“It’s life, Jim, but not as we know it.”

Mike Arnautov, 2020

What is Life? It depends on whom you ask. Having worked with biologists, geneticists and bioinformaticians (roughly, scientists who study biological data), I know what they would say: they would shrug and claim the answer to be well known, pointing one at some fat volumes of biochemistry. Some would talk about cellular respiratory cycle, and some would point out that we can actually assemble a bacterium from scratch (sort-of[1]).

On the other hand, the proverbial Man on the Clapham Omnibus would probably say that life is something one has obviously somehow mislaid: “Get a life, man!” may be his considered advice, “We know life when we see it, that’s all there is to it!”

But this is a philosophical gathering, so before even attempting to answer the question, we should look at the question itself: do we understand what it means and hence what sort of an answer can be given to it?

Firstly, we ought to be sure that “life” is a term which actually has some meaning. I think most people would agree that it does and I do not propose to explore the contrary view. So the question I am tackling is really: “there is something called life, what is it?”. But that still leaves us with two distinct ways in which we can consider the matter.

We can take the realist view: here’s life as an ontic something, let’s see how it is to be defined. In this interpretation one considers the distinction between ontic categories “life” and “non-life” and attempts to draw the boundary between the two.

Various definitions of such a boundary can be (and have been) proposed, fuelling arguments on topics such as “Are viruses alive?”. I confess I find this approach unsatisfactory. Are viruses alive? Who cares! What changes if we answer the question one way or another? It is like accepting that there is a distinction between a hill and a valley and then in all seriousness attempting to pin down the exact point where the valley ends and the hill begins. Yet nothing at all hangs on the answer. Neither of the two categories (hill/valley) depend on it.

Alternatively, we can take the conceptualist view, rephrasing the question thusly: “we have this concept called life, which undoubtedly captures something pragmatically meaningful about how the world is; what is the shape of that concept and where would we put its boundaries?” In other words, what is it that makes/permits Mr Spock to utter his immortal words: “It’s life Jim, but not as we know it.”?

This is the approach I prefer, because I take the neo-Kantian view of the world. To paraphrase Niels Bohr, we have no way of knowing what and how the world is, but only how it behaves. Looking at the world, I can certainly see that butterflies are very different from pebbles, so I can reasonably posit that there is some qualitative difference involved. We call that difference “life”. In this understanding “life” undoubtedly really exists (to some value of “really”), and yet “life” is a human concept and as such its shape and boundaries are for us to draw and, if needs be, re-draw.

---

1 Craig Venter’s achievement was to synthesise the nucleus of a bacterium, but not its outer envelope – that was taken from a natural bacterium.

That being the case, we would be perfectly entitled to declare that "life" is life as we know it – self-replicating biology with DNA. Anything else, however life-like, could be declared to be just that: "life-like" rather than life, which would make Mr Spock guilty of a category error.

And yet, our intuition clearly disagrees. In discussions about so far hypothetical extra-terrestrial life, it is generally accepted that there are two distinct possibilities: either it will be a variant on Earthly, DNA-based life or – by far the more interesting case – it will be something else. In fact, would-be xenobiologists have been known to ponder the possibility of life based on silicon rather than carbon and/or on fluorine rather than oxygen[2]. So, clearly, we are prepared to accept that "life" is not limited to "life as we know it". The latter case being seen (rightly, in my view) as merely one data point in the conceptual space of possibilities, and as we all should know, one cannot generalise from a single data point.

So it would seem that we are quite prepared to allow Mr Spock to be right, without being able to say (at this stage, anyway) what it is that allows him to be right. I suggest, therefore, that in the conceptualist reading of our theme, the question becomes: how far could (or should!) we stretch the boundaries of the concept of life, consistent with our intuitive notion of life?

From this angle "life" is seen as a human concept, and thus any answer one can give to that question is necessarily a contingent one, shaped by the present moment – by one's knowledge and experience, embedded as they both are in one's wider cultural context. By the very act of interpreting the question that way, I implicitly accept that our conceptualisations (as reflected in language) do not necessarily amount to some objective representations of the world. Language is a tool we use to grapple with the world, it is not a representation of the world.

There are historical parallels of such pragmatic stretching of concepts. For example, "energy" used to mean just mechanical energy – kinetic and potential. Nowadays, though, we include thermal energy, binding energy (chemical or nuclear) and the whole notion of the entropic cascade of energy types. Similarly, the concept of a number, which started of as counting labels, got generalised to share ratios (fractions), measurements (real numbers), complex numbers (indispensable for physics since Maxwell's unification of electricity and magnetism) onto the abstraction of quaternions and octonions. Why shouldn't the concept of life be similarly expandable?

Let me reiterate, to avoid any misunderstandings... I am not proposing with nominalists that there is in reality no such thing as life. I accept that the concept of life is pragmatically useful because it does capture something real, but in doing so, I submit that the shape and the boundary of applicability of that concept is entirely down to us and is guided by two factors: our intuition on one hand and pragmatic uses of the term on the other -- with a degree of dynamic interplay between them.

This talk is about what my intuition has to say on the subject. As Americans say "your mileage may vary", but that's fine – it is only by comparing our intuitions and testing them against a variety of scenarios, real or hypothetical, that we can have some chance of answering the question posed by this weekend's theme.

So what were my immediate responses to that question? Well, firstly, I was almost sorry to see omitted from the list, so kindly supplied by Tim in his call for speakers, the following classic answer attributed to V.I.lenin: "Life is a form of existence of protein matter." There, problem solved. Please do pause with me do admire the vacuous pomposity of this pronouncement, so

2 For a far more extensive list of possibilities see Isaac Asimov's contribution *Not as We Know it: The Chemistry of Life*: http://www.bigear.org/CSMO/HTML/CS09/cs09p05.htm

characteristic of regimes in which the Big Boss must be, to use a recent phrase, a very stable genius. But I digress...

My initial, automatic response to the question was to invoke Darwinian evolution, with some additional constraints, but as I thought about it, doubts started seeping in. I certainly expect that if we ever decide that an ex-terrestrial something qualifies as “life”, it will be a product of such evolution. However, it’s hard to see such a call being made on the basis of a definition of life based on evolution. If anything, it’s the wrong way around: evolutionary background is more likely to be inferred, once the label “life” is agreed to be appropriate.

Plenty of SF literature[3] agrees: on meeting an alien creature protagonists infer the presence of the ecology responsible for its existence, rather than seeking to a proof of such an ecology first, before granting the title of “life” to the creature in question.

Imagine a reverse case: some extra-terrestrial intelligences encountering me and trying to decide whether I qualify as “life”. Am I capable of Darwinian evolution? Well, er… not by myself, no. Never mind evolving, I could not even survive entirely on my own, outside the biosphere, which had enabled my existence.

Clearly, we have to be careful in distinguishing “life” from “being alive”. The ecology as a whole is “life”, capable of Darwinian evolution, but a single creature is merely “alive”. The connection between the two concepts is clear enough: to be alive one has to be a part of life, hence the inference from a creature to the ecology from which it springs. Thus it is our conclusion (however motivated) that something is alive, which allows us to infer the existence of life.

So far so good, but here’s a problem. We are already capable of creating living systems which are not a product of Darwinian evolution. When we create a bacterium in which one “normal” pair of amino-acids is replaced by an artificial pair, which does not occur in nature, isn’t such a bacterium alive? After all, there is no Darwinian ecology underpinning its existence.

The obvious answer is that while there may not be such an ecology, there might have been one. Perhaps the whole of our biology could have been based on the anomalous set of amino-acids. But that’s a counter-factual argument and counter-factuals are a tricky ground. Undoubtedly such an anomalous ecology is conceivable, but let us not fall into the trap (*a la* “philosophical zombies”) of conflating what is conceivable with what is possible. We simply do not know enough to assert that such an ecology could evolve naturally. In fact, we cannot be even sure that it could be created artificially, even though to the best of our very limited knowledge there is no reason why it could not be.)

And then, of course, there are recent experiments with “biological robots” assemblages of living cells, put together artificially, in a way that makes no evolutionary sense, to perform particular tasks – so far very simple ones. Are we prepared to say that such assemblages of living cell are not alive? Or that they are, but do not represent “life”. It all gets rather muddy.

You may be tempted to say that it is *our* evolution that gives rise to any such artificial life, but that makes things worse. Such all-purpose explanations actually explain nothing. It really is no different from invoking God as an all-purpose explanation. I submit that the inevitable conclusion is simply this: Darwinian evolution is in practice useless as a criterion for deciding what is and what is not life.

---

3 These days SF stands for Speculative Fiction. The old term Science Fiction is now deprecated as manifestly inaccurate. And “sci-fi” abbreviation is now downright derogatory.

OK, time to re-think. Perhaps metabolism should be used as such a criterion? This seems promising. Roughly speaking, metabolism is a set of chemical reactions (or perhaps more generally, of processes), which allow organisms to sustain their existence in defiance of the Second Law of Thermodynamics[4].

Unfortunately, there is a problem with this approach too. If you look up a definition of metabolism, it will always include some reference to life, living beings, organisms or similar. E.g. Wikipedia defines it as "is the set of life-sustaining chemical reactions in organisms". The reason for this is simple: without such a qualification, exothermic reactions sustaining a flame would qualify as metabolism. A flame "eats" its fuel, eliminates the waste product (e.g. ash or CO_2) and in doing so sustains its continuing existence, in defiance of the 2nd Law. Are we prepared to accept fire as life? Obviously not!

So the metabolism criterion fails, because it is essentially circular. Can we salvage anything from it? Perhaps there is something special about biological metabolic processes which would do the trick?

You may recall that at the beginning of this talk, I noted that at least some biologists may refer one to the cellular respiratory cycle, which is common to all life forms (other than viruses, should they qualify as life). For me, the cycle is a thing of sheer beauty – a feat of quantum-molecular engineering to marvel at[5]. What it does is to shuffle protons (hydrogen nuclei) across biological membranes, thereby creating an energy store, from which cells can draw to perform their functions. It is the basis of all life on earth.

However, as argued above, we should not restrict "life" to Earthly life. In more abstract terms, what the molecular machinery of the respiratory cycle does is to transfer a proton across a membrane, effectively extracting energy from the environment, and then *returns to its original configuration*, ready for dealing with the next proton.

Stuart Kauffman uses a more abstract view of the cycle to define life[6]. Unfortunately, as you might guess by now, I can see a problem. That description matches what physicists call the Carnot Cycle, originally developed to describe how steam engines work. Kauffman is fully aware of this and in fact defines life as a system which performs at least one (Carnot) work cycle. As I understand him, in doing so he simply brackets out mechanical contraptions (e.g. steam engines) which satisfy the same criterion. That's not exactly useful for devising a philosophical definition of life.

Well, perhaps I am over-thinking the problem. Rather than attempting to look "under the hood" of the phenomenon of life, maybe it would be more useful to take note of the glaringly obvious distinction I already noted above: butterflies are very obviously different from pebbles in their behaviour. In fact, there is an old saying in biology: "Under carefully controlled laboratory conditions, a biological sample will do as it damn well pleases." Might this difference in behaviour be the key? Butterflies act, while pebbles are acted upon.

To make this stick, we need to be able to define "action" as opposed to "behaviour". Clearly enough, a thermostat does not "act", but these days we do have examples of artefacts which at least appear to be able to act. When a machine beats the world champion in a game of Go, it is hard to avoid the impression of action as it selects its moves. Admittedly, Alpha-Go did not physically place

4 As an aside, the Second Law is often misunderstood as necessitating a progression from order to disorder; however, entropy is not the same as apparent order – e.g. growth of orderly crystals out of disorderly liquid is very much in line with the Second Law.

5 If you are interested in the details, Nick Lane gives probably the best, highly approachable description of it in his book "The Vital Question".

6 Stuart Kauffman *Investigations*

counters on the Go board, but for roboticists that would have been a fairly easy and quite pointless elaboration and in any case, as philosophers we understand that choice itself is an act.

Watching Go expert commenting on the game at the time, it was striking that they kept slipping into expressions such as “it wants”, “it intends”, “it believes”, “it plans”, “it has decided” etc… They were very clearly strongly tempted to adopt what Dennett calls the intentional stance[7], implicitly assigning Brentano-style intentionality to the machine. Is it sensible to do so?

Philosophical opinions differ. For example, John Searle differentiates between original and derived intentionality[8] and argues that only the derived, inauthentic kind can be ascribed to AI contraptions. It is humans, not machines who display original intentionality and thus can be thought to act in the agency meaning of the word. Dennett disagrees and claims that there is no original intentionality. All intentionality is conferred by observers adopting the intentional stance, as being the most productive for understanding observed behaviour.

I happen to agree with Dennett, but for argument’s sake let’s assume that Searle is right. Can Mr Spock measure or deduce the kind of intentionality is being displayed by the behaviour of whatever it is he happens to observe? I cannot see how. If anybody has any ideas for constructing an intentionality-meter, please let me know. In the meantime, I have to accept that intentional action may or may not be a hallmark of life, but as a criterion for answering the question “is this life?” it is pretty useless.

It is striking that once again I get unstuck on the fact that a proposed criterion fails to differentiate between life and machines. Is this in fact significant? Can a mechanical contraption be alive (i.e. represent “life”, though not as we know it)? My intuition rebels against this notion and SF literature suggests that I am not alone. “Space opera” SF tales often posit two distinct types of intelligence in our galaxy: biological and mechanical While I may be unable to draw a boundary tightly enough to exclude all non-life, can I at least declare this to be a limit beyond which that boundary of the concept of “life” cannot be stretched? This may be a weaker result than the one I was seeking, but a partial result is still better than nothing!

I am inclined to stick with that, at least provisionally. True, some SF writers tell stories in which even that boundary becomes debatable[9]. And in any case how exactly does one draw a line between machines and non-machines? When does a cyborg, a mechanically augmented human, cease to be a living being? When does a biologically augmented machine cease to be a machine? What about a “bacterial robot” bio-engineered to deliver drugs into cells – is it a machine? It would seem that I have merely shifted the question instead of answering it. Maybe I should simply read less SF. Or maybe that proverbial Man on the Clapham Omnibus has it right after all: we know life when we see it. But how?

What we seem to have is an odd situation: while there is a number of plausible ways of differentiating life from non-life, on closer examination they fail to capture fully our concept of life. Combining some of them does not appear to solve the problem either, because their individual inadequacies do not seem to cancel each other.

Does this amount to a defeat? I do not think so. Put this way, the situation should sound familiar. It is no different from what Wittgenstein found in his “Philosophical Investigations” when attempting to define what we mean by “game”. His solution was to say that the concept of game was held

---

7 Dennett, D. 1987 *The Intentional Stance* MIT Press

8 Searle, John.1980 “Minds, Brains, and Programs.” *Behavioral and Brain Sciences*

9 For example, Stanslaw Lem, in his book *The Invincible*.

together not by any strict criteria, but by the notion of family resemblance. To quote Wikipedia on family resemblance[10]:

> "It argues that things which could be thought to be connected by one essential common feature may in fact be connected by a series of overlapping similarities, where no one feature is common to all of the things."

I submit that the same applies to our concept of life. If we ever do come across life as-we-do-not-know-it, we shall know it by the degree of family resemblance to life as-we-do-know-it. And in doing so, we may stretch our notion of life beyond its previous boundaries. For all I know, in the end our descendants may even stretch it to cover machines (however defined) – that will be up to them. Which is as it should be – language is mutable and evolves to reflect our needs, priorities and preoccupations.

So what is it that allows Mr Spock to say "It's life, Jim, but not as we know it!"? I am afraid you'll have to ask Mr Spock – his reasons may well be different in each individual case.

Thank you for listening.

- o O o -

10 *https://en.wikipedia.org/wiki/Family_resemblancelife*