using cellular automata, which is something I did about a decade ago. It was hard to interpret what was going on in the cellular automata, even when they looked like they were computing, because the patterns are not obvious a-priori, but that was only because I did it

half-heartedly, being more excited at the time about the computational patterns in modern cells.

* Criticism of other ideas

The idea of RNA world assumes RNA can form. RNA has a sugar in it's backbone, and it has different bases, and it's much much too complicated to make RNA abiotically. By contrast, proteins are dead-simple to make, you can't avoid making amino acids from methane, carbon dioxide, and water. So it is obvious chemically that proteins are earlier than RNA.

Further, RNA can't self-replicate. That's really good, because if it could, it would kill the computation like a cancer, but this is what is assumed in RNA world--- some sort of self-replicating RNA.

The ideas of Dyson on cells-first suffer from the problem of no-computation. If you don't start with a computing automaton, you have no computing automaton inside the cells--- they are too small. They are unlikely to have a diverse enough collection of species to make a computation, and even if they did, it's potential for evolution is too small, becaue it is isolated, so it has a limited memory. These ideas are reasonable for the emergence of cells once the computing soup has evolve to a good enough point to package the machinary in isolated compartments.

The ideas of Thomas Gold on the importance of petrolium and deep-vents, archaea first if you like, I think are ok, but they are completely compatible with the view I am pushing here.

3. - How much data is stored in our bodies?

Question: How much data is stored in our bodies?

If you were to upload our bodies onto a computer. Exactly how much storage would we take up?

This cannot be answered fully today--- this is the question of what is the bit-content of the relevant computation that is going on in the body. If you know the bits to simulate the computation efficiently, excluding the bits that are effectively random, you could say.

Whenever you have a biological system, some bits are doing important things, like the bits that tell you what domains are bound to an active protein, and some bits are useless, like the bits that tell you the precise orientation of a subpart of the cytoskeleton, or the bit that describes the orientation of some water molecule. When you want to store the data in a person, you are generally speaking about the biologically relevant data. If you store this data, and destroy the organism, and then restore the organism from new molecules, arranged according to the instructions in this irreducible data, the organism will behave in

a statistically indistringuishable way from the original.

One can say some general things about the size of this data:

1. It isn't infinite

There are some people who claim that the computation is like an analog computer (or as close as one can come given quantum limitations). This means that there are molecules or large objects in the cell that store analog data in their positions.

This is not true at all, and this fallacy is persuasive enough that one must argue against it. When you have a system in a thermal bath, there is always diffusion going on between interactions. If you have a molecule which is storing data in some way relevant to the biology, it must store this data in a way that can be retrieved by the remaining computation effectively. If this data is randomized by diffusion, it is not effective storage, and this bit may be discarded and replaced by a random number generator.

If you have a diffusing protein with interactions with other proteins every time Δt on average, and with diffusion constant D, the protein will randomized into a Gaussian of size $\Delta x = D\sqrt{(\Delta t)}$ before the next interaction. This means that it is pointless and wasteful in terms of storage to specify the position to more accuracy than this size, and the number of possible positions is on a lattice of size Δx. The number of points grows as the log of ΔX, so the number of bits a position can encode is bounded by the log of this, and it only grows logarithmically. This means that no matter how absurdly quickly you try to make the protein interact, the number of bits it can store in the position is never very high. It would be better off adding 5 binding domains rather than trying to localize more precisely.

The result applies to all other continuous storage mechanisms you can dream up--- the position of untethered molecules diffuses, the angles of proteins randomize, the concentrations of atoms are only relevant to the extent that a localized channel or protein can discriminate between different concentrations. In all biological systems, the spatial resolution cutoff for all motions is coarse enough to make the dominant information storage mechanism molecular binding.

2. It isn't that large

The binding of molecules at first glance includes a large number of bits, since each protein is a long sequence of amino acids. But this is also a false bit content. In order to be dynamical data, capable of computing, the data has to change in time. Even if you have a very complicated protein, if it's only action is to bind to a ligand, then it has exactly two states, and carries one bit of information. If it binds to a polymer, then it can have many different bits, but these bits should be associated to the binding sites of the polymer. The protein has two states, the polymer has many different states of protein binding.

To store this protein state efficiently on a computer, I just have to name the proteins with a unique name (this only takes a few bits), and give the state of each one of each type. The proteins which carry 1 bit of information will be fully specified by the number of 1-state

proteins and the number of 0-state proteins. To specify this number only requires log growing information, so it has negligible bit content.

This is absurd, of course. If you include rough position information, you will always need about 10 bits per proteins (to name a billion locations in the cell where it could be) to really specify the state. So there aren't going to be 1 bit proteins which are not tethered to one spot. But the point here is that the dynamical data that is doing the computation is far smaller than the data encoded in the protein amino-acid sequence, because this data is ROM, it isn't RAM, and it can be specified ahead of time, you don't need to simulate to know the types of proteins that are running around

3. It's very small in proteins

There is a simple formalism (see here, I authored this: http://arxiv.org/abs/q-bio.MN/0503028) which allows you to estimate the bit-capacity of binding proteins. The formalism is also useful for describing how proteins bind. It turns out to be related to D. Harel's higraphs, but it extends the formalism nontrivially to include polymerization (Harel was only interested in finite state automata, and didn't consider binding molecules).

The point of this formalism is to easily estimate the bit-capacity of protein networks. The estimate is simply by multiplying the bit-capacity of a typical protein by the total number of proteins. Excluding proteins of known function, metabolic stuff, and so on, you find that there is a range of bit-values in a human cell from as low as 10kB to as high as 1MB, but the higher end is extremely optimistic, and assumes that every different binding state is functionally discriminatable, so you can tell apart every phosphoryllation state of P53 from every other just by looking at the future dynamics. This is clearly false, and I tend to believe the lower estimates.

4. It's rather big in RNA

On the other hand, RNA strands in the cell can do much more. RNA is self-binding in complementary pairs, and to predict the future behavior of a strand, you need to know the sequence. This is because the sequence can find another complementary sequence and bind, and this bound sequence can attach to a protein, and so on, in a closed loop computation.

In order for this information to contribute to the bit-content, one must assume that there is a vast undiscovered network of interacting RNA. I will assume this without any compunctions. This explains many enduring mysteries, which I will not go into.

The memory capacity in an RNA computer in the cell is easy to estimate, it's twice the number of nucleotides. Unlike for proteins, each nucleotide is carrying RAM, not ROM, if the interactions with other RNA is sufficiently complex. This isn't true of mRNA, this isn't true of tRNA, it isn't true of ribosomal RNA in isolation, so it requires more roles for RNA than were ever imagined. This is a prediction which does not surprise biologists anymore.

5. It all hinges on how much RNA computing is going on in the cells

The estimate for RNA is the memory content of DNA, which is on the order of 10^9 bits. The RNA can be an order of magnitude greater, two orders of magnitude greater, but no more, since you will run out of space. This is 10^11 bits per cell.

6. The biggest thing is the brain

In the brain, there is more genetic material than anywhere else. If you just take the weight of the brain and consider that it is all RNA, you get a reasonable estimate of the bit content of a person. It is approximately 10^22 bits per person, or 109 Terabytes, a billion terabytes in RNA, much less in everything else (so the Wikipedia estimate is probably ok for everything else). This is the correct estimate of the memory capacity of a person, since it is riduclous to think that the information in the vast unknown RNA in the brain is decoupled from the neuron activity, considering that the neuron activity is otherwise completely computationally pathetic.

I described this idea in more detail here: [Could we build a super computer out of wires and switches instead of a microchip?.](#)

4. - Could we build a supercomputer out of wires and switches instead of a microchip?

Question: Could we build a supercomputer out of wires and switches instead of a microchip?

I thought of this question; would too much wiring make a computer burn down? Or can you build an actual full-speed supercomputer computer using switches and wires?

In other words, could simply adding wires and switches into a computer already composed of switch and wire, can you achieve a calculation speed which is limited only by the speed of light, or perhaps 10% of this?

This is an important question of parallelizability. Can you write any program with a distributed computer composed of switches and wires, as you say, that is efficient, that gets the same type of computation done as a microprocessor computer, with cache and memory.

The first point is that Turing universality guarantees that nearly any nontrivial system that can store data in nonlinear switches and make other switches flip based on the contents, can be capable of universal computation with an appropriate initial state and appropriately adjusted switch interactions. For example, just by wiring the switches and wires to play Conway's game of life, or Wolfram and Cook's cellular automaton 110, you

get a full computer. This part is easy, and it isn't your question. Your question is about efficiency.

These computers have a bit-capacity which is roughly equal to the number of switches, give or take a factor of 2. This means that to get a supercomputer's worth of data, 10^12 bits (100 Gigs), you need a million by million square of switches. This gives you an idea of the scale. You can do it in a 10,000 cube, again, this would be enormous with ordinary size switches. You didn't specify the size of the switches, but if even if they are a mm across, you need a 10m cube. If they are neurons, you can pack approximately this many neurons in 10cm cube, and this is a human brain.

Your question is therefore one that is relevant in biology, because it is asking whether the neurons in a brain can function as a wire and switch computer, toggling other neurons. Nearly all current models of brain functions postulate that this is the way the brain works. This idea is the neural net model, related to Hopfield nets, which are a variation on the long-range random-coupling Ising model. In this model, the working memory at any one instant in the brain is contained in the pattern of firing and non-firing neurons, and the computation is performed by toggling the firing/non-firing state of other neurons in response.

This idea gives a bit-capacity of the RAM in the head roughly equal to the number of neurons, and a processing rate which is about 1 step per millisecond, distributed, meaning each neuron is computing independently of any other. This is the typical computational density in the brain in normal models of brain function. This working bit capacity is a hard wall, it is only modified in superficial way by neuron potentiation.

Neural nets are too slow

The neural net model, while fine in terms of the computational ideas, is clearly wrong from immediate experience with brains. I must pause here to say that neuroscientists by and large do not agree, and many of them believe that neurons strictly do a net computation still, although this opinion is likely changing, because it is obviously false.

The first observation is that the memory capacity of the brain is extraordinarily limited, in both time and memory. For exmaple, if I glance at a street, and I see a car, it takes about 1/10th of a second for the visual processing to complete, and for me to identify the car. In that process, there are only *100 cycles* of neural activity possible, in other words, the neurons can only fire 100 times. Each neuron holding a bit, even with the most efficient computation you can imagine, there is *absolutely no way* that you will identify the car and recall properties of cars, like "driving, there is a wheel, on the road, right side up" and all the innumerable little dormant things, from a trillion bits in 100 steps of a millisecond.

What you *can* do in this time is produce a unique pattern of firing that serves to uniquely identify the observation of a car, and bin the firing into the appropriate bin.

This then can be used by *something else* to do the rest of the computation in thinking.

This is the problem of too-low computational capacity of neural nets. It cannot be improved by making brains bigger, because it is a problem of depth, not of breadth.

Neural nets forget

The second problem is that neural nets can't remember anything. Even if at one instant, the net sees the car, it must remember the pattern for the car at subsequent steps. This requires crazy "resonant circuits" which store the computational data. The resonant circuits means that the neurons excite other neurons and so on, in a loop, which stays active when you turn off the stimulus.

The loop idea leads to a serious problem--- neural nets with loops, as they are usually made, are unstable either to runaway activity, or to shutting off. If you activate neurons, and these activate others, the stable state is that all the neurons are either turned on or turned off. In order to get over this, you need a global control on neuron activity, which restricts the number which are turned on, and this global control is difficult to imagine.

In order to get around this, artificial neural nets just forbid loops. They make layers, where each neuron tells the next layer what to do. These layers are also observed in visual cortex, and they exist, but they are clearly impossible for storing memories. This only works for a quick run-through-once neural computation from input to output, not for steady-state thinking in closed loop.

The instability problem has not been satisfactorily adressed, although it is theoretically possible to do so. You can make complicated sum-rules for total firing, and try to get the computation to proceed naturally with these sum rules. But here, it is next to impossible to imagine how these resonant circuits recall distant memories, or do anything more than store the last immediate stimulus for a short time.

Neural nets are expensive

Storing and transmitting a bit by making a neuron fire is *ridiculously expensive* on the biological scale. You need to pump ions to keep the neuron at potential, let these ions leak, and expend a huge amount of energy to pump the ions out at each firing. This requires mountains of ATP per bit, a cellular level of energy usage. The brain is already metabolically expensive, but in terms of energy cost per bit, it's thousands of ATP's per bit per millisecond, because to keep the bit in working memory, it must be kept going in a closed loop of firing. There is no permanent storage in this model which does not require horrendous energy expenditure.

The mechanisms of bit storage in the genetic level are very cheap in comparison. DNA stores lots of memory reliably for years with no energy expenditure. RNA stores

memory reliably on the hour time scale, with no energy expenditure, and rewrites and writes are accoplished using only a few ATP per bit, and then only once.

This leads one to expect that the actual memory storage in the brain in intracellular, not cellular, and based on RNA. This idea is not in the recent biology literature. The closest is a paper by John Mattick on RNA in the brain from 2010 or thereabouts.

Neural nets have no DNA programmable instincts

If you have RNA as the active component of memory, you can easily understand instinctive knowledge--- things that are not learned, but pre-programmed. In humans, this pre-programmed stuff includes face-recognition, walking, smiling, visual processing, certain language facilities, and a billion other invisible things that direct internal senses and processing.

If you have to encode these computations on the neural network level, you fail, because the neural net is removed by several layers from DNA. The DNA has to make RNA which makes regulatory RNA and proteins, which work to place the neurons, which then do the expensive neuron level switch and wire computation. It is obvious that the layers of translation required reduce the fidelity of the information, so that if there are only a few billion bits in the DNA, only a few kilobytes of instinct would remain. This is insufficient for any reasonable model of biological instinctive behavior.

This is another place where the current model is clearly wrong.

RNA ticker tape

It is clear that there is hidden computation internal to the neurons. The source of these computations is almost certainly intracellular RNA, which is the main computational workhorse in the cell.

The RNA in a cell is the only entity which is active and carries significant bit density. It can transform by cutting and splicing, and it can double bind to identify complementary strands. These operations are very sensitive to the precise bit content, and allow rich full computation. The RNA analogous to a microprocessor.

In order to make a decent model for the brain, this RNA must be coupled to neuron level electrochemical computation directly. This requires a model in which RNA directly affects what signals come out of neurons.

I will give a model for this behavior, which is just a guess, but a reasonable one. The model is the ticker-tape. You have RNA attached to the neuron at the axon, which is read out base by base. Every time you hit a C, you fire the neuron. The recieving dendrite then writes out RNA constantly, and writes out a T every time it recieves a signal. The RNA is then read out by complementary binding at the ticker tape, and the

RNA computes the rest of the thing intracellularly. If the neuron identifies the signal recieved RNA, it takes another strand of RNA and puts it on the membrane, and reads this one to give the output.

The amount of memory in the brain is then the number of bits in the RNA involved, which is about a gigabyte per cell. There are hundreds of billions of cells in the brain, which translates to hundreds of billions of gigabytes. The efficiency of memory retrieval and modification is a few ATP's per bit, with thousands of ATP's used for long-range neural communication only.

The brain then becomes an internet of independent computers, each neuron itself being a sizable computer itself.

Mattick's experimental compilation

John Mattick has recently pointed out that the biggest component of brain by weight, other than water, is genetic material. This is noncoding RNA which is actively transported up and down from dendrites, and is clearly doing something important. While it is not stated explicitly in his paper, it is clear that he expects the RNA to function along the lines suggested above.

This idea is not accepted in neuroscience. RNA memory was originally proposed in the 1950s, in a worm model, when it was found that RNA was capable of transmitting memory from worm to worm. This idea floated around in science fiction circles in the 1970s and 1980s, since the idea of RNA memory led writers to imagine a pill which stored knowledge in RNA, and then you take the pill and learn something. This is described on the Wikipedia page on RNA memory, and it is unfairly discredited. There is no substitute.

Distributed computing

Outside of biology, the goal of parallel computing in the late 1980s was to produce computing devices which were as parallel as the brain was imagined to be. With this goal in mind, Thinking Machines produced a line of supercomputers in the late 1980s and early 1990s, with a goal of having a great number of very simple processors that worked together to produce the computation.

Engineeringwise, the method was not great, because of the expense of the communication overhead. The better solution was to have powerful processors, and to use communication only when necessary. Modern supercomputers are now made from clusters of commercial stand-alone computers, with communication between processors only accounting for a small fraction of the bit content.

This solution was puzzling, because it seems the brain found another solution. But I am saying that this is not so, that the brain operates in exactly the way the the human

engineered computers were found to be more efficient.

So the basic answer is no--- the distributed solution is not the efficient way to build a computer, you want to have small components with a lot of processing, and do long-distance wire communication only when necessary to link the independent computers together. This is the solution found both by biological evolution and the evolution of human engineering.

A note on these ideas: while I knew that RNA is the major computational component in the cell, because it is the only way to explain the missing information paradox in genetic regulation and evolution, I had no idea to suspect it should be important to brain function. Gaby Maimon suggested that there is a paradox of missing information in the brain which should be solved the same way, and I was skeptical of this for a week or so, but he was right. The ideas about brain RNA were all developed in conversation with Gaby, who is my brother.

5. - Are cosmic muons causing mutations or even influence evolutionary rate?

Question:

As there are experiments studying influence of cosmic rays on organisms, http://www.ncbi.nlm.nih.gov/pubmed/11541768

I ask my self, if there any influence to DNA from atmospheric muons on the Earth surface as well?

The flux is well known, http://arxiv.org/abs/hep-ph/9803488 but, what is the energy needed to change the DNA? Do cosmics actually have higher probability to change DNA despite of cellular reparation mechanisms? If one takes greater timescales into account, could muons have been one of the causes of biological evolution?

Edit: Broader context is to approximate, if fluctuations of high energy primary cosmic rays (considering galactic flux, not the sun activity) are correlated to biological evolution. Apart from muons, other direct feedback from cosmic rays to living organisms is the C-14 production and incorporation.

Specific question is to approximate, how little is the volume in a cell (whether it is DNA it self or RNA mechanisms modifying DNA), how often does it effectively hit by a secondary cosmic, how high is the probability to hit the cell, when it's most susceptible for mutation.

I think, rate of successful genetic mutation can be calculated as $fc=(\tau1+\tau2)\cdot f1\cdot f2$, where $f1$ and $\tau1$ would be the cosmics rate and duration of passage, and $f2$ and $\tau2$ the sensitivity "rate" or probability and time window for a DNA change

This is a biology answer, not a physics answer, but you must not look to direct modifications of DNA by cosmic rays as a source of mutations. This is a ridiculous idea, and it is basically unsupported, and I am certain that it is completely wrong (although among biologists, I might be in the minority).

Generally, the cell has error correcting features, and mutation is not simple cleaving of DNA and rejoining with a point error. This type of mutation is rare. The idea that this is the origin of mutation comes from the 1950s, when it was noticed that stressing organisms with ionizing radiation both breaks DNA and leads to consistent increase in the mutation rate.

But there is absolutely no reason to believe the mechanism is direct stress of DNA, the biology is vastly more complicated. One can get germ-line mutation by exposing flowers to heat, by placing them next to a flame (Darwin noticed this and reports it in the Origin). One can also get mutations by various stress mechanisms in different ways that are not affecting the DNA in a cleaving sort of way.

The notion that mutations are primitive undirected random events belongs to the prehistory of biology, before 2001, and predates the recognition that there is a big brain of RNA in the eukaryotic cell, directing everything that is going on. I am not sure how many biologists believe this, but it was proposed by John Mattick in 2001, and as far as I am concerned, it is certainly correct.

The RNA brain is as important to the cell as your brain is important to you. It directs alternative splicing, gene shutting on/off (by sending out micro-RNA's), transcription and translation, and it also just does thinking, by computing with gigabytes of data.

Given this (although it is possibly still a minority view), the only correct mechanism to attribute germ-line mutations to is to active RNA editing of the DNA in germ-line, both in response to complicated RNA messaging, and in response to heavy gigabyte-sized (or, in an egg, terabyte sized) internal RNA computation, and through the direction of crossing-over. The RNA is sensitive to radiation in both getting cleaved and getting bent, but it is also temperature sensitive, and generally as fungible as you would expect of RAM. The DNA is kept stable and is corrected, and is as immutable (except for direct editing) as you would expect of ROM.

The result of the RNA editing is major modifications in non-coding regions, and these control everything we think of as complicated biology, including embryogenesis, neurogenesis, neuron activity, everything sophisticated. The proteins are as complicated as muscle and bone, doing mechanical and chemical things, but no significant computation. The mutations in genes are mostly meaningless random-drift spot-mutations that are neutral as far as selection is concerned (protien coding genes have not been the major determinant of evolution at least since we diverged from something slightly more complex than a worm).

Given this point of view, the background of ionizing radiation is no more significant for mutation generation than the thermal background of random motion--- it is one more source of error that the computational mechanism of RNA evolves to correct. If you have a high radiation environment, organisms will adapt to it by adding more error correction (there are such organisms around), and if you have a low radiation environment, you will still mutate at the same rate, since the mutations are introduced from inside, they are not external to the system.

The above is what I believe, but I haven't been following the literature to see how mainstream this view has become in the last few years. I am sure it is more mainstream now than in 2003, when it was considered lunatic crazy.

6. – Nanorobots. What stops us from producing them yet?

Question: Nanorobots. What stops us from producing them yet? [closed]

If we can already predicts accuratelly motion on molecular levels, what stops us from developing small robots to, for instance, navigate through our blood vessels looking for cancerous cells and destroying them? What are the predictions for when we will able to do this?

If someone is very interested on this goal and wants to participate on research, what should exactly he study? Programming, chemistry, physics, quantum mechanics? What could help at all? Books/resources welcome.

I don't believe that the materials we use for engineering on the macroscopic scales, gears and wheels and metals and so forth, will be appropriate for nano-devices. Macroscopic devices are usually built out of chemically homogenous materials, which do not afford flexibility in knowing where to cut and splice to design new shapes, or how to assemble the parts into a working machine. For example, when you build a mechanical gear out of cut metal, you need to mechanically make a precision measurement of length from an edge to know where to cut, and this is just difficult to do on a microscopic scale, because things jitter, and you can't see what you're doing.

So in order to engineer devices on a nano-scale, the easiest way is if you can sniff out where your parts are, by chemically labelling the positions. This is easiest in chemically completely inhomogenous materials, where every molecule and submolecule is unique and identifiable by its active chemical properties and chemical enviroment.

Designing such materials from scratch would give any chemist nightmares. Fortunately, nature already did it for us! We have bacteria and single-cell eukariotes, which are already doing all sorts of mechanical and chemical things with interchangable and engineerable parts. It is much more plausible that we will build biological devices on the scale of biological cells, with the ability to perform complex tasks which exhaust

your intended domain.

Biological things are environmentally friendly.

Using biological materials has the added benefit of complete bio-degradability, and it is potentially complete environmentally cost-free. Science fiction authors like to imagine a human-designed super-bug, which accidentally outcompetes the natural ones, it is almost certain we would have to work pretty hard to come up with a bug designed from scratch which would even be able to survive in the world. We can a mouse that glows in the dark. You can just imagine how long that mouse would last in a forest.

It is very difficult to imagine an accidental super-bug, although it is possible that one might be designed on purpose, of course.

If you are very concerned about non-mixing of artificial and natural life, it is theoretically possible to make all artificial living things of opposite chirality molecules. Such molecules will be alien to life on earth, but chemically indistinguishable in a mirror image sense, so they work the same. The drawback, of course, is that you have to get all the chemicals from scratch, you can't eat biological things to extract spare parts. But once you build a mirror-image plant, and a mirror image nitrogen-fixing bacterium, etc, you can set up a mirror image plot of land, and nothing from the regular biological world can contaminate, because a jungle of opposite chirality life will seem to be a wasteland from the point of view of ordinary life.

Obstacles to artificial life

The main obstacle is that we do not have a good enough mechanistic understanding of the processes in natural life. The main biochemical molecules are known, the metabolic processes are worked out, but the interesting control apparatus is not understood in even the most basic way.

The reason is that the actual information in the molecules is not the chemical structure, but in their dynamical conformations, and bindings. These conformations and bindings carry data about the state of the cell, and the number of state variables in a single cell carried by the proteins alone potentially rivals a small 1980s microcomputer. This is much larger information capacity than that in ordinary dynamical systems, and with good reason. Biological systems are themselves information devices, they are computers, and the dynamics is only properly describable in the proper level of abstraction, which matches the level at which the cell itself carries and manipulates the information.

The protein conformations and bindings can be enumerated and mapped out in a given organism, but to artifically design new proteins, you have to have a protein-domain library each part of which has a known function in a given protein context. Protein

domains are parts of proteins you can mix and match in order to provide new function. We do not have a protein domain library.

RNA computing

The second obstacle is that the majority of the computation in eukaryotic cells is certainly done by RNA, and the mechanism of the computation is unknown. This is not believed by the majority of biologists, but there are more who do now than before. The RNA world is not understood even at the most basic level. But I am sure that it is the products of the RNA world in eukaryotic cells that are most significant for controlling the cell.

All of this is essentially biology, not physics, so I will stop here.

7.   - Mathematically-oriented Treatment of General Relativity

Question: Mathematically-oriented Treatment of General Relativity

Can someone suggest a textbook that treats general relativity from a rigorous mathematical perspective? Ideally, such a book would

1. Prove all theorems used.

2. Use modern "mathematical notation" as opposed to "physics notation", especially with respect to linear algebra and differential geometry.

3. Have examples that illustrate both computational and theoretical aspects.

4. Have a range of exercises with varying degrees of difficulty, with answers.

An ideal text would read a lot more like a math book than a physics book and would demand few prerequisites in physics. Bottom line is that I would like a book that provides an axiomatic development of general relativity clearly and with mathematical precision works out the details of the theory.

Addendum (1): I did not intend to start a war over notation. As I said in one of the comments below, I think indicial notation together with the summation convention is very useful. The coordinate-free approach has its uses as well and I see no reason why the two can't peacefully coexist. What I meant by "mathematics notation" vs. "physics notation" is the following: Consider, as an example, one of the leading texts on smooth manifolds, John Lee's Introduction to Smooth Manifolds. I am very accustomed to this notation and it very similar to the notation used by Tu's Introduction to Manifolds, for instance, and other popular texts on differential geometry. On the other hand, take Frankel's Geometry of

Physics. Now, this is a nice book but it is very difficult for me to follow it because 1) Lack of proofs and 2)the notation does not agree with other math texts that I'm accustomed to. Of course, there are commonalities but enough is different that I find it really annoying to try to translate between the two...

Addendum (2): For the benefit of future readers, In addition to suggestions below, I have found another text that also closely-aligns with the criteria I stated above. It is, Spacetime: Foundations of General Relativity and Differential Geometry by Marcus Kriele. The author begins by discussing affine geometry, analysis on manifolds, multilinear algebra and other underpinnings and leads into general relativity at roughly the midpoint of the text. The notation is also fairly consistent with the books on differential geometry I mentioned above.

The Physics work in this field is rigorous enough. Hawking and Ellis is a standard reference, and it is perfectly fine in terms of rigor.

Digression on notation
If you have a tensor contraction of some sort of moderate complexity, for example:

$K\_{rq} = F\_{ij}^{kj} G\_{prs}^i H^{sp}\_{kq}$
and you try to express it in an index-free notation, usually that means that you make some parenthesized expression which makes

$K=G(F,H)$
Or maybe

$K=F(G,H)$
Or something else. It is very easy to prove (rigorously) that there is no parentheses notation which reproduces tensor index contractions, because parentheses are parsed by a stack-language (context free grammar in Chomsky's classification) while indices cannot be parsed this way, because they include general graphs. The parentheses generate parse trees, and you always have exponentially many maximal trees inside any graph, so there is exponential redundancy in the notation.

This means that any attempt at an index free notation which uses parentheses, like mathematicians do, is bound to fail miserably: it will have exponentially many different expressions for the same tensor expression. In the mathematics literature, you often see tensor spaces defined in terms of maps, with many "natural isomorphisms" between different classes of maps. This reflects the awful match between functional notation and index notation.

Diagrammatic Formalisms fix Exponential Growth
Because the parenthesized notation fails for tensors, and index contraction matches objects in pairs, there are many useful diagrammatic formalisms for tensorial objects. Diagrams represent contractions in a way that does not require a name for each index, because the diagram lines match up sockets to plugs with a line, without using a name.

For the Lorentz group and general relativity, Penrose introduced a diagrammatic index

notation which is very useful. For the high spin representations of SU(2), and their Clebsch-Gordon and Wigner 6-j symbols, Penrose type diagrams are absolutely essential. Much of the recent literature on quantum groups and Jones polynomial, for example, is entirely dependent on Penrose notation for SU(2) indices, and sometimes SU(3).

Feynman diagrams are the most famous diagrammatic formalism, and these are also useful because the contraction structure of indices/propagators in a quantum field theory expression leads to exponential growth and non-obvious symmetries. Feynman diagrams took over from Schwinger style algebraic expressions because the algebraic expressions have the same exponential redundancy compared to the diagrams.

Within the field of theoretical biology, the same problem of exponential notation blow-up occurs. Protein interaction diagrams are exponentially redundant in Petri-net notation, or in terms of algebraic expressions. The diagrammatic notations introduced there solve the problem completely, and give a good match between the diagrammatic expression and the protein function in a model.

Within the field of semantics within philosophy (if there is anything left of it), the ideas of Frege also lead to an exponential growth of the same type. Frege considered a sentence as a composition of subject and predicate, and considered the predicate a function from the subject to meaning. The function is defined by attaching the predicate to the subject. So that "John is running" is thought of as the function "Is running"("John").

Then an adverb is a function from predicates to predicates, so "John is running quickly" means ("quickly"("Is running"))("John"), where the quickly acts on "is running" to make a new predicate, and this is applied to "John".

But now, what about adverb modifiers, like "very", as in "John is running very quickly"? You can represent these are functions from adverbs to adverbs, or as functions from predicates to predicates, depending on how you parenthesize:

(("very"("quickly"))("Is running"))("John")

vs.

(("very")(("quickly")("Is running"))("John")

Which of these two parenthetization is correct define two schools of semantic philosophy. There is endless debate on the proper Fregian representation of different parts of speech. The resolution, as always, is to identify the proper diagrammatic form, which removes the exponential ambiguity of parenthesized functional representation. The fact that philosophers have not done this in 100 years of this type of debate on Fregian semantics shows that the field is not healthy.

8. - Software to simulate and visualize atoms?

Question: Software to simulate and visualize atoms?

Not sure if this is a physics or chemistry question. But if the motion of atoms and it's particles can be described by quantum mechanics, then is there a software that simulate full atoms and it's boundings, in a way you can visualize them, and that can be used, for instance, to throw 2 molecules together and watch them reacting?

EDIT: Car Parrinello Works!

The criticism below applies only to the type of molecular dynamics done using molecular potentials, as exemplified by CHARMM. This is the only molecular dynamics I had been exposed to, and it is total crap.

There is a second kind of Molecular dynamics which includes the valence electrons, a core potential, and an clever algorithm to update the valence electron fluid. This is the Car Parrinello method. The Car Parrinello method contains just the right amount of information to do a quantitatively correct simulation, it is the ideal computational dynamics, and I am stunned to learn about it. Thanks to Richard Terrett for pointing it out.

EDIT: Answer to Question

Use CPMD. If your molecule is in solution, use CPMD plus a brownian random force representing the effects of the solute.

Criticism of everything else

There is software which claims to do this, it is called (non CPMD) Molecular Dynamics, and Molecular Dynamics is often used in Chemistry and material science to produce simulations which the proponents claim are quantitatively accurate to predict the detailed microscopic motion of say organic molecules.

The principles behind this software are fundamentally deeply flawed, and, without huge modifications, it simply does not give quantitatively accurate results. Molecular dynamics doesn't quantitatively work for anything at all beyond the homogenous non-conducting solids and small molecules it is calibrated to match.

I consider this a big problem, since a lot of research money goes towards producing this software and calibrating it, and there are a lot of people involved in promoting this stuff, which is a total waste of time and money.

What is Molecular Dynamics?

The idea is to model each atom as a point, and to give a force-law between adjacent atoms,

of each different type. So that there is a C-C force law, which models the preferred bond-angles, and a H-C force law, and and O-C and an O-O force law. Then you add up the forces. Then you correct the model for three-nucleus interactions, like a H-H-O interaction which fixes the water bond angle, and so on, until you get a reasonable match to a large swath of experimental data.

Then you take a molecule where you know the structure, and you simulate the atoms using these force laws derived from simple molecules and monatomic solids. The idea is that you are getting some approximate picture of the dynamics from the structure only. This allows you to get a dynamical picture of chemistry

Molecular Dynamics Is A Complete Fake

The reason Molecular Dynamics fails is that the relevant electronic and electromagnetic forces between atoms and molecules are not at all local, not even to a good approximation, and the fundamental approximation in Molecular dynamics is that you can model them using local forces, local pushes and pulls that don't depend on the global positions of far away atoms. These forces are assumed to be mechanical, like the forces in a tinkertoy model, and this means that forces transmitted like sound waves between atoms, while in actual molecules the relevant forces are often purely electronic and grossly nonlocal.

It is simple to illustrate this using symmetric molecules: if you have a Benzene ring, the electrons propagate around the ring, like a wire conducting a current, and these electron currents give the ring rigidity, like a microscopic stiff metal wire. The mechanical properties of Benzene cannot be approximated by a model which does not take into account the delocalized electrons in the ring. The electronic currents is disrupted by moving the carbons, and they transmit forces at the speed of electron-density-variations (the orbital speed of electrons in the Bohr model, much greater than the sound speed).

This is not an atypical situation. The molecular backbones of DNA, RNA contain delocalized electrons, and the mechanical properties of the molecules are determined by the regions where the electrons delocalize. You just can't predict the stiffness of DNA by knowing the stiffness of two-atom or three-atom sub-parts, without knowing the easy paths for electron delocalization.

This problem is nearly impossible to fix, because adding more nonlocal forces requires sophisticated calculations which are fundamentally different type than 2-body and 3-body potentials. They require a seperate fluid flow model over the atoms at the very least, which will give the delocalized electron flow.

Is it fixable?

I think that the fundamental idea of MD is wrong, so I don't hold out any hope for it being useful at any time, ever. But a different idea might do the same thing correctly.

In order for something like MD to work, it must take into account at least:

- Delocalized electrons in molecules
- long-range electromagnetic interactions

Most importantly, even if you know the exact forces, the idea of simulating molecular motion by accelerations and collisions is simply idiotic. The simulation should be at the very least stochastic, not deterministic, so that the momentum flows through each local region is found by Monte-Carlo, while only the slow global variables (like the overall shape of the molecules) are simulated by stochastic dynamics. The little things are always either frozen out (like electrons) or in thermal equilibrium (like jiggling nuclei or small molecules).

- MD needs to take into account local statistical equilibrium.

CPMD fixes the main issue, of electronic delocalization. This issue is so pressing, because without it, molecular simulations are basically fraudulent.

9. - What objective criteria distinguish between valid science, fringe science and pseudoscience in physics?

Question: What objective criteria distinguish between valid science, fringe science and pseudoscience in physics?

Plenty of research activity in physics have been vigorously opposed by their opponents as pseudoscience or fringe science, while other research are mainstream. It is possible some topic is pseudoscience if the experts claim it is so, but they could potentially be biased. For the nonexperts out there, short of appealing to authority, what objective criteria can be used to distinguish between valid science, fringe science and pseudoscience, i.e. topics which are a waste of time compared to topics worth pursuing.

I would like to take a stab at this, because I think that this is where we are most privileged. This is perhaps the last moment in time when science is the exclusive domain of the specialist. I am optimistic that soon every person will either directly possess, or personally know some people who together possess, enough expertise to access all human knowledge. Public dissemination makes science into what it was always intended to be: an activity of masses, not elites.

In such an environment, bogus ideas don't stand a chance.

Scientific criteria
Pseudoscience and fringe science are pejorative political terms, which are never used by the proponents of the ideas. I don't like the labels--- I prefer simpler labels, like "right" and "wrong". One can ask, what distinguishes a correct idea from an incorrect idea?

For this case it's simple. The only criteria are

1. Internal consistency
2. Theoretical elegance
3. Experiment

Of these, the third is by far the most important--- without it, we would get stuck in ruts all the time, we would run out of new ideas, and research quality would be determined by pure politics, and we know by bitter experience that in politics, Aristotle beats Aristarchus. Further experiment suggest new theories, and new ideas, more often than theory suggests new experiments.

But experiment is not the sole arbiter, because experiment requires theory to interpret, and experiments can be flawed--- consider early studies on ESP, or more recent OPERA results So you need to consider 1 and 2 when evaluating theories and experiments, to see if they make sense in the framework of knowledge already accumulated. But you can't go overboard--- if experiment tells you something, and careful analysis shows that it is correct, then that's that.

It is my opinion that these three principles by themselves suffice to distinguish right ideas from wrong ideas, without any need for political labels like pseudoscience, fringe science, and so on. This way you don't take any chances dismissing an idea politically. If the idea is wrong, it should be trivial to refute it by finding an internal contradiction, a theoretical weakness like a huge amount of unnatural parameter tuning, or a flat out contradiction with experiment.

In today's media climate, it takes less time and effort to fully refute a wrong idea than it takes to come up with it in the first place. So wrong ideas today have a negative multiplication factor, and it should be a short while before they are all extinct. This is a pity, in a certain sense, because it means that future generations will never know an Archimedes Plutonium, or an Alexander Abian, they were quite possibly the last and best of their kind.

Evaluating theories

When evaluating theories, it is important to not restrict yourself to "right" and "wrong", although this classification is still important. You also need to consider "fruitful" vs. "dead-endy", "interesting" vs. "boring", "original" vs. "derivative". These criteria are more human, and more subject to error than the rigorous scientific standard of right and wrong, but they are necessary, because even wrong ideas can often be tweaked into correct ones, and you need to know where to tweak.

Most of the ideas that are traditionally labelled pseudoscience are generally easy to dismiss even in a theoretical evaluation without direct experimental evidence, because they are clearly boring, clearly derivative, and clearly dead-ends. One doesn't have to classify them as pseudoscience to see this--- one can label them boring, derivative, and dead-endy without distinguishing them from their peer-reviewed cousins which are just as boring, just as derivative, and just as dead-endy. I don't see much difference between a numerological analysis of the standard model that appears in a peer reviewed journal and a numerologically motivated analysis of the standard model that somebody publishes on a

personal web-page.

Vindication

Many ideas which were dismissed as pseudo-science turned out to be correct. So it is best to ignore the labels and consider the ideas on their merits. Here is an incontrovertible list:

- Meteorites: yes, rocks do fall from the sky, and this is the reason that the moon and other rocky bodies are pockmarked with craters.
- Non-Mendelian inheritence: It turns out the Mendel's genes are not the main story in heredity, that most of DNA is non-coding and that the non-coding DNA is transcribed to mystery RNA whose function is to regulate proteins. The major evolution of organisms is in non-coding regions, the coding regions mostly undergo selection neutral mutations which act as a clock for species divergence, and the exact mechanism of RNA function is not known (although I have my own ideas)
- Radiation hormesis: this idea might not be right, the jury is still out, but the statistical evidence does not dismiss the idea that small doses of radiation might not promote cancer, but might counterintuitively halt the production of cancers. The linear-no-threshhold model that cancer rate is linear in the radiation dose was held up as scientific consensus, although the experimental evidence is essentially nonexistent, and the model is now more theoretically dubious than ever.
- Continental drift: this idea was dismissed as pseudoscience for far too long, considering the overwhelming statistical evidence in the matching fossil records at corresponding locations along the African and South-American coastlines
- Jumping Genes: This was dismissed for a short while only, but won the Nobel prize.

There are further examples of theories which were dismissed as vague or ill-defined, perhaps not quite pseudoscience:

- S-matrix theory: It is incredible to me that this stuff was dismissed as rubbish, because this theory is more mathematically sophisticated than anything in physics before or since. It is fasionable again, now that the internet allows people to learn the details.
- String theory: This is an outgrowth of S-matrix theory, but again, it was dismissed for a long time, perhaps because it was too good to be true. Well it's too good, and its true.

In addition, I personally find the following vague spiritual ideas impossible to refute, and perhaps they might gain statistical evidence in their favor with time, although they are currently not part of science:

- Synchronicity: this is the idea that people's minds are correlated to a greater extent than one would expect from random chance. Some classical examples are flawed, because they require supernatural agency, but the basic idea is that if you have an idea at a certain point in time, somebody far away is very likely to have the same idea too, and likely at the same time. This can be a turn of phrase, an idea for a scientific theory, a fasion scheme, anything at all. This link between minds suggest that a superstructure is active in human experience, a collective mind of some sort,

whose logical positive manifestation is this nonlocal correlation between the independent thoughts of individuals

- Perceptual Auras: Some people claim that they are able to see an "aura" surrounding other individuals, and their perception is so strong, that they try to make machinery detect the aura. This idea is clearly bogus as physics, there are no energy fields surrounding people, but the perceptual effect is real--- I have seen the auras myself, mostly in settings such as public performance which are designed to maximize the perception of an individual. It remains to be seen if these perceptual auras are inter-objective, whether two people will agree on the perceptual hue and shape of an individual's aura. James Joyce uses the color term "heliotrope" to refer to the color of a perceptual aura, a sort of golden glow which looks like nothing at all, but directs the mind's attention, and reveals internal activity. The halos of ancient and mideval art are also clear representations of an aura around the holy folks, and the "orgone energy" of pseudoscience is a similar thing. Again, it is clearly perceptual, not physical, but it is clearly perceptually real.
- Hypnosis and Sexual Suggestion: This is the popular idea that there are people with the ability to control others through the force of will. Mesmer was a proponent of such ideas, and surrounded it with the trappings of science, by claiming that the mechanism of control was magnetic fields. This is clearly incorrect--- magnetism has no effect on higher cognition. But the phenomenon of natural hypnosis can be seen in human interactions all the time, especially in relations of women to men. I have seen even at an early age, that women are able to suggest actions to men which they feel compelled to perform, even if these actions are silly. A female friend of mine once turned to me at a party and said "watch this", and then commanded the man standing next to her to stand on the table, which he promptly did. She then made him dance around for a while, and then asked him to come back down (I hope he's not reading this). This suggestion mechanism is clearly related to authoritarian power structures in society, and fetishized sexual submission/domination. But the entire phenomenon, including the simple existence of meditative trances, was dismissed as pseudoscience, it maybe still is.
- Musical synesthesia: this is the perception of colors in the source of music. I have, on rare occasions, seen colored streaks from a guitar during a session, and I have seen artists depict colors eminating from the sources of music. I find it fascinating that the brain places the color at the known source of the music, rather than as a general hue, or at random spots, like when you are hit in the eye. This is clearly high-level perception, but can it be replicated in different people? The notion of musical pitch is also interesting--- it has a perceptual height. Is this pitch/height relationship subjective, or is it the same between different individuals?

Why do wrong ideas persist?

I believe that any persistent wrong idea, especially in today's media environment, is only alive because it is serving a social purpose which is not apparent. Once a good substitute correct idea is formulated which can serve the same social purpose, the idea can die. Alternatively, the social purpose might be rendered obsolete, by an evolution of society.

- There is a persistent idiotic idea that global warming is a hoax. This idea is

preposterous, especially since global warming was proposed already in the Silent Spring era, when it was called "the Greenhouse Effect". A quick back of the envelope calculation shows that human CO2 emissions match the rise in CO2 levels, and a quick back-of-the-envelope estimate of heat trapping dynamics also shows that it is plausible that it would be responsible for the warming we see today.

This, along with the published results of detailed models for the distribution of warming, ice-core data on the correlation between CO2 and temperature, and projected predictions that each successive year will be warmer than the last (predictions which are each exceedingly unlikely considering how hot current global climate is, and which are uniformly correct), make it certain that the world is warming due to the activity of humans. The only strike against this theory, as far as the media is concerned, is that it was predicted by left-wing activists in the early 1970s.

So there is a social force at work here. The obvious reason to deny warming is to prevent a carbon tax, or emissions cap, and this is supported by big money. So you have denials, but there is no sound science behind the denials. But global warming denial, like genocide denial or big-bang denial, is on the wane.

- The idea that the Earth is young seems to never die out. The social reason is obvious--- people want to believe the bible. The creation story in the bible is found here (I translated this, so I know it's accurate: http://en.wikisource.org/wiki/Bible_%28Wikisource%29/Genesis). It basically depicts the Earth as a flat disk with a big water-covered sky-dome which leaks rain.

I think in this case it is important to recognize that the wrong ideas are persisting because of the important life-altering effects of religious experience, which has motivated people to allow reforms, even to the point where they will die for their beliefs.

The notion of religion, that it is possible to gain experience of consciousness different from that of the individual person, is to my mind entirely reasonable. But the authoritarian claim that the Bible is a good source of ethics is entirely bankrupt.

- Prayers/Chants/Magical spells: The idea that an incantation will work to produce a physical effect is manifestly absurd.

But people believe it, because these chants and prayers have a perceptual effect on the people who perform the ritual, and those that come in contact with them. This effect is real, but to what extent it is inter-objective remains to be seen. It is a powerful way to direct people's attention, see Allen Ginsberg's magical "Om": http://www.intrepidtrips.com/pranksters/ginsberg/ . These things transcends their emptiness as physics, because of their mental resonance.

Wrong ideas in science
In many fields of science there are flat out absurd ideas which gain traction, mostly because nobody thinks about them enough to realize that they are absurd. It is hard to name many of these, because I am just as blinded as anyone else by my place in time, but I

will try:

These are a counterpoint to false ideas by non-scientists. These are false ideas promoted by scientists. I would call them pseudoscience, but unfortunately, they are mainstream positions today.

- Germ line mutations are random: This is a fairy tail which should not by taken seriously by anyone. The idea here that mutations are molecular readjustments caused by x-ray particles hitting DNA, or by mistakes in copying. That this is absurd was already pointed out by Wolfgang Pauli--- consider how long it would take to random mutate your way from a mouse to a human, and compare with the actual time taken by evolution. It has recently been established that bacteria can control their mutation rate, and that stress can induce mutations, not through molecular damage, but through cell signalling. The proper source of mutations is not known at present, although I have my own ideas.
- Modern Synthesis: This is the idea that Mendelian inheritence coupled with Darwin's natural selection can account for the evolution of higher organisms. The idea is that each gene equilibrates according to a separate population genetics model, and the independent selection pressures lead to organism drift with time. This idea is completely insensitive to the detailed computational structure of genetic and RNA networks, and does not allow for coherent system-level evolution. The mechanism is clearly more complicated, and more authorly, involving an actual computing system internal to germline cells, not a blind random number generator plus a blind selector. A parable: once, there was only one book, a cookbook, which described how to make macaroni and cheese. The book was transcribed by scribes, and they sometimes made errors, a letter dropped here, a duplicated passage there, and so on. After many years, behold, we have the Library of Congress! The argument is demolished by the absurd time scale required for the errors to accumulate to produce "War and Peace", if there even is such a timescale at all, considering that you would have to pass through a mountain of gibberish to get there. The existence of collective selection on a system level is not taken seriously in population genetics, likely because such effects are difficult to model.
- Group selection (other than kin selection) is forbidden: There is the following argument against group selection: if a mutation favors the group, but disfavors the individual, then this mutation will be selected against by natural selection, just by the mathematical process of elmination of individuals carrying the gene. This argument ignores the existence of sexual selection, the traits for which it is selecting are themselves selected for. Sexual selection traits can evolve in a group to select for traits which favor the group, but which are not beneficial to the individual. In my opinion, this is the purpose of sexual reproduction--- to allow group selection through sexual selection. But the consensus view is the bullet--- no group selection. It is belied by altruistic scouting behavior in prarie moles, which is beyond the extreme limit predicted by kin selection (although no-one seems to have suggested that this is a sexually selected trait).
- Fever kills germs: This is the nutty idea that the 3 degree change in body temperature is designed to slow down bacterial or viral replication in the body. This is completely preposterous in the case of bacteria, since they are completely

insensitive to such changes-- they live in the real world. It is a little more reasonable for viruses, but it is still absurd. The obvious reason is that the temperature change is used by the body to regulate internal temperature sensitive mechanism. This mechanism is not known at present, but I have my own ideas.

In physics, I compiled some of the persistent wrong ideas here: Common false beliefs in Physics

Cold Fusion

One of the most deplorable cases where the label "fringe science" and "pseudoscience" is thrown around a lot is in the field of cold fusion. Many dozens of groups, mostly in unpopular research centers, although a few in well financed respectable laboratories, have reported nuclear effects in deuterated Palladium. The political structure of physics dismisses these claims, but in my opinion, there is no sound theoretical argument against them.

I have not seen a convincing refutation of cold fusion, but I have seen papers with very solid evidence that nuclear effects are happening. The papers continue to trickle out, and cold fusion is now acceptable again to the American Chemical Society, although not yet to the American Physical Society, or the National Science Foundation.

It would be nice to see a debate of cold fusion on its merits, without political labels getting in the way. Only for some specific theoretical models, are there clear refutations.

10. – Does the universe follow a general physical and computational principle of evolution?

Question: Does the universe follow a general physical and computational principle of evolution? [closed]

Evolution is a principle in biology, whereby organisms evolve their ability to replicate and multiply in numbers over successive generations. From a computational point of view, the organisms employ a common programming (DNA, RNA, etc.) and runtime environment (cell biology). This evolution has eventually evolved a more powerful ((in a restricted sense) computational environment, the hierarchy of biological brains, with the currently most powerful class instance being the human brain. A collection of human brains is now working on evolving a quantum computer. Perhaps one day the the quantum computers will produce a….

From an anthropic point of view, we are here to think about these things, because our universe creates environments such as our planet earth, which support the evolution of life.

So my question is whether there is a general physical and computational principle at work here, demonstrated by the following incomplete, in parts almost surely incorrect, and highly speculative hierarchical chain of evolution:

1. Within the Multiverse, universes are instantiated following particular physical laws originated and adapted from their embedding universe

2. Each instantiated embedded universe follows computational rules (the particular physical laws of this universe) and create a finite number of embedded universes (the only candidate seems to be for singularities of black holes to correspond with a white-hole, inflationary, embedded universe), each of which evolve in the same general manner

3. Each universe produces a finite number of computational schemes to build organisms that follow certain classes of program patters. These organism evolve according to the theory of biological evolution.

4. Biological organisms evolve biological brains, a higher class computational scheme

5. biological brains build computational devices, which at some point attain the property of replication

6. The replicating computational devices build higher classes of computational devices, following the general principle of evolution…

This is the fecund universe idea, due to Smolin. The original form assumed that a new universe formed every time a black hole appeared (as a sink for the information loss that relativists believed in back then), and then the universe is tuned to maximize the number of black holes formed, constrained by the condition that life is possible.

These types of ideas are anthropic, and they are hard to make testable. Even if the universe is replicating itself and changing, it is not life. Life is not about replication. Fire replicates itself, and tries to maximize cumbustible consumption. Fire isn't life.

Life is when you have a computer in nature. That's not the case for universe-forming processes, because the universe is just not that complicated on the elementary scale. You can see the universe wasn't designed, and evolution and design are synonyms. All design is a process of evolution in your head, and all evolution in a complex system can be equally well called a process of design in a disembodied computational entity formed by all the evolving creatures.

Since evolution is a property of complex systems, and there is no complex system here, just black hole formation from galaxies (this doesn't allow a universal computer), you don't have evolution as I see it. You just have, at best, something replicating, like fire.

The theory is also incorrect because black holes don't lose information, and don't make new universes. The current universe we are in is also not particularly tuned for black hole

formation. Further, the measure which tells you how to maximize is not at all clear: should you maximize the total number of black holes the universe will ever form? Does it matter if they form early or late? What's the weight? These questions are, to my mind, an abuse of language in the sense of Carnap--- they are positivistically meaningless.

11. - Does human evolution have a direction or purpose, and if so, does it have an end state? [closed]

Question: Does human evolution have a direction or purpose, and if so, does it have an end state? [closed]

Evolution seems to be commonly associated with genetic change which leads to the improved reproductive success of a species. This concept has itself evolved with the work of Richard Dawkins and others to possibly include ideas, cultural traits, or even business practices. In this question, I'm making the assumption that either genes, memes, and ideas evolve and produce reproductive success, or they do not.

A key tenet of Darwinian evolution is the random change of genetic material minimally in response to an environment. These changes in genetic make-up can either support reproductive success or reduce reproductive success, but only in interaction with the environment. For example, thick white fur may not be as useful for making babies in the desert.

But ideas are not always random responses to environmental affects. Ideas, cultural traits, or business practices are developed in an attempt to improve reproductive success. They are purposeful.

Is it correct then to claim evolution now, with humans, has a purpose — and therefore a direction?

Edit: In the original phrasing of the question I jumped too quickly from genetic evolution to 'memes' and tie in the word 'purpose'. One idea that generated the question was that the 'purpose' of evolution, for both genetic material and 'ideas-memes' is reproductive success. The goal of evolution is 'more babies.' These can be living babies or ideas.

But as I thought a little futher the 'idea' that memes can evolve and succeed - loosely a la dawkins - also seemed to support the genetic idea. So even though memes do not 'evolve' under the same pressures as genes, they do reproduce in a kind of success pattern of continuing the spread or use of the original idea. I think the question this is leading to: are 'memes' now a 'factor' influencing the genetic evolution of humans like 'hot, cold, wet, dry' were millions of years ago? If so, is evolution still 'random' or directionless - apparently directionless? Then to continue the question, returning to genes, is selecting traits for 'success' another 'influencer' off evolution?

I don't see much of a difference between evolution and design. What people call design is an evolutionary process of ideas in the head, while evolution itself is very complicated, consisting of mutations that are directed by sexual selection, crossing over, and complex RNA networks whose computational capacity is in the gigabytes or terabytes (terabytes for an egg-cell).

The question of whether this system strives toward a limit is the question of whether biological entities have tendencies to evolve toward a certain state. It is possible that the answer is yes, but it is difficult to decide when one only sees one planet, since our presence conditions evolution to end on us.

The way to gain evidence for the proposition that evolution has an end state is to simulate a truly evolving system. This is not done with usual selection models, since the evolution in the models is of a very primitive random-mutate/select sort, which does not include sexual selection (which is very high level), directed mutations using RNA networks, and genetic crossing over. Without this, you are missing the most important components.

So the answer must be that we don't know, although my personal feeling is that this is the case, that evolution is directed to maximize the computational ability of the system, by exploiting all possible resources as efficiently as possible to have as much RAM as possible.

12. - What does Ron Maimon want his legacy to be?

Question: What does Ron Maimon want his legacy to be?

I think the biology stuff is the main thing, the recognition of the computational structure of cellular processes, and the associated origin of life business, which allows you to do a whole bunch of stuff that was remote before. For example, it is possible to make a complete catalog of all possible chemical systems that admit life anywhere in the universe, with no bias coming from Earth biology, and to roughly estimate the relative progress of these systems relative to one another, just from the computational capacity of each system, and it's processing speed. Not that I did this, but it's possible along the same lines as the origin of life business.

But now, it is clearer and clearer that the actual computation in modern cells is 98% RNA, that's where nearly all the RAM is, and the protein computations are just the residual stuff from a previous era, so the main work I did on this, on protein network diagrammatic language, is very dated and of limited usefulness. and the RNA algorithms are completely mysterious, because data is just trickling in today, although they will be worked out soon.

I just want to make as many contributions as I can, and make sure that this is done in an open environment. Many people want this too. I guess it's the same as anyone else, you just want to be a part of the scientific process.

13. – What is theoretical biology?

Question: What is theoretical biology?

Theoretical biology is the study of biology from first principles. The goal for me is to make a full mathematical model of the behavior of a cell which is capable of predicting the behavior in a general situation, at least statistically (meaning, the range of possible behaviors).

The idea is a little difficult to make clear, because there are two immediate things one thinks:

1. Isn't this just physics

Physics will predict the behavior of the atoms in the cell, so it solves the problem of biology in principle. But the biology is NOT just the physics, because the physics is absolutely enormous, and mostly irrelevant. Which water molecule is where is just not an interesting biological question, and at most contributes only a little stochasticity to the biologically relevant events. Removing information is the main idea.

2. Isn't it intractible

The other respose is that the problem is impossible, because there is no meaningful simplification to be made. The cell is in some sense maximally complex, it is Turing complete. A Turing complete system admits no reduction in description beyond a certain point.

The second point is what I got hung up on in the 1990s, as Chris Henley told me in 1995 informally "What can be done in biology that has the same depth as, say, the renormalization group in physics". This question was difficult, from this question I understood that the Turing completeness stood in the way.

Turing completeness meant that you couldn't reduce the cell to a simple iteration system, like the renormalization group reduces the Ising model. It meant that any simplification of the sort that physicists like, the kind that makes the behavior obvious in a simple model with tractable behavior, was going to be futile in biology, because it would be ignoring the Turing complete aspects.

But then, you can see that we have a whole field that is devoted to artificial Turing complete systems, and they make progress, This is computer science. How do they make theory. To a large extent, their theory consists of ways of making a language which more succinctly expresses the structure of algorithms and their properties.

So proceeding in this way, you can make a language for the information transformations in the cell, and so long as you keep making the language higher level in the proper way, you will produce a more and more elegant description of the program of the cell. This is what theoretical biology means to me: it is the program of describing the computational

algorithms the cell is performing with a more and more high level Turing complete language appropriate to the domain. When you do it right, you have a minimal description, which captures the function in a minimal number of bits. At this point, you can start to make predictions about behavior, and produce artificial biological circuits comparable to natural ones.

This was the motivation for my own work in 2001 to 2005, producing a language for proteins. The major lesson from this is that proteins are computationally peripheral outside of bacterial cells, and the major story computation wise is in noncoding purely computing RNA.

14. - How much of our DNA is 'junk' DNA and why?

Question: How much of our DNA is 'junk' DNA and why?

At most 20% or so of the genome is non-functional, and structural. The vast majority 80-90% is functional, although only a few percent is coding, and only about 10% more is obviously directly regulating the coding regions.

The structural elements are telomere ends, some structural code at the meeting point of the X of the chromosome, and repetitive elements that are scattered throughout (although these seem like they might be used to count something). These repetitive elements are over-estimated by Encode's algorithms, but they still correctly conclude that 90% is functional, as John Mattick bravely did a decade earlier.

The function of most of the RNA is purely computational, it is not used to make proteins directly, or to regulate proteins directly. The RNA is transcribed, and complementary binds with other RNA and either splices and resplices it, or else joins and rejoins with new strands, so as to make a little brain in the cell, a brain capable of regulating protein function. The regulation comes in short RNA signals which go in the cytosol and regulate protein expression, and also in other signals that are used to tag locations and so on.

The RNA brain is also making "knowing" modifications in the DNA through the action of reverse transcriptase. The reason I say "knowing" and put it in quotes is because the computation is pretty heavy but it's not anywhere near a brain. It's a few gigabytes in a typical cell, a few more gigabytes in a neuron (which has more RNA), and a few terabytes in an egg cell (which is larger and has more RNA still).

The purpose of this computation is the same as in the brain--- computation evolves for it's own sake, so as to produce all the intelligence needed for a eukaryote to survive. Bacterial cells are degenerate and have either not yet developed these RNA brains, or else, as I suspect, they just evolved them out. It all depends on details of bacterial evolution that are shrouded in the mist of archeology, as they happened long before we have any molecular clues. The philogeny of archaea, eukaryotes and bacteria might be able to sort out which at

some point.

Bacteria have no RNA brain beyond some anti-phage resistance. They share plasmids and conjugate to make an enormous collective which seems to evolve largely by modern synthesis methods. While they were the first organisms whose genetics were understood, largely because it was so simple, for the same reason they do not make a good model for the main story in eukaryotes.

The way I personally came to this idea is just by comparing the information capacity of the known biological processes in proteins and messenger, ribosomal, and transfer RNA to the information in the raw genome and to the computation revealed in the behavior of eukaryotes. Even those who are untrained in genetics or in the methods of science can see that the modern synthesis form of evolution doesn't pass the smell test, the criticisms were frequent and started in the 1950s--- the evolution of computing systems cannot be by small changes in code, it is as ridiculous as claiming that you can evolve a new version of firefox by mutating the ascii codes until it works better. You can, in principle, it just takes forever longer than the age of the universe. There is a persistant group that keeps claiming God did it, by writing the genome by magic, because the complexity is too great, and the networks are too coordinated. Their criticisms are not nonsense (although their conclusion is), it is essentially the same criticism levelled by Leslie Valiant in his Turing Award lecture last year. A computing system cannot evolve without computing mutators that co-evolve with the system, otherwise the thing quickly gets stuck in a rut, which is the maximally optimized accessible code.

The way it happens in real life, outside of viruses and certain bacteria, is for the mutation mechanism to be aware of what it is doing when modifying the genome. Most of the modifications are in the DNA which controls the information structure, the nuclear brain of the cell, not in the proteins, which are structural elements which have largely been fixed since multicellular organisms first arose.

The evidence for this is cumulative, because it is a new idea which makes predictions. I will recapitulate the main points of evidence, although now it makes no difference, because sequencing machines produce as much evidence as you need.

1. The noncoding genome is correlated in size with the tissue complexity and emryogenesis program of the organism. The coding genome is fixed.

Worms and humans basically have the same coding genes, but the program for humans is enormously more complicated, as we need to make a spleen, a liver, a computing brain, a neck, stomach, gall-bladder, etc, etc, and worms only need to make a gut and some nerves. The idea that the program has not increased in size since then except by a factor of 2 or so is laughable, the program required if you try to write it would strain the best modern supercomputer.

2. The noncoding genome is structured and transcribed.

If you look at the RNA in a nucleus, you will find transcription is active all over the genome,

or at least 90% of it. This is the clincher--- transcription is regulated and metabolically expensive.

3. The noncoding genome contains conserved regions

There are highly conserved regions which do not correspond to proteins, to obvious regulators, or to associated anything, they are just conserved.

4. The developmental program relies on non-coding RNA

long noncoding RNA are found at all stages of development at critical positions, for example, there are long non-coding RNA at the head region of a fish (I think) egg.

These points were made in Mattick's 2001 paper. Since the prediction can be purely made just by computational considerations, with no regard to biological data, these are all things that are surprising or unexpected in traditional models of biology, and are obvious in the RNA view.

You must consider the evolution of alternative splicing. It is not imaginable that you can evolve a mechanism of splicing without an RNA computaiton already there, since you need to splice and resplice and when the intron changes, the intron-splicer needs to change in a coordinated manner. The same is true for chromosome break and join evolutions, you need to allow these to breed for a little while with non-break animals.

This makes reinterpretations of the action of all parts of the genome: the retrotransposons are not "parasitic", they are simply data copied from one part of the genome to another in the course of computation. The endoretroviruses are functional and endogenous, and likely used to communicate genetic information from cell to cell, rather than being frozen infections.

All these things are things that biologists got wrong, because they were going by stupid dogma. Too bad for them. Busted.

15. - Is so-called "junk DNA" really useless? What is it doing?

Question: Is so-called "junk DNA" really useless? What is it doing?

There is very little "junk" DNA, this is a purely theoretical dogma derived from a primitive model of biology, which is based on a brainless non-computing evolution.

The experimental evidence that the DNA is junk is pretty close to nil, it is hard to establish lack of function in a biological system, the discovery of function only extends in one direction, towards more functions being discovered. In order to declare that something does not have function, one needs a good model of the cell which works, at which point,

the cell is understood, and whatever is left without function is known to be function free.

Yet there is a theoretical model of mutation which produces, as a predicted consequence, the idea that junk accumulates in the genome. This is the model that DNA is an inert dumb molecule, carrying information which gets duplicated, mutated at random, and then accumulates a bunch of useless cruft which never does anything. This is the point of view of neo-Darwinism.

One must separate this unsupported lunacy from Darwinian evolution, which does not make claims that the genetic material is so stupidly evolved. Darwinian evolution by common ancestry and selection pressure is a scientific fact. That the process of mutagenesis and production of sequence is a stupid one does not follow, because the rewriting system co-evolves with the system.

This idea, that DNA mutation and evolution is unintelligent in the sense of computation, is entirely baseless theoretical speculation. It is ruled out by models of computation, as explained most clearly by Leslie Valiant--- the distance between optimal functional codes in a Hamming sense (the number of mutation steps to get from one good code to another) increases as the code complexity increases, and in order to make evolutionary steps that are effective, the rewriting mechanism needs to co-evolve with the complexity of the code.

While the DNA system does not have access to a computer the size of a human brain, or even an insect brain, the DNA can transcribe gigabyte of distinct genomic RNA, and this RNA has a base-pairing complexity which allows it to actively compute with gigabytes of RAM. This computation is capable of rewriting the genome in a coherent way, and it definitely easily can close the gap. Such a mechanism can be said to be required on purely theoretical grounds--- you need a computer to evolve a program.

Unfortuantely, this is exactly what creationists say. They identify the computer with God. While I don't want to knock them too much, because the arguments the creationists make regarding intelligent design are parallel to those above and to Leslie Valiant's, the computation we are toalking about here is not godlike at all, it's some gigabytes or terabytes in egg, it's not an infinite all knowing mind. Although, in a philosophical sense, it can be thought of as collectively approaching this in the infinite genome limit.

The computation in RNA makes a bunch of predictions. The most significant is that nearly all the genome should be transcribed in a regulated and controlled way. This has now been demonstrated, it was already clear by 2001. This observation is completely at odds with standard views

It further predicts that this RNA must either rewrite itself actively and produce new unmappable sequences during the course of computation (this produces the biggest computer), or it must complementary bind using chaperone proteins into enormous complexes whose complementary binding structure contains the computing information (this produces a much smaller computer). These predictions are not safe, they are really sticking the neck out, and they are probed today with sequencing data of nuclei. If this is correct, you have to find complementary binding networks of RNA, and probably novel RNA

which has been rewritten so much it has no DNA corresponding to it.

What you learn from the computational function of junk DNA?

The most important thing is the origin of complexity in embryogenesis. The program for development is entirely regulated by long noncoding RNAs. This is the focus of Mattick's work, and he has accumulated an essentially infinite amount of confirming data on this. One must remember that the developmental program of human and worms are not comparable in complexity, although the genetic networks are essentially the same, up to a duplication or two here or there. The lack of correspondence between the computation in the protein network and the complexity of the organism was a major clue that the noncoding DNA was heavily regulating with a computation. It's exactly like inferring that an animal has a brain from the behavior--- it is difficult, because for any one behavior you can always postulate a robotic circuit. But for evolvable general behavior you need an evolving computation.

Yet another thing is that you increase the computational capacity of the actual brain a billion-fold. The RNA in the brain is mostly noncoding, mostly strange, and it is associated with genomic modifications such as transposon activity. The natural hypothesis here is that the brain has networked RNA in many cells to achieve a very networked computation. This is larger in capacity than the neural network model by nine orders of magnitude at least, but it requires a channel membrane protein that is capable of writing RNA based on voltage, and another which is capable of reading RNA and producing action potentials. These have not been observed yet, although nucleotide gated channels have, but it is a firm prediction.

Yet another is a complete reversal of the role of transposons and endoretroviruses. The ERVs are functional, and the origin of retroviruses is from ERVs not the other way around. The transposons are back-cribs from the RNA into the DNA, and allow the cellular RNA brain to store data long-term for retrieval. The amount and known function of transposable elements is incompatible with a neutral role.

Another thing is you understand the temperature regulation in mammals. The RNA brain is temperature sensitive, due to relying on RNA-RNA complementary binding, and it is optimal around 40 degrees, give or take. The RNA binding transition temperature. Since the molecule is long, it prefers temperature regulation to near this point.

Basically, it's the holy grail of eukaryote biology. There is a political resistance to this, because it supports a more computing view of evolution, and this is associated with religion. Damn skippy. But it's not a creationist religion, that's for sure, it's just an appreciation of computing networks in nature.

Regarding the bladderwort, there are exactly two possibilities: either the folks sequencing made a mistake, and there are easy mistakes to make to miss whole enormous sections of genome by bad methodology.

1. Their contigs might be actually distant and falsely overlap, due to a genome with a

repetitive structure.

2. The bladderwort might have an extremely variable genome which gets so much transposon activity, that it is highly variable from cell to cell, so that the only reproducible contigs are from coding segments.
3. Just plain incompetence--- it's one result that flies in the face of every other genome project.

The other possibility, if the result is correct as they interpret it, is that the noncoding bladderwort RNA is just reproduced directly from RNA. This means purely maternal inheritance in bladderwort (if it is sexual, possibly not, plants are often asexual). The sequencing of RNA in bladderwort MUST reveal exactly the same exact complement of noncoding RNA as in related plant species. The bladderwort is not magic, it is an anomaly, and when confronted with a paradox, your job is to get to the bottom of it, not jump to political conclusions.

16. – Why does nature need RNA?

Question: Why does nature need RNA?

In organisms today, DNA is the ROM, RNA is the RAM. You can't do anything without RAM.

17. – Is there any scientific proof which follows the scientific methods, of changing of species?

Question: Is there any scientific proof which follows the scientific methods, of changing of species?

A scientific proof is simple: you make predictions that are explicable naturally under the hypotheses of changing species, and are complete coincidences otherwise. Once you have a 1 in a million coincidence, you're convinced. For example, the Higgs boson was declared "discovered" from a bump in the number of scattering events which would be a 1 in a million coincidence without the Higgs boson.

You could make up a bunch of different ideas to explain this bump after the fact. But you don't do that. You make predictions ahead of time, and compare honestly.

For evolution of species from divergence from ancestors, you have a bunch of classical predictions: species forms should make a phylogenetic tree--- they do (check, 1 in infinity coincidence in the absence of change). Species on isolated Islands should fill niches with forms that are more closely related than on a big mainland (check, one in infinity

coincidence), DNA should make the same phylogenetic tree as form (check, with a few modifications to the tree, due to our inability to intuit plant relationships well, one in infinity coincidence). Fossils should occur in dated strata in forms that are capable of branching out to later forms (check, one in infinity), and early forms should just look primitive (they do--- just look at a trilobyte, or early fish--- it's primitive, nothing looks like this, although this is harder to quantify). No modern species should be found in early strata (check, one in infinity coincidence).

This is the scientific method, it's Baysian confidence, and honest comparison of prediction to theory, where you make the prediction ahead of time. The creationist idea does not predict any of these ahead of time, it can only justify it by saying there was a plot by God to plant fossils and radioactive elements to make the Earth look old.

Also, since the creationist story is a bunch of social hooey, there is no other competing scientific hypothesis. A scientific hypothesis is something you would come up with yourself from looking at the data, not something a guy in robes reads to you from a big tome and tells you to believe.

18. - Who is the one person alive today, that is most likely to make the biggest contribution to humanity in his/her lifetime?

Question: Who is the one person alive today, that is most likely to make the biggest contribution to humanity in his/her lifetime?

My money is on John Mattick and Craig Venter. If they sort out RNA and engineering biology, the whole world will completely change. You will grow chairs out of the ground from chair seeds, and horses will eat sugar, have doors and won't poop.

19. - Opinions: Which views of yours seems obvious to you but not to most people?

Question: Opinions: Which views of yours seems obvious to you but not to most people?

All views seem obvious once the arguments are internalized, and the evidence is overwhelming. For ones where this is in disagreement with majority opinion, for me:

1. 9/11 truth: the 9/11 attack was staged by one government official, or a small group. The physical evidence is overwhelming, so it is obvious. Most truthers believe in a huge conspiracy, and most of the public rejects the huge conspiracy and believes the cockamamie government story. I believe in a teeny-tiny one-person (or two

person, or three person) conspiracy, so I think I understand where both positions are obviously coming from, and how the real deal went.

2. Cold fusion is real: the tritium was detected in many independent labs, and is a smoking gun. I explained how it happens on Stackexchange, and once the nuclear fragmentation predictions were in accordance with Iwamura's data (and the other trace transmutation data), it became obvious to me that it was right.
3. Marlovian authorship: the stylometries compiled by Farey and others are completely persuasive, in conjunction with "The Shakespeare Guide to Italy" and just reading Marlowe for myself and verifying that it's the same guy (it's easy, it's obvious once you knw).
4. Abiogenic petroleum/coal: The evidence here is summarized in Gold's book. The most important thing is the elemental contamination which points to mantle origins, the helium, the radioactive elements. The chemistry is what got me to this, there is no way to convert life to oil, but it can be done in mantle as has been reproduced in pressure anvils.
5. String theory is correct: This is a difficult to explain position, because it is something most physicists believe, but the general public can't be persuaded because it requires a crapload of intricate arguments (much more intricate than the other ones) to explain why, and you need to believe first that quantum mechanics is correct and General Relativity is correct, something the public hasn't caught up on.
6. The origin of life is through Von-Neumann/Wolfram/Conway automaton: Chemistry can produce these types of systems, and they spontaneously and naturally evolve, in the sense of Darwin, right from the start.
7. Mattick's RNA networks are the bulk computation in modern eukaryotes: this is obvious just from the information capacity of the thing.

Other things which I am uncertain on (the above have enough evidence to be completely obvious, the ones below are not):

1. Free market capitalism and equality as in socialism are compatible: this is because of the predictions of idealized free market models. It has never been seen in the real world, but from the pattern of Soviet wages (which were competitive) and Yugoslavian entrepreneurship (which was competitive with the west without gross inequality), I am convinced that it is possible to reconcile capitalism and socialism to maintain the best features of both models simultaneously, without onerous government oversight.
2. Quantum mechanics MIGHT BE wrong: as 't Hooft has written, we don't know for sure quantum mechanics is correct on highly entangled states, like quantum computers.
3. Black holes emit nonthermally: the classical solutions for rotating/charged black holes emit matter that falls in after only a little bit of singeing on the Cauchy horizon, so this should work quantum mechanically. But it requires a gluing which hasn't been thought up.
4. The brain is linked RNA computation: this is more speculative, because the mechanism of linking up the Mattick networks of different cells. This looks more certain now, after Mattick's 2010 review of long noncoding RNA in the brain, but it is not obvious yet, unlike point 7 above.

There are more, these are the ones that are on my brain most days, because they are constantly challenged on quora, and are easy to propagandize.

20. - Is physics about to provide a fundamental theory of life that makes Darwin's a special case?

Question: Is physics about to provide a fundamental theory of life that makes Darwin's a special case?

The way life arises is from a natural Turing complete automaton formed from chemical systems. This is most likely non-aromatic abiogenic amino acids from the Earth's early atmosphere plus abiogenic petroleum combining, to make the full set of amino acids, in an environment sufficiently basic (or whatever, maybe salty) to ensure that the proteins polymerize spontaneously.

The result is that you get a computing automaton, and this begins to evolve IMMEDIATELY, without self-replication put in. Self-replication is the kiss of death, because the replicators fill out the system.

I have explained this in more detail in my answer here: How did life begin? There really is no theoretical mystery left. You probably just need to mix petroleum and amino acids in the right proportion in the right environment to start the process off, and see how the earliest stages look.

The evolution is through competition between different large-scale computing systems which digest non-like like an immune system. It does not require fine-tuning, or magic, it is inevitable in the proper chemical environment, and a Turing complete automaton is necessary and sufficient.

21. - Is biology the study of replicating information?

Question: Is biology the study of replicating information?

To a certain extent this is a useful definition, but I would add two modifications. First, the information can't just be replicating, the essential thing is that it needs to be computing--- this means it needs to act on itself to make a general purpose Turing computer, (to be precise, Turing complete with randomness, so there's a non-computable oracle). Second, the notion of replication is not essential, certainly not at the beginning, the replication doesn't have to have a high fidelity copy-generation. The precise information doesn't have to replicate precisely at all, it can replicate in a very weird way, by negative replication, just

by preserving characteristics of itself, and digesting away non-self molecules, like an immune system.

An example is some prebiotic set of low-alanine enzymes that digests all molecules with a greater than 50% content of alanine (random example, it probably doesn't exist, but there are other characteristics that do exist). This digestive system acts to negatively-replicate itself, in that it digests away stuff that is incompatible with its own characteristics, replenishing more molecules of its own general class, including the very enzymes tha maintain this class. This type of low-fidelity replication of vague character is itself ia form of modest replicaiton, and it is all you have when you start out a computing system.

But a computing system with a character set evolves to make more and more characters, as the forming and digesting enzymes narrow themselves down to a smaller compatible set. The parts of this system acquire incompatible characteristics (for example, a mostly lysine set and a mostly leucine set), and the competition between these subsystems is already Darwinian long before a replication machine evolves.

The computation and specialization in such systems is a form of evolution that does not require any special adjustment, or any self-replication to be put in by hand. The biology then evolves replication at a later stage, as the competition in such computing character-building soups is fine for making a Darwinian struggle, even without precise replication.

The essential characteristic is the Turing completeness of the system, the existence of information carrying molecules which are capable of transforming one another in a way that can be programmed to run any algorithm in principle. This is the main characteristic of life, the Turing completeness, and it is present in both the protein and nucleic acid component of modern life. In the proteins, it is clear, in the nucleic acids, less so, because the full set of nucleic acid transformations is not known today.

22. - Is otherwise respected biologist Stuart Kauffman as off-base as I'm sure he must be for saying "Information Theory Does Not Apply To The Evolution Of The Biosphere"?

Question: Is otherwise respected biologist Stuart Kauffman as off-base as I'm sure he must be for saying "Information Theory Does Not Apply To The Evolution Of The Biosphere"?

Kauffman's is off base here, but he makes a good point.

Kauffman claims that you can't a-priori decide which bits in a physical description will end up being biologically relevant, so you need to include the entire physics. His example is the locations of certain electrons in a configuration of the chlorophyll molecule, which end up being important to the process of photosynthesis. You wouldn't know that these electrons

are somehow going to be important a-priori, before the molecule is put to this use. So he claims that you need to know everything about the molecule to extract the biologically relevant information.

This is a reasonable argument, but it's just false. The reason is that the states which are available to the molecule are always defined by the collective of other molecules it can interact with, and in the case of the chlorophyll molecule, the electronic displacement would be completely unimportant if there weren't another molecule available to accept the electrons and use the extra energy to eventually add a phosphate to ADP. The molecular cycle in question is defined by the available molecules for extracting the energy.

In this context, the interaction of chlorophyll with light is important only inasmuch as it triggers a chemical transformation in the next molecule downstream. This is an information change which can be described in a simple diagram, the kinematics and dynamics both, and without the other molecules, it carries no information.

The information capacity of a molecular system can be defined circularly using the concept of a relevant bit. A bit that can be read out by the system is a relevant bit, and it can be read out when it can transform another relevant bit. I know this sounds terribly circular, but it's not hard to see which bits are relevant--- you start with some states you know are relevant, for example, an enzyme that cleaves and destroys some protein you know is doing something in the system. Then a bit is relevant when it can affect the function of this enzyme in some way, through some cascade of events. The collection of bits which can affect one another and which are stored in the molecular configuration define the functional diagram of the molecules.

From the diagram, you can figure out which bits are relevant, from the possible interactions between all the states of the molecules with other molecules. So Chlorophyll's electronic excitation is relevant precisely because there is an adaptor which will take this excitation and do something with it, and change the state of an ADP molecule to ATP, which now carries a bit of information (and some energy), and can modify other bits. The network defines the relevant information.

This definition does not start at the physical description exactly, but abstracts out those bits which are capable of affecting other bits in the environment. The result is a closed system, and the possible interactions can be mapped using pairs of molecules, you know what possibilities are possible without having to probe the entire electronic structure of all the molecules all the time. You only have to understand those changes which can be read out by other molecules available in the system.

This is the main point of the functional approach to biomolecules, you don't need to know everything, and the set of relevant data is defined by the possible interactions of your molecule with the other molecules already present. it is circular, but benignly so, it defines the biological information self-consistently by its interaction with itself, by the computation, by those transformations which are capable of impressing their states on other molecules whose state is relevant.

23. - Does the quantity of DNA scale on a linear basis based upon the size of an organism?

Question: Does the quantity of DNA scale on a linear basis based upon the size of an organism?

Noncoding genomes vary widely from species to species. As this variation is mostly due to polyploidy, not just multiple copies of chromosomes, but any multiple copies of identical data which leaves a lot of closely related duplicate information in the genome which does not increase computational (Kolmogorov) complexity. This is a failure of garbage collection, or a hack to produce large size with some minimal variation, it doesn't increase the complexity very much.

So you get an occasional monster genome, plant genomes can be especially enormous due to polyploidy, but duplication exhausts the method by which the canonical c-value examples happen:

Since there is no need to repeat, here is a paper from John Mattick's group: The relationship between non-protein-coding DNA an... [Bioessays. 2007].

As they explain on page 2, the minimum size of a genome in a group of related species, say a taxonomic order, is obviously and directly related to the complexity of embryogenesis, and increases in size through insects through vertebrates. The charts summarize existing data. There is no C value paradox, there is strong evidence from C-value for computational function of the noncoding RNA. The C-value here is a good measure of the nonredundant genome, or equivalently the smallest genome in a philogenetic clade (the smallest genome is a measure of how nonredundant the genome can get and still work).

Single cell amoeba are relatively complex in thier behavior, and one should not presume simplicity because they are small. The traditional explanations of large genomes as accumulation of junk is unsupported political nonsense.

24. - If this were the best of all possible worlds, how would Ron Maimon hope his work, presented in "Computational Theory of Biological Function I - Kinematics of Molecular Trees", would be used?

Question: If this were the best of all possible worlds, how would Ron Maimon hope his work, presented in "Computational Theory of Biological Function I - Kinematics of Molecular Trees", would be used?

I did this stuff because I knew it was completely original from beginning to end. I figured

"Whoa, dude, you just freakin' wrote the Principia, yo! Good job." while everyone around me thought "How sad. It seems that Ron went crazy." I have follow ups that I didn't put out, because Arxiv is shitty now, it requires you to get an endorsement to post. They can go stuff their endorsements you-know-where.

The other papers deal with loops, replacing the stupid infinite polymers with complexes (as is the case in nearly all protein complexes), and explains how to estimate the protein computational capacity, and how to do verbs, molecular transformations. The main theorem is in the second paper, which is not up, but which I presented at a conference in 2004, which provides an algorithm for unpacking a diagram along different states into larger diagrams, going all the way down to all the species that are present. It is completely efficient given the formalism, and so there is an efficient packing and unpacking procedure for the formalism, which can make as detailed an accounting of the complexes as you like.

It was used internally in Gene Network Sciences inc, where I was a member from the beginning until Jan 2005. We used it internally, as a language for describing how proteins network up inside the cell. The focus was on signal transduction, since this is what is screwed up in most cancers.

There is another diagrammatic formalism, described by Kohn around 2000, but I think it is primitive and arbitrary compared to the thing I described, and I don't think it works well at all, the formalism I give is kind of optimal and more or less unique, as I try to explain in the guts of the paper. The predecessor work is Harel's Higraphs (which I didn't know about in 2001, but I cited him in 2005, once Vipul Periwal showed me the paper. Harel's business from 1988 is the non-recursive version of the formalism), It worked well for the purpose, you could say what you wanted to say, and the result was that the company had quite a bit of a leg up on others as far as making models of various bio-networks. This was useful for getting deals and grants for a little while.

Between 2001-2003, we used it to make a model for protein interactions of about 500 proteins, with all the interactions our biologists could get out of the literature. It was this big chart that explained how the proteins interacted, to the extent this is known. I was mostly using it to get a sense of how biological computation works, but the company wanted to use it further to predict how the cell would respond to various combinations drugs in cancer stages. You could sort of do it straight from the diagram, and this seat-of-the-pants things was really what was most useful for prediction, but there was also a quantitative model, which was really to sound all fancy for the investors.

The formal model involved a program to take the diagram and turn it into a mathematics model, which the company liked to do as differential equations for chemical concentrations, not as individual molecules. I thought this was stupid, but they didn't listen to me, I was just the crazy guy who did the diagrams.

This proto-model from GNS was a snapshot of the cell's regulatory machinery, and it really gave a picture of what was going on inside. In this respect, it was extremely revealing of the internal processes in cells. But the diagram was proprietary, and it has hardly ever seen the light of day, despite my repeated urgings to publish the diagram along with some seat of

the pants predictions. No. They wanted to publish quantitative model results and keep their diagram proprietary, that's business for you.

Some model results were published in the mid 2000s by the company, they were so-so, they sort of work when you fiddle with the parameters, but only because the diagram is qualitatively right, not because the quantitative simulation is right. My guess is that it has rotted from lack of peer review and no keeping up with the literature. It was essentially haphazardly compiled by just reading the entire literature in 2001, the literature has grown since, and they moved on to other things, using Vipul Periwal's "network inference" methodologies, which are pure Baysian models without an attempt to get at a mechanistic understanding.

The diagram and language (arguably) is still the property of Gene Network Scieces inc. at least until 2020 or so. So no one is allowed to use it in the US. In Europe the patent was sensibly denied as a software patent. Serves me right for writing a patent, it was against my better judgement, but I was crazy enough to be loyal to those cocksuckers, and did what they wanted, even if it conflicted with my own certain knowledge of what the right thing to do is. I don't listen to anyone anymore.

Since it is proprietary, in the best of all possible worlds, people will ignore it until 2020, or whenever the patent expires, then subsequently rewrite the biology books to use this formalism to describe the signal transduction networks. People suggested doing this in 2003, when I first put it out, but they didn't do it once the patent got slapped on it, and I don't blame them. So long as it is proprietary, it should be left to wither and die, like all the other proprietary shit.

Anyway, the point of this formalism is to understand the computation in proteins. The main take-home lesson that became clear already in 2002-2003 is that it's not a very big computation, about a byte per protein, and that makes the total computation in the cell order kilobytes. Since this is clearly absurd for embroygenesis, or wiring up a nervous system, you have a missing information problem.

This problem is solved by RNA. The RNA network language is completely different, and depends heavily on the sequence of the RNA molecules. This was the big conclusion, but this conclusion was scooped by John Mattick slightly earlier.

Once you understand that RNA is doing most of the computation, the focus shifts to finding a language appropriate to describing the interaction of nuclear sequences that do not have the sequence derived from the DNA. This project is ongoing, but it explains all the mysteries of biology, including temperature regulation, brain function, embryogenesis, (real) epignetics, everything. Basically, if there's a question in biology, more likely than not, the answer is RNA.

25. - DNA is a chemical mixture. So hypothetically let's say I keep on arranging the molecules and finally create a DNA. My question is: When does that chemical compound become a living organism?

Question: DNA is a chemical mixture. So hypothetically let's say I keep on arranging the molecules and finally create a DNA. My question is: When does that chemical compound become a living organism?

The compounds become an organism when they are embedded in a cell which is capable of reproducing and evolving these things, at least in an environment.

Chemistry by itself is not life by itself, it becomes life when it can evolve to indefinite complexity. This, in my mind, requires for sure a Turing complete chemical system around the genetic molecules, and usually other systems to allow data sharing, so either sexual reproduction or plasmid sharing.

You could build life in the context of other life by surrounding the DNA with a whole collection of proteins, ribosomes, and so on, until you have an artificial cell. But then the DNA will not be the most important thing, the remaining molecules that give it context are.

I should point out that 'reproduce' here is used in an abstract sense of reproducing the algorithm in some way. It doesn't necessarily mean "reproduce by the usual channels". A donkey is sterile, but it can still get cancer and the cancer cells can be immortal in a dish, forever reproducing a version of the program encoded by the donkey's DNA with appropriate evolved modifications, into the future.

26. - Is perception measurable as energy? IF so does it differ from biological evolutionary measurements?

Question: Is perception measurable as energy? IF so does it differ from biological evolutionary measurements?

Perception is measurable as computation. Most energy, as physicists define the term, is just heat, to the degree that you can use the body's internal computation to control metabolism, you can control heat to a certain extent, but that's not usually what people mean by "energy", they mean influence of one computing entity on another.

Computation can be measured by focusing on known inputs and outputs, and identifying the internal data required to reproduce the inputs and outputs. All the spiritual stuff can be understood this way without any supernatural anything.

God is a little more, because it requires considering an infinite limit of computing things,

going to infinite complexity.

27. – Cell Biology: How can a human's immune system fight off a virus and keep the human alive?

Question: Cell Biology: How can a human's immune system fight off a virus and keep the human alive?

There are two mechanisms here, both not fully understood. One is innate immunity, the ability of cells to sense they are infected wth a virus and warn their neighbors. They do this through the interferon system, an inflammation response transmitted from cell to cell, and when cells detect interferon in the environment, they get ready for infection, and they shut down many systems which viruses can hijack to survive, and they also secrete more interferon to warn other cells.

This allows the body to gain time, because each infected cell warns its neighbors, the most likely target for the daughter particles. It is not clear to what extent this type of immunity is adaptive, and to what extent it is "innate", as the name implied. It works through the Rig-I system and related proteins, which (probably) identify double stranded cytosolic RNA or foreign RNA, and then cascade to the nucleus, where they activate certain genes and trigger a refractory state, which lasts for a long time, at least a day or two. The genes activated by Interferon are classified and known today, and their interactions with other cellular networks can be inferred relatively cleanly through classical systems biology method. I am personally curious whether the double stranded viral RNA is used for more than triggering the interferon response, one can make a wild speculation that it might be also used to make a template to detect other RNA of this type. There's no evidence for this.

The other mechanism is acquired immunity. Ultimately in the acquired immunity system the white blood cells learn to identify the virus with antibodies and break it up whenever it is in the blood or lymph, or anywhere in the intracellular spaces of the body. Once the probability of a virus meeting an acquired-immunity antibody-carrying white blood cell is greater than it's probability of meeting a susceptible cell, the virus can no longer replicate productively.

This system doesn't work so well when it is the acquired immunity cells, the T-cells that are the target, as happens in HIV. But most of the time, if you get the flu, you recover and have antibodies that make you immune to this strain in the future.

28. - What does the drop in the cost of DNA sequencing allow that was not possible before?

Question: What does the drop in the cost of DNA sequencing allow that was not possible before?

Everything! First, it can allow a complete elucidation of the RNA networks in cells, by massive sequencing of RNA in all cell types. You can look at the Encode project to see what is possible--- you can get transcription factor binding sites with CHIP-seq, you can get protein expression data using mRNA-seq, you can get noncoding RNA by indiscriminate sequencing with nuclear separation.

This can produce a complete mechanistic understanding of biology, as RNA interactions are not well elucidated using traditional biochemical techniques--- they are just too computationally rich to be determined in any way other than massive sequencing.

29. - If you had unlimited wealth, what scientific experiments and programs would you fund?

Question: If you had unlimited wealth, what scientific experiments and programs would you fund?

The three areas where there is immediate technology requiring no breakthrough:

1. Peaceful nuclear explosions (here money is not the main issue):

PACER: cheap unlimited fusion energy (H-bombs in salt).
ORION: dirt cheap interplanetary travel. (H-bomb rocket)

PACER is a self-cleaning breeder, but ORION pollutes. In order to make ORION fallout neutral, Perhaps have each launch carry an equivalent load of radioactive elements up and out, and dump it in interplanetary space. This requires a research project to extract all sorts of hot isotopes from the environment cheaply, but, hey, money is no object.

2. Cold fusion--- I would just ask some Pd/d cold fusion folks what experiments to fund, but I would definitely run a series of experiments on Pd/d in alpha radiation, and specifically tuned x-rays. Cold fusion doesn't replace ORION because it is useless for rockets, and it doesn't replace PACER either, because so far, it is expensive and limited because it consumes Palladium.

3. Anti-chiral biology: I would develop all the standard biochemistry tools on molecules of opposite chirality, so as to have a safe arena for artificial life: it

wouldn't be able to interbreed with our life, or eat our food, and any outbreak will be contained. But this requires simple tools to make anti-chiral DNA/RNA, proteins, and so on, until you have an anti-chiral cow, .

I would also personally fund research in completely technologically useless areas, combinatorics and logic, string theory and astronomy, all the fields where technological payoff is remote.

30. – Why are an abundance of physicists moving to theoretical biology?

Question: Why are an abundance of physicists moving to theoretical biology?

The funding is good, the experiments are excellent, and the theory has so far been relatively primitive. There are also good experiments which need quantitative people, and the culture is becoming super-honest as good work drives out bad. Also, a paper in biology will be read, understood, and cited, while good physics papers are often ignored for a long time.

One thing that is not a factor is competition with other physicists. Great physicists are always an inspiration, they are not really competition--- they always expand the field. When Einstein came around, it's not like the number of great physicists suddenly diminished because Einstein was taking away all the good stuff, the number exploded because the field expanded. In the same way, Witten's work has led to the acceptance of string theory and dozens of fruitful extensions, and the only effect is to draw thousands of good people into physics.

On the other hand, bad science does drive people out. in 2000, there was flat out fraudulent work in physics that was taking all the attention of the phenomenological field, and which was extremely demoralizing. I can't describe how frustrating it is to have to explain why large extra dimensions are nonsense again, and again, and being ignored or dismissed, because nobody can imagine that all these clever people are so stupid to miss renormalization constraints and neutrino masses, and gloss over the wrong fixes in the literature. "It can't be that simple" is the stupid thing one heard most often. When your field has gone crazy, it's time to leave.

Large extra dimensions was incompetent to the point of fraudulent, and the refocusing of physics around it made it that suddenly everything I was interested in was out, in favor of clearly wrong nonsense. All the really good young people other than Gubser (who was lucky enough to make a major contribution in 1997, right before the shift) were driven out of physics or out of the country, Simeon Hellerman went to Japan, Zureb Kakushadze went to Wall St., the ones that were left were doing nonsense, and ignoring the holographic revolution.

The draw in 1990s physics was holography, I wanted to understand the holographic principle, to figure out black holes. Around 2000, it was clear that this problem, that I thought would take 40-50 years in 1995, was actually largely solved, and this meant that the major mystery that drove me on was finished, and Susskind was the major player here, but also Maldacena, Polyakov, Witten, all the usual suspects. That's a great thing, but it took the wind out my sails personally.

But at the same time, in biology, there is a simple insight that is a key to the field (at least I think so). The principle is that biology is computation in nature. It emerges when a system becomes Turing complete, it starts Darwinian evolving immediately without any self-replication put in, and the computation means that the theory is like in computer science, it involves methods of extracting the algorithm out of the physics.

I started doing that, to and around 2001, I started doing a protein language. By 2003 I was done, and the startup company I was at mapped a chunk of the cell's known protein interactions. Extrapolating to 20,000 proteins from the 500 we had, it was a little paradoxical, because the total computation was at most 100 kilobytes of data, more like 1-50 KB, and there was nothing particularly complex there. I couldn't imagine putting together a human with a Apple II. I couldn't do it with a supercomputer.

But at that time, RNA was "molecule of the year" in Science or Nature, and this cover made it clear where the hidden computation was--- it's in the RNA. This is a huge deal, suddenly as a physicist, you understand where most of the computation in the genome is, resolve the missing information puzzle, figure out gene regulation (it wasn't so well understood then as today), and temperature regulation, all theoretically! This is the type of prediction that one can not make in particle physics today, it's like not knowing about the neutron, and then you know about it.

It's extremely exciting in biology right now, because the knowledge is expanding exponentially. Once you know how to do theory, the methods become sophisticated. Further, the science is teaching more about nature than any other field.

It is also clearly going to be the major technology of the 21st century, you should have a sugar outlet next to the electrical outlet, and plug in artificial biological technology made out of artificial cells. To plan these requires a complete method of describing biological cells, a precise model of all the processes, so that you can make artificial ones, and it produces a type of precise control on single-molecule chemistry that makes chemists drool.

It's a good field, so I think it's a good call.

31. - Genomics: Is the ENCODE project legitimate or wrong?

Question: Genomics: Is the ENCODE project legitimate or wrong?

The ENCODE project came to perfectly valid conclusions that were obvious to some people for a decade before the project was started. Their results in fact do lend strong support to the thesis that there is very little junk DNA, that 80% of the DNA at least is transcribed in cell nuclei, even though only a very small fraction of this has a role in protein synthesis or known regulation.

Despite the claims of another answer here, there is absolutely no reason to expect junk DNA to get expressed at all. This is not noise, nobody who didn't have an idea of what the junk DNA was doing predicted that junk DNA would be carefully expressed or even expressed at all, aside from those who knew that it was functional.

Genetic expression is regulated by multiple elements in each gene, it requires proteins to bind, another protein to take the sequence into RNA, and it is carefully controlled with chaperone molecules which modify and tag the transcript in the cases where we know the function. The process is not a haphazard accidental binding of a single transcription factor to a random domain leading to some ridiculous accidental RNA sequence which gets degraded, it requires activating complexes to form, and it is tightly regulated.

The expression and regulation-factor data revealed by the ENCODE project is extremely strong evidence for the full real relevant function of 80% of the genome. One must remember that these patterns of expression of noncoding genome are stable and reproducible from cell type to cell type, they are controlled programmatically during embryogenesis, as established by Mattick, and their function is demonstrated simply by the fact that the cell cares to transcribe them at all, as this does not happen when you insert random gibberish in genomes. Eukaryotes are not bacteria.

But the noncoding RNAs are not just transcribed, the are transcribed in regular levels which are stable in embryogenesis from cell type to cell type, and change in time in ways that reflect the ongoing developmental program. This type of regular transcription is a smoking gun of function, it is just that the function is not of the standard type.

So what is the function of these RNAs? It is not established in the scientific literature, but I would like to give a hypothesis that I believe with a confidence of certainty. Knowing this, none of the ENCODE conclusions were surprising at all.

The RNAs are transcribed to take part in pure RNA-RNA computation, mediated by chaperone proteins that complementary bind and modify the RNA to make a closed loop Turing complete computation, for the sole purpose of thinking about what to do (although with a gigabyte computation, so the thinking is limited). Mattick proposes that the RNA network is complementary binding in the nucleus, but I am confident that the RNA is actively rewritten in this process, the sequences munching away at the data to produce new RNA byte data. The reason to claim rewriting rather than just complementary binding is just that the computational capacity goes up by a factor of at least 10-30 (depending on the

details) if you allow resplicing and rewriting, as opposed to fixed complementary binding, because the binding is by domains which are fixed and unchanging of length 10-30 bases by specificity, while rewriting allows the full RNA sequence space to be RAM. The principle that a biological system maximizes it's information capacity is pretty accurate. But it might by what Mattick is thinking too, then it is much harder to test.

I must point out that in this theoretical picture, the transposable elements and endoretroviruses of the genome are not at all junky random insertions, but deliberate insertions of a tagged sequence into the DNA. The tags can be modified by RNA computation, but they contain conserved parts for the insertion to be properly traced to its network. An alu transposon is just a tag to identify the transcript so it can be identified upon production, and the insertion is a long coherent mutation of sorts, it has modifed the computation programmed by the DNA in a permanent fasion. There is no evidence that these elements are parasitic other than they appear in many places in the genome, they replicate, and "Crick says so". The C-number paradox is not a paradox at all in this view either, as it is first not true, most genomes do correlate with naive complexity, aside from polyploidy, and second, it is related to transposable elements, so it is just a manifestation of an organism evolved to store more or less of the computational data in the genome.

The firm test of this is simply to find completely unmappable squences or partly mappable sequences that have been heavily edited by the computation. If the computation stores data by resplicing and rewriting, you cannot avoid finding a ton of completely inexplicable sequences in non-polyA selected nuclear RNA-seq.

We already know biological systems are capable of evolving computers of high sophistication, our brain is an example. The claim here is that the cell has evolved such a computer much earlier, although necessarily of a much lower complexity.

So as not to repeat, I will link another answer of mine here: Ron Maimon's answer to Is so-called "junk DNA" really useless? What is it doing?

32. – What is your greatest contribution to the field of Science?

Question: What is your greatest contribution to the field of Science?

I think it's the computational stuff in biology--- it resolves the origin of life issues, at the very least in outline, and makes correct predictions for RNA function in some detail, some of these predictions remain predictions, because they haven't been conclusively discovered. Others RNA things aren't predictions anymore, they were discovered. This stuff dominated my thinking since 2001.

But I also have this theory of cold fusion, which I believe is correct, so maybe that's more important in the near term. There is less competition here, I know for sure this stuff is

original. Everything else in the theory end of this field is made up nonsense.

I discovered some other minor things too, but mostly simultaneously with others, or sometimes slightly ahead, sometimes slightly behind. That can't be too great, because someone else was thinking the same thing.

33. - How can atheists imagine that the staggeringly complex human, containing an amazing brain, was made without a shred of assistance from anything with intelligence?

Question: How can atheists imagine that the staggeringly complex human, containing an amazing brain, was made without a shred of assistance from anything with intelligence?

Yes, this is a big clue that there is an intelligence at work inside the genomes of organisms. This intelligence is not external to the system, obviously, it must be composed of the information carrying molecules in the cell, in this case, it is obviously RNA.

But this intelligent design means that you can make a scientific prediction, namely that the RNA in egg cells and complex eukaryotes is involved in networks of splicing and rewriting that actively compute with the tens of gigabytes of RAM available in the genetic sequence. It predicts that sequencing machines will find completely unmappable RNA in the cell, which does not correspond to anything in the genome. It predicts that mutations will consist of reverse-transcribed sections of this RNA, and work based on a plan, as consistent not with an infinite intelligence, but with a gigabyte-scale computation.

This is being discovered as we speak. This is the lesson of intelligent design in biology, not supernatural nonsense from the 10th century BC. But it's still intelligent design in a sense, just not an infinite intelligence, a finite relatively large computer compared to the modern synthesis.

34. - Is there any evidence that evolution within a species (microevolution) stops short before it becomes evolution to a new species (macroevolution)?

Question: Is there any evidence that evolution within a species (microevolution) stops short before it becomes evolution to a new species (macroevolution)?

Yes, there is overwhelming evidence for this, from selective breeding experiments conducted by breeders over many generations. By selective breeding, you can produce all

sorts of crazy different dogs from a common ancestor, but when you try to produce a dog the size of a bear, you hit limitations that stop you long before you even get close. The dogs become sickly, because their organs don't coordinate well together, and the body plan reaches the limits, you get hip-displasia, the rate of producing viable offspring gets smaller from inbreeding, various other things. At some point, you can no longer breed animals, and you have reached the edge of the variations you can produce from this genetic stock.

This is universal to breeding of animals over domestication scales, you reach limits which are constrained by the pre-existing range of variations in the genetic material you are rearranging.

This evidence, however, should not be used to support creationist claims, simply because creationism is stupid. The reason it happens is because you are introducing variations and selection pressure which are two orders of magnitude bigger than anything that happens in the wild, and the mutations can't "keep up" with the selection, so as to produce a continuously evolvable path. It's sort of like taking over a capitalist economy with a Soviet system, first there is rapid progress, because you are optimizing pre-existing things, but then you realize you stopped coming up with new things, and you are stuck There is good evidence that if you do the selection slower, with normal selection pressure, you do produce a dog the size of a bear--- because this is what a bear is.

There are two time scales in selective breeding, the scales at which selection operates to select variations of pre-existing genetic stuff, and the scales at which you produce new viable genetic stuff which can be selected. The second scale is the bottleneck in artificial breeding, as it requires a whole bunch of mutations in the genome to first optimize all the genetic networks for the gradually changing body plan, and so on, over many thousands of generations, then adapting the new plan again.

Still, this is the strongest claim that anti-evolutionists have, and it must not be dismissed, because it actually is true. It doesn't demonstrate creationism, because creationism is wrong, not because it is a false observation.

35. - Is the complete connectome enough to model the brain in silico? If not, what else is needed?

Question: Is the complete connectome enough to model the brain in silico? If not, what else is needed?

You should take the following with a grain of salt, because it is not (yet) supported by experimental evidence, but I state it, because it is the only possibility I can see given the theoretical evidence.

Likely what is needed is the complete RNA sequences in all the cells of the brain, most importantly the synaptic RNA, because this is where all the actual computation is

happening. Not the RNA for making proteins, not the RNA for regulating them, but RNA that is used purely for thinking.

This RNA must be linked up to the electrochemical network to produce different pulses for different sequences, and it must read out the electrochemical pulses and convert them to sequence. The result is that each cell is an RNA computer with gigabytes of RAM, linked by a 3000 baud modem to a few thousand other cells, which read out the sequence of inputs and link up the computations as such.

This model is new, it is an original idea, you won't read about it anywhere else.

In order to make this work, you need to have a protein RNA complex which is sensitive to action potentials and transcribes neuron signals directly into sequence. It also requires an intracellular RNA-RNA computation, but this required for other reasons.

The resulting mess means that the connections are only sufficient for simulating the communication overhead of the computation, it is missing the bulk of the computation. The amount of bulk computation is order 1 gigabyte per cell, and there are 300 billion cells, so it is staggering, far, far beyond any current machine.

The reasons to predict this (it is a prediction, this hypothesis is not supported by direct evidence) is that brains are able to store memories, initiate action potential sequences that are coordinated, and their processing speeds are not consistent with the overhead processing speed of the communication between cells. The cellular level models limit the brain's active memory to a number of bits equal to the number of cells, giving C. elegans a memory capacity of 300 bits, which is ludicrous. For Drosophila, it's 100,000 bits, still ludicrous. There is no way to explain why it is ludicrous without getting an intuition for what a 300 bit computer can do, so I can't go further, except urge the reader to experiment with a 300 bit machine until full understanding of the range of behavior comes (it doesn't take long), to see that it doesn't do anything, it's less than one cell nucleus.

36. – What proof did Darwin cite that was so revolutionary?

Question: What proof did Darwin cite that was so revolutionary?

The evidence Darwin cited was of two types:

1. Comparison between artificial selection and natural changes in the fossil record
2. A reinterpretation of the phylogenetic tree as a historical product of divergence.

These two pieces of evidence were sufficient to be certain that some form of selection pressure plus variation were sufficient to explain the history of life.

Number 2 is important, because the phylogenetic relationship between life forms had long

been established--- life forms lie on a tree. What this means is that there are frogs, and within frogs there are different types of frogs, and there are bats, and within bats there are different types of bats, but there are no creatures with the head of a bat and a body of a frog, the organisms are separated into classes, and these classes split into subclasses, with no overlapping structure.

The interpretation of this is that the species diverged from a common ancestor. This leads to a tree of life.

The problem is that people knew about this for a long time, and came up with bogus philosophical explanations for this. They said "obviously things are organized in classes which don't overlap--- that's how things organize! That's just plain common sense." The stupid people who said this included Aristotle, so it was a major problem. It required imagination to see this as something that requires explanation, rather than an obvious feature of the world.

Number 1 was Darwin's first major contribution. It's in Chapter 1 of the Origin of Species, and it establishes that you can change animals by a large amount due to selection pressure only, by breeding. This established not only that it was possible to make changes this way, but exactly how long it takes to make changes of a given magnitude under a given amount of selection pressure.

By comparing the changes in morphology in the fossil record to the rates of artificial selection, Darwin could estimate how much time would be needed to evolve a fish to a human. This allowed a quantitative estimation of rate, and this quantitative estimate was good enough for the biologists to overrule the physicists, who at the time believed the Earth was at most a hundred million years old.

The remaining evidence consists of the speciation observed in small islands, where you could see that a few closely related species had filled niches which were reserved for very different kinds of species on the mainland. This was obviously because only one type of finch migrated to the island, and speciated, not because of a separate creation event for different finches.

This type of evidence was important for persuading people who believe in supernatural creation, because it was obvious that a supernatural creator wouldn't make different finches fill niches that on the mainland were filled by different types of birds. For people who didn't care about supernatural nonsense, this is not so important--- obviously these species diverged, but the question is mechanism.

The 20th century saw a retrenchment of Darwin's ideas by atheists, who wished to use the ideas for monstrous social rearrangements. This use of Darwin is not so optimal. Sure, God didn't create hand-create life, it evolved, but the mechanism is very complex, and teleological in certain ways, it is not a simple competition of selfish genes. These ideas are stupid and were ruled out before they were proposed.

37. - What scientific evidence exists for abiogenesis?

Question: What scientific evidence exists for abiogenesis?

I think the best evidence for this is the existence of complex computing autamata, spontaneously generated with no effort, from any combination of parts capable of storing memory and producing transformations on this memory. This was discovered by Nils Barricelli, John Conway, and really most strikingly driven home by Stephen Wolfram, following Barricelli's approach.

The existence of natural computers from simple building blocks means that there are systems where Darwinian evolution begins to happen spontaneously, long before any precise self-replicating entity is even close to being produced. The evolution can proceed to produce life, and I would define the origin of life as the moment of production of the computing automaton, not the self-replicating stage, which is rather late in the game.

I explained this in more detail in the answer here: How did life begin on Earth? .

38. - Is the "new" (Page on Arxiv) evidence for a cosmic origin of life on Earth convincing?

Question: Is the "new" (Page on Arxiv) evidence for a cosmic origin of life on Earth convincing?

No, it is not convincing, because the genomes don't follow a straight line in complexity, nor are the last points correct--- the genome complexity is not meaured by "functional non-redundant genome", it is probably best measured by compressing the genome using bzip2.

The dating of prokaryote divergence is busted, because we have no idea what the genomes 3bn years ago look like. This is comparing modern genome complexity. Further, there is no good model for how a self-replicating molecule can evolve to do anything other than become the most efficient self-replicator, let alone get longer, so the idea that genome complexity is linear in time is unsupported and ridiculous.

The origin of life on Earth is described in an answer of mine here: Ron Maimon's answer to How did life on Earth begin? . It has nothing to do with self replicating molecules, it is a network of proteins in conjunction with petroleum energy resource.

39. - Creationism: Did life on Earth happen as a series of random events?

Question: Creationism: Did life on Earth happen as a series of random events?

It is impossible for something so sophisticated to form "randomly", but it was self-directed, in the sense that the material events did not require supernatural intervention. We know that, because there are no supernatural events.

The origin of life is explained in my answer to this question: How did life begin on Earth?

The events in a Turing complete soup can in no way be said to be purely random, they stop being random after the first events. As the soup becomes more sophisticated through time evolution (which is Darwinian evolution), it becomes more intelligent in the precise computational sense. The result is teleological, the structures evolve toward greater complexity, and they stay teleological and directed as long as the computing soup stays computing.

40. - What is the statistical likelihood that modern man (Homo sapiens) would evolve from a single celled life form to what we are today over the last 500M years (from the Cambrian Period)?

Question: What is the statistical likelihood that modern man (Homo sapiens) would evolve from a single celled life form to what we are today over the last 500M years (from the Cambrian Period)?

It would take forever, it wouldn't happen, because the evolution people talk about using DNA, point mutations and so on accumulating favorable mutations, cannot proceed past a certain point--- it gets to the limit of evolvability very quickly. You can see this in simple evolving replicators, like viruses.

The reason evolution happens anyway is because it is not a simple process of point random mutation. There are networks of computing RNA in animals which computationally munch the DNA, and insert new ideas in, to try, in each successive generation. These noncoding RNA networks are also required for embryogenesis, and are observed in oocytes (John Mattick postulated these, and put together the evidence for their occurence).

Once you have a good mutation mechanism, the RNA networks providing the mechanism, the too-long-a-time paradox of modern-synthesis evolution evaporates. Then you can form a human from Cambrian animals in the alotted time frame, because the mutations are sensible and authorly.

The paradox has been noted by Wolfgang Pauli already in the 1950s, he opposed the modern synthesis for this reason. It is not a flaw in Darwin, but in 20th century folks who made a stupid unintelligent model of evolution, with no basis in evidence, and no plausibility, as a way of denying intelligence in nature, in effect, to deny God.

41. - Could viruses have been created by extraterrestrials?

Question: Could viruses have been created by extraterrestrials?

The origin of viruses is actually somewhat mysterious, because they are extremely limited in their evolvability. Viruses can only speciate through cross-species infection and wandering around a little bit in genome-space, and they are not evolvable enough to cross types, so an RNA virus can never become a dsRNA virus, or a ssDNA virus, or a retrovirus, or anything other than a slightly different RNA virus. This is in stark contrast to eukaryotes, where there is a smooth evolvability path linking an oak tree to a baboon (through a single celled common ancestor).

This is the limit of evolution of self-replicating entities with errors. It is very different from entities which are truly alive.

Since viruses cannot really evolve, if new ones can't be formed, you would think they should all be extinct. As soon as one type is gone for good, nothing can replace it. One can make two hypotheses here:

* There are on their way to going extinct, but it's taking forever, so the viruses are remnants of early viruses which were around at the origin of life.
* There are new viruses produced from scratch from something else.

I prefer hypothesis b, because I can't imagine that dinosaurs got influenza, or hepatitis. There must be a source.

The source then is in self-packaging genetic material which can leave and enter cell bodies, which has nothing to do with virus infections. Such entities do exist (surprisingly), the endo-retroviruses are an example.

Endoretroviruses form a significant fraction of the human genome, and they produce a coat, a reverse transcriptase, and they can even become virulent under certain conditions, such as cancer. There is a Koala endoretrovirus event underway right now, as Koalas are all getting a retroviral insertion in their genome.

The retroviral properties make it natural to assume that they are serving an important role in the body, perhaps to communicate RNA from one cell to another, so as to link the different computations in different nuclei, and perhaps incorporate into the DNA. They might also be used to transfer RNA between different individuals for all we know. This is

speculative, but it is a very easily tested hypothesis--- look for endoretroviral particles in the bloodstream of a healthy animal, and see if they insert themselves into genomes of distant cells.

If the retroviruses are used for intercellular communication, this case, one can postulate that new retroviruses form when endo-retroviruses are accidently virulent and replicating. This is not hard to imagine, and it allows new types of viruses to emerge, with no relation to previous viruses. This claim, as far as I know, is original, it is not present in the virology literature I have read. It might be false, but at least it's original (no, actually, it's not--- I just didn't know the relevant literature--- see Adriana Heguy's answer on this page)

The source of RNA viruses is more difficult to see. It is possible that these evolve in special places, perhaps in bird genomes. If the viruses come from the complex animals, each virus, one must find the originating animal, and it must have homologous packaging proteins.

I should point out that if this is true, AIDS did not have to result from somebody screwing a chimp (or eating a chimp). The chimps could have independently generated the same infection from their own endoretroviruses. I don't know if this is viable, although I could test this by myself, just by comparing the sequence of SIV and HIV to see if they differ by more or less than the ERVs of humans and chimps. I didn't test yet, I am just asking.

42. - Evolutionary theory is based on the fact that life previously originated on Earth and we all evolved from a universal common ancestor. Where did life itself originate?

Question: Evolutionary theory is based on the fact that life previously originated on Earth and we all evolved from a universal common ancestor. Where did life itself originate?

Life emerges whenever a system with bulk computation is Turing complete. This is not a rare or strange event, as Wolfram shows--- many random cellular automata are automatically Turing complete, it simply requires that they have semi-stable structures that store data stably, and interact predictably.

In addition, you need bulk noise, to evade complexity bounds. The theory is described here: Ron Maimon's answer to How did life on Earth begin? and nowhere else (with the exception of other quora answers linking this one).

43. - Is it possible that life originated from protein and not RNA?

Question: Is it possible that life originated from protein and not RNA?

It is not only likely, it is impossible that life originated from RNA, because RNA cannot in any way plausibly form under prebiotic conditions. Further, a replicator doesn't work for the origin of life, you need a Turing complete soup. This is provided by random proteins in conjunction with petroleum, and RNA doesn't do anything metabolically and is useless in this regard. You can read more about it here: Ron Maimon's answer to How did life on Earth begin? .

44. - What would be the minimum set of elements, particles, cells and rules, which if programmed to run indefinately by itself, will show life like behaviour and evolve into higher life forms?

Question: What would be the minimum set of elements, particles, cells and rules, which if programmed to run indefinately by itself, will show life like behaviour and evolve into higher life forms?

This is the same as the question of what is the minimal Turing complete automaton. I will ignore 1d, because the Turing behavior there is very degenerate, requiring a very long time to do any meaningful computation. So the simplest such system is Conway's Game of Life, seeded with infinite random pattern, and provided with a randomness source, either the random pattern, or a small probability of a cell flipping color.

The minimal requirements are

* Turing complete behavior
* randomness
* Potentially infinite memory, and fast enough processor speed.
* Efficient local storage of information.
* Sufficiently fast bulk processing.

If Conway's system doesn't work well, it would only be because it is 2d. There might be obstacles to efficient shuttling of information long distances in a bounded size 2d cellular automaton (although I doubt it). Then the answer would be any simple 3d cellular automaton, or a polymer automaton in 3d.

45. - Has anyone used information theory to study the origin of life, and if so what were the findings?

Question: Has anyone used information theory to study the origin of life, and if so what were the findings?

Yes, there is such a formalism, unique up to some silly details, described in this paper of mine: [q-bio/0503028] Computational Theory of Biological Function I . It is a way of describing the interactions in a complex molecular system, from a computational point of view. The computational point of view is describing the information in the molecules, it's what you mean when you say "information theory", although the precise usage of the term in the literature is not quite what you are talking about. This research is the outgrowth of an origins-of-life hypothesis I described for the first time here: How did life begin on Earth?

The main point for the origin of life is that all you need is a Turing complete molecular soup, and you're done. This solves the problem permanently and persuasively.

The hypothesis is amenable to test by both simulation and experiment. The simulations simply simulate Turing complete automata, like Wolfram's, except in 3d, and with a bulk computing capacity (so potentially infinite strings at each cell with rewrite rules according to nearest neighbor or power-law decaying inteactions, modeling proteins). As Wolfram noted, every such system is Turing complete, unless it is random or trivial.

The experimental test is simply to make an amino acid and petroleum/water mix under conditions where polypeptides spontaneously form. The system should, under certain conditions of salinity and Ph, form a computing system, and the result will then start evolving immediately, as described in the linked answer.

This idea is original, so there are no references, other than the ones I linked here.

46. - Could life evolve in space?

Question: Could life evolve in space?

As explained here: Ron Maimon's answer to How did life on Earth begin? the only thing required to form life is the formation of a Turing complete system spontaneously, with bulk computing and essentially infinite memory. In space, it is hard to form such a system, but one cannot rule it out without a good survey of all the phenomena that happen in space.

In order to store memory, you would need stable structures. If you are using molecules, they would have to withstand the harsh cosmic rays for long enough to impress their structure on other molecules before they are destroyed. This is not likely in dilute space. It you are in a dense cloud, you can use ripples in density to store data, but the

transformations are usually either chaotic or regular, you either get a mess or a solar system. The chaotic systems tend to fall apart, and in any case, the amount of memory and processing in such a system is incredibly tiny in relation to the memory and processing in a watery soup of proteins and hydrocarbons.

Other places are planetary atmospheres in gas giants, which do admit stable structures, like the red-spot in Jupiter. Is Jupiter's atmosphere Turing complete? Possibly, but it's not likely. The Navier Stokes equations generate a downward cascade which has a stable equilibrium structure in fully developed turbulence.

Perhaps you can store data in molecules other than water and carbon-chains. This can be investigated experimentally, since one can easily determine when a collection of molecules make a Turing complete set, and determine the memory and processing from experiments. So far, only a protein, water, hydrocarbon soup can make a big enough computer, but it is an open question to catalogue all Turing complete chemical systems.

Perhaps life can evolve on the surface of a neutron star. Such life would be interesting, because it's time-scale would be as much faster than ours than ours is faster than geological scales.

47. – Is evolution still a theory?

Question: Is evolution still a theory?

There are several parts of the theory of evolution, and one must distinguish them because they have different degrees of support:

1. Creatures are descended from ancient common ancestors, so that people once looked more like monkeys, and previously like rats, and earlier like yeast, and earlier less differentiated still.

This is not a theory, but a fact. The reason it is a fact is because it makes enormous numbers of highly nontrivial predictions regarding life today. It predicts that living things form a tree, that the DNA sequences should show homology in the same way as that tree constructed from morphology, that fossils show convergence of genuses into one common ancestor, that the geological record and the fossil record conform to the DNA record, and all of these highly nontrivial tests are aced by the theory, with no real contradiction.

It is impossible to sanely dispute this given the evidence, and neither the Pope, nor Behe, the father of intelligent design, dispute this. This is only disputed in certain political environments where scientific truth is put on the back burner.

2. The mechanism of evolution is a form of natural selection, meaning some living things pass on their forms preferentially, while others die out.

This idea is less strongly supported than the first, but it is nearly certain too. The reason is that the selection pressure is observed in nature, we see that some animals pass on genes and others don't, and we can measure the degree of selection. We also have put animals under selection pressure of various strengths using artificial breeding, and the level of variation we have produced using artificial breeding, is comparable to the morphology change in the fossil record, and the natural changes proceed at roughly the rate one would expect from extrapolating the artificial selection to the levels of selection pressure we observe in nature. This means that natural selection fits the changes well, and one can be reasonably certain this is the major mechanism.

This is as far as one has to accept the doctrine of evolution from empirical evidence.

3. The mechanism is by competition as to which animal lives longest.

This is completely uncertain. The mechanism in animals might be predominantly sexual selection, as we can see that humans are still evolving relatively quickly, and the only mechanism in humans is sexual selection. But sexual selection in humans is very strong, selecting for height, big head, big penis, and wider hips (all are related for obvious physiological reasons), and more abstractly, for more efficient intelligence, social maneuvering, stuff like that.

4. The mechanism is through spot mutations of a DNA sequence which is replicated with mistakes. The mistakes are random and caused by cosmic rays, or some other environmental noise. The genes reassort and evolve independently and selfishly.

Although viruses work this way, for the most part, for actual living things, I think it is safe to say today that this is complete nonsense. The more we learn about mutation mechanisms, the more it is clear that they are not random and they are under the control of the gametes. Random error is not a way to evolve complex code, as there is a bottleneck in evolving algorithms if the algorithm mutation method doesn't co-increase in complexity along with the algorithms themselves.

Selfish gene theory predicts that genes are going to be the major evolving things, the protein coding sequences, and this is not born out by sequencing data. The majority of evolution seems to be in noncoding regions of DNA, that are purely regulatory. Humans have essentially the same genes as nematodes, with some duplications, but vastly more noncoding regulatory DNA.

So I think you don't have to accept selfish gene. In fact, you probably have to accept the exact opposite, although perhaps not certainly yet.

5. The evolution has no purpose or goal, and has nothing to do with religious doctrines.

This is not a scientific question so much, but a philosophical one. When you look at natural systems, some of them, like yourself or your cat, have a purpose and a goal, and others, like a rock or a hurricane, don't. Since all biological systems are Turing complete and have a

large store of memory and fast processing speeds, they are more like a brain than a rock, they have the computational spark. The collective phenomena in evolution produce a larger computation in the collective, so a larger intelligence, and there is no reason to think that this enormous natural computation doesn't have a goal. It is also possible to say it doesn't. But this is not a scientific question, so it isn't served well by the pontifications of scientists, most of whom have a sociological agenda when pontificating.

6. The doctrines of evolution should be emulated by society, to make a system which allows people to suffer and be weeded out mercilessly through unregulated exploitation through the result of competition.

This is not a scientific question either, but this is the main issue religions have with evolution--- it looks like a self-serving way for people to reimpose the Darwinian society of the Roman empire on the modern world. The Christian religion worked hard to topple this order, and this gain must not be reversed.

Whether or not nature works by cutthroat competition with no mercy, and I don't think it does, it has no bearing on how compassionate humans should be to one another. There is no reason to immitate nature if nature is abhorrently cruel.

48.- What are the chemical origins of life? How did non-living chemical compounds generate self-replicating, complex life forms?

Question: What are the chemical origins of life? How did non-living chemical compounds generate self-replicating, complex life forms?

The origin of life is from random proteins (amino-acid polymers, including branched polymers) near random hydrocarbons (petroleum). The hydrocarbons are common in the early Earth, they are still around today, we call them "petroleum". These things are called "fossil fuels" by moronic petroleum scientists in the pockets of oil companies, but it has long been understood that they have nothing to do with life, that they are ubiquitous in the solar system, and that they form under abiotic conditions. This was discovered in the Soviet Union, under the influence of Dmitry Mendeleev, and it was also patiently explained and promoted in the west by Thomas Gold, but it is not appreciated by the public or by scientists, despite the evidence being overwhelming. Their stupidity is not my responsibility.

Simiarly, amino acids are ubiquitous in the universe, since they are naturally made by CO_2, H_2O, and NH_3. These are the most common small molecules on early Earth's atmosphere, as they are on Jupiter, and they produce amino acids in great quantities, and the hydrocarbons are always there. We are sure hydrocarbons and amino acids both form, and both mix, in proto-oceans of hydrocarbon and water.

The life begins when proteins form under ocean conditions that favor peptide bonds

between amino-acids to form spontaneously. Then you get a collection of proteins. The important thing is that this collection is Turing complete, in the sense of Wolfram's automata: they produce any computation if you fiddle with them. They also have abundant and growing food (petrolium), so they can learn to extract energy from hydrocarbons. Once you have a Turing complete collection with a food source, this is sufficient to produce Darwinian evolution, and from this point, the system will engineer stronger and stronger mechanisms to store information spontaneously and retrieve it, producing an RNA-like molecule for denser information storage (it doesn't have to be RNA), a DNA-like molecule for permanent storage, and then producing ribosomes, and cells. The evolution starts at the moment you have a "class 4" Wolfram automaton.

This process of Darwinian competition and high-level replication of properties begins long before there is any kind of exact replication of low-level molecules, it occurs much as in immunology, by weeding out thins that are incompatible with the environment.

In fact, this is the correct definition of Darwinian evolution. Replication with modification is a bastardization which is appropriate only for systems which are barely alive like viruses, or for the parts of living things that are no longer being innovated on a regular basis, like proteins in the bodies of mammals.

The idea is explained in more detail in my answer here: [How did life begin on Earth?](#) .

The point of view invokes Turing computation, which in the common-sense computational view is the definition of intelligence, and the action of this collective intelligence to engineer the information molecules of life by trial and error and collective evolution of thought, meaning computation, in the soup, so it is a form of intelligent design. Except the intelligent designer is the molecular soup. This idea is kind of obvious since Wolfram noticed that automata are easily Turing complete, so it always amazes me that I was the first to think it up. Right now, Chaitin is going around South America making similar claims for origin of life from Wolfram, but this is an old idea, due exclusively to me (although I am surprised I was the first, it's completely obvious) and I put it on stackexchange before he wrote anything at all about this.

It is also easy to test this idea in computer experiments, by simulating a Turing complete system of an enormous size, and watching it evolve. The predictions are that you will produce Darwinian competition immediately, with absolutely no obvious replication.

49.- If carbon, an abiotic thing, was the basis of life on earth, what's the fundamental difference between a biotic being and an abiotic thing? How did the first genes come into existence? How did they form?

Question: If carbon, an abiotic thing, was the basis of life on earth, what's the fundamental difference between a biotic being and an abiotic thing? How did the first genes come into existence? How did they form?

A living system is Turing complete and embedded in a larger Turing complete system with an essentially infinite memory. The main difference is the difference between a computing cellular automaton and a non-computing one, except here the information is stored in the molecular bindings and configurations.

The idea is not in the literature, but I described it here: How did life begin on Earth? .

It's self-evident and it solves the problem, as it is both true that Turing complete systems evolve in the Darwinian sense, that other systems called "Darwinian" are not Darwinian, since they are not computing (for example, self-replicating molecules), and these systems are absolutely incapable of generating life. The resolution also clarifies what to look for regarding evolution in modern life, and I think every other answer to this question is obsolete now that I have explained this idea. Sorry, folks, I figured it out. You didn't. Nyah, nyah.

I should also add that if you think of the big computer made of molecules as a gigantic disembodied mind forming, which is perfectly fine, since computation is essentially synonymous with mind, then this is basically saying "God did it", except in a testable way that has nothing to do with established religion.

50.- Presuming that life appeared as a result of a natural event, what could have been this event?

Question: Presuming that life appeared as a result of a natural event, what could have been this event?

It is my opinion, I think shared by nobody else with the possible recent exception of Chaitin, that this event is the formation of a soup of molecules which together make a Turing complete system, with a potentially infinite memory due to the large size and processing distributed over the whole volume, and with random chance introducing random disorder into the system.

The model for this is the cellular automata of Von-Neumann, Conway, and Wolfram. These

model Turing complete systems which appear in simple simulations of information flows based on rules.

The processes of catalytic transformation of random proteins are exactly analogous, and certainly form a system of the automaton type, and when these proteins form a computing automaton, as opposed to a stable, periodic, or random automaton, they begin to evolve spontaneously, and they will eventually generate some form of stable information storage and modern life, though a not-yet-replicating form of biological evolution.

I should point out that the replicating parts of modern biology are generally things that were evolutionarily sorted out long ago, and modern evolution bears more of a resemblence to this pre-biotic automaton evolution than to the modern synthesis idea of SNPs and fixation. The process of evolution must never be separated from the idea of computation, because the two go hand in hand.

The natural event, forming a Turing complete system out of molecules, can occur in many chemical contexts, and is possible even in gaseuous systems, and perhaps on non-molecular substrates, like neutron-star surfaces or interiors. The only requirement is that information can be stably stored without randomizing, and that some clumps of information can reliably transform other clumps of information with a rate that is not too slow and that is uniform in a volume. The randomness required for evading complexity bounds on the computation always comes for free in physical systems, since they always have some thermal jitter.

51. – What's wrong with the theory of evolution by natural selection?

Question: What's wrong with the theory of evolution by natural selection?

There is nothing wrong with the idea that all life came from a common ancestor over a few billion years of descent and modification, and there is almost surely nothing wrong with the theory of evolution by some form of selection, but the theory of evolution by natural selection, meaning that the selection process is predominantly through the action of early death by predation or disease, has some uncertainty.

There is a serious issue for testing this idea, in that by itself, it does not give an estimate for the time-scale for speciation for advanced forms of life (meaning, eukaryotes, not viruses or bacteria or archaea) from the basic mechanism alone. You need a theory of mutation genesis and selection landscapes. Darwin gave a theory that tells us that if you have a bunch of rats, they will turn into elephants over a period of about 10 million years, but the only reliable estimate for how long this takes is from observing how long it took, which means that we don't have a test of the detailed mechanism, and the mutation genesis mechanism and selection landscape are generally taken to be primitive.

I am actually lying a little in the above. Darwin did give a method for estimating the time

scale for evolution by natural selection, which was simply the time scale for evolution by artificial selection, by human breeding. He observed that over 6000 years, we could take a wolf to a chihuahua, a crab-apple to a farm apple, and a zebra-like animal to a pony. So this gave him an upper bound for the rate of change of morphology, and this estimate was more than enough to account for the changes in the fossil record.

You can't argue with Darwin's argument about artificial selection, this does produce very fast changes that go about a hundred times faster than natural evolution, and if you consider the morphological changes in anatomy over geological time, they are consistent with a slower version of artificial selection, a process Darwin identified as natural selection. But this estimate is not very quantitative.

The estimate from artificial breeding is given under extremely harsh selection pressure, essentially perfect selection. Selection by predators is imperfect, and selection by disease and early death is also imperfect. But the natural process takes longer than the artificial process too, so it works out, more or less, enough to give good confidence that some sort of selection is at work. But the fit is not particularly quantitative, so you don't know exactly what kind of selection is operating.

In the 20th century, the introduction of Mendelian genetics made the modern synthesis, which modelled evolution as a process which begins with a set of genes, which mutate by random deletions and insertions, or occasional duplications, and then the most fit gene fixates in the population independently of other genes. This gene-individualistic model was supported by the observation that crossing over shuffles genes arond, so you can think of the fitness of each gene in isolation to a certain extent, because it will wander about at random through the organisms.

This individualistic model of gene selection is clearly false, it doesn't work to explain macro-evolution, because the amount of morphological evolution between distantly related animals has no counterpart in the selection of the genes involved. We have almost the same gene set as simple worms with only a dozen tissue types, and the difference in coding genes between humans and chimps is essentially zero. Further, the mutations in genes accumulate at a steady rate, in an apparently rather uncoordinated manner that seems to have no selection pressure at all controlling it, that is called "neutral evolution", and the neutral model gives an accurate molecular clock that can track species divergence.

Further, there is a theoretical problem with the idea that genes evolve randomly and independently to produce evolution, in that the time-scale for any sort of functional change goes to zero in such a model very quickly--- there is no chance of producing long term evolvability. This is on general principles--- when you have a computational program, and you evolve it by a given fixed collection of mutations, there is a quick saturation where you find the best local minimum, and then you stay there, until the selection landscape changes. This kind of stuff, the primitive kind of minimum-finding evolution, is not consistent with producing new structures and new tissues of advanced collective genetic function, it doesn't make any more sense than saying that books are written by a process of copying previous authors, with occasional errors introduced by careless scribes. That's not how new books are written, they are authored by an intelligent process.

There are exceptions to the neutral mutation rule--- some single nucleotide mutations in protein coding regions are lethal, and some are very deleterious, like the mutation for sickle-cell anemia. But the general rule is that you don't find a difference in coding regions which is at all consistent with the morphological difference you see in higher animals, and also, you see a tremendous amount of non-coding region variation which is much more difficult to account for, and the total length of non-coding RNA does correlate with the morphological complexity of the animal, so that animals with a large number of tissue types, like human, have a gigabyte of genome, while animals with a small number of tissue types, like worm, have tens of megabytes of genome, even though the genes are largely homologous.

The clear inference is that the non-coding genome is the part that is relevant for evolution, and it is functioning to control gene expression and activity in much the same way your brain controls your body, so that it is computing at a reasonably high level of complexity, and is therefore intelligent to some extent. Many of these RNA mechanisms have been discovered in the last decade, and the RNA-brain hypothesis, usually associated with John Mattick, although I don't want to put words in his mouth, is now not as controversial as it used to be.

Knowing about the noncoding RNA regulation, and given the much higher bit-density of RNA as compared to protein, it is manifest that evolution proceeds in an authorly way by rewriting non-coding RNA during replication. This is the low level mechanism, and it co-evolves with the high level stuff, so that the mutations RNA cleavage and insertion introduces in a complex genome are sensible, and compatible with the already existing complex program encoded by the genome.

This mutation mechanism is required, and it is sufficiently different from the textbook picture of blind mutation that I think it is closer to intelligent design than to the modern synthesis. It is not yet accepted that such a thing is required, because biologists associate complex mutation mechanism with a rejection of the blind, mechanical, idea of the modern synthesis, ideas which were ruled out when they were first proposed, by the absurd time scale in estimates to go from mouse to human by blind mutation. But now that we can see the genomes, the modern synthesis is completely busted. You can see retrotransposons, endoretroviruses, copy-number variations in non-coding regions, general complex non-coding templates with unknown functions, gene-like silenced sections, and all sorts of stuff that is obviously showing that the standard story is a tall tale.

The other problem with evolution by natural selection was already pointed out by Darwin--- this is sexual selection. Once you have sex and brains, there is a new mechanism of selection by mate choice, and this selection can produce peacock tails, things that reduce the predator fitness of the organism.

The existence of sexual selection allows things which are forbidden by ordinary natural selection. Mate choice is itself evolved, so that if a subspecies of a given species of rabbit decides to sexually select for fast running, and another subspecies decides to select for fluffier tails, the group that selects for fast running will outcompete the group that selects

for fluffier tails as they diverge and speciate, and can drive the other group to extinction.

This type of thing is called "group selection", and sexual selection allows you to have group selection, because the basket of sexually selected traits is the group-selected through competition with other groups which have a different basket of sexually selected traits. Group selection is completely forbidden by ordinary natural selection, because traits which are beneficial to the group, but harmful to the individual, will be weeded out. Sexual selection allows you to compensate for this, by sexually preferring traits which are beneficial to the group.

The existence of group selection through sexual selection is not recognized in mainstream evolutionary biology, but it explains certain mysteries. In prarie moles, there is an observed scouting behavior, where the animal will raise its head to look for predators, and alert the group if it found a predator. This trait puts the animal at risk individually, by a quantifiable amount (you can see how often the predator gets the risk-taker) and by an amount which makes it difficult to see how this behavior could have possibly evolved. Kin-selection was proposed as an idea which can account for this in principle, but the kin-relatedness of the pack is not sufficient to account for the scouting behavior by itself, because the pack is too big, the genetic relatedness is not enough to explain the amount of risk. The idea I am proposing is then that this behavior must be sexually selected.

The evolution of animals might be entirely through sexual selection, and hardly at all through natural selection. The natural selection would then only operate on a higher level, selecting between different groups that have evolved different traits to a great degree very fast through sexual selection. This requires that cheetahs sexually select for speed, that skunks sexually select for smell, that giraffes sexually select for long-necks, and in general that every trait we associate with natural selection could be a product purely of sexual selection, with natural selection operating on the group level only.

The only case I have any experience with is that of humans, which is because I happen to be a human. In our case, it's our extraordinary brain which is the most distinctive trait, and language which is the most extraordinary behavior, and these are insanely strongly sexually selected. It is difficult for humans to reproduce without a long bout of talking first, and the talking is not necessary for reproduction. The brains don't protect you from predators at all. So I can see with my own eyes that humans evolve brains by sexual selection, and not by natural selection. Whether this is true for everything other species, you need data to say that.

This idea is also not in the evolutionary biology literature. In evolutionary biology, the consensus is, by blind authority, that group selection is forbidden, because without a sexual selection mechanism, it is impossible (aside from the small correction provided by kin selection).

52. - Are there any scientists out there who can or do make a 'good' case against Evolution and can anyone also make a case for a younger earth (i.e. not billions of years old)?

Question: Are there any scientists out there who can or do make a 'good' case against Evolution and can anyone also make a case for a younger earth (i.e. not billions of years old)?

You can't make a case for a younger Earth, because we have mountains of evidence that the Earth is about 5 billion years old, for example, the amount of radioactive decay products we find in naturally occurring Uranium. Uranium has several decay channels, and they are all consistent with an old Earth. The picture from different radioactive elements is consistent, and consistent with the fossil record. You could say the Earth was created 6000 years ago with fossils in the ground and radioactive elements in strange proportions in rocks, designed to look like an old earth, but you can equally say that the Earth was created 10 minutes ago, for that matter, or will only be created 10 minutes from now. This type of thing, creating an Earth which looks ancient, is logically positivistically meaningless, you would be abusing the word "created".

You also can't make a case against the thesis that all life on Earth comes from an ancient common ancestor by descent, modification, and selection, because this is completely obvious when you look at the common genetic heritage of life. Just looking at genetic data, it is obvious that the organisms are related in a tree, and that the tree matches a fossil record of species splitting from each other by decent and modification. You can also identify the process of separation of species as selection pressure, due to different rates of success of different lineages in the background of other life and other things, because we have artificial breeding, and we can get a sense of the time-scale required from the rate of artificial breeding, and it's consistent with the fossil record being a product of a form of natural selection. This point is Darwin's and it is also unassailable.

But the motivation for such questions are religious--- can you make a case that life on Earth is the product of a deep intelligence, which can be thought of as analogous to the infinite intelligence which religious people intuitively identify as the source of ethics? Can you make a case for the concept of God in the Bible, that thing that tells people how to behave as historical agents, as being somehow relevant to both biology and to human social organization? And can you rule out certain forms of evolution from these ideas, not just the vague Darwinian concept of natural selection, but more specific mechanisms which lack in intelligence?

In this case, I think the answer is yes, although nobody makes such a case today. So I will make such a case below.

The main issue is that scientists often present the view that evolution completely lacks intelligence, that it is a blind process. This is the thesis Dawkins promotes, and one can identify this with the modern synthesis.

Evolution is clearly a natural process, it doesn't require supernatural intervention, nothing we observe requires supernatural intervention. But that doesn't mean evolution is an unintelligent process, or that it doesn't have a goal. Your brain also doesn't require supernatural intervention, but it is intelligent, and your brain can formulate goals. It is difficult to clearly say which natural processes are intelligent and which are not in such a way that brains come out on one side and everything else goes on the other.

Alan Turing distilled the essence of the notion of intelligence into the mathematical notion of computation, and I will accept this: the intelligence of a naturally occurring computer, like a brain, (technically, this requires a random number generator, a random oracle in Turing's way of stating things, but the notion of a random oracle is predicated on a consistent notion of probability, or working in a Solovay universe, but forget these technical considerations--- a random number generator is what you intuitively think it is) is the total number of megabytes and processor cycles required to simulate the process in the most-efficient possible representation. When the number is small, then the thing is not intelligent, and when the number is big, then its intelligent.

Not everything requires a big computation to simulate the outcome. For regular deterministic systems, the computation doesn't grow linearly with time as is required to predict a computation, to predict the far-future of a regular system requires only logarithmically growing computational effort. For a chaotic system, you don't need growing computation either--- at some point you switch out to a random number generator choosing a point at random from the attractor according to the stable measure. Both these kinds of dynamical systems are relatively stupid. To simulate the atmosphere of the entire Earth to best-possible precision requires no more than a few gigabytes of velocity data, none of which is stable enough to compute anything at all, the bits randomize each other and forget their initial state within a few weeks. To simulate the moons of jupiter (to the extent that the motion is regular) requires a lot of data, but the data is not interacting to make a full computer.

But for other natural systems, like protein networks in cells, you can have complex computational interactions that form an actual computer, with dozens of kilobytes of data in each cell. Likewise, RNA networks in cells can have gigabytes of data, all interacting in that way identified by Turing as leading to full computation--- with switching behavior and stable memory. This is sometimes given a proof, but I prefer to baldly state it, because if you are familiar with how little is required to produce Turing computation, you can see that it is obvious.

Given this computational behavior, it is difficult to say whether the process of evolution is intelligent or unintelligent. It all depends on how much of the RNA and protein computation is relevant to the process of producing new genomes.

Within a widespread view of evolution, the modern synthesis, evolution is not very intelligent. In the modern synthesis, mutations are produced by random SNPs, duplications, and deletions, and then selection works as usual. This process is very difficult to make work, because it is unintelligent. Scientific consensus for a long time was that this is the entire story in evolution, but this consensus was not strongly supported by any evidence,

because we didn't have enough genetic data to see what was going on in detail. It was also very strange to imagine that this is what is going on, because it seems like it would take eons and eons to make evolution go from a rat to a human by SNPs, deletions, and duplications, but it only took a few tens of millions of years in nature, which is a very short time to SNP/delete/duplicate mutate a gigabyte genome between distant forms.

This idea, the idea that evolution is guided by unintelligent processes, is to my mind completely busted by modern biology.

The proper hypothesis is that there are large networks of RNA in the cell, that these networks make a computation of many gigabytes size, and that this computation is important for creating mutations and crossing over in egg cells. The intelligence is not enormous--- it is small compared to a brain--- but it is not negligible, and it shows that the evolutionary process is far more sophisticated than the blind process Dawkins imagines.

The networking of these genomic computations through their interactions in the wild then produces an even larger system, which is hopeless to simulate except using an enormous computer, with trillions and trillions of genomic gigabytes interacting in complex ways. This is the size of the intelligence in life on Earth, and it is pretty darn close to infinite, at least by comparison to a laptop, or even to a human brain.

So it is perfectly sensible to ascribe intelligence to the process of evolution, and even to abstract away the limiting intelligence and call it God, and then the process of natural evolution, the processes of social evolution, and the processes in your own head that lead to deciding between different actions are all consistently thought of as the product of the limiting conception of computation of infinite size, with all the goals, aspirations, dreams, and more that our brains, which are of finite size, can be seen to possess.

Then the Biblical texts are simply an expression of this intuition regarding computation, in an era where all computation was seen as a supernatural phenomenon. It is important to keep this intuition because it is not a false intuition, but it is also important to understand how it is realized in nature.

Cognitive Science

53. – Are the experiences of the "flow of time" and of "cause and effect" qualia?

Question: Are the experiences of the "flow of time" and of "cause and effect" qualia?

Physics and biology have no answer why or how we personally experience the color red and we say that the experience of the color red is a quale.

Physics also has no answer why or how we experience the flow of time. So is the experience of the flow of time also a quale?

Since the flow of time is so closely related to cause and effect (i.e. the cause cannot come after the effect), is our experience of cause and effect also a quale?

Yes, the flow of time is exactly of the nature of a qualia, but I dispute the idea that physics has no answer about how or why we experience it. The thing that one must accept first is the physicist's positivism: once you answer every question about the measurable attributes of the brain and behavioral states, you have given the answer all questions about the qualia, even though the map between the qualia categories and the measurable attributes might be complex.

The theory of mind that physicists usually accept without reservations is the computational theory of mind, which was proposed by Alan Turing in the 1940s. The computational theory of mind begins with the observation that a computer can simulate the behavior of any physical system, including a human being, and then makes the positivist claim (this shouldn't even have to be said) that if you have something behaving indistinguishably from a conscious person, it is a conscious person. Turing codified the positivism in the Turing test--- if the communication with the computational entity is indistinguishable from the communication with a person, one identifies the computational entity as a conscious thing, like a person. This is common sense, and it is not something I consider worthy of reasoned debate.

The computational theory of mind identifies the flow of time with the direction of the computation, the relation between inputs and outputs. In any physical computer, the computation must dump heat into the environment in order to keep its information from randomizing, so for a regular computation in time, this is the same as the entropy arrow of time. The identification of the entropic and conscious arrow of time follows.

The "feeling" of time passing is a high level property of the brain's computation, and it is a property of the software, or the mind, or the soul, I use all these terms interchangably because they are equivalent in the computational view. The software is abstract data, and the manipulations of this data can be described using larger qualia, which are just shorthands for classes of data stored in the software.

The embedding of the software into the physics matches the conscious time sense with the physical time sense. The main complicated stuff is in the software, not in the laws of physics. That the flow of time is not physics is manifested by its dependence on the mental state of the observer--- having a traumatic experience requiring immediate action, taking hallucinogenic drugs, falling ill, each of these can change the perception of time in drastic ways.

The feeling of time flow is so universal, and the changes in this flow in unusual states is so jarring to the individual, that many people look to physics for confirmation of the feeling that time is something fluid and unreal. They find false confirmation of this in relativity (which has nothing to do with the essentially nonrelativistic events in the brain's computation) and string theory, and this makes these topics more popular than type II

superconductors, although these are equally interesting and both are about equally relevant to consciousness.

Anyway, the person who removed time flow from physics explicitly and put it into the mind is Ludwig Boltzmann, and once he did so, he was able to ask questions about Boltzmann brains. The concept of a Boltzmann brain (a brain emerging fully formed from a thermal state by a fluctuation, and feeling continuity with previous experience vastly removed in physical time) requires a separation of the sense of cause and effect or of time from the laws of physics. This is standard physics philosophy, it reappears in the many-worlds interpretation, which is controversial for reasons that have nothing to do with its computational theory of mind.

54. - What is the scientific consensus on Penrose and Hameroff's recent paper claiming that their Orch-OR theory is the best theory of consciousness proposed thus far?

Question: What is the scientific consensus on Penrose and Hameroff's recent paper claiming that their Orch-OR theory is the best theory of consciousness proposed thus far?

It's an honest attack on the computational theory of mind. The idea is that the brain is not a computer, because computers are limited by certain undecidability results and Penrose thinks we're not. So he needs non-computability in the brain, and he decided to find it in microtubuoles.

This requires not only quantum physics in the brain, but unknown and uncomputable physics which can influence consciousness. I should point out that there is a known uncomputable process which influences consciousness--- generating random numbers. Adding random numbers to a Turing machine makes it a "hypercomputer" of a certain simple sort, it can compute a random number, and a Turing machine can't.

But that's not enough for Penrose, he wasn't something that can solve the Halting problem in the brain. Since this is impossible in current physics, he needs new physics, and new crazy physics of an unimaginable complexity, which cannot even be simulated.

This is exactly what you need to reject the computational theory of mind, and I commend Penrose for his honesty. But this is what compels one to just accept the computational theory of mind. For a person who accepts the comptuational theory, the paper is not useful, as it is certainly nothing to do with consciousness, and this is scientific consensus.

I should add that any quantum behavior in microtubuoles can't possibly be quantum coherent at room temperature over scales of the brain. This is the technical objection to quantum computation in the brain, and nothing Penrose and collaborators have done overcomes this objection.

55. – What do neuroscientists think about Douglas Hofstadter's lecture/theory on "Analogy as the Core of Cognition?"

Question: What do neuroscientists think about Douglas Hofstadter's lecture/theory on "Analogy as the Core of Cognition?"

I am not a neuroscientist, but I think he has hit on an important factor in cognition. The issue I have with the model is only in possible incompleteness. I think it is spot on for the classical problems he is applying it to.

The incompleteness, I think, stems from cases where you actually have an internal computing simulation of a system. For example, when I say "I have water flowing past a barrier, and vortices are shedding in pairs. I place a wire behind the barrier, how do the vortices move in response to the wire?

The resulting cognitive computation is done by mentally imagining the simulation, and picturing the flow field, and the response to the wire sticking in blocking the vectors, and slowly, your picture improves until you can see how the vortex will slide this way or that (I haven't done it personally, I just made up the example right now).

This type of cognition is how the most mentally challenging discoveries are made, not through the simpler analogical process that Hofstadter correctly identifies as lying behind the simpler arguments, such as the Einstein examples in his book. The thinking involved in constructing elaborate arguments is often a complete mental simulation built up from peices that are hard to articulate, and only communicated well in pictures and computer simulations, not from analogies with well defined objects and arrows.

This is the limitation in his system, in my opinion, the emphasis on structural relations that can be modelled in 1980s style knowledge base systems, with arrows and blobs. But within those structures, I think he is dead on accurate, as best as can be described at this coarse level of detail.

The book is a quick read, but I am only 2/3 through.

56. – Answer: What is Ron Maimon's theory of internal evolution?

Question: What is Ron Maimon's theory of internal evolution?

It's nothing particularly extraordinary, it's just the observations that the processes in a conscious mind is an evolution, involving competing ideas that survive and vie for attention, and some of them win, and those are the ones you think about, while a whole

bunch of related ideas don't win and swim around in a vast pool of little algorithms in your subconscious.

Sometimes you have a coup, where one species takes over completely from another, and that's like a "a-ha" moment.

The analogy is simply between the anarchic computation in the brain (as Daniel Dennett puts it) and the only roughly coordinated computation in neurons. The idea comes from the realization that the brain is a networked computation, not a distributed computation, so that each neuron has a large internal computer that is relevant for the overall function of the brain. So that there might actually be a grandmother neuron (this was a joke in neuroscience in the past), although it would have a bunch of other related neurons which are similar, in case it dies or something.

This point of view is described very well in Daniel Dennett's recent articles and writings about consciousness, and I defer to him for the long-winded philosophical blah-blah-blah. He thinks about things deeply and honestly. It's his idea, but I didn't get it from him. It's obvious once you realize the depth of each neuron's individual computation, that they are not computing predominantly in a network, but individually, and the network is like a crappy ethernet for them.

The main point of this for me is that it puts the old "nature" vs "nurture" saw to bed for good. When you are talking about evolution, the product of the evolution has very little to do with either nature or nurture, it evolves. The forms that come out are not directed by nature or by nurture, but by the internal evolutionary struggle inside the system.

So for example, on the planet Earth, the "nature" of the Earth was a primodial goop, with petrolium oozing from the mantle, and water, amino acids raining from the sky. That's it for the nature.

The nurture was the sunlight, and occasional asteroid impacts. That's all the nurture.

Can you predict from this what dinosaurs looked like? Obviously not. These are the products of evolution.

So when you have an evolving system, you don't look for an explanation of the ouputs in the inputs, or in the correcting mechanisms of nurture. You look at the actual details of the stuff that is evolving. That's what's going on in brains, evolution, not nature, not nurture, at least not a significant amount of either.

Question: What effect has the computer had on philosophy in general and philosophy of mind in particular? How have theories of computation and computer science shaped brain research and understanding of intelligence and big questions we are currently asking?

The computer obsoleted everything that came before, there is hardly any philosophy before Turing, it is mostly babble with a few exceptions. The exceptions are Leibniz, who arrived at the idea of formal computation already from Newton's laws, Russell, who defined a computing system in the Principia Mathematica, although he didn't think in computational terms yet, and Mach, who worked hard to found philosophy on the mental states of the brain, which is the only computation he had access to.

The majority of philosophy after Turing is also ignorant babble, because most of it, with the exception of the logical positivists and the functionalists are ignoring Turing and trying to reverse the progress. This is inexcusable, because Turing already made the advance, it is just burying the head in the sand and trying to pretend nothing happened.

So what did the computer do exactly?

1. It precisely define logic and mathematics

The laws of logic are defined today as a computer algorithm that can deduce the consequence from any system of axioms. The process went the other way, where first the logical algorithm was defined by Hilbert and Godel, following Boole, Frege, Whitehead & Russell, and other foundational logicians, and then the computer was abstracted by Turing as the simplest machine which can do arbitrary logical deductions using the Hilbert-Godel algorithm.

The formalism of first order logic is a standardized example of a formal algorithm, a computer program, that can produce all the deductions of a given set of axioms. This defines what it means to have a precise axiomatic system, it defines mathematical thought.

2. It can simulate anything

This is often called the "Church Turing thesis". It is a law of nature which states that any algorithm which is precisely defined can be implemented on a Turing machine.

That means that whenever you can precisely define any algorithm, for example a completely different kind of computer, like an analog machine, or another type of logic which is also precisely defined algorithmically, like modal logic, you can always simulate

the other computer or the other logic using an ordinary computer.

That this is true is both an input and an output of physics--- all physical laws which describe nature are computable (aside from possible randomness), so that any physical process can be simulated on a computer. So that the outcome of anything we can see can be generated by a regular ordinary computer.

3. It can think

This is a corrolary of 2, because a computer can simulate a person.

This conclusion was so shocking to people that there were arguments against it. But there is no argument possible--- the computer can simulate a person, therefore it can think. No further argument is necessary. The only caveat is that there is randomness in the universe, so you might need a random number generator.

If you wish to say computers can't think, you need to find a noncomputable physical law, so string theory is wrong, and there is some quantum gravity weirdness which requires uncomputable processes, and these then somehow influence human decisions in a different way than randomness.

4. Computers can occur spontaneously in nature

This is something that is clear after the work of Von Neumann, Conway, and Wolfram. There are extremely simple cellular automata that produce full computation, the complex automata. These natural computers do not require careful design by human engineers.

5. Biology is computation in nature

The complex systems can appear and evolve without being seeded with replication. The replication is high-level, of the algorithm, it isn't simply copying with errors, and the computation in biology becomes highly sophisticated, and networked, and cooperative/competitive in different degrees, as it is in your brain. This is how Darwinian evolution works, not in a primitive copy-error brainless way, but in a computing way which is sophisticated.

This allows you to effortlessly make many predictions in biology that are half-way between the modern synthesis (simple gene algorithms--- no significant computer) and intelligent design (God--- infinite computer).

These are the main points. But the computer and logic also found philosophy, in logical positivism. The logical positivists combined the positivism of Mach with the logic of Russell, and produced a philosophy which was able to ground traditional notions like "cause/effect", "meaning", "use" entirely in terms of "observations" and "correlations", things which can be processed by a computer. This program answered or mooted all the classical questions of philosophy, one by one, very quickly. The mooting was done by Carnap in the 1940s.

Since this stuff is encroaching on the philosophers realm, and it is ultimately coming from the sciences, it is actively resisted within the field of philosophy. Nearly all the major philosophy since 1970 has been devoted to attacking this point of view one way or another, either attacking positivism (later Wittgenstein), attacking computation as the foundation for thinking (Chalmers, Searle), or attacking the notion that computation is a good foundation for science (Popper). These attacks are reactionary, and wrong. These gains are irreversible, they are a ratchet, the classical philosophical problems have been solved.

There are also a few people who maintain that there is nothing seriously wrong with the computational foundations, and these are Daniel Dennett, Douglas Hofstadter, and most people in computer science or physical sciences.

58.- What is the state of the art in the quantum mind/brain hypothesis in a broad sense (not only consciousness)? Which is the relevant, updated literature? Is it a dead idea?

Question: What is the state of the art in the quantum mind/brain hypothesis in a broad sense (not only consciousness)? Which is the relevant, updated literature? Is it a dead idea?

It's a speculative idea, because nobody has a really plausible way that you can make a coherent large superposition type quantum state in the brain. It's at room temperature.

If there were a property of the brain that couldn't be explained by the hypothesis of a deep classical computation, then it would be worth investigating, but it seems that most of the motivation is just from a misunderstanding of the depth of the brain's computation, or from misunderstanding the implications of Godel's theorem. The computer in the brain has a random number generator, so it is not strictly a Turing computer, but a Turing computer with a random oracle, and these are more powerful, for one, they can compute the random oracle!

The Turing limitations are very intuitive, and seem to correspond to philosophical limits of knowledge that define the boundary of knowledge, not to some mystical idea that you can transcend computation itself. In any case, reasoning of this sort cannot lead to any conclusion that the brain is quantum, because normal quantum mechanics is also computable, only slowly. In order to make a non-computable brain, you would need new physics which is noncomputable AND also a mechanism to make it relevant to the brain. Penrose has speculated in this direction, but I think it is better to bite the bullet and just say the brain is doing classical computation.

Since there is no evidence otherwise, it is important to investigate with this hypothesis as the default. Since I think it is true, I think it is a waste of time to go along the quantum brain route. The idea of a noncomputable physical process (other than true randomness) is

philosophically difficult, and there is no evidence for it, certainly there is no argument for it from anything we have observed so far

59. - How do intuitionists and other constructivists feel about the probabilistic method?

Question: How do intuitionists and other constructivists feel about the probabilistic method?

It's not completely compatible with intuitionism formally, because the notion of "random construction" isn't the same, for constructions with infinite numbers of steps, as Turing computable in the intuitive model. The intuitionist logic will construct the object you claim exists from the proof that it exists, and this is a deterministic computation in the intuitive model.

But this is why Erdos's method is deep--- it is still a construction in the sense of a type of computation, but the model of computation for an object constructed by the method is a probabilistic computation. For finite constructions, it doesn't matter, but when you are talking about infinitely many random choices, it makes a difference.

A probablistic computer is the same as adding a Turing oracle which is a "random real number", but this concept cannot be completely well defined in usual set theory, because there are non-measurable sets. Cohen and Solovay showed that it is fine anyway to have random real numbers, and even without this advance, you can define random computation in the usual set theory, it's just a little bit more of a headache, and you are explicitly referring to Borel sets and measures the whole time, so it looks like a terribly complicated and totally non-constructive procedure, referring to all sorts of crazy kinds of uncountable sets.

Because the probabilistic model of computation is well defined, and seems to be the correct moel of computation in the physical universe, I personally am happy to call proofs using this method constructive. But a stickler intuitionist might insist that a constructed object must be Turing computable.

60. - Could the Internet already be conscious without us realising it?

Question: Could the Internet already be conscious without us realising it?

Not independently of the people involved, because the information is generally static, it isn't computing without humans intervening. If there is a consciousness, it is the same

collective consciousness of society that Jung identified in a way, and also the religious folks who sometimes identify this type of collective social activity as something which can be modified by the action of God.

61. - What is the place of logic in cognitive science?

Question: What is the place of logic in cognitive science?

Logic is at the foundation of computer science, it shows you how to build a universal machine. The role of computers is as a foundational in-principle solution to the problem of cognition, and Turing completeness is an in-principle foundation to biology and brain simultaneously. Turing completeness is ultimately a property of logic--- the computer was originally defined as the simplest definable machine which could do logical deduction on arbitrary axioms.

But the mechanisms of higher cognition are not easily modelled with formal logic, because this is equivalent to modeling them using a computer program. The computer program needs to be sophisticated, comparable in size to the irreducible processing in the brain, which means that modelling sophisticated thoughs using primitive logical sentences is hopeless.

This observation is the basis for the rejection of simple-minded logical AI schemes in the 1980s. It's not an in-principle repudiation of AI, it is just the observation that the computation is much more complex than initial attempts admitted.

62. - Could the functions of the human brain ever be described mathematically?

Question: Could the functions of the human brain ever be described mathematically?

There is no such thing as a truly analog computer, the analog models are always an approximation which can be replaced without loss by a digital model of high enough fidelity. The nature of the limit that gives rise to analog behavior is the large number of atoms acting together, and in the regime of brain operation, when you are dealing with single molecules or molecules with very small copy-number, the statistical flucuations make it impossible for the system to be analog.

To explain: suppose you wanted to make an analog system solve Laplace's equation with a delta function source in an infinite three-dimensional half space with reflecting boundary conditions. You can inject a dye at a certain point in a large pond of water, and then the

concentration profile is, to good approximation, the solution to Laplace's equation.

But you are relying on the fact that there are Avogadro's number worth of atoms there to get a reliable value of the analog system. If you could only inject one atom at a time, you would need to make an Avogadro's number's worth of position measurements before the result would statistically reproduce the solution of the equation.

Because we are macroscopic, we have a false intuition regarding this, thinking that analog systems are natural. When you are dealing with molecular events, you are in the opposite limit, where a deterministic analog response is the anomalous thing, most things are not capable of acting reproducibly enough to act as an analog system.

So when you have molecules diffusing, and having chemical reactions, the diffusion randomizes the position, and this randomization dumps the data in the position of the molecule into random motion of water, losing the information irreversibly. So any analog information in the position is only relevant to the degree that you measure the molecule's position, and this can only be done realistically by binding it to another molecule, whose position is also uncertain.

You can tether the molecule to a scaffold, and localize it like this, but then you are only getting the discrete information as to which component of the scaffold you tied it to, not the precise analog information about the position of scaffold. This information is washed out by Brownian motion of the scaffold.

Likewise, when you have some dynamical change, like a protein conformation change, this change occurs at some time, and lets say it's precisely defined. This time cannot be known to the biological system, only the first time that some other molecule noticed the change, by binding to the new molecule, or by trying to bind and failing. This introduces a time-discreteness scale, and if you make time discrete at a spacing smaller than the interaction time between molecules, and store the positional information only to the degree that it is not washed out by Brownian motion, you are left with a discrete number of bits.

Further this number of bits is dwarfed by the much larger number of bits in the molecule's binding conformation. So that the major data carrier in the cell is through the binding and unbinding of molecules, not in any analog source.

The voltages and electrical gradients in the cell are only functionally relevant when the computation in the brain can tell the voltages apart, and this introduces a discretization. No matter how fine you make it, it is always a far smaller number of bits to describe the electrical potential everywhere in the brain than to describe the conformation and binding of all the molecules of the brain, and the latter is discrete data.

So there is never a barrier to a complete digital description of the processes in the brain, and it is a mistake, often repeated, to claim that the brain is effectively an analog computation. It just isn't. At the molecule scales, it is impossible to maintain analog computation in a way robust to the Brownian randomization. It is false intuition from the

everyday world extrapolating to the micro-world, ignoring Einstein 1905, not the relativity work, the Brownian motion work.

63. - Is there some physical activity in the brain which could not be in principle implemented by a machine?

Question: Is there some physical activity in the brain which could not be in principle implemented by a machine?

No matter what the structure of the brain internal machinery, assuming there are no new laws of physics involved, and this is a pretty safe bet, then it must be simulatable by a computer, a regular Turing machine. If quantum mechanics is involved in a nontrivial way, which is theoretically all but ruled out, it will require a quantum computer. If quantum mechanics is involved just to generate random noise, it will require a computer with a random number generator, a Turing machine with a random oracle. It is not clear whether a true random oracle can be replaced by a psuedo-random number generator without loss. But it is not difficult to make a machine that outputs true random numbers, you just heat up a memory chip and read out the results. In none of the reasonable standard possibilities is it plausible or arguable that the brain has non-computable magic.

The implication of this is that ordinary computation is just as magical as all the stuff going on in your brain. Even a limited encounter with computers supports this assertion, as the behavior of the programs is unpredictable and maximally complex. The model appears good so far.

It must be added that some people don't like this conclusion, in particular Roger Penrose, and these people are forced to introduce convoluted ways in which a non-computable effect can affect the brain. Penrose has suggested that microtubules are sensitive to quantum gravity effects, but it is just because he doesn't accept that computation is the right language for thoughts.

But under any reasonable interpretation of current knowledge, computation is the right language. There is not even an in-principle way you can make a non-computable effect in the brain. But the argument, although trivial, seems to have been too hard for philosophers to grasp, so they continue to debate the issue.

64.- Are systems of aesthetics, ethics, values and moral law Turing-computable?

> **Question**: Are systems of aesthetics, ethics, values and moral law Turing-computable?

Everything is Turing computable, as we understand it, because computers can simulate molecular and atomic systems with high fidelity, and reproduce their behavior to arbitrary time, including enough atoms to make a person (in principle-- the simulation becomes difficult for larger systems).

For a good-enough classical approximation, the simulation isn't even hopelessly large, because you can model things in thermal equilibrium, like water, as thermal jitter on the stuff that isn't in thermal equilibrium, like bio-molecules. Further, not all the bio-molecule is doing computation, and you can simulate most of the effects using stochastic binding, unbinding, and diffusion-reaction, with a system size that is not much larger than the number of binding configurations. The main bottleneck information-wise is the number of conformations and bindings of biomolecules, not the space positions or the thermal stuff.

This type of simulation is not impossible to actually do, it requires gigabytes for a cell, many Terabytes for small multi-cellular animals, and 10^{21} bytes for a human. These are not impossible computational sizes, although they are still many orders of magnitude bigger than current supercomputers.

So not only can a computer do these things in principle, unless we are missing a law of nature, it can do it in practice, with achievable machines, although we are extremely far away right now, being about at the level of single-celled organisms.

This is why Penrose, who denies computation is a good model of thought, is required to postulate new physics relevant to the brain. He is aware of the fact that simulations exist, and he uses this to make a testable prediction--- that a hypercomputing quantum gravity is relevant to the physics of certain brain structures, in his model, microtubules. This idea wil be tested as we simulate neurons, and we will know if it is true, but I think it is safe to presume it will be false, because it has no support other than the intuition that computations are not rich enough to model internal experience, and this intuition is false on closer inspection, because computations are just as sublime as you expect for a good model of internal experience.

65. - Is intelligence hereditary or environmental? Why are some people smarter than others?

Question: Is intelligence hereditary or environmental? Why are some people smarter than others?

It's not nature and it's not nurture. It's evolution.

The thing going on in the head is evolution, competing little algorithms for doing stuff. The end result of this evolution has very little to do with the starting point (the DNA), or with the input of others (the nurture), because it is mostly going on within itself. The ideas are competing in the collective of other ideas, and the strongest algorithms survive, and make new thinking in the head.

You can guide it some from the outside, but the inputs and outputs are always vastly smaller than the thing itself. We only get a few kilobytes of data in text we read of talking we are exposed to, and the internal stuff is enormously larger. So the internal stuff has very little to do with either. It is self-generated, and the nurture (or nature) is just guiding the thing.

That doesn't mean you can't have a genetic defect screw it up, or complete social isolation, or cultural impact which directs some ideas to become prominent and others to fade away, but the thing itself has little to do with either. If you toss away all nurture and ignore your nature, you can think whatever the heck you want.

To see that there is something else, other than nature or nurture, use this analogy: try to predict the outcome of evolution on Earth from the Earth's nature (the primordial soup) plus it's nurture (the sunlight, occasional meteor impacts). You will not be able to predict what dinosaurs look like from this data.

With humans, because we have language, we can transfer some blocks of thought from person to person. This stuff has to hit good local flora and fauna to have an impact on the person, and there's no way to make that happen except by encouraging the evolutionary process and throwing it the idea again and again. But it gets easier with good communication technology.

The idea of IQ is nonsense, it's measuring the biodiversity of life on Earth using the height of the tallest animal. Sure, tall animals come later, but it's a bad measure.

66.- If a cell inside human body had consciousness, would it be aware of the larger consciousness inside which it is living?

Question: If a cell inside human body had consciousness, would it be aware of the larger consciousness inside which it is living?

This is essentially the same question as how a human being can become aware of the gods. You could ask the same question about ants--- if you are an ant, can you become aware that there is a colony?

Consciousness is a vague term, and it is hard to know what to make of it. I don't know what "self-aware" means exactly, because it's hard to imagine an objective test for it, but if you have one, give it, and then you can say if cells pass it.

On the other hand, we do have a good model for computation, and computation of a certain type is what is going on in the brain, since it is the most that is allowed by physics, or even by general philosophy, as to what a natural system can do. Computation is rather mystical and unknowable (despite the common perception, which is due to the rather primitive types of computations people see their laptop doing), it is the definition of complexity when it gets large, and the existence of computation in the brain is rather obvious, since computation was defined as the simplest abstraction of the operations a human being can do which still contained the ability to do all logical deductions, and therefore do mathematics.

There is a certain complexity limitation in computation, in that the complexity of a fixed computer program means it can never do certain things which are of a higher complexity, like prove the Kolmogorov complexity of a string whose Kolmogorov complexity is larger than the program. But the real computations in nature have access to a random number generator, and there is no bound on the complexity of a computation with a random oracle, the random oracle, the random number generator, will produce an uncomputable sequence of arbitrarily large Kolmogorov complexity. So using randomness, you have no obvious complexity limits, and when I say "computation", I mean "computation with access to randomness", which is slightly different than "computation with a fixed program".

The issue with cells is that we know how many bits of data can fit in a typical cell, about 1-10 gigabytes, and that's certainly not enough for a human-style consciousness. There is no way this computation is sufficient to read "Hamlet" or to compose, or even listen to, a Beethoven symphony.

But let's pretend.

Under these circumstances, the cell can become aware of the larger organism by noticing the constraints on its behavior, and providential action. For example, imagine this dialog between two cells in my liver:

Liver-cell Louie: I am sick of being a liver cell. I want to be cancer.

Liver-cell Lisa: No! don't become cancer. It's dangerous. I heard a liver cell became cancer last year, and Immune-cell Ingmar ate her!

Louie: I know ingmar. He has a crappy antigen set, I can evade him.

Lisa: But you know you're not going to be a tumor. No one has ever been a tumor before.

Louie: I am sure I can make it. Here I go.(transmute) RAWR! From now on, you call me CANCER CELL CARL!

Carl: I'm replicating already. FEEL THE POWER!

Lisa: I don't know about this Louie, uh, I mean Carl. I have a bad feeling you will be punished by Ron.

Carl: Ron! Ron! That stupid superstition? Have you ever seen Ron? I don't believe in Ron.

Lisa: But look at the blood stream supplying us with nutrients, the nerve cells supplying us with signals. Is that not evidence of Ron?

Carl: It is nothing of the sort, that's Harry the Heart cell and his friends, supplying us with blood, and he won't stop just because I'm cancer. In fact, I have assurances from Stem-cell Stanley and his buddies that they will make a whole new set of blood vessels in case I make a big tumor, and we'll have more nutrients than ever! Even you.

Lisa: I still think this is a bad idea. Ron is going to punish you somehow. I don't know how.

a few months later, Carl is a big tumor.

Carl: Look at me now! I made it, I am a huge tumor, I have vast influence. Blood cells come to me from new vessels specially constructed. Aren't you feeling stupid now for being a liver cell?

Lisa: I still think it's not right. What if everyone became cancer? Wouldn't the whole blood system collapse? Wouldn't we stop getting nerve signals and the bile production would drop... it doesn't feel right.

Carl: Forget about the other cells. I am unique! I have always known it. I always knew I was meant to be cancer.

Then I go to the doctor, and he tells me I have liver cancer. So I have an operation, where the tumor is removed.

Carl: Oh no! What's happening? We are being severed from the blood stream! All is lost.

Lisa: I told you it wasn't a good idea to become cancer, we are being punished by Ron.

Carl: I don't believe in Ron! It's just bad luck.

So the sign of being embedded in a larger collective with consciousness is the traditional religious notion of providence and punishment, that events which seem uncoordinated will provide evidence of a larger intelligence which serves to correct actions which harm the collective.

It is not clear to what extent human collectives are more intelligent than individuals, and the notion of God is even more subtle still, since it imagines a limit of ethical behavior extending upwards through even larger collectives.

These stories I am telling about Lisa, Carl and so on, are parallel to some of Jesus's parables.

67. - How do I believe that the humans around me actually possess consciousness, and it's not just me who is conscious? What if the people around me are merely programmed to act that way?

Question: How do I believe that the humans around me actually possess consciousness, and it's not just me who is conscious? What if the people around me are merely programmed to act that way?

This question is meaningless in logical positivism, it makes no difference to observation, and you could ask the same question of yourself--- how do you know if you are not just programmed to delude yourself into thinking you are conscious? Your perspective makes it that it is difficult for you to imagine this possibility, but you can't make objective evidence for it, except by throwing your hands in the air and saying "but I AM conscious!"

So you should trust other people when they throw their hands up in the air and say "but I AM conscious!", and so you should trust any other programmed agent that sufficiently resembles a human being in responses, and where you can sense a complicated evolving internal computation, and where you can talk to it, and when you ask it "how do I know you are conscious?" it throws it metaphorical hands in the air and says "but I AM conscious!".

This is the insight of Mach and the logical positivists.

It is the ability to decipher and respond to your words, which requires a massive amount of computing on top of computing, with evolving algorithms, with the attendant ability to sense the connotations and subtle implications, and compute the consequences of those, that is the logical positive deinition of what it means to be conscious. There is no reason to suppose that the quality of consciousness is anything more than this computation, that's what it "feels like" to have such a massive computation.

The size of the computation required to simulate consciousness is staggering, it is on the order of 10 gigabyte per cell, or for a human brain, with order 100 billion cells, it's 1 trillion gigabytes, 10^{21} bytes. This is simply the total weight of RNA in the brain times 2 bits per base. This is a staggeringly huge amount of random access memory, it dwarfs every computer we have, it is comparable, but slightly larger, that the total data on all the hard drives on earth. This is the information content of a single brain.

This is assuming the not-at-all mainstream hypothesis that I take for granted, that the computation in the brain is intracellular and done by RNA. Some arguments supporting this position have been compiled by Mattick, but I don't want to put words in his mouth, this position is not really in the recent literature. The reason to believe this is simply that it gives the right model for the level of computation in the brain, and it is evolvable from cellular mechanisms, it provides a mechanism for memory and learning, and it is the right order of magnitude for the computation we can do, unlike other models which might be

more popular, but which fail at accounting for even the simplest cognitive tasks human beings do, like recognizing a bicycle at a glance, and remembering the recognition for some minutes.

A computation of the sort the brain does, when it consists of 10^21 bytes networked together in 10^11 clumps of 10^10 bytes, when it is active and responding to stimuli, is the definition of consciosness, as far as a logical positivist is concerned. It makes no more sense to say that it is there without consciousness than it is to say: "How do I know my computer is running Microsoft Windows, or simply simulating Microsoft Windows?" Simulating Microsoft Windows is the same as running it, and simulating consciousness through programming is equivalent to consciousness.

The reason that this is not intuitive is that the actual data we get from people which gives us evidence of their humongous internal computation is always vastly smaller than the 10^21 bytes in their heads. When we talk or type, we get a few bits per word, so that even a long communication is only a few kilobytes of information.

Because of this gap in information between the few kilobytes we see and the unimaginable amount of data inside, we can easily imagine a preprogrammed computer, with only a few kilobytes, that stores the answer to the questions we ask our friends, and feed these pre-prepared answers to us. Such a computation would be trivial, and such a computer is a zombie. This is what it would mean to be "preprogrammed and not conscious", the computation would be tiny.

But the problem with this perspective is that you can't predict in advance what the interaction is going to be! If you want to fool someone, you need to have pre-prepared kilobyte answers to any of the possible kilobyte questions, and the number of kilobyte questions is 2^8000, a number which is close enough to infinity. To make a database of all possible questions and all possible responses might be imaginable, but then the next question depends on the answer to the previous one. For example:

> What do you think of Shakespeare's Hamlet?
> He's a whiner! (preprogrammed)
> Why?
> ??? (can't respond--- need an even bigger database)

So even a very short interaction with a person is evidence of a computation vastly larger than the size of the communication overhead, because it comes out coherent, no matter what the question! To gain evidence of a computation of large size, you don't need more than a little communication with the computation, to see that there is internal processing going on.

It is just plain statistically impossible to get a zombie computation with a table-base lookup to do an interaction over any length of time, and that length is ridiculously short, even a few questions will trip it up:

> Wazzup?

> Not much. Hanging.
> doing what?
> Chilling watching TV.
> What's on?
> It's stupid, it's some talk show with this guy that thinks he's an alien.
> What does he think aliens look like?
> You know, green, big eyes.

Already the unpredictable nature of the questions makes it that any stupid program will fail. This type of thing is a Turing test, and we do it all the time online. You can't be fooled by a bot.

This is the essence of the Turing test, and why it is controversial. The Turing test says that to verify consciousness, you only need a few kilobytes of interaction which verify coherent responses to arbitrary questions. This is counterintuitive, because the questions and answers are relatively short amount of data compared to the richness of internal experience, a few kilobytes compared to 10^{21} bytes, so our intuition is that a computer can be pre-programmed by a trick to respond to these without having 10^{21} bytes inside. But it's not true, as even a few kilobytes is enough variability to be essentially infinite.

68.- Are there any structures in the brain completely unique to homo sapiens?

Question: Are there any structures in the brain completely unique to homo sapiens?

No there aren't any that we know of. The main differences in human brain tissue is the larger size relative to the body, and the noncoding RNA expressed. Mammals have retrotransposons which are active in brain tissue, and humans have a different set of retrotransposons than chimps, as these are very quickly evolving unstable genomic features, and they are likely dynamically changing during the lifetime of an individual. The molecular signatures don't show up as anatomy, but they change the character of the computation involved, if that computation is intracellular, as I am sure it is. This is where one should look for the source of the extraordinary cognitive abilities of humans.

Music

69. - Answer: How did Einstein's musical practice inform his scientific work?

> **Question**: How did Einstein's musical practice inform his scientific work?

> There is no connection, except for the fact that playing music requires you to invent and interpret structures constantly, and keeps your brain limber. Heisenberg was a nearly concert-level pianist, and there are many other physicists who use music to relax. Composing is very mentally demanding, and resembles the creative process in physics, except the constraints are totally different, coming from human psychology and not matching nature.

70. - What is the relationship between music and math?

> **Question**: What is the relationship between music and math?

> There aren't much. Music is an art in patterns, while math is an art in different kinds of patterns at a much higher level of abstraction. Schubert and Beethoven weren't mathematicians.
>
> But there is a formal language aspect to music, in that the 12 tone scale is like integers mod 12, with a certain harmony pattern. The stuff you read in music theory is useless, so I'll tell you the real deal.
>
> To do this, you need a notation. I will name the 12 tones with single capital letter names:
>
> CJDKEFLGHAIB
>
> The letters A-G are the standard English note names: C = do D= re and so on, while the letters HIJKL have been added, keeping the name "H" for the semitone between G and A, as was traditional in Bach's time. This way you don't distinguish the tones from one another by calling some of them "sharps" or "flats", they are all symmetrical, as the are on the scale.
>
> You also need a notation for relative tones, to describe tones which are different from a given tone. I will use the following: abcdefghijk for +1-+11, l for +12 (octave), then la lb lc for the next octave up. To go down yxwvutsrqpon represent -1 to -11. m is -12, while z is zero, or just another name for the tone itself.
>
> These little letters are relative to the big letter, so that

Cdg

is a major triad in C

Ddg

is a major triad in D, and so on. To indicate that tones are simultaneous I use <> brackets to enclose the simultaneous tones (written vertically when not in ascii).

<Ddg>

is a major D chord

<Dcg>

is a minor C chord.

To indicate rhythm, I use ordinary parentheses to divide time, and + to indicate when the tone extends over the next slot. So that

CCGGAAG. | FFEEDDC. | GGFFEED. | GGFFEED . | CCGGEEG. | FFEEDDC.|

Is "Twinkle Twinkle Little Star". I hate the standard musical heiroglyphs. I usually write it this way:

[C][!!!!!!!!] zzggiig. | eeddbbz. | ggeeddb. | ggeeddb. | zzggiig. | eeddbbz. |

The brackets means don't make a sound, just tones are relative to the key (C in this case), and the whole thing is in eight quarter notes between the bars, and the bars are nothing, just like the divisions of measures in a staff.

I use scoping rules, so that <Ddg> allows you to play a D major chord without changing what "z" means in the outer context. <D> for a D note without changing the conext (the stuff inside the brackets is a different local scope for the relative tones). If you want to have a melody externally and a D major chord specified absolutely, you say <<Ddg>d> where the innermost d is relative to D, so it's actually L, the inner g is relative to D, so it's A, while the d played simultaneously is relative to C (or whatever the outer context is). You need some links to match the "+" to what it is extending, but it works ok, you can label the + using a previous tone too, and I use - for a glide. It's just a replacement for the silly ancient notation people use, because I don't want to buy special paper, or use special software. Ascii is fine to express the ideas.

The basic unit of time for a letter is a quarter notes (or whatever you decide the unit of time is). Then, to divide time, use parentheses for integer division of time, and + signs to indicate the extension of the tone into adjacent spaces:

[!!!!] C (CD) (+ F) (C .)

or in relative way:

[C][!!!!] z (zb) (+g)(z.)

The . means a rest. It's ASCII friendly music notation.

The frequencies go up by a factor of twelfth-root of two every step. There are two critical tones relative to any tone: e and g. The g is the so-called fifth, and the e is the fourth. The frequency ratio lz/e is equal to g/z, that's what makes them special.

These have overlapping harmonics, because they are close to integer ratio of 3/2 in frequency. The major third is "d", and sort of has overlapping harmonics, while the minor third is "c", which is basically just neutral with "z". The overlapping harmonics thing is overblown, it basically only is needed to understand e and g.

The remaining tones are basically all equal, except for a,f, and k. These are the most dissonant tones.

To make scales, you just avoid the three dissonant tones, a,f,k. So that a major scale is

zbdegi and k

k is special, because it is dissonant, and used to produce the "tension before release" in European classical music. You can use a and f for a similar effect, but a "goes the other way" (meaning the melodies are going down).

The minor scale is the tones

zcdeghj

the "real scale", meaning what people understand when they are playing an instrument in a key is the union of the major and minor omitting a,f,k.

z bcd e g hij

z e g are the main anchors, the bcd are the three non-dissonant tones between z and e, while hij are the three nondissonant tones between z and mg.

All these 9 tones, the scale minus a,f,k, are all not dissonant with z, and allow our ear to keep imagining z is droning on underneath. When a,f,k appear, to a lesser extent d and h, there is a certain amount of dissonance, and these are often used as a transition tone to z (the tonic), the disrupt the imagined drone with dissonance.

There are symmetric scales:

z bc e g ij

This is one of the "modes" (I don't like the way these are explained--- basically, take z eg and fill in a random selection of bcd,hij, and you get a reasonable scale, the modes are just cyclic rotations of zbdegik to make the tonic elsewhere)

Which is nice, because the intervals going up can be turned into intervals going down from the same tonic. The major scale is symmetric with respect to an inversion of order around the tone "d". This allows you to write a melody, and then reverse it and play the reverse intervals relative to d, without leaving the major scale for z.

There are the equal-interval scales, like the 6 tone scale:

z b d f h j

Which is a completely equal inteval thing which is rather dissonant, but it's like the Simpson's theme, very avant garde sounding.

There is also the circle of fifths, which can be understood best this way:

0:[z] 1:[eg] 2:[bj] 3:[ci] 4:[dh] 5:[ka] 6:[f]

This is arranged in terms of distance from the tonic, in steps of fifths or fourths. The closest to z are e and g, and then the e and g of e and g, which are b and j, then c and i, then d and h, then k and a, finally f. These represent how far away from the tonic you are harmonically.

To arrange it in a circle, do this: fahcjezgbidkf. The middle is z, and going out you get more distant fifth-steps (going right) or fourth steps (going left)

To make musical tone patterns, it's not mathematical thinking, rather it's understanding the aesthetic relationship of the tones to one another, and being able to transition from one key to another, reverse a melody, make harmonies between simultaneous melodies, and divide time in strange ways, with polyrhythms. These are the simple skills you learn in composition school. To make music using these skills is an art like abstract painting, the mathematics ends with the simple observations above.

71. - How did Einstein's musical practice inform his scientific work?

Question: How did Einstein's musical practice inform his scientific work?

There is no connection, except for the fact that playing music requires you to invent and interpret structures constantly, and keeps your brain limber. Heisenberg was a nearly concert-level pianist, and there are many other physicists who use music to relax. Composing is very mentally demanding, and resembles the creative process in physics, except the constraints are totally different, coming from human psychology and not

matching nature.

Book Recommendations:

72. – What is the best way for a physicist to learn biology?

Question: What is the best way for a physicist to learn biology?

A physicist only needs a tiny amount of background to read the literature immediately, the basic principles can be learned from any elementary biochemistry book, and Watson's is a good one. Then read the literature! It's not so hard.

It helps to have a point of view coming in. My own point of view was to elucidate the information flows in the biomolecules, because I was interested in the computation in the molecules. This turned out to be fruitful enough for a lifetime, so I didn't look for anything else.

73. – What are Ron Maimon's favorite books?

Question: What are Ron Maimon's favorite books?

Standard ones: Polyakov's "Gauge Fields and Strings", Milgram's "Obedience to Authority", 't Hooft "Under the Spell of the Gauge Principle" (reprint collection), Polchinski's "String Theory", Green Schwarz Witten "String Theory", Mandelstam and Yourgrou "Variational Principles", Cohen's "Set Theory and the Continuum Hypothesis", Sade's "The 120 Days of Sodom", "The Misfortunes of Virtue", "Eugenie de Franval and Collected Stories", Lang's "Calculus" and "Algebra", Connes "Noncommutative Geometry", Mandelbrot's books, Fadeev's book on path integrals, Parisi's "Statistical Field Theory", Anderson's "A Career In Theoretical Physics" (collected papers), Linnik "The Dispersion Method in Binary Additive Problems", Bogoliubov (the younger) and co "Inverse Scattering Method and Correlation Functions". Wheeler, deWitt, Everett "The Many Worlds Interpretation of Quantum Mechanics" (Everett's thesis), "Metamagical Themas" (Douglas Hofstadter), Marx "The Communist Manifesto", "Capital".

There are great collected papers collections for Landau, Godel, a bunch of others you can find on the shelf at libraries. I honestly don't remember all the stuff, and I am probably leaving out great stuff. I liked turbulence books, but I can't remember any specific titles now. Kraichnan has a bunch, and he is great.

The best stuff is primary scientific literature. The indispensible journals from before the internet were Nuclear Physics B, Physical Review, Nuovo Cimento, JETP Letters (Soviet era), Reviews of Modern Physics. There are great preprint collections by Dyson on SU(6), on Conformal Field Theory from the 1980s, on string theory.

I liked a bunch of novels, 1984 was probably the most important to me growing up, I had a major crush on Julia. I liked Bukowski's "Pulp" tremedously, it was a hipster book in the 1990s, and it was excellent. I read "Ulysses" when it was declared the best novel of the 20th century, I liked it. I only got halfway through Finnegan's Wake, I am not as erudite as Gell-Mann. I liked plays "Beckett's "Endgame", "Animal Farm" and "Down and Out in Paris and London" were great too. Lessing "Golden Notebook" when she won the Nobel prize, I don't remember others.

But my favorite thing was science fiction, because this is stuff you need to make up from scratch, including setting. Here, I liked Judith Merrill's science fiction collections (and her short story --- "That Only A Mother"), Isaac Asimov for sure (his short stories mostly, I read Foundation though and liked it), Doris Lessing "The Making of the Representative for Planet 8", and Kurt Vonnegut's "The Sirens of Titan", and "Cat's Cradle". These are standard classics, I don't like to talk about stuff that's already famous.

Nonstandard books: Dyson's "Origins of Life", Gold's "The Deep Hot Biosphere", Stuart Kauffman's "Origins of Order", Stephen Wolfram's "Cellular Automata and Complexity" (also A New Kind of Science), Mizuno's "Element Transmutations: The Reality of Cold Fusion" (the literature here is A library of papers about cold fusion).

These are most important, because they are marginalized, yet they have correct insights inside. This is where I think you can make progress.

There are books I liked when I was a kid that I am not sure I would like today,
But there are books I didn't read when I was young, that I skimmed as an adult, and thought "Why didn't I ever go through this? it would have saved so much time!" Landau and Lifschitz series especially.

74. - What classics Ron Maimon recommends reading?

Question: What classics Ron Maimon recommends reading?

I don't recomment reading things that are too old without translating them to modern language. I read all of Archimedes works a long time ago, but I didn't actually read them, I just looked at the theorems and tried to figure out what the heck he was doing from modern perspective (the methods were too antiquated to follow). I did the same with Euclid, and tried with Appolonius, but I got bored.

The same method with the Principia is good, but harder to do, because you need to rederive the cycloid properties. I put some things online to help. But Archimedes and Newton you should know, although slogging through the ancient stuff is a drag. Anything from before 1900 is a drag, people a long time ago were just kinda stupid.

There is a summary of some classics in Barbour's "Absolute or Relative Motion". It's a book about history, but you see the history of Mechanics, especially stuff about Huygens, something about Leibnitz, the stuff usually ignored by modern writers.

But some things, old or not, you can just read straight through:

Galileo: Diaolog and Two New Sciences.
Euler: Analysis Infinatorum
Maxwell: The original article on the Maxwell Boltzmann distribution is good, the electrodynamics work needs to be translated to modern notation, like Newton. I only skimmed this.
Einstein: All the scientific work, every paper, every book.
Dirac: The Principles of Quanutm Mechanics, the original articles are good too.
Bohr: The classic papers, you can skip BKS theory, read Heisenberg instead.
Heisenberg: All of hte papers (I read some of them).
Landau: Collected papers, they are great. Landau/Lifschitz series is a comprehensive education in pre 1960 physics.
Fermi: His little book on thermodynamics is great. I haven't read any of his papers.
Bethe: The Bethe Ansatz paper is good, it is covered again in a Russian book on the Inverse Scattering Method and Correlation Functions.
Feynman: All the papers (there aren't so many), all the books--- they are excellent. Quantum Mechanics and Path Integrals is one I didn't like as much, the 1948 Reviews of Modern Physics article was in less dated notation and clearer in my opinon.
Schwinger: "Quantum Electrodynamics, a Reprint Collection" is great, there are classic papers there, although it is difficult to read because it is so formal. A good modern introduction makes this stuff accessible.
Pauling: His quantum mechanics books is a classic, this is the chemists' orbitals.
Dyson: He wrote a great review of the SU(6) papers in a preprint collection called "Higher symmetries" or something like that. I liked his papers a lot, although I haven't read all of them.

Dyson is alive, from here on out, the classics are sort of recent. But I don't think there is much before 1900 that is really super-worthwhile.

I like Adler, Anderson, Brout, Callan, Coleman, DeWitt, Frohlich, Glashow, Gell-Mann, Gross, Lee, Nambu, Parisi, Skyrme, Veltman, 'tHooft, Weinberg, Yang, Bardeen, Von-Neumann, Scherk, Schwarz, Yoneya, Polyakov, Van-Niewenhuisen, Witten, Polchinsky, Mandelstam, Wilson, blah blah, all of these are famous names. I copied from Wikipedia's "Quantum Field Theory" template, which I wrote, but it's been edited by politics since. All these people wrote classics, and everyone should read everything they wrote (but I didn't read everything, sadly).

I am overlapping with people active today now, so I'll stop. Most of this stuff I read decades ago, I don't remember all of it, I am going by what stuck in my mind. I forgot Mandelbrot's "Fractal Geometry of Nature", for example, that was something I liked, and a bunch of historical math books I don't remember. And Hawking! All his articles, all his books. And Penrose too, I can't do this, too many people I'll forget.

75. – What are examples of great scientific books that introduced the public to revolutionary concepts?

Question: What are examples of great scientific books that introduced the public to revolutionary concepts?

For books aimed at general audiences, recently, there was Wolframs "A New Kind of Science", which despite the overselling and overreaching applications, did introduce the public to Wolfram's automata and the origin of complexity. The results popularized in this book are foundational to modern biology, as I see it.

In the 1990s, Thomas Gold's "The Deep Hot Biosphere", introduced the public to the Soviet abiogenic petroleum geology. In a completely different field, Mizuno's "Nuclear Transformations: The Reality of Cold Fusion" introduced important experimental results which could not be published in ordinary venues.

In the 1970s, Mandelbrot's "The Fractal Geometry of Nature" introduced the concepts of Levy flights, recursive fractals, elementary renormalization, and Haussdorf measure, and irregular geometry in general.

But my favorite one, for sheer originality and profundity, is from the 1960s: Milgram's "Obedience to Authority" which introduced the public to a scientific result which didn't even have a field at the time of publication, it was something like experimental sociology, it still doesn't have a name. This experiment was so successful and illuminating, it was banned. It probably should have instead been made compulsory.

76. – I study physics and mathematics with full interest but it lacks when it comes to chemistry. Does anyone who have a remedy for this?

Question: I study physics and mathematics with full interest but it lacks when it comes to chemistry. Does anyone who have a remedy for this?

Read Pauling, and the interest will come. Also Crick, and the other structural chemists of

the 1950s. The structure of biomolecules, DNA and certain proteins, is a terrific motivator, and it requires some deep chemistry to understand fully. A great popular account appears in "The Eighth Day of Creation".

After that, organic synthesis also becomes interesting, and inorganic chemistry can be used as a probe into biomolecules, so some inorganic chemistry becomes exciting. For a more recent motivator, there is density functional codes, like the Car Parinello method, which allow you to see the electronic structure of biomolecules on a computer in real time.

77. - What is the general feeling about Douglas Hofstadter's line of thought?

Question: What is the general feeling about Douglas Hofstadter's line of thought?

I read Godel Escher Bach as a teenager, I don't think this work is his best, because while it is fine as literature, the main theorem is mystified rather than explained, and the main thesis is murky and isn't so deep when unraveled. It is mostly an artier version of Nagel's "Godel's Proof", and it doesn't make the theorem as transparent as it should be.

To prove Godel's theorem takes about 20 seconds once you know what a computer is: given an axiomatic system S, write a program which

1. Prints its code into a variable R
2. Deduces all consequences of S
3. If it finds the theorem "R does not halt", it halts.

Each of these can be done by a computer program, each is a simple exercise. GEB obfuscates the issue by using the traditional approach of Godel numbering, and formal logic, and by avoiding computers. But the "Typographic number theory" with it's pushes and pops, is a cute way of doing logic, and the various puzzles make the implications of Godel's theorem more intuitive, and it doesn't come to wrong conclusions, so I can't complain. I liked it when I read it.

The main thesis in the book is that the self-awareness of human consciousness somehow acts as a form of reflection, like an axiomatic system becoming stronger by adding "this system is consistent" again and again (I am paraphrasing). This idea is interesting, because it is a semi-algorithmic view of the brain. Adding layers of self-reflection to a system adds strength, and the human mind, Hofstadter is theorizing, is doing this type of self-reflection constantly, in the way it examines its own thinking. It is an interesting idea, and it is possibly true, but it requires a way of making the mind capable of somehow increasing the power of a fixed formal system. Perhaps this is possible because the brain has a randomness source, since computation with a random number generator is non-algorithmic in the strict Turing sense, and it might be able to evolve systems that reflect higher and higher, by evolving stronger systems.

The main idea is illustrated using self-referencing artists and self-referencing constructions. It's a new philosophical idea, and it has some merit, but it is only nebulously stated in the book, in an arty way. Hofstadter made it clear that he meant that the self-reference in Godel is analogous to the self-reflection that humans do, and to the formal reflections in both works of art and in axiomatic systems which are more powerful. But his training in logic is limited, so the language is vaguer than it could be (but he was also writing 30 years ago).

The really interesting new stuff for me is in Metamagical Themas and subsequent work. In the last chapters of this Metamagical Themas, Hofstadter isolated and defined the completely original concept of superrational decision making. This is a great original contribution to mathematical philosophy, full stop.

Superrationality is the first mathematically precise statement of the golden rule, and it is, when appropriately extended, the solution to the problem of cooperation in one-shot prisoner dilemma type games, a cooperation which is observed in human behavior, Nash be damned. It is, when extended properly, formulating monotheistic religion independently of intuition, emotion, or sacred texts. This I think is extremely important, as important as Kant and Kierkegaard and Plato, at least once you understand the further implications, because it allows you to define a notion of God independent of religious revelation or the specific human condition.

The extended notion comes when you try to make a consistent superrational system which extends to arbitrary games. To do this, you need to assume that the superrational strategy exists and is self-consistent over all situations in all games, associating a utility function to each play in every concievable game. The result is tantamount to an infinitely wise agent watching over everybody, knowing the personal circumstances of all play, and making decisions about what is best for the infinite future, the resulting agent so constructed might as well be called God. The intuition is clearly identical with that of religious believers in the monotheistic tradition past the 1st century AD. This is how I understood what all those religious people were talking about, and since the idea is mathematically precise, you can't argue that it is nonsense, and you can see it has nothing to do with magic, or creating the universe, and it makes sense in positivism.

The rest of his work is in making models of analogy-making in human cognition, a project which he explains in "Fluid Concepts and Creative Analogies" and in his latest book. The Fluid Concepts era is more traditional AI, trying to write code to make simple analogies, like copycat and seek-whence. These programs make analogies in limited domains, and with limited success.

In analyzing these programs, and making comparison to human cognition, Hofstandter began to bite the bullet in true AI, making a machine that thinks like a human. The main barrier, as he explains, is making these intuitive analogical leaps that brains do so effortlessly. The computational models here are terribly inadequate, as he shows, because each of the analogical paths are in very fuzzy loose ways that are not well captured by a set of formal sentences. In this, he is trying to break free of the simple models of formal

language that have been used with no real success to model human cognition in the 1960s, and which gave logical positivism a bad name.

The logical models are broken because the cognitive categories are far more complex than the boolean models which assign truth to a hard logical sentence. When you say "I applied to Harvard, I got an offer from Purdue, but Purdue has a much nicer department anyway." and someone says "Sour grapes." (Hofstadter's example), you are making a very sophisticated high level analogy, where many different pieces are set up in a very complex analogy in the two situations. Hoftstadter doesn't deny that ultimately this analogy is done by a form of computation in the brain, but he denies that it is fruitful to model it as a formal analogy between lines and arrows in a formal model, rather one has to have a very loose categorization scheme which allows for vague linkages of various strengths, with a model of computation which is very different from the classical one.

These insights are born of hard struggles with the problems of true AI, facing failure with intellectual honesty, and struggling with the understanding that the real processes of cognition are vastly more computationally deep than what has been acknowledged so far within neuroscience or computer science models of cognition. This insight was important to me personally, and it should serve as an example of what to do when you DON'T have all the answers, you go forward as best you can, without losing sight of the goal, and without pretending you understand things more clearly than you do. Hofstadter writes honest things in a field which is usually dominated by politics, and for this alone, I think it is worthwhile to read everything he writes.

78. – Which technology startups have the best Wikipedia entries?

Question: Which technology startups have the best Wikipedia entries?

This has to be Pacific Biosciences. They had one of the major technological breakthroughs of sequencing, and their machines, while not yet competitive in precision, are superior in many respects to existing technology, and are unfairly politically maligned, it seems so as to drive them out of business.

The big genius here is Steven Turner, who I met at Cornell. His idea to turn zero-mode waveguide into a new technology was brilliant and inspiring, and I hope these guys stay in business long enough to realize their potential, which is really to revolutionize sequencing. They are able to get the pure nucleotide content of a single strand of DNA. The same method can be used to see ribosome action, all sorts of things, at the single-molecule level.

79. - Which 10 books would you recommend that your children (or anyone) read throughout their life?

Question: Which 10 books would you recommend that your children (or anyone) read throughout their life?

I ordered it randomly, because rank-order doesn't make sense for this stuff.

1. Paul Cohen "Set Theory and the Continuum Hypothesis"
2. Alexander Polyakov "Gauge Fields and Strings"
3. Benoit Mandelbrot "The Fractal Geometry of Nature"
4. Brian Kernighan and Dennis Ritchie "The C Programming Language"
5. Freeman Dyson "Origins of Life"
6. Thomas Gold "The Deep Hot Biosphere"
7. Stanley Milgram "Obedience to Authority: an Experimental View"
8. Douglas Hofstadter "Metamagical Themas",
9. Marquis de Sade "The 120 Days of Sodom"
10. Charles Bukowski "Pulp"

80.- Are people from India good at Mathematics?

Question: Are people from India good at Mathematics?

Creative mathematics comes from God, and by God I mean the supremum of large countable computable ordinals. You can't appreciate these ordinal structures without a culture that transmits infinitary thinking to you, so to produce creative mathematics, you need to have a theology, and you absolutely need someone to introduce you to infinity.

The reason India produced mathematicians is largely because India has an original locally developed theology, Hinduism, and an independently created mathematical tradition of infinity, the Kerala school, dating back to the 14th century, which developed results on infinite series well before Europeans did. They produced their own theology of infinity, roughly analogous to the infinite theology that developed in Europe in the 17th century, this is the theology of calculus.

The modern theology of infinity is Cantor's ordinals, which is superior in every way to the previous calculus theology, but it is more abstract. Cantor knew he was doing theology--- he explicitly recognized the concept of ordinals as corresponding to the religious concept of God. He was also criticized by theologens for introducing countable and uncountable infinity, which seemed to contradict the unity of God (two different kinds of infinity, at least). But Cantor's ordinals are the essence of mathematical theology, and they contain the fundamental logical thing that makes infinity important, and it became clear relatively quickly. It is senseless to cling to the old notion of infinity when there's a better one, so

mathematicians threw out the calculus theology in the late 19th century.

But the infinite series thinking includes a lot of the ordinal thinking inside it, except disguised a little. You can think of a convergent infinite sum of positive terms as generating a sequence of points, the partial sums, and the various ordinals as corresponding to various orders of points on a line. So you can rearrange the sum to be a sum of sums, and ordinals just abstract out the logical thing in this scheme, the rearrangement of infinite sums of sums, which is the thing that produces new math.

Ramanujan achieved his original results entirely within the pre-ordinal theology. To do this, he first learning some 18th and 19th century European mathematics, and he incorporated the local Kerala style calculus-type infinite series manipulations which he picked up by osmosis from the surrounding culture. He then produced a bunch of new results from his own tremendous creative effort. So he did a lot of infinite series manipulations, and infinite continued fractions, stuff that produces new mathematics, but which was out of fasion in Europe in the early 20th century, because ordinals were introduced and people were going in another direction, towards logic and computers.

Modern mathematicians call unjustified manipulations of infinite series "analytic function theory", and this justifies the manipulations within a rigorous framework, and Ramanujan was quickly brought up to to speed on analytic functions by Hardy. The other thing Ramanujan liked, continued fractions, is now called "SL(2,Z)" and "modular forms", and this just took over mathematics in the late 20th century, making Ramanujan more relevant than ever. So when Ramanujan died, Hardy thought he was a great talent, but out of touch with modern mathematics, meaning Ramanujan's theology was old-fasioned infinite series and continued fractions, not ordinals. But now people think of him as at least a century ahead of his time, because the modular ideas became so important for proving Fermat's last theorem.

People do not appreciate what it means to grow up in a culture with an embedded mathematical tradition, the theology is essential. As a child, one is given puzzles and games, one is encouraged to think about these puzzles, and one is exposed to various formal manipulations and philosophical conjectures regarding infinity which are just not presented in books. In American culture, you just don't get exposed to this stuff, and if you try to think about it, you are heckled.

But American culture encourages people to go off on their own, so there are individuals who pick up a lot of mathematics anyway, despite the culture. Often they immerse themselves in a foreign mathematical culture which has an independent history of mathematics. Since new mathematics is a product of individuals or very small teams of individuals, culture becomes irrelevant now. We have an internet, there's no need to squirrel knowledge away.

Disqus Comments

- Who is Lucy the Australopithecus and why was Obama
  allowed to touch that fossil?

Ron Maimon→AlchemistRienhart

6 years ago

Darwin made no claim about the mutations being "genetic damage". This hypothesis is the result of the fruit-fly experiments in radiation from the 1950s, where it was discovered that mutation rate is proportional to chromosome damage. But the mutations themselves, when they aren't protein mutations are NOT random, they are controlled by RNA networks, and the result is much closer to intelligent design than to modern synthesis, because the mechanism of rewriting DNA is authorly, proceeding by inserting functional chunks into the genome, not by any random process. The random hypothesis is simply impossible, the IDers are right about this.

Ron Maimon→leopardo71

6 years ago

Mutations are not "completely random", they are very purposeful, biologists have been misled by looking at protein coding regions, where incidental rare mutations (like sickle cell) occur. Most mutations are in non-coding regions, and very tightly regulated by RNA networks. That's not accepted yet, but it explains the main source of friction between biologists and anti-evolutionists. It's a form of intelligent design, but the designer isn't God. It's RNA.

Ron Maimon→Edward Clayton

6 years ago

The question is how the crossing over happens in those individuals when they try to reproduce. It requires new biology to line up 2 chromosomes with one. I am not denying that it happens, I am pointing out that it requires new biology to make it work, because you need a way of organzing the chromosomes in partnered pairs along a split/merge and then decide how to produce the offspring crossing over. This requires a brainier RNA network in egg cells than what is normally assumed by biologists, the RNA has to line up corresponding chromosomes with each other correctly, and this is not a trivial feat for the network. I agree with you on the facts, although I am not sure about the viability statistics of the offspring of chromosome-merged and chromosome normal humans.

Ron Maimon→Edward Clayton

6 years ago

No, I just didn't know about the "vestigial telomeres", your comment is the first time I heard of it. But I should look at the sequences myself, because there is a lot of "telomere-like" sequence in the middle of chromosomes, there are lots of embedded repeats, and it is possible that the split occured in the telomere-like region instead. I don't know if anyone else proposed this, but it must be investigated without bias. If the telomere regions are more likely from a merge, I have to change the assessment of course.

The common-ancestry is not the issue with the chromosome merge, but the mechanism of merger is completely unknown. To merge chromosomes, you need a mechanism that can do it, and the presently acknowledged mechanisms are too dumb to do it, as they assume chromosomes blindly line up. The real mechanism must involve non-coding RNA linking up the chromosomes in corresponding parts, and it needs to work with split and merged chromosomes, because we know this has happened in history. It's the strongest argument they have, because it really requires new mechanisms of biology, all the rest of the stuff they say is drivel.

Ron Maimon→thegoodstuff 24

6 years ago

Oh yeah, sure. You can check out reviews on "DNA editing" and "RNA editing". For ERVs, just google "HERV" (Human endo-retro-virus), this is 30% or so of the genome, maybe 40%, look it up, I forget. The interpretation is original to me, for all I know, but maybe not.

Ron Maimon→PammyPooh

6 years ago

Ha ha! Yes, my ideas are facts, in those cases where they are correct. The way to check is to compare to actual DNA and RNA sequences, to real data. You aren't looking for facts, you're looking for authority. Sorry, I won't bite. I have no authority, and I don't believe in authority. I didn't call you names.

Ron Maimon→thegoodstuff 24

6 years ago

Oh my, somebody has missed out on the revolution in biology in the last decade. DNA inherently mutates by duplicating/silencing, shortening and lengthening and editing non-coding sequences, and back-writing entirely novel RNA from RNA networks using actual reverse-transcriptase, through mechanisms like reverse-transcribing ERVs (endo-retro-viruses--- they're 40% of the human genome). These modifications produce very intelligent mutations.

Ron Maimon→Dan Cameron

6 years ago

There is a real issue with Darwinism, which Dawkins doesn't understand, that it requires a sort of "intelligence" to produce new mutations. This was understood by Darwin, but Dawkins denies it. The intelligence involved here is not infinite, it's the finite but relatively large intelligence of computing RNA networks, they are about as powerful as a modern laptop.

Some people who "deny evolution" are just denying modern synthesis evolution, not the process itself. These critics are not just reasonable, they are right. The religious critics derive their science often by misinterpreting the sensible critics.

Ron Maimon→James F. McGrath

6 years ago

I should say, Michael Behe also has a case to make, but he is not identifying God in the cell, he is seeing the intelligence of RNA networks. That's more brains than modern synthesis, but infinitely less brains than a supernatural God.

You can quibble about the Maccabees, I am not sure, as Carrier is not sure, and after reading his tour-de-force I now trust him completely and slavishly in all his historical conclusions.

Ron Maimon→Edward Christophersen

6 years ago

That's not where mutations come from. Mutations are written into DNA by endo-retroviruses, among other mechanisms. These are the major source of evolution today, protein evolution stabilized 300,000,000 years ago, and is now mostly a clock for species divergence.

Ron Maimon→thegoodstuff 24

6 years ago

This is false, evolution doesn't happen because of external factors only. It is driven largely through production of new traits in the DNA itself, without any external changes at all.

Ron Maimon→Dan Cameron

6 years ago

Do you really think that the population of 10,000 that were this common ancestor left bones around we can find? The chances are 0%, we'll never find fossils of them, we aren't that lucky, and even if we did, we wouldn't know it for what it was. The inference of a common ancestor is obvious by looking at DNA sequences (as well as homologous characteristics, other fossils, and so on). It's not a bluff, it's an obvious inference if you look at protein sequences and time their divergence. Stop being retarded.

Ron Maimon→PammyPooh

6 years ago

The RH negative blood line was common in Chimps until 10,000,000 years ago, when a small population of entirely RH positive chimps killed their neighbors, including the last RH negative survivor. It was a tragic day. Come on, stop being stupid, this is not a mystery if you know how DNA evolves. A better question you should have asked is how the Human Chromosome split into two in Chimps--- this is a bit stronger mystery.

- The Turing Lecture

Ron Maimon

9 years ago

The Turing lecture presents a very important idea: that it is impossible to build a convincing sophisticated computational circuit using point mutations in any reasonable time scale. This point was made intuitively by Pauli in the 1950s, when it was a response to modern synthesis evolution. It has been made by nonscientists as well, who find it hard to buy the idea of evolution happening by random point mutation.

The idea that there is a learning mechanism for modifying the genome is obviously true, and it is fantastic that someone with some clout has finally taken this position. But the learning hypothesis is not very persuasive because it is missing the main point. The memory capacity of the proteins is just too small to account for the computational complexity of the DNA rewriting, the number of protein combinations is generally vastly smaller than the number of DNA combinations. This paradox was very clear to those who studied protein networks, it makes it difficult to imagine any model of evolution which operates on DNA using proteins alone. To modify gigabytes of functional data, You need a molecule which can encode gigabytes of random-access data in a read/write way, in a dense encoding. The only real candidate, excluding DNA methylation and DNA conformation (which are recently emphasized hacks with similar function) is RNA.

The lack of direct communication or compatibility between the protein and DNA level information demands that the information about future mutations must be stored in self-modifying computing strands of RNA. This is a firm prediction, it must be so, and yet it is not accepted fully within biology.

This prediction allows one to predict that the DNA in the cell is mostly noncoding, since the coding region is controlled and evolved by the noncoding parts. It requires that the noncoding part is transcribed into RNA despite being noncoding. This is also now known to be true. It provides the only convincing role for the massive amount of long noncoding RNA in the nucleus. This RNA is computing, making a nervous system for the cell, and this computation provides the only plausible mechanism for machine-learning in the genome.

The computation in RNA requires that the RNA can distinguish different strands from one another, and this is done through complementary binding, and splicing/resplicing events in RNA. The hypothesis is that these events are making closed-loop computation with each other, without any need for translation to proteins, using only proteins available in the nucleus.

The RNA/RNA events and RNA reverse transcription (which is required for the evolution to be influenced by the RNA) are the "strange events" which are described in the video, but they are not all that random, they are constrained by the allowed mutagenesis in the RNA computing system in the nucleus.

The communication between RNA and DNA predicts that there are reverse transcriptase genes in the genome (this is also true, in ERVs), and that these are transcribed and active in certain cases. This has been observed in human tissue in cancer cells, where full functional HERVs are produced, including the polymerase, which allows for reverse transcription. The reverse transcription functionality of the human genome is important, since it allows you to couple the RNA computation back into the DNA.

The contains complementary matching motifs which allow self binding (true and surprising) and that that it is composed of interpretable domains with features which are not random. The "evolvability" condition then presented here is much too strict, the evolution can include rewriting of the code which is sensible, directed by the RNA networks in an egg.

This completely resolves the paradox of evolution presented in the video, but it introduces vastly more computation into the cell than is known at present. It demands that the RNA computation is sufficiently complex to essentially have a model of the protein production and the noncoding RNA production in the cell, so that it can sensibly modify the DNA for future generations.

These predictions are biologically surprising, and they are the only plausible way to allow for evolution to produce observed biological complexity, yet are not accepted fully.