

Why Life Can't Be Defined

An excerpt from: *The Quest for a Universal Theory of Life* (forthcoming)
by Carol Cleland

2.1 Overview

Faced with the question 'what is life?' many scientists and some philosophers advance definitions of life. But is a universal theory of life the sort of thing that could be captured in a definition? This chapter and the next explain why the answer is 'no'. There are significant logical differences between definitions and scientific theories, and these differences have consequences for the pursuit of universal biology.

Section 2.2 begins with a survey of the most influential contemporary definitions of life: thermodynamic, metabolic, evolutionary, and autopoietic. As will become apparent, all of these definitions, which come in a variety of incarnations of varying degrees of abstractness (including purely informational), face robust counterexamples and attempts to resolve them invariably generate new problems. It is thus hardly surprising that there is no consensus among scientists or philosophers about which definition of life is correct. ...

2.2 Popular definitions of life

The history of attempts to define life is very long. Some scholars believe that it dates back to the work of Aristotle. Scientific interest in defining life was rekindled in the latter part of the 20th century by astronomer Carl Sagan's now classic essay in *Encyclopedia Britannica* (1970). Writing during a period of rapid advances in molecular biology, computer technology, and space science, Sagan reviewed the then most widely accepted scientific definitions of life—physiological, metabolic, biochemical, thermodynamic, and Darwinian (which he somewhat misleadingly called "genetic")—and showed that each faces robust counterexamples, either including things that are not alive (e.g., candle flames and quartz crystals) or excluding things (e.g., dormant seeds and spores) that most biologists would be reluctant to classify as nonliving. Sagan's essay produced a flurry of new activity on definitions of life. Some of the definitions that he discussed (e.g., physiological) were abandoned. Others were reformulated in hopes of undermining the strongest counterexamples advanced against them. New definitions of life, e.g., autopoietic, were also introduced. The upshot is that we currently have an impressively large and diverse collection of definitions of life.

Yet despite these efforts, there remains no widely accepted definition of life. Although reworked and refined in light of contemporary scientific advances, the most popular contemporary definitions (thermodynamic, metabolic, evolutionary, and autopoietic) still face serious problems, ranging from robust counterexamples to being too *ad hoc*, i.e., designed expressly for the purpose of fortifying a favored view on the nature of life against what is *prima facie* a devastating counterexample. A brief survey of these definitions and their problems is illuminating.

2.2.1 *Thermodynamic definitions*

Thermodynamic definitions of life are popular among biophysicists. Commonly associated with the writings of physicist Erwin Schrödinger (1944), a basic thermodynamic definition might characterize a living system as one that creates and maintains local order (organismal bodies and ecosystems) by extracting energy (“drinking order”) from the environment over an extended period of time. As Sagan noted, however, many inanimate physical systems create order by extracting energy from their environment. Mineral crystals, such as quartz or diamond, provide good illustrations.

2.2.2 *Metabolic definitions*

Metabolic and thermodynamic definitions of life are closely related, the primary difference between them being that the former focus on the flow of material (vs. energy, interpreted statistically as local order) between a physical system and its environment. Sagan’s nuts and bolts version characterizes life as “an object with a definite boundary, continually exchanging some of its materials with its surroundings, but without altering its general properties at least over some period of time” (Sagan 1970). As Sagan observes, however, this definition faces compelling counterexamples. It includes nonliving phenomena such as candle flames, which have a definite shape and maintain themselves by continually exchanging material with their environment (extracting oxygen and combining it with wax, and releasing carbon dioxide and water in the process), and excludes viable seeds and spores, which may lie dormant for hundreds or even thousands of years before germinating.

More sophisticated chemical metabolic definitions have been developed in efforts to ward off counterexamples of these sorts. Some restrict the possible chemistries of metabolic processes to organic (carbon containing) molecules. Basing a definition of life too closely on the chemistry of familiar Earth life is problematic insofar as it excludes forms of life that are conceptually and perhaps even scientifically plausible. The Horta of Star Trek fame provides an especially colorful illustration. In the episode *The Devil in the Dark*, the Horta, an amoeba-like blob of rock, confounds Captain Kirk and his crew by burrowing through solid rock and behaving in a life-like manner. The Horta turns out to be a truly alien, chemical form of life (based on silicon rather than carbon); Mr. Spock discovers, via a Vulcan mind meld (of course), that it is a “mother” defending her eggs from miners.

It is of course an open question whether life forms differing this radically from Earth life in basic chemical composition are biologically possible but surely one doesn’t want to rule them out merely “by definition.” Besides, there are theoretical reasons for thinking that silicon-based life forms such as the Horta may not be as scientifically implausible as sometimes supposed. Schulze-Makuch and Irwin (2008) argue that under physical and chemical conditions very different from those on Earth (viz., little free oxygen or liquid water and very low temperatures) silicon may be able to form longer and more complex molecular structures; the latter has long been thought to be a requirement for life (Pace 2001). They conjecture that there might even be silicon-based life forms on Titan, a moon of Saturn. As biochemist William Bains

observes, however, Mr. Spock's experience with Titanians wouldn't be very pleasant. Even a small whiff of their breath would stink to high heaven, and if one were beamed aboard the starship Enterprise it would boil, burst into flames, and kill everyone in the transporter room (Heward 2010).

A less extreme but far more worrisome illustration of how a metabolic definition of life based too closely on the biochemistry of familiar Earth life could hinder the recognition of extraterrestrial life is provided by the Viking spacecraft missions to Mars in the 1970s. The purpose was to search for microbial life on Mars. To this day they remain the only dedicated *in situ* search for extraterrestrial life that has actually been conducted; the European Space Agency plans to launch another mission of this sort (dubbed ExoMars) in 2018. The life-detection experiments performed robotically on Mars by the Viking lander module were at least implicitly based upon a chemical-metabolic definition of life; Gilbert Levin, the principle investigator of the "labeled release" (LR) experiment, which produced the most promising and yet perplexing results, claims that the biology teams explicitly agreed upon such a definition.

The Viking instrument package contained three metabolic experiments along with a gas chromatograph mass spectrometer (GCMS) for detecting organic molecules. In the LR experiment, martian soil was robotically introduced into a test chamber and injected with a carbohydrate solution labeled with radioactive carbon-14 that is readily metabolizable by cultured Earth microbes. To the delight of the Viking biology team the soil started evolving $^{14}\text{CO}_2$ —just what one would expect from Earth microbes. The martian soil sample was subsequently heated to 160°C for three hours, more than enough to kill Earth microbes, and the reaction ceased, strongly suggesting the initial reaction was biological. When another martian soil sample was given a second helping of nutrients, however, the anticipated burst of new activity from hungry martian microbes not only failed to occur but even more mysteriously $^{14}\text{CO}_2$ left over from the initial reaction began disappearing. The biology team was stunned. They couldn't explain these results in terms of any known chemical process, biological or non-biological. In the end they concluded that it must not have been biological on the basis of the anomalous (for Earth microbes) results and, most importantly, the failure of the Viking GCMS to find any organic molecules in the soil to its limits of detection with sample heating up to 500°C .

Non-biological, theoretical explanations for the mysterious Viking results were developed and a consensus quickly reached that they were produced by an exotic oxidation state of iron. To this day, however, there is no direct empirical evidence that the martian surface is strongly oxidizing. Furthermore, it has been shown that the Viking GCMS was not very accurate and would not have detected as many as 10^6 bacterial cells per gram of soil. Yet NASA has not placed a high priority on examining the martian surface for the presence of a powerful inorganic oxidant. Why? As one NASA scientist opined, because we were looking for life and what we found wasn't life. In other words, the definition undergirding the Viking life-detection instrument package was not satisfied, and hence: no life. This illustrates both the potential power of an accepted definition of life to blind scientists to alternative forms of life and the dangers of basing a definition of life too closely on the biochemistry of

familiar Earth life: We might not recognize an alternative form of life if we search for life with an inadequate definition of life.¹

2.2.3 *Evolutionary definitions*

Evolutionary accounts currently dominate scientific and philosophical thought about the nature of life. Most evolutionary definitions of life are founded upon Darwin's theory of evolution by natural selection.

The most abstract versions of the Darwinian definition of life follow Dawkins (1983) and Dennett (1995) in holding not only that the capacity to undergo evolution by natural selection is universal to life ("universal Darwinism") but also that it is independent of the material stuff that instantiates life. Darwinian definitions of this type are especially popular among artificial life (ALife) researchers, who frequently construe Darwinian evolution as an algorithmic process (reducible to a system of equations). As an illustration, Rasmussen (e.g., 1992) and Ray (e.g., 1992) contend that a sufficiently complex computer simulation of life exhibiting Darwinian evolution would truly be alive; indeed Ray believes that we have already created such purely informational living creatures! On this version of the Darwinian definition, it is not the computer that is alive but rather the computational processes themselves; the computer's role is analogous to that of glassware that may someday be used in the artificial synthesis of organic life in a laboratory.

Most biologists and biochemists, however, find the idea of an immaterial computer simulation of life actually *being* alive counterintuitive. Few biochemists believe that a computer simulation of biochemistry is biochemistry itself. A computer simulation of photosynthesis does not produce authentic carbohydrates (chemical compounds of carbon, hydrogen, and oxygen); at best it may be said to yield simulated carbohydrates. Analogously why would one think that a computer simulation of life is actually alive, as opposed to being merely a simulation of life? The only unequivocal example of life available to us is familiar biological Earth life, and it is chemical. One cannot of course rule out someday discovering that the essential characteristics of life are independent of the stuff that comprises it. It is important to keep in mind, however, that there currently exists no empirical support whatsoever for this claim. All that a computer simulation of life establishes is that some of the higher-level characteristics of life can be successfully emulated on a computer, which is hardly surprising because computers can simulate the purely structural and functional characteristics of virtually any physical system, e.g., bridges, landslides, hurricanes, and the human circulatory system. In this context, it is especially important not to be seduced by the starry eyed claim that computers are on the threshold of providing us with the first truly alien forms of life. For to recapitulate, the claim that such simulations are truly alive rests upon the assumption that the essential properties of life are fully extractable from their material embodiments, and this is exactly what is at issue! We will return to this issue shortly.

¹ I am not of course claiming that the Viking spacecrafts found martian life; on my view, the jury is still out.

In light of these concerns, many biologists and biochemists explicitly restrict Darwinian definitions of life to chemical systems. The “chemical Darwinian definition,” which became NASA’s official “working definition” of life in the 1990s (Joyce 1994), is the prototype and remains very popular (NRC 2007). According to the chemical Darwinian definition “life is a self-sustained chemical system capable of undergoing Darwinian evolution” (Joyce 1994: xi-xii). As stated, the chemical Darwinian definition is not restricted to organic molecules, and hence is open to the possibility of silicon-based life on Titan, or perhaps even sulfur-based life in the sulfuric acid clouds of Venus (Grinspoon 2004: 283-286). Nevertheless it is often tacitly limited to organic chemistry. As discussed in the next section (2.3.2), it is difficult to design life-detection instrument packages for spacecraft missions on the basis of the chemical Darwinian definition. But this is merely a practical problem. A far more serious problem, which it shares with other evolutionary definitions, concerns the status of individual organisms as living things. Biological evolution characterizes populations. Individual organisms (fertile and reproducing, fertile and non-reproducing, sterile members of fertile species, or sterile hybrids) do not evolve, and hence are not encompassed by purely evolutionary definitions. Most proponents of evolutionary definitions concede that this is a shortcoming. Various strategies are proposed for dealing with it. One of the most popular involves pointing out that individual organisms are the product of evolution. A more sophisticated approach, suggested by Tibor Gánti (1971), who somewhat ironically doesn’t embrace an evolutionary definition of life, argues that the capacity to reproduce (a pre-requisite for evolution) is only potential, and hence that even sterile hybrids, such as mules, should count as living things. Both solutions fall short, however, when confronted with the very real (but thus far hypothetical) possibility of artificial microbes synthesized in a laboratory to be intrinsically sterile. Viewed from this perspective, it is hardly surprising that other proponents of evolutionary definitions (e.g., Bedau 1998) simply bite the bullet and reject the claim that individual organisms are primary living things. All of these strategies are extremely *ad hoc*, however, that is, motivated solely by the desire to protect a preferred definition of life from what would otherwise be intractable counterexamples.

2.2.4 *Defining life as self-organized complexity*

In recent years, newer definitions focusing on the self-organizing properties of organisms have been developed. The most influential are autopoietic definitions and Gánti’s chemoton model of life (1971). Kauffman’s definition of life in terms of autonomous agency should also be included. For as Luisi observes, all three modes of definition “...share a common view about minimal life as a distributed emergent property based on an organized network of reactions and/or processes” (2003: 53). They differ primarily on how they understand self-organized complexity. Kauffman cashes it out in terms of physics, the product of a novel principle of thermodynamics. Autopoietic definitions and the chemoton model, in contrast, regard it as a logico-mathematical property: Luisi (2003: 53) explicitly states that autopoiesis is concerned with the “internal logic” of life, and Gánti (1971) develops a formal model of life, as a “fluid automata”, described by a system of equations.

The most obvious difficulty with all three approaches is that the vague concepts of self-organization (e.g., “emergence”, “self-referentiality”, “self-sustainability”, and “auto-maintenance”) invoked to explain life are treated as primitive, raising the hoary old problem of goal-directed or self-causation. The basic idea seems to be that at a certain level of organization novel functional properties spontaneously arise, either as a matter of fundamental physics or logic, rendering the system capable of developing increasingly greater organized complexity. Advocates seem to think that this is an asset, but it is actually a weakness. In truth, these approaches to explicating the nature of life have much in common with older views such as vitalism and British emergentism, with the downside that less attention is paid to the physical and metaphysical problems posed by the notion of self-causation.

In addition to the problem of making good sense of self-causation, autopoietic definitions, which have received more attention than the other two, are notable for their permissiveness. They classify many systems that would otherwise be considered paradigmatically nonliving (and hence, compelling counterexamples) as living things. Some provocative illustrations are whole eco-systems, human institutions (such as corporations), and even Earth itself²; computer simulations are also commonly included on the grounds that the difference between life and nonlife is merely a matter of how the components of a system are organized and integrated. That a proposed definition admits such *prima facie* nonliving entities (seeming counterexamples) as living individuals would normally count against it. Turning the tables and insisting that they are living things *just because* a favorite definition says so is logically problematic.

As earlier, it is important to keep in mind that there is nothing problematic about attempting to develop a coherent concept of autopoiesis and exploring its potential for shedding light on unambiguous cases of life. Furthermore, research within the context of an empirically successful scientific theory sometimes changes classification schemes in significant ways. It is thus foolish to insist that our current classifications of entities as living and nonliving are the final word. But such a project is very different from formulating a logically rigorous definition that (in virtue of supplying necessary and sufficient conditions for life) fixes what is and is not a living thing in one fell swoop.

2.3 What exactly are definitions?

That the most popular contemporary definitions of life are defective does not of course establish that no definition of life could be successful. Efforts to formulate more satisfactory definitions of life are ongoing. Some attempt to resolve problems with earlier definitions by combining the collective (evolutionary) and individualistic (metabolic) aspects of life, instead of trying to “reduce” one to the other. Nevertheless it is striking that even with the remarkable advances in biology and biochemistry of the past couple of centuries we still lack a consensus on a definition of life. This suggests that there may be something wrong with the project of defining life.

² The claim that the Earth itself is a living individual is known as the Gaia hypothesis; it was first proposed by (Lovelock 1972) and co-developed with Lynn Margulis (Lovelock and Margulis 1974).