Complexity, Artificial Life, and Artificial Intelligence

Carlos Gershenson
State University of New York at Binghamton
School of Systems Science and Industrial Engineering
Universidad Nacional Autónoma de México
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Centro de Ciencias de la Complejidad
cgg@binghamton.edu

Abstract The scientific fields of complexity, Artificial Life (ALife), and artificial intelligence (AI) share commonalities: historic, conceptual, methodological, and philosophical. Although their origins trace back to the 1940s birth of cybernetics, they were able to develop properly only as modern information technology became available. In this perspective, I offer a personal (and thus biased) account of the expectations and limitations of these fields, some of which have their roots in the limits of formal systems. I use interactions, self-organization, emergence, and balance to compare different aspects of complexity, ALife, and AI. Even when the trajectory of the article is influenced by my personal experience, the general questions posed (which outweigh the answers) will, I hope, be useful in aligning efforts in these fields toward overcoming—or accepting—their limits.

Keywords
Complexity, emergence, self-organization, balance

1 Introduction

The best way to understand man is by creating him.
—José Negrete-Martínez

Complexity[1] has been studied since antiquity. Just to mention a few examples, Aristotle's concept of "more than the sum of its parts" is related to emergence (see section 6), the Sanskrit term *tantra* (interwoven) has several parallels with complexity, and ecology has always been inherently complex. There are several historical examples of what would later be called Artificial Life (ALife),[2] with a mild surge after the publication of Mary Shelley's *Frankenstein; or, The Modern Prometheus* in 1818 (Ball, 2016; Taylor, 2024; Taylor & Dorin, 2020). There have been many artificial creatures, first as ancient myths, then with automata (possible with clockmaking technology required to precisely measure time for making accurate maps as Europeans navigated around the planet), and, in the previous century, with the development of the first digital computers (Barricelli, 1954; Berlekamp et al., 1982; von Neumann, 1966).

Still, the modern scientific study of complex systems and the field of Artificial Life (under that name) can be traced to the 1980s around the Santa Fe Institute (SFI; founded in 1984) and the nearby Los Alamos National Laboratory (LANL; created for the Manhattan Project) in northern

1 The word *complexity* comes from the Latin *plexus*, which could be translated as "entwined." We can thus say that complex systems are those whose elements are difficult to separate (De Domenico et al., 2019). This is because there are relevant *interactions* among them (Gershenson, 2013b). Thus the traditional reductionist approach that simplifies and isolates in order to predict is inadequate to study complexity (Gershenson, 2013a).

2 Artificial Life applies the synthetic method to biology (Steels & Brooks, 1995): building systems that attempt to reproduce properties of living systems to understand them better (Aguilar et al., 2014).

New Mexico. SFI (celebrating its 40th anniversary) was the first research institution to use the term *complexity*, even when there were several places where similar research had been carried out. The first conference on Artificial Life (1987) took place in Los Alamos, while the second (1990) and third (1992) were in Santa Fe, all three organized by Chris Langton (who coined the term *ALife* and worked at SFI for some years) and others. In 1991, Francisco Varela, Paul Bourgine, and others, organized the first European conference on Artificial Life, with a perspective tending more toward cognitive science. Eventually, the two "schools" converged.

I will not attempt to provide a historical account of complexity, ALife, or artificial intelligence (AI). My purpose is to notice the similarities and differences between the three fields as they share conceptual, methodological, and philosophical approaches.

In the next section, I review the historical and technological circumstances that predated the development of complexity, ALife, and AI. In section 3, I mention common limitations that these fields face, along with the expectations they have generated. In subsequent sections, I relate the concepts of interactions, self-organization, emergence, and balance to complexity, ALife, and AI, before closing the article with open questions.

2 Computers as Telescopes

Where there is an observatory and a telescope, we expect that any eyes will see new worlds at once.
—*Henry David Thoreau*

Why did complexity as we know it and "life as it could be" (Langton, 1989) become popular in the 1980s and not before or after? The answer is personal computers. Before then, digital computing was restricted to the few research institutions that could afford the expensive equipment. (Thus there were few developments that would now be considered as complexity or ALife. In the case of AI, though more projects were funded by governments, companies had fallen into an "AI winter" because of unfulfilled expectations; Floridi, 2020.) PCs changed everything. The number of people who could exploit and explore new possibilities in information processing suddenly exploded.

As already mentioned, there were a few examples of what could be considered Artificial Life (e.g., Barricelli, 1954; Berlekamp et al., 1982; von Neumann, 1966), while Alan Turing (1950), John von Neumann (von Neumann & Morgenstern, 1944), and others were interested in the potential ability of computers to model the human mind. We can say that *cybernetics* (Ashby, 1956; Heylighen & Joslyn, 2001; Rosenblueth et al., 1943; Wiener, 1948)[3] set the basis for the scientific study of complex systems, intelligence, and life. This is because cybernetics was the first transdisciplinary effort to study phenomena independently of their substrates. Systems were studied in terms of their *organization*, rather than in terms of their components. And because organization (Ashby, 1962; Atlan, 1974; Rupe & Crutchfield, 2024; von Foerster, 1960) can be described in terms of information (Prokopenko et al., 2009; Shannon, 1948), it became clear that the technology capable of increasing information processing (i.e., computation), storage, and transmission would be essential.

Something similar happened with fractals, which were named only in 1975 by Benoît Mandelbrot (1982). Still, some examples were already proposed in the late 19th and early 20th centuries by Cantor, von Koch, Sierpiński, and others. Even more, Gaston Julia and Pierre Fatou had studied iterative functions, which can be used to construct fractals. Still, these were mainly forgotten. But Mandelbrot had a huge advantage: access to computers that could *draw* fractals, as he worked at IBM Research.[4] *Then*, the interest in fractals exploded.

Before telescopes, no planet beyond Saturn could be detected, and our moon was the only satellite known. Galileo was able to see Jupiter's four largest satellites with his telescope. More planets followed. Other galaxies were observed only less than a century ago, as more powerful telescopes

3 Also known as "control and communication in animals and machines" (Wiener, 1948).

4 Well, he was also a student of Julia. And his uncle Szolem (who knew Sierpiński) had suggested that he work on iterative functions. And he was extremely smart.

became available. The first exoplanet was detected in 1992. Now more than 5,000 exoplanets have been confirmed in more than 4,000 planetary systems. It is only because of these observations of exoplanets that we now know that most stars have planetary systems, even if we have yet to detect their majority.

Before microscopes, doctors were taught that disease was caused by the imbalance of "humors" (or astrological influence, from which the term of *influenza* comes). It took more than two centuries for the germ theory of disease to be broadly accepted. But without *seeing* pathogens, how could we attempt to prevent and cure the diseases they caused? Leeches, one might suggest.[5]

Before computers, we did not have the proper tools to study complex systems. Just like our vision is limited to perceiving the macro and the micro, our limited cognition restricted us to dealing with only a few variables, even if we had huge blackboards. As Heinz Pagels (1989) noted, computers are like telescopes for complexity—and Artificial Life, and AI. All three have *information processing* at their core. Thus we could begin to study them only once information technology had reached a level where enough information could be stored, transmitted, and processed to simulate intelligence, life, and complexity (Simon, 1996).

Why New Mexico, the "land of enchantment"? This is a trickier question. Better said, attempting to answer it has to be more subjective. Still, I can speculate that at the time, there was enough talent (some Nobel Prize winners) and freedom of research at LANL (e.g., arXiv was created there by Paul Ginsparg in 1991). Unfortunately, as several colleagues who have worked at LANL told me, the situation changed at the laboratory for different reasons, resulting in limited creativity and fewer people being attracted to it. Nevertheless, it seems to me that, "back in the day," it was remote enough so that nonmainstream ideas could be explored, but not so remote that the successful ideas that were developed could not spread.

3 Promises and Limits

Every man takes the limits of his own field of vision for the limits of the world.
—Arthur Schopenhauer

The limits of my language means the limits of my world.
—Ludwig Wittgenstein

One could naively think that we only need enough computational power to completely model and understand intelligence, life, and complexity. Many promises were made: robots smarter than humans in all domains, diseases cured, genomes controlled, brains understood, futures predicted precisely—even though there have been some notable advances, results have not fulfilled many expectations. Still, some researchers are still hopeful of achieving these goals simply with better models and faster computers. And projects with these expectations are still being funded.[6] Nevertheless, even before the first electronic computers were built, this approach was "doomed" by the limits of formal systems as proven by Gödel (1931), Turing (1936), Chaitin (1974, 2004), and others.

In the late 19th century, Georg Cantor proposed set theory (for which he was ridiculed and ostracized), which later became the basis of modern mathematics. Paradoxes arose. Whitehead and Russell (1910, p. 13) attempted unsuccessfully to overcome them. David Hilbert launched a program to try to prove that mathematics is complete (all statements can be proven true or false), consistent (no contradictions), and decidable (questions posed within mathematics could be answered). A young John von Neumann, then a PhD student of Hilbert, was working on this topic, and probably that is why he was the only one who understood when Kurt Gödel presented his results proving that formal systems powerful enough to express Peano arithmetic cannot be both

5 Certainly the history of pathology is much more complex than that (Mukherjee, 2022).

6 I am not suggesting that the failed attempts will never be achieved, nor that relevant progress has not been made. My argument, explained subsequently, is that we will not achieve them with the limited methods we have now, although this does not imply that new methods may eventually be developed that might overcome the present limits.

consistent and complete. Later, Turing proved that mathematics is not decidable, for which he defined the concepts of the Turing machine and computable numbers. The implications of these results are that formal systems are limited in ways that have yet to be completely understood in many fields where we rely on formal systems. A sign of this is that we still attempt to use formal systems for tasks that would require going beyond those limits (e.g., artificial general intelligence, personalized medicine, financial forecasting). This is because, in all of these cases, there is a certain "creativity," in the sense that conditions change in unpredictable ways because of novel information generated by interactions (also known as nonstationarity, described later; Gershenson, 2013a), and our previous solutions become obsolete. Still, in many cases, partial success is better than nothing at all, especially because we have yet to find a suitable alternative.

Even when adaptation is widely used (Ashby, 1947a), there are always parts of systems (axioms in the formal case, hardware or hard code in the engineering case) that cannot be changed. Still, we might argue that "real" intelligence, life, and complexity cannot change the laws of physics or chemistry, so in a sense, they are also limited.

Independently of our definitions of intelligence, life, and complexity, we can say that artificial systems have yet to exhibit behavior as rich as the behavior of natural systems. Could this be because of the limits of formal systems, or simply because we have yet to understand how nature changes itself?

Moreover, it might be that we *want* artificial systems to be simpler than natural ones. This is because we can attempt to better understand less detailed versions of natural systems. For example, it would be difficult to sell robots that are not consistent (1 + 1 should always be 2). But it seems that we need to allow for some inconsistencies if we want to simulate broadly human intelligence.

In the case of Artificial Life, these limits have been evident in the study of *open-ended evolution* (Pattee & Sayama, 2019; Standish, 2003; Taylor et al., 2016). As Hernández-Orozco et al. (2018) showed, undecidability and irreducibility (which might be considered as desirable or undesirable but are precisely some of the limits of formal systems) are conditions for open-endedness.

For complexity, a relevant case is *emergence* (Abrahão & Zenil, 2022; Bedau, 1997; Bedau & Humphreys, 2008; Schmickl, 2022) (to be expanded in section 6). There are several notions and flavors of emergence. In general, it can be said that emergent properties are those present at one scale (usually lower or faster, but not necessarily) and not at another scale (usually higher or slower) (Gershenson, 2023b). In particular, some see "strong emergence" as problematic, because it usually implies *downward causation* (Bitbol, 2012; Campbell, 1974; Farnsworth et al., 2017; Flack, 2017). This means that emergent properties at a higher scale have a causal effect on elements at a lower scale. We have yet to find a formalism that properly describes downward causation, while some argue that it does not even exist (downward causation might be apparent, an epiphenomenon, but the laws of physics explain everything).[7] Could it be because of the same limits of formal systems? Nevertheless, for practical purposes, does it really matter? Even if, *in theory*, everything could be reduced to physics, *in practice*, it is not. So, in any case, we do need descriptions at all levels to understand and face complexity.

For AI, several limits have been identified, one of the most relevant being that of *meaning* (Mitchell, 2020; Rota, 1986). In principle and practice, machines can simulate in a very sophisticated way our cognitive abilities. Still, do they really understand (Harnad, 1990; Searle, 1980)? We might say that pragmatically, it does not matter. But it should, as a feature of human cognition is the ability to change meanings arbitrarily and adaptively, which again seems limited by formal systems used to implement AI systems. There have been impressive advances within information theory, but methods for creating semantics and understanding are still at an early stage. Some people (e.g.,

7 One example comes from personal conversations with David Wolpert, who does not believe in downward causation but concedes that it might be that in some cases, in practice, it might be easier to predict lower-scale phenomena from higher-scale properties, similar to one-way functions used in cryptography: in reality, the higher scale is caused by the lower one, but in practice, it is not computable. Another view is that speaking about causality between scales is a conceptual mistake, because, independently of observers, phenomena occur at *all* scales (Gershenson, 2007, p. 31). It is only our descriptions that represent limited aspects of phenomena at particular scales.

Agüera y Arcas, 2022) have argued that the surprising capabilities of recent large language models could be considered as understanding, although this is still hotly debated (Mitchell, 2023).

It might be that these limits are actually a feature, not a problem. We need only accept them to be able to exploit them, rather than fighting against them. Imagine that mathematics (or any powerful formal system) was consistent, complete, and decidable, as Hilbert and others hoped for. Yes, we would have "absolute truths" and certainty. But would we have creativity? Innovation? Serendipity? It seems to me that many of the features of our world (without which we would not be here) *require* the limits we have been so eagerly trying to eliminate.

Even when there have been impressive advances in the scientific study of complexity, Artificial Life, and AI, several open problems may be related to the inherent limits of formal systems. Will we be able to go beyond them?

4 Interactions

The aim of science is not things themselves, as the dogmatists in their simplicity imagine, but the relations among things; outside these relations there is no reality knowable.
—Henri Poincaré

Neither from itself nor from another,
Nor from both,
Nor without a cause,
Does anything whatever, anywhere arise.
—Nāgārjuna, Mūlaamadhyamakakārikā 1:1

Reductionism is correct, but incomplete.
—Murray Gell-Mann

Etymologically and conceptually, we can say that the most relevant feature of complex systems is *interactions* (De Domenico et al., 2019; Gershenson, 2013b). Complexity comes from the Latin *plexus*, which means "entwined," and has some similarities with the Sanskrit *tantra*. In both cases, interactions make it difficult to study or describe elements in isolation, just like threads in a fabric (which is the literal meaning of *tantra*). We can say that this is related to the concept of *tendrel* (Tibetan; Sanskrit *pratītyasamutpāda*) from Buddhist philosophy, which could be translated as "interdependent origination," "dependent arising," or simply "causation." *Tendrel* notices that phenomena arise in relation to other phenomena. Nothing can be isolated, nor be caused only by itself or out of nothing. Thus everything is related, directly or indirectly (Garfield, 1995; Gershenson, 2023a; Rovelli, 2021).

Traditional science and philosophy (since the times of Galileo, Descartes, Newton, and Laplace) have been reductionist in the sense that within this paradigm, we try to simplify and isolate phenomena to predict and control them (Heylighen et al., 2007; Morin, 2007). In other words, we aim at finding *fundamental* "laws" and at using them to obtain a priori knowledge (predict the future), *reducing* phenomena to the fundamental laws used to describe them. This has been extremely successful and has led to impressive advances in engineering, medicine, and more. Still, this does not imply that reductionism does not have its limits or that there might be more suitable descriptions of the world for certain purposes. Precisely when we have relevant interactions, reductionism is inadequate, as it neglects interactions and their implications.

Interactions have several implications (Gershenson, 2013b), but I can say at a general level that the main one is that interactions may produce *information* that was not present in initial or boundary conditions. This inherently limits predictability (Gershenson, 2013a), as we cannot know a priori which information will be generated. This is known as *computational irreducibility* (Hernández-Orozco et al., 2018; Wolfram, 2002; Zenil, 2013): There is no "shortcut" to the future, as information should be processed through interactions to reach it.

It should be noted that, *in practice*, computational irreducibility might not pose such a challenge as it does *in theory*. If we are interested only in a particular context, we could potentially explore exhaustively, or at least systematically, all or several possibilities, then a posteriori be able to describe and predict the future of complex systems, including their emergent properties and variables. Still, if we are dealing with *nonstationary problems*,[8] then even if we have a "full" understanding of a specific complex system, if the problem changes (which is not rare precisely because of interactions), it might be that new relevant information will arise and our understanding will be obsolete.

The fact that traditional tools (from reductionist science) are insufficient to study complex systems has led some researchers to seek alternatives (Heylighen et al., 2024; Kauffman & Roli, 2023), in part because we seem unable to address global challenges precisely because of their complexity.

The relevance of interactions and limits of predictability have been discussed mainly concerning complex systems, but they are relevant for ALife and AI as well. Interactions in ALife and AI systems can also generate novel information, limiting predictability, for better or for worse. There have been several attempts with varying degrees of success, but we still lack a general, common framework to describe, understand, and control complex systems. And it might be that such a framework could be developed within ALife or AI, then generalized for all complex systems.

Interactions limit the predictability of complex systems. Thus, in many cases, future information can only be known a posteriori (because of computational irreducibility). This implies that traditional reductionist approaches and methods seem insufficient to properly understand complexity, life, and intelligence. Still, we have yet to develop widely accepted methods that show the desired sufficiency.

5 Self-Organization

The beauty of a living thing is not the atoms that go into it,
but the way those atoms are put together.
—Carl Sagan

Nature holds several examples of self-organization (Camazine et al., 2003): flocks, schools, swarms, herds, crowds, and so on. In these examples, no leader or external source tells individuals what to do, but the properties of the system are a result of the distributed *interactions* of individuals. Thus the study of self-organization is closely tied to complexity and the information technology necessary to model it. Also, the term *self-organizing system* has its origins in cybernetics (Ashby, 1947b, 1962; Lendaris, 1964; von Foerster, 1960). Nevertheless, there have been several examples of self-organization in physical and chemical systems (Bak et al., 1987; Eigen & Schuster, 1979; Feltz et al., 2006; Haken, 1981; Nicolis & Prigogine, 1977; Schweitzer, 1997).

A system can be described as self-organizing when its components interact to produce a global pattern or behavior (Gershenson & Heylighen, 2003). This description can be useful when we are interested in relating multiple scales (elements and system, micro and macro) and how changes in one might affect the others (e.g., changes in individuals affect a society or changes in a society affect individuals).

If we are dealing with a complex problem, novel information can make it nonstationary; that is, the problem changes. If the problem changes faster than the time required to find a novel solution through optimization or other traditional techniques, then the solutions will be obsolete. Self-organization can be a viable approach to developing adaptive solutions that are able to face nonstationary problems, because when the problems change, elements can adjust through their interactions (Frei & Di Marzo Serugendo, 2011; Gershenson, 2007).

Self-organization has been used broadly in ALife: for software (digital organisms), hardware (robots), and wetware (protocells) (for a review, see Gershenson et al., 2020).

In AI, self-organization has had a more limited use. Still, it could be argued that most artificial neural network models are implicitly self-organizing (Gershenson, 2010), as their weights (interactions) are modified during the training phase. And explicitly, Kohonen networks are self-organizing

8 A stationary problem does not change in time, so once a solution is found, it will remain valid. A nonstationary problem does change in time, so novel solutions should be found, ideally as fast as the problem changes (Gershenson, 2007).

(Kohonen, 2000). Also in robotics, self-organization has been relevant, implicitly or explicitly (Pfeifer et al., 2007).

Self-organization can be useful when multiple scales are modeled at the same time. It has been a relevant concept for complex systems and ALife, with a potential in AI that has yet to be fully explored.

6 Emergence

> *You could not have evolved a complex system like a city or an organism—with an enormous number of components—without the emergence of laws that constrain their behavior in order for them to be resilient.*
> *—Geoffrey West*

The concept of emergence has certain analogies with Aristotle's "the whole being more than the sum of its parts," where the "more" is the emergent bit. Emergence was popular in the 19th century (McLaughlin, 1992; Mengal, 2006) but fell out of favor in the early 20th century owing to the success of reductionist approaches. But when information technology allowed the scientific study of complex systems, emergence became relevant again (Bedau & Humphreys, 2008).

Still, emergence probably caused most of the confusion and skepticism around complexity in the 1980s and 1990s, in part because some people described emergent properties as "surprising" or "unexpected." Then, emergence should be a measure of our ignorance, because once we understand these properties, they are no longer surprising or unexpected.

Nevertheless, there is nothing mysterious about emergence if properly described (Anderson, 1972). In a general way, emergent properties are those present at one scale but not at another (Gershenson, 2023b). For example, a bar of gold has color, conductivity, malleability, and so on. Still, its components (gold atoms) do not have these properties, so we can call them emergent. In a similar way, it is accepted that cells are alive, but they are composed of molecules that are not alive. Whatever our definition of life, we can say that it emerges out of the interactions of molecules. It is accepted that a human is intelligent, but we are composed of cells that are not intelligent (in the same way). Whatever our definition of intelligence, we can say that intelligence emerges out of interactions of cells.

Emergence comes in different flavors, some less controversial than others (for a review, see Gershenson, 2023b). For example, *weak* emergence (Bedau, 1997) is about properties described by an observer, such as gliders in the Game of Life (Beer, 2014; Berlekamp et al., 1982). Still, gliders do not change the rules of the Game of Life, and these rules are sufficient to compute the future states of the system. *Strong* emergence (Bar-Yam, 2004; Schmickl, 2022) would be when having all information at one scale is not enough to derive information at another scale. In many cases, strongly emergent information or properties have a causal effect on the elements that produced them. For example, molecules form cells, but living cells make molecules that cannot be produced without biospheres. Also, individuals create social norms, and these norms promote and constrain the behaviors of individuals.

One could say that weak emergence is "in the observer," while strong emergence "is real." Some (reductionist) people do not believe that strong emergence exists (e.g., Weinberg, 1993), as it implies downward causation, and for them, only "fundamental" phenomena described by physics are real. Independently of our notion of reality, *in practice*, the laws of physics are not sufficient to describe, explain, or, even less, predict phenomena at higher scales (even in fluid dynamics and chemistry, we do not have to go to life, intelligence, and culture).

I conjecture that strongly emergent properties are not computable *in practice*, and that is why a lower-scale description is not enough to predict them. If there is no practical way in which the properties of one scale can be described in terms of the "laws" of another, then we should validly describe those properties as emergent. Of course, this cannot be proven, for reasons similar to why a number cannot be proven to be random (Chaitin, 1975) or Kolmogorov's complexity is not computable (in theory) (Delahaye & Zenil, 2012). Note that this approach does not rely on downward causation, but it does not prevent it (Hoel et al., 2013).

For example, a person can be melted by the words of a loved one, but this cannot be derived from the laws of physics that describe the melting of solid matter, no matter how detailed a description one might have at the "fundamental" level. Certainly the laws of physics are not being violated. They are simply not enough, as there is no *meaning* in physics (Farnsworth et al., 2013; Gershenson, 2012).

Emergence has been a central concept for complex systems and Artificial Life (Gershenson, 2023b). Many ALife models have been used to better describe and understand different flavors of emergence (e.g., Bersini, 2006; Beuls & Steels, 2013; Hidalgo et al., 2016; Martínez et al., 2012; Moreno & Ruiz-Mirazo, 2009; Roli & Kauffman, 2020; Torres-Sosa et al., 2012; Walker, 2014; Watson et al., 2011).

In AI, emergence has been less relevant. Still, unpredictable capabilities of large language models have been recently described as emergent (Wei et al., 2022), sparking some controversy.

Emergence can be a useful concept when information is not present at one scale but is present at another. Even when it is prevalent, we lack the conceptual and formal tools to precisely speak and measure emergence in complex, living, and intelligent systems.

7 Balance

Everything tends to a balance.

In recent years, I have been developing a narrative of "balance"[9] to bring together concepts of the scientific study of complex systems and to communicate them to a general audience. There are several historical examples of balance from ancient cultures, and those examples show that it has been a common, long-standing practice to try to avoid extremes. Still, *criticality* (Adami, 1995; Bak et al., 1987; Balleza et al., 2008; Chialvo, 2010; Mora & Bialek, 2011; Muñoz, 2018; Pascual & Guichard, 2005; Roli et al., 2018; Sánchez-Puig et al., 2023; Stanley, 1987) can be seen as a type of balance between order and chaos (Kauffman, 1993; Langton, 1990). Life (and computation) needs some stability (order) to keep on functioning. But too much stability limits adaptability. At the other extreme, too much variability (chaos) loses useful information. At "the edge," evolution, life, and intelligence can emerge.

More generally, balance is a tautology, because we describe a posteriori phenomena that survived and evolved as balanced, between "too few" and "too much" *change*. Certainly there can be "dynamic balance," where the precise trade-off varies and systems need to adapt (as exemplified by the slower-is-faster effect; Gershenson & Helbing, 2015). Also, interactions, perturbations, or noise can increase the change in a system, for which *antifragility* (Pineda et al., 2019; Taleb, 2012) is desirable. And we have recently shown that *heterogeneity* can "extend" the "balanced" region of systems (López-Díaz et al., 2023; Sánchez-Puig et al., 2023).

In AI, a well-studied balance is that between exploration and exploitation in search (Downing, 2015; Hills et al., 2015), also known as search in breadth or depth, respectively (when solution spaces are represented as trees). In other words, to try to find the best solution to a problem, one can exploit current solutions and try to improve them or explore completely novel solutions with the hope that some might be better than current ones. Because the best strategy cannot be predetermined, as it depends on the problem space (Wolpert & Macready, 1995, 1997), the precise balance between exploration and exploitation will depend on the particular problem space that is searched.

Balance also offers a promising narrative by which to study evolution (natural and artificial) (Jablonka & Lamb, 2006; Kauffman & Roli, 2021; Wagner, 2005), as, by definition, that which evolves needs to be balanced.

Phenomena that endure tend to avoid extremes, so they can be called "balanced" a posteriori (once they have endured). Still, this tautology can be useful to bring together common concepts in complex systems, ALife, and AI.

9 We can roughly define *balance* as that which avoids extremes.

8 Inconclusion

Being ill defined is a feature common to all important concepts.
—Benoît Mandelbrot

I have mentioned conceptual similarities and challenges among complexity, ALife, and AI. Still, many open questions remain.

There are no agreed-upon definitions of complexity, life, or intelligence. But perhaps this is more a feature than a problem. If we could define one of these precisely, then we would not have so many open questions about them. And we do because their richness goes beyond our current ability to understand. It remains to be seen whether we "only" need a revolution in science (Heylighen et al., 2024; Kauffman & Roli, 2023) to be able to understand them properly. Or it might be that some aspects are inherently beyond understanding as we know it (Wolpert, 2024).

In practice, several relevant, recent advances have generated great expectations. Whether we consider novel forms of life, either by exploiting current ones (Blackiston et al., 2021, 2023; Gibson et al., 2010; Kriegman et al., 2020) or by exploring novel ones (Čejková et al., 2017; Muñuzuri & Pérez-Mercader, 2022; Rasmussen et al., 2008), we will take important steps to understanding life on Earth and on other planets (Walker et al., 2018).

Historically, AI has had its cycles of expectations (summers) and disappointments (winters). We have had several years of building expectations. For example, autonomous vehicles are still "2 years away" after more than 15 years. Deep neural networks and large language models have achieved impressive performance, but in the end, they are "only" ad hoc statistical engines. It is not clear that by following the same approach, something like "understanding meaning" could be achieved (Mitchell, 2020). Still, for many practical purposes, this is not relevant. Nevertheless, there are limits to what current approaches will be able to do.

As for the scientific study of complex systems, perhaps its success will be achieved when most disciplines complete integrating their concepts and methods and adopt them as their own, so few people would speak about "complexity economics" or "biological complexity," simply because most people would be familiar with the relevant concepts and methods. Still, there will always be a narrow space for studying complexity per se, as the study of the interactions in systems at all scales.

The limitations outlined in this article might be overcome. We have no clear idea of how this will be possible, but there are several promising explorations. If further research helps to better delineate the limits of science rather than going beyond them, this will certainly be useful and will allow us to make better decisions, even if it is only by knowing what we have no way of knowing.

Acknowledgments

I am thankful for comments and feedback from Jan Dijksterhuis, Mario Franco, Stuart Kauffman, Amahury López-Díaz, Andrea Roli, David Wolpert, members of the Foundations of Information Science mailing list hosted at the Universidad de Zaragoza, and two anonymous referees. Steve Spero helped with proofreading the paper.

References

Abrahão, F. S., & Zenil, H. (2022). Emergence and algorithmic information dynamics of systems and observers. *Philosophical Transactions of the Royal Society A*, *380*(2227), Article 20200429. https://doi.org/10.1098/rsta.2020.0429, PubMed: 35599568

Adami, C. (1995). Self-organized criticality in living systems. *Physics Letters A*, *203*(1), 29–32. https://doi.org/10.1016/0375-9601(95)00372-A

Agüera y Arcas, B. (2022). Do large language models understand us? *Daedalus*, *151*(2), 183–197. https://doi.org/10.1162/daed_a_01909

Aguilar, W., Santamaría-Bonfil, G., Froese, T., & Gershenson, C. (2014). The past, present, and future of Artificial Life. *Frontiers in Robotics and AI*, *1*(8). https://doi.org/10.3389/frobt.2014.00008

Anderson, P. W. (1972). More is different. *Science*, *177*(4047), 393–396. https://doi.org/10.1126/science.177.4047.393

Ashby, W. R. (1947a). The nervous system as physical machine: With special reference to the origin of adaptive behavior. *Mind*, *56*(221), 44–59. https://www.jstor.org/stable/2250675, https://doi.org/10.1093/mind/LVI.221.44, PubMed: 20285973

Ashby, W. R. (1947b). Principles of the self-organizing dynamic system. *Journal of General Psychology*, *37*(2), 125–128. https://doi.org/10.1080/00221309.1947.9918144

Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman and Hall. https://doi.org/10.5962/bhl.title.5851

Ashby, W. R. (1962). Principles of the self-organizing system. In H. V. Foerster, & G. W. Zopf Jr. (Eds.), *Principles of self-organization* (pp. 255–278). Pergamon.

Atlan, H. (1974). On a formal definition of organization. *Journal of Theoretical Biology*, *45*(2), 295–304. https://doi.org/10.1016/0022-5193(74)90115-5, PubMed: 4844620

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/*f* noise. *Physical Review Letters*, *59*(4), 381–384. https://doi.org/10.1103/PhysRevLett.59.381, PubMed: 10035754

Ball, P. (2016, April 16). Man made: A history of synthetic life. *Distillations Magazine*. https://sciencehistory.org/stories/magazine/man-made-a-history-of-synthetic-life/

Balleza, E., Alvarez-Buylla, E. R., Chaos, A., Kauffman, S., Shmulevich, I., & Aldana, M. (2008). Critical dynamics in genetic regulatory networks: Examples from four kingdoms. *PLoS ONE*, *3*(6), Article e2456. https://doi.org/10.1371/journal.pone.0002456, PubMed: 18560561

Barricelli, N. (1954). Esempi numerici di processi di evoluzione. *Methodos*, *6*, 45–68.

Bar-Yam, Y. (2004). A mathematical theory of strong emergence using multiscale variety. *Complexity*, *9*(6), 15–24. https://doi.org/10.1002/cplx.20029

Bedau, M. A. (1997). Weak emergence. In J. Tomberlin (Ed.), *Philosophical perspectives: Mind, causation, and world* (Vol. 11, pp. 375–399). Blackwell. https://doi.org/10.1111/0029-4624.31.s11.17

Bedau, M. A., & Humphreys, P. (Eds.). (2008). *Emergence: Contemporary readings in philosophy and science*. MIT Press. https://doi.org/10.7551/mitpress/9780262026215.001.0001

Beer, R. D. (2014). The cognitive domain of a glider in the Game of Life. *Artificial Life*, *20*(2), 183–206. https://doi.org/10.1162/ARTL_a_00125, PubMed: 24494612

Berlekamp, E. R., Conway, J. H., & Guy, R. K. (1982). *Winning ways for your mathematical plays: Vol. 2. Games in particular*. Academic Press.

Bersini, H. (2006). Formalizing emergence: The natural after-life of Artificial Life. In B. Feltz, M. Crommelinck, & P. Goujon (Eds.), *Self-organization and emergence in life sciences* (pp. 41–60). Springer. https://doi.org/10.1007/1-4020-3917-4_3

Beuls, K., & Steels, L. (2013). Agent-based models of strategies for the emergence and evolution of grammatical agreement. *PLoS ONE*, *8*(3), Article e58960+. https://doi.org/10.1371/journal.pone.0058960, PubMed: 23527055

Bitbol, M. (2012). Downward causation without foundations. *Synthese*, *185*(2), 233–255. https://doi.org/10.1007/s11229-010-9723-5

Blackiston, D., Kriegman, S., Bongard, J., & Levin, M. (2023). Biological robots: Perspectives on an emerging interdisciplinary field. *Soft Robotics*, *10*(4), 674–686. https://doi.org/10.1089/soro.2022.0142, PubMed: 37083430

Blackiston, D., Lederer, E., Kriegman, S., Garnier, S., Bongard, J., & Levin, M. (2021). A cellular platform for the development of synthetic living machines. *Science Robotics*, *6*(52). https://doi.org/10.1126/scirobotics.abf1571, PubMed: 34043553

Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2003). *Self-organization in biological systems*. Princeton University Press.

Campbell, D. T. (1974). "Downward causation" in hierarchically organized biological systems. In F. J. Ayala & T. Dobzhansky (Eds.), *Studies in the philosophy of biology* (pp. 179–186). Macmillan. https://doi.org/10.1007/978-1-349-01892-5_11

Čejková, J., Banno, T., Hanczyc, M. M., & Štěpánek, F. (2017). Droplets as liquid robots. *Artificial Life*, *23*(4), 528–549. https://doi.org/10.1162/ARTL_a_00243, PubMed: 28985113

Chaitin, G. J. (1974). Information-theoretic limitations of formal systems. *Journal of the ACM*, *21*(3), 403–424. https://doi.org/10.1145/321832.321839

Chaitin, G. J. (1975). Randomness and mathematical proof. *Scientific American*, *232*(5), 47–52. https://doi.org/10.1038/scientificamerican0575-47

Chaitin, G. J. (2004). *Irreducible complexity in pure mathematics*. ArXiv. http://arxiv.org/abs/math/0411091

Chialvo, D. R. (2010). Emergent complex neural dynamics. *Nature Physics*, *6*(10), 744–750. https://doi.org/10.1038/nphys1803

De Domenico, M., Camargo, C., Gershenson, C., Goldsmith, D., Jeschonnek, S., Kay, L., Nichele, S., Nicolás, J., Schmickl, T., Stella, M., Brandoff, J., Salinas, Á. J. M., & Sayama, H. (2019). *Complexity explained: A grassroot collaborative initiative to create a set of essential concepts of complex systems*. https://complexityexplained.github.io

Delahaye, J. P., & Zenil, H. (2012). Numerical evaluation of algorithmic complexity for short strings: A glance into the innermost structure of randomness. *Applied Mathematics and Computation*, *219*(1), 63–77. https://doi.org/10.1016/j.amc.2011.10.006

Downing, K. L. (2015). *Intelligence emerging: Adaptivity and search in evolving neural systems*. MIT Press. https://doi.org/10.7551/mitpress/9898.001.0001

Eigen, M., & Schuster, P. (1979). *The hypercycle, a principle of natural self-organization*. Springer. https://doi.org/10.1007/978-3-642-67247-7

Farnsworth, K. D., Ellis, G. F. R., & Jaeger, L. (2017). Living through downward causation: From molecules to ecosystems. In S. I. Walker, P. C. W. Davies, & G. F. R. Ellis (Eds.), *From matter to life: Information and causality* (pp. 303–333). Cambridge University Press. https://doi.org/10.1017/9781316584200.013

Farnsworth, K. D., Nelson, J., & Gershenson, C. (2013). Living is information processing: From molecules to global systems. *Acta Biotheoretica*, *61*(2), 203–222. https://doi.org/10.1007/s10441-013-9179-3, PubMed: 23456459

Feltz, B., Crommelinck, M., & Goujon, P. (Eds.). (2006). *Self-organization and emergence in life sciences* (Synthese Library Vol. 331). Springer. https://doi.org/10.1007/1-4020-3917-4

Flack, J. C. (2017). Coarse-graining as a downward causation mechanism. *Philosophical Transactions of the Royal Society A*, *375*(2109), Article 20160338. https://doi.org/10.1098/rsta.2016.0338, PubMed: 29133440

Floridi, L. (2020). AI and its new winter: From myths to realities. *Philosophy and Technology*, *33*(1), 1–3. https://doi.org/10.1007/s13347-020-00396-6

Frei, R., & Di Marzo Serugendo, G. (2011). Advances in complexity engineering. *International Journal of Bio-inspired Computation*, *3*(4), 199–212. https://doi.org/10.1504/IJBIC.2011.041144

Garfield, J. L. (1995). *The fundamental wisdom of the middle way: Nagarjuna's Mulamadhyamakakarika*. Oxford University Press. https://doi.org/10.1093/oso/9780195103175.001.0001

Gershenson, C. (2007). *Design and control of self-organizing systems*. CopIt Arxives. https://copitarxives.fisica.unam.mx/TS0002EN/TS0002EN.html

Gershenson, C. (2010). Computing networks: A general framework to contrast neural and swarm cognitions. *Paladyn, Journal of Behavioral Robotics*, *1*(2), 147–153. https://doi.org/10.2478/s13230-010-0015-z

Gershenson, C. (2012). The world as evolving information. In A. Minai, D. Braha, & Y. Bar-Yam (Eds.), *Unifying themes in complex systems* (Vol. 7, pp. 100–115). Springer. https://doi.org/10.1007/978-3-642-18003-3_10

Gershenson, C. (2013a). Facing complexity: Prediction vs. adaptation. In A. Massip & A. Bastardas (Eds.), *Complexity perspectives on language, communication and society* (pp. 3–14). Springer. https://doi.org/10.1007/978-3-642-32817-6_2

Gershenson, C. (2013b). The implications of interactions for science and philosophy. *Foundations of Science*, *18*(4), 781–790. https://doi.org/10.1007/s10699-012-9305-8

Gershenson, C. (2023a). Complexity and Buddhism: Understanding interactions. *Buddhism Today*, *52*, 44–48.

Gershenson, C. (2023b). Emergence in Artificial Life. *Artificial Life*, *29*(2), 153–167. https://doi.org/10.1162/artl_a_00397

Gershenson, C., & Helbing, D. (2015). When slower is faster. *Complexity*, *21*(2), 9–15. https://doi.org/10.1002/cplx.21736

Gershenson, C., & Heylighen, F. (2003). When can we call a system self-organizing? In W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim, & J. Ziegler (Eds.), *Advances in Artificial Life, 7th European conference, ECAL 2003 LNAI 2801* (pp. 606–614). Springer. https://doi.org/10.1007/978-3-540-39432-7_65

Gershenson, C., Trianni, V., Werfel, J., & Sayama, H. (2020). Self-organization and Artificial Life. *Artificial Life*, *26*(3), 391–408. https://doi.org/10.1162/artl_a_00324, PubMed: 32697161

Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z. Q., Segall-Shapiro, T. H., … Venter, J. C. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, *329*(5987), 52–56. https://doi.org/10.1126/science.1190719, PubMed: 20488990

Gödel, K. (1931). Über formal unentscheidbare sätze der principia mathematica und verwandter systeme I. *Monatshefte für Mathematik und Physik*, *38*(1), 173–198. https://doi.org/10.1007/BF01700692

Haken, H. (1981). Synergetics and the problem of self-organization. In G. Roth & H. Schwegler (Eds.), *Self-organizing systems: An interdisciplinary approach* (pp. 9–12). Campus.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1), 335–346. https://doi.org/10.1016/0167-2789(90)90087-6

Hernández-Orozco, S., Hernández-Quiroz, F., & Zenil, H. (2018). Undecidability and irreducibility conditions for open-ended evolution and emergence. *Artificial Life*, *24*(1), 56–70. https://doi.org/10.1162/ARTL_a_00254, PubMed: 29369710

Heylighen, F., Beigi, S., & Vidal, C. (2024). *The third story of the universe: An evolutionary worldview for the noosphere* (Working Paper). CLEA/Human Energy.

Heylighen, F., Cilliers, P., & Gershenson, C. (2007). Complexity and philosophy. In J. Bogg & R. Geyer (Eds.), *Complexity, science and society* (pp. 117–134). Radcliffe.

Heylighen, F., & Joslyn, C. (2001). Cybernetics and second order cybernetics. In R. A. Meyers (Ed.), *Encyclopedia of physical science and technology* (3rd ed., Vol. 4, pp. 155–170). Academic Press.

Hidalgo, J., Grilli, J., Suweis, S., Maritan, A., & Muñoz, M. A. (2016). Cooperation, competition and the emergence of criticality in communities of adaptive systems. *Journal of Statistical Mechanics: Theory and Experiment*, *2016*(3), Article 033203. https://doi.org/10.1088/1742-5468/2016/03/033203

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*(1), 46–54. https://doi.org/10.1016/j.tics.2014.10.004, PubMed: 25487706

Hoel, E. P., Albantakis, L., & Tononi, G. (2013). Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(49), 19790–19795. https://doi.org/10.1073/pnas.1314922110, PubMed: 24248356

Jablonka, E., & Lamb, M. J. (2006). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. MIT Press.

Kauffman, S. A. (1993). *The origins of order*. Oxford University Press. https://doi.org/10.1093/oso/9780195079517.001.0001

Kauffman, S., & Roli, A. (2021). The world is not a theorem. *Entropy*, *23*(11). https://doi.org/10.3390/e23111467

Kauffman, S. A., & Roli, A. (2023). A third transition in science? *Interface Focus*, *13*(3), Article 20220063. https://doi.org/10.1098/rsfs.2022.0063, PubMed: 37065266

Kohonen, T. (2000). *Self-organizing maps* (3rd ed.). Springer. https://doi.org/10.1007/978-3-642-56927-2

Kriegman, S., Blackiston, D., Levin, M., & Bongard, J. (2020). A scalable pipeline for designing reconfigurable organisms. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(4), 1853–1859. https://doi.org/10.1073/pnas.1910837117, PubMed: 31932426

Langton, C. (1989). Artificial Life. In C. Langton (Ed.), *Artificial life* (pp. 1–47). Addison-Wesley.

Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D*, *42*(1–3), 12–37. https://doi.org/10.1016/0167-2789(90)90064-V

Lendaris, G. G. (1964). On the definition of self-organizing systems. *Proceedings of the IEEE*, *52*(3), 324–325. https://doi.org/10.1109/PROC.1964.2905

López-Díaz, A. J., Sánchez-Puig, F., & Gershenson, C. (2023). Temporal, structural, and functional heterogeneities extend criticality and antifragility in random Boolean networks. *Entropy*, *25*(2). https://doi.org/10.3390/e25020254, PubMed: 36832621

Mandelbrot, B. (1982). *The fractal geometry of nature*. W. H. Freeman.

Martínez, G., Adamatzky, A., & Alonso-Sanz, R. (2012). Complex dynamics of elementary cellular automata emerging in chaotic rules. *International Journal of Bifurcation and Chaos*, *22*(2), Article 1250023. https://doi.org/10.1142/S021812741250023X

McLaughlin, B. P. (1992). The rise and fall of British emergentism. In A. Beckerman, H. Flohr, & J. Kim (Eds.), *Emergence or reduction? Essays on the prospects of nonreductive physicalism* (pp. 49–93). Walter de Gruyter. https://doi.org/10.1515/9783110870084.49

Mengal, P. (2006). The concept of emergence in the XIXth century: From natural theology to biology. In B. Feltz, M. Crommelinck, & P. Goujon (Eds.), *Self-organization and emergence in life sciences* (pp. 215–224). Springer. https://doi.org/10.1007/1-4020-3917-4_13

Mitchell, M. (2020). On crashing the barrier of meaning in artificial intelligence. *AI Magazine*, *41*(2), 86–92. https://doi.org/10.1609/aimag.v41i2.5259

Mitchell, M. (2023). AI's challenge of understanding the world. *Science*, *382*(6671), Article eadm8175. https://www.science.org/doi/abs/10.1126/science.adm8175, https://doi.org/10.1126/science.adm8175, PubMed: 37943939

Mora, T., & Bialek, W. (2011). Are biological systems poised at criticality? *Journal of Statistical Physics*, *144*(2), 268–302. https://doi.org/10.1007/s10955-011-0229-4

Moreno, A., & Ruiz-Mirazo, K. (2009). The problem of the emergence of functional diversity in prebiotic evolution. *Biology and Philosophy*, *24*(5), 585–605. https://doi.org/10.1007/s10539-009-9178-6

Morin, E. (2007). Restricted complexity, general complexity. In C. Gershenson, D. Aerts, & B. Edmonds (Eds.), *Philosophy and complexity* (pp. 5–29). World Scientific. https://doi.org/10.1142/9789812707420_0002

Mukherjee, S. (2022). *The song of the cell: An exploration of medicine and the new human*. The Bodley Head.

Muñoz, M. A. (2018). Colloquium: Criticality and dynamical scaling in living systems. *Reviews of Modern Physics*, *90*, Article 031001. https://doi.org/10.1103/RevModPhys.90.031001

Muñuzuri, A. P., & Pérez-Mercader, J. (2022). Unified representation of life's basic properties by a 3-species stochastic cubic autocatalytic reaction-diffusion system of equations. *Physics of Life Reviews*, *41*, 64–83. https://doi.org/10.1016/j.plrev.2022.03.003, PubMed: 35594602

Nicolis, G., & Prigogine, I. (1977). *Self-organization in non-equilibrium systems: From dissipative structures to order through fluctuations*. John Wiley.

Pagels, H. R. (1989). *The dreams of reason: The computer and the rise of the sciences of complexity*. Bantam Books.

Pascual, M., & Guichard, F. (2005). Criticality and disturbance in spatial ecological systems. *Trends in Ecology and Evolution*, *20*(2), 88–95. https://doi.org/10.1016/j.tree.2004.11.012, PubMed: 16701348

Pattee, H. H., & Sayama, H. (2019). Evolved open-endedness, not open-ended evolution. *Artificial Life*, *25*(1), 4–8. https://doi.org/10.1162/artl_a_00276, PubMed: 30933631

Pfeifer, R., Lungarella, M., & Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *Science*, *318*(5853), 1088–1093. https://doi.org/10.1126/science.1145803, PubMed: 18006736

Pineda, O. K., Kim, H., & Gershenson, C. (2019). A novel antifragility measure based on satisfaction and its application to random and biological Boolean networks. *Complexity*, *2019*, Article 10. https://doi.org/10.1155/2019/3728621

Prokopenko, M., Boschetti, F., & Ryan, A. (2009). An information-theoretic primer on complexity, self-organisation and emergence. *Complexity*, *15*(1), 11–28. https://doi.org/10.1002/cplx.20249

Rasmussen, S., Bedau, M. A., Chen, L., Deamer, D., Krakauer, D. C., Packard, N. H., & Stadler, P. F. (Eds.). (2008). *Protocells: Bridging nonliving and living matter*. MIT Press. https://doi.org/10.7551/mitpress/9780262182683.001.0001

Roli, A., & Kauffman, S. A. (2020). Emergence of organisms. *Entropy*, *22*(10), Article 1163. https://doi.org/10.3390/e22101163, PubMed: 33286932

Roli, A., Villani, M., Filisetti, A., & Serra, R. (2018). Dynamical criticality: Overview and open questions. *Journal of Systems Science and Complexity*, *31*(3), 647–663. https://doi.org/10.1007/s11424-017-6117-5

Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of Science*, *10*(1), 18–24. https://doi.org/10.1086/286788

Rota, G. C. (1986). In memoriam of Stan Ulam—the barrier of meaning. *Physica D*, *22*(1–3), 1–3. https://doi.org/10.1016/0167-2789(86)90228-9

Rovelli, C. (2021). *Helgoland: Making sense of the quantum revolution*. Riverhead Books.

Rupe, A., & Crutchfield, J. P. (2024). On principles of emergent organization. *Physics Reports*, *1071*, 1–47. https://doi.org/10.1016/j.physrep.2024.04.001

Sánchez-Puig, F., Zapata, O., Pineda, O. K., Iñiguez, G., & Gershenson, C. (2023). Heterogeneity extends criticality. *Frontiers in Complex Systems*, *1*. https://doi.org/10.3389/fcpxs.2023.1111486

Schmickl, T. (2022). Strong emergence arising from weak emergence. *Complexity*, *2022*, Article 9956885. https://doi.org/10.1155/2022/9956885

Schweitzer, F. (Ed.) (1997). *Self-organization of complex structures: From individual to collective dynamics*. Gordon and Breach.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–424. https://doi.org/10.1017/S0140525X00005756

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3–4), 379–423, 623–656. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). MIT Press.

Standish, R. K. (2003). Open-ended artificial evolution. *International Journal of Computational Intelligence and Applications*, *3*(2), 167–175. https://doi.org/10.1142/S1469026803000914

Stanley, H. E. (1987). *Introduction to phase transitions and critical phenomena*. Oxford University Press.

Steels, L., & Brooks, R. A. (1995). *The Artificial Life route to artificial intelligence: Building embodied, situated agents*. Erlbaum.

Taleb, N. N. (2012). *Antifragile: Things that gain from disorder*. Random House.

Taylor, T. (2024). An afterword to *Rise of the self-replicators*: Placing John A. Etzler, Frigyes Karinthy, Fred Stahl, and others in the early history of thought about self-reproducing machines. *Artificial Life*, *30*(1), 91–105. https://doi.org/10.1162/artl_a_00424

Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., McMullin, B., Packard, N., Rasmussen, S., Virgo, N., Agmon, E., Clark, E., McGregor, S., Ofria, C., Ropella, G., . . . Wiser, M. (2016). Open-ended evolution: Perspectives from the OEE workshop in York. *Artificial Life*, *22*(3), 408–423. https://doi.org/10.1162/ARTL_a_00210, PubMed: 27472417

Taylor, T., & Dorin, A. (2020). *Rise of the self-replicators: Early visions of machines, AI and robots that can reproduce and evolve*. Springer. https://doi.org/10.1007/978-3-030-48234-3

Torres-Sosa, C., Huang, S., & Aldana, M. (2012). Criticality is an emergent property of genetic networks that exhibit evolvability. *PLoS Computational Biology*, *8*(9), Article e1002669. https://doi.org/10.1371/journal.pcbi.1002669, PubMed: 22969419

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, Series 2*, *s2–42*(1), 230–265. https://doi.org/10.1112/plms/s2-42.1.230

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

von Foerster, H. (1960). On self-organizing systems and their environments. In M. C. Yovitts & S. Cameron (Eds.), *Self-organizing systems* (pp. 31–50). Pergamon.

von Neumann, J. (1966). *The theory of self-reproducing automata* (A. W. Burks, Ed.). University of Illinois Press.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Wagner, A. (2005). *Robustness and evolvability in living systems*. Princeton University Press.

Walker, S. I. (2014). Top-down causation and the rise of information in the emergence of life. *Information*, *5*(3), 424–439. https://doi.org/10.3390/info5030424

Walker, S. I., Bains, W., Cronin, L., DasSarma, S., Danielache, S., Domagal-Goldman, S., Kacar, B., Kiang, N. Y., Lenardic, A., Reinhard, C. T., Moore, W., Schwieterman, E. W., Shkolnik, E. L., & Smith, H. B. (2018). Exoplanet biosignatures: Future directions. *Astrobiology*, *18*(6), 779–824. https://doi.org/10.1089/ast.2017.1738, PubMed: 29938538

Watson, R. A., Mills, R., & Buckley, C. L. (2011). Global adaptation in networks of selfish components: Emergent associative memory at the system scale. *Artificial Life*, *17*(3), 147–166. https://doi.org/10.1162/artl_a_00029, PubMed: 21554114

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). *Emergent abilities of large language models*. ArXiv. https://doi.org/10.48550/arXiv.2206.07682

Weinberg, S. (1993). *Dreams of a final theory: The search for the fundamental laws of nature*. Vintage Press.

Whitehead, A. N., & Russell, B. (1910–1913). *Principia mathematica*. Cambridge University Press.

Wiener, N. (1948). *Cybernetics; or, Control and communication in the animal and the machine*. John Wiley.

Wolfram, S. (2002). *A new kind of science*. Wolfram Media.

Wolpert, D. H. (2024). What can we know about that which we cannot even imagine? In M. Streit-Bianchi & V. Gorini (Eds.), *New frontiers in science in the era of AI* (pp. 301–331). Springer. https://doi.org/10.1007/978-3-031-61187-2_15

Wolpert, D. H., & Macready, W. G. (1995). *No free lunch theorems for search* (Technical Report No. SFI-WP-95-02-010). Santa Fe Institute.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, *1*(1), 67–82. https://doi.org/10.1109/4235.585893

Zenil, H. (Ed.). (2013). *Irreducibility and computational equivalence: 10 years after Wolfram's a new kind of science*. Springer. https://doi.org/10.1007/978-3-642-35482-3