

Investigating Biological Assumptions through Radical Reimplementation

Joel Lehman and Kenneth O. Stanley

To appear in: *Artificial Life* journal. Cambridge, MA: MIT Press, 2014.

Abstract

An important goal in both artificial life and biology is uncovering the most general principles underlying life, which might catalyze both our understanding of life and engineering life-like machines. While many such general principles have been hypothesized, conclusively *testing* them is difficult because life on Earth provides only a singular example from which to infer. To circumvent this limitation, this paper formalizes an approach called *radical reimplementation*. The idea is to investigate an abstract biological hypothesis by intentionally reimplementing its main principles to diverge *maximally* from existing natural examples. If the reimplementation successfully exhibits properties resembling biology it may better support the underlying hypothesis than an alternative example inspired more directly by nature. The approach thereby provides a principled alternative to a common tradition of defending and minimizing deviations from nature in artificial life. This work reviews examples that can be interpreted through the lens of radical reimplementation to yield potential insights into biology despite having purposefully unnatural experimental setups. In this way, radical reimplementation can help renew the relevance of computational systems for investigating biological theory and can act as a practical philosophical tool to help separate the fundamental features of terrestrial biology from the epiphenomenal.

Keywords: Artificial Life, Biological Relevance, Evolution, Development, Artificial Intelligence

1 Introduction

A jet engine is not among the endless forms crafted by natural evolution, but its absence is not surprising. Biological evolution is not an engineer, but a *tinkerer* constrained to tweak previous designs incrementally rather than holistically revise them [47]. Thus it is natural that the attractors for biological design, even across divergent lineages, differ from those of more directed design methodologies such as human engineering. Furthermore, while evolution expands to fill reachable niches [90, 13], it is unlikely to exhaustively explore every possible way of filling them [47, 63, 13]. In this way, historical contingency may lead natural evolution to converge to a solution for a particular biological problem, leaving other plausible possibilities unrealized [10, 6]. In total, the effect of these constraints biases evolution to evolve solutions from a systematically reduced subset of all that are possible. Interestingly, because of this systematic bias in how nature generates solutions to biologically-relevant problems, features that are merely incidental to a deeper phenomenon may appear ubiquitous. The result is that inferences based on observing only biological examples may be misleading [53, 101, 85].

This insight is a central motivation for the field of artificial life (alife; [53, 11]), which aims to synthesize and explore life in all possible instantiations, thereby shedding light on the general regularities of life. A common approach is to seed computational simulations with simple self-replicating organisms in the hope of facilitating open-ended evolution of increasingly complex and diverse organisms [53, 86]. Yet a danger in such an approach is that directly transplanting the mechanisms of terrestrial life into computational mediums may not help isolate the most general principles of life. The problem is that the transplanted mechanisms themselves might be incidental to biological evolution rather than universal across all *possible* forms of evolution. In other words, artificial life may sometimes not be artificial enough.

As a helpful analogy, consider an isolated man on an island trying to understand flight. One avenue towards such understanding would be inference from biology. If he observed only birds, the man might conclude that flappable wings with feathers are essential to flight. However, after seeing a bat flying, that same man could safely conclude that because bats do not have feathers, feathers must not be essential after all. Yet it might still seem to him that *flappable wings* are [65]. Thus even with multiple examples of the same phenomenon, direct inference can remain deceptive.

To validate his understanding of flight, the man might synthesize a flying object made of inorganic material from his current best understanding of flight’s principles (i.e. flight as a result of flapping wings). Yet even if such synthesis was successful, it would not isolate the general principles of flight, which run deeper than wings. To take this particular example to its logical conclusion, imagine instead that the man was able to synthesize a working helicopter. The helicopter’s mechanism of flight, by its radical divergence from previously observed examples, provides evidence that the man has arrived at an understanding of flight deeper than wings.

This example illustrates the value of divergent implementations for isolating and validating basic principles. Importantly, inferring the deeper principles of aerodynamics, which is easier given the helicopter, helps to unify understanding of all artificial and biological instances of flight. At the same time, the helicopter also serves to vali-

date aerodynamics by demonstrating what engineering with such knowledge enables: agile powered flight for human transportation.

Returning from the analogy, the implication is that synthesis of life in divergent media may not always illuminate general principles of life beyond only what has been observed on Earth; indeed, synthesized models often include many mechanisms directly imported from biology. These insights thus motivate an approach to investigating biological principles through alife that can be called *radical reimplementation*. The main idea in radical reimplementation is to craft an abstract hypothesis about the crux of a given biological principle, and then to test this hypothesis by reimplementing the abstraction in a way that is maximally different from how it exists in nature, yet that still exhibits biological similarity in its effects. Just as a helicopter’s significant divergence from a bird yields valuable perspective on the nature of flight, a successful radical reimplementation of any other biologically-relevant principle may also more clearly expose the core of the underlying phenomenon than would considering additional biological examples; hence the need for investigations outside nature.

While alife is a relevant field of research for such investigations, a difficulty for alife researchers seeking to be relevant to biology is to demonstrate that their models relate convincingly to biological truth [101, 11, 45]. To achieve such relevance, researchers often argue that the abstractions made by their models are principled, i.e. that the biological details filtered out through abstraction were non-essential. In other words, alife researchers often minimize and defend deviations from nature [106, 29].

Thus another way to view radical reimplementation is as a principled alternative to the tradition within alife of *minimizing deviations from nature*. While the differences in models derived from nature must usually be well-motivated, explicitly *maximizing* divergence can be equally as principled as minimizing and justifying it. Interestingly, although an unofficial motto of alife is “life-as-it-could-be,” to explore beyond “life-as-it-is-known,” this “could-be” is not often a radical departure from natural mechanisms.

In this way, the advantage of the radical reimplementation approach is that it escapes the underlying restrictiveness of attempting to mimic nature directly. It is not necessary to claim that a model is intrinsically biologically plausible; instead the idea is to provide radically novel examples of phenomena usually associated with biology. In short, the radical reimplementation approach challenges researchers to maximize the divergence from the workings of terrestrial biology while instantiating qualitatively similar output. When successful, such reimplementations indirectly provide evidence for the underlying principles guiding their creation.

While some reimplementations may be guided only by domain-specific principles (e.g. a child folding a paper airplane is unlikely to be inspired by the core mechanisms of flight), if such underlying principles are general enough also to encompass natural examples, they can provide a hypothesis about the root mechanism of a biological phenomenon. Radical reimplementations can then provide evidence for or against the underlying hypothesis. That is, a maximally-divergent reimplementation of a plausible yet *entirely false* biological hypothesis is unlikely to produce results qualitatively similar to biology purely by coincidence. For example, incorrectly inferring rules of flight from birds is unlikely to lead to engineering a passenger jet. In this way, beyond providing an additional and less-biased example from which to *infer* general biological

principles, a well-motivated radical reimplementation may itself also *provide* abstract evidence for a particular biological hypothesis.

In this paper the radical reimplementation approach is illustrated by reviewing four examples of existing research in the spirit of radical reimplementation. Examples are guided by abstractions general enough to explain both biological and artificial instances of a particular biological phenomenon. In this way, the reviewed examples can be interpreted as counterexamples designed to probe the validity of fundamental biological assumptions.

In particular, the examples review radical reimplementations of diverse biological phenomena, including intelligence, self-reproduction, development, and evolution. In each case, the radical reimplementation perspective allows exploration of biological questions through biologically implausible models. In this way the examples showcase the ability of radical reimplementation to ask questions about seemingly fundamental biological phenomena and to potentially help to illuminate the most general truths underlying specific biological processes on Earth. Importantly, deeper understanding and progress in science often results from reassessing and overturning widely-accepted assumptions. Thus new philosophical tools for investigating such assumptions in a principled way, like radical reimplementation, may facilitate such progress.

2 Background

To better understand the context motivating the radical reimplementation approach, the traditional approach to alife and its reception within the field of biology are reviewed in the next sections.

2.1 Artificial Life

The field of alife is diverse; it studies the potential for creating novel forms of life in all mediums. Its ambitious aims are to study “artificial systems that exhibit behavior characteristic of natural living systems,” to explain “life in any of its possible manifestations,” and ultimately to “extract the logical form of living systems [53].”

Interestingly, biologists, especially philosophers of biology, also have aimed to uncover the most general rules governing life and its evolution. However, while alife aims to *synthesize* new life to illustrate such rules, biologists generally proceed instead by *extrapolating* from terrestrial organisms. A potential obstacle to accurate extrapolation is known as the $N = 1$ *problem* [101, 53]: From a statistician’s point of view, biologists must infer general rules for life everywhere based on a sample size of only one (life on Earth), which may be misleading. As an analogy, for someone unfamiliar with a deck of playing cards it would be impossible to infer the structure of the entire deck, i.e. all of the various suits and ranks, from only observing the four of diamonds. The implication is that if life varies dramatically over its possible realizations, our observation of Earth alone may not be sufficient to extrapolate life’s general structure. In this sense, alife offers a means of expanding this sample size through synthesis: “The ideal experimental evolutionary biology would involve creation of multiple planetary systems, some essentially identical, others varying by a

parameter of interest, and observing them for billions of years. A practical alternative [...] is to create synthetic life in a computer [85].”

This idea of synthesizing and simulating life within a computer is one of the central threads within alife, underpinning such experiments as self-reproducing computational entities [52], three-dimensional virtual creatures [93], and simulated petri dishes of digital organisms [19, 79, 62]. The general approach is to craft algorithmic abstractions of biological systems, to run virtual experiments, and then to investigate the products. One advantage of computational alife is that it is not limited to mathematically tractable models, which often force unrealistic assumptions onto models in fields such as theoretical biology. Thus some propose that alife is best applied in the spirit of biological models but with such unrealistic assumptions relaxed [75].

Yet this approach to alife, by constraining it to be *like life as it is on Earth*, inherently conflicts with one of alife’s central objectives: to be the science of “life as it could be” [53]. Indeed, a common dismissal of some alife experiments is that they are too engineered [104], i.e. not feasible to have *emerged* naturally. This criticism of engineered approaches overlooks the potential advantage of testing hypotheses through situations impossible in nature and through counterexamples. While if too much is engineered with a top-down approach, the results may have little to do with natural *emergence*, they might yet still imply much about life. That is, designating an alife model to be biologically relevant only when a similar system could have feasibly arisen on its own from nature is unnecessarily limiting: Human-engineered life may operate in vastly different ways than naturally occurring life without loss of generality. After all, diverse examples of a phenomenon may more easily reveal the deepest principles underlying it than more homogeneous examples; this observation motivates considering engineered alife approaches as a means of probing fundamental questions about life. Thus radical reimplementation can be seen as a rationale and defense for engineering-based alife, which is often dismissed as biologically implausible. For example, a traditional alife approach might show that if specific conditions similar to those observed in nature are met, that an evolutionary algorithm will lead to the evolution of cooperation. In contrast, an engineering-oriented radical reimplementation might instantiate conditions entirely divergent from nature yet in which cooperation *also* evolves.

Of course, many of the grand philosophical questions alife seeks to answer were investigated first by philosophers of biology. Furthermore, the achievement of alife’s most ambitious goals would deeply impact our understanding of biology. Because of this synergistic overlap, one might expect the study of alife to be generally well-received and perceived as important within biology. However, as reviewed in the next section, there is continuing controversy in biology over alife’s relevance.

2.2 Perception of Artificial Life within Biology

The opinions of scientists in a particular field about the rigorousness or importance of the science done in other fields may sometimes derive from implicit generalizations. For example, one relevant such criterion is whether the object of study is physically real or if it is instead abstract or simulated. Scientists studying “real” phenomena may be skeptical of experimental results from simulations or abstract models because

while the simulated result is dependent upon the integrity of the designed simulation, the integrity of the real world itself is rarely in question. In other words, an extra step is necessary to justify simulated results.

It may be for this reason that many researchers in biology who are accustomed to measuring physical quantities associated with real organisms tend to criticize alife, which often measures quantities that are the result of a computational simulation. This sentiment is illustrated by a quote from John Maynard Smith: “I discuss below a particular example of a dynamic system, ‘Turing’s morphogenetic waves,’ which gives rise to just the kind of structure that, as a biologist, I want to see. But first I must explain why I have a general feeling of unease when contemplating complex systems dynamics. Its devotees are practicing fact-free science. A fact for them is, at best, the output of a computer simulation: it is rarely a fact about the world. [70]” Dawkins has similarly discussed the difficulty involved in creating biologically accurate simulations, although he concludes with more optimistic comments on the ultimate prospects of alife [29].

However, while validating alife experiments requires an additional step not often needed in biology experiments, this additional burden in itself is not sufficient to dismiss alife outright. That is, it would be parochial to suggest that how science is done in one’s particular field must necessarily imply that any deviation is wrong. For example, a physicist could similarly criticize biology on the same grounds because physics experiments generally do not depend on geography, but a biologist’s results could, due to different distributions of wildlife. So a biologist’s results might require an additional step (from the physicist’s view) of normalization or justification. The particular methodology, as long as it is sound, is not the core of the issue. Ultimately what is important about any scientific experiment, regardless of the field, is whether it is rigorous and principled; rules of thumb like physicality versus simulation thus should not displace less formulaic human analysis of the quality of the experiment.

Indeed, a stronger attack on the relevance of alife is that its researchers potentially generally practice sloppy science. Alife’s reliance on simulations may make conducting principled experiments more difficult because of the challenge in justifying and validating such simulations. That is, there may be hundreds of parameters in a simulation that must be carefully chosen and defended. Yet it is not impossible: Dawkins has argued the merits of the NetSpinner models [29], and the Avida models of computational organisms are generally well-received within biology [62, 61, 36, 22, 23, 79, 19, 106]. Furthermore, the use of simulations derives from the use of models in general, which are prevalent in theoretical biology. The main difference is that alife simulations are almost always too complex for rigorous mathematical analysis. However this complexity is becoming less problematic: Increases in computational power have enabled empirical analysis of alife simulations through large numbers of independent trials with varying parameters [1, 78]. Such large-scale experiments can help demonstrate the robustness of results, i.e. that the conclusions of the experiment do not result trivially from overtuning of specific parameter settings.

However, it is true that some work in alife, just as in other fields, might fall prey to the temptation to promote conclusions overreaching the true implications of the presented evidence [101]. Such overreach is particularly easy in a field like alife in which there is often no inherent hard *grounding* of the science to reality. For

example, alife researchers can make abstract claims about what the results from their simulations represent, while a scientist who claims to have a method for real-world robotic bipedal locomotion would eventually have to back such claims with an actual walking robot.

In conclusion, the overall difficulty facing researchers in alife seeking relevance to biology is in producing relevant conclusions and sufficiently justifying their simulations and conclusions in a convincing way. Oftentimes the simplifications and assumptions made in complex alife systems have a particularly ad-hoc feel. While the results of such models may still be interesting and thought-provoking, the link to reality and biology is easily seen as tenuous, especially within the field of biology; it is unclear how the results should be interpreted or what (if anything) they generally imply. However, one unorthodox way to circumvent this issue is *not to claim* that a model reflects terrestrial biology or that it is biologically plausible. That is, a model might shed light on hypotheses related to biology even if its mechanisms are implemented in a way unrelated to or inconsistent with biology. While this step might seem a strange path towards establishing biological relevance, the next section explores heuristics for judging abstractions by this kind of indirect evidence.

3 Heuristics for Evaluating Abstractions and Reimplementations

Because a central motivation of this paper is to aid the search for deep and elegant abstractions of natural phenomena, it is important to describe clearly what criteria are considered to evaluate them. In addition, because the radical reimplementation approach requires reimplementations to diverge significantly from nature, it is also important to describe what constitutes radical divergence.

3.1 Evaluating Abstractions

Intuitively, a deep abstraction concisely captures the crux or essence of a phenomenon. It neither abstracts away what is essential nor includes details that merely coincide with examples of the phenomenon. In contrast, a more superficial abstraction focuses on non-causal (but perhaps seductively intuitive) features of a phenomenon that may *mask* more general principles.

An alternative way of viewing a deep abstraction is as a hypothesis for a root causal mechanism that underlies all examples of a particular phenomenon. For example, lift, a deep abstraction of flight, can also be seen as an abstract mechanism that potentially unifies all biological and artificial instances of flight. That is, whatever the specific concrete mechanism, wherever lift overpowers gravity’s pull, the result is flight. Importantly, because an overly coarse abstraction might trivially unify examples of a phenomenon, the value of an abstraction or root mechanism cannot be judged only by its ability to unify examples, but also by how well it captures the crux of the phenomenon. However, quantifying this intuitive concept is difficult. To circumvent such difficulty, the approach in this paper is motivated by the idea that

synthesis through engineering can demonstrate understanding. In this way, engineering can enable indirect empirical investigation of the quality of a proposed highly abstract mechanism.

More concretely, the suggested approach for evaluating abstractions relies on simultaneously maximizing two competing criteria: A deep abstraction of a phenomenon (1) facilitates thorough exploitation of that phenomenon through human design and (2) elegantly unifies natural and artificial examples of the phenomenon. The first criterion suggests that an abstraction’s quality can be judged by the engineering possibilities enabled through exploiting it, relative to other abstractions. For example, understanding aerodynamics hints at powered flight beyond the natural example, like passenger jets, which are not similarly enabled by viewing flight as a product of flapping wings. In a similar way, engineering or synthesizing divergent examples of other biological principles concretely demonstrates deeper understanding by grounding the guiding abstractions in reality. The second criterion is that a deeper abstraction will concisely explain more examples of that phenomenon (both natural and artificial) when compared to competing abstractions. Importantly, the requirement of ranking highly on both heuristics circumvents limitations in each taken separately. For example, abstracting away too much detail may trivially unify all examples of a phenomenon, and yet exploiting such an uninformative abstraction is unlikely to lead to greater engineering possibilities. Conversely, if a particular engineering approach enables new possibilities but does not also explain biological examples, the implication is that there may exist a yet deeper principle underlying both.

In this way, existing biological examples and divergent engineered ones can together provide evidence that a phenomenon is well-understood and that the abstraction exploited to accomplish such engineering is deep. Naturally, engineering-based exploitation of abstractions typically serves to bend the underlying phenomenon to meet human needs. Thus the quality of a particular abstraction may be estimated by the degree to which it facilitates such bending in many contexts while also explaining natural instances of the phenomenon. Note that if a particular phenomena has no practical implications, it can still potentially be bent through human engineering to demonstrate that its cause is well-understood.

3.2 Evaluating Divergence between Implementations

An important practical consideration for the radical reimplement approach is defining what constitutes radical divergence between implementations. While intuitively one can appeal to the difference between biological flight and a helicopter as an example of what is meant by radical divergence, more explicit definitions can serve to reduce subjectivity and clarify basic assumptions. For this reason, this section elucidates such measures of divergence.

Such elucidation should help to clarify a central claim of this paper, that alife reimplementations of biological phenomena may often not diverge significantly enough from terrestrial example to provide insight into universal biological principles. In particular, a common strategy in alife is to abstract biological systems into simpler, more tractable models. As more details are abstracted away, the model increasingly differs from the underlying system being modeled. In this way, a model’s level of

abstraction can be viewed as one dimension of distance between the model and its biological inspiration: A more abstract model of a system diverges further from the underlying system than does a more concrete one.

The motivation for abstract models is to filter away as much non-essential detail from the underlying biological system as possible, in order to isolate its essential features. Importantly, such an approach may isolate the essential features of the particular studied *biological* system while still failing to isolate the essential features of the underlying *general* phenomenon. In terms of the flight analogy, one might isolate and understand the specific principles of winged animal flight without grasping the greater generalities of universal flight.

To take an example more relevant to alife, artificial neural networks (ANNs) are computational abstractions of biological brains that apply simplified models of neurons [71]. The idea is to realize intelligent behavior by interconnecting many instances of the simplified neuron models. One popular ANN model abstracts a biological neuron as a weighted-sum unit [71], ignoring many specific biological features such as ion channels and neuromodulatory chemicals. In this way, such a model further diverges from biological example than does a model with greater biological detail. However, a weighted-sum unit abstraction of a neuron still implements key features directly imported from biological neural networks (e.g. neural networks are a collection of interconnected simple processing units), which naturally limits the generality of questions about universal biology that such a model can reliably help answer. Thus, any ANN model would likely be poorly suited to investigate assumptions about the necessity of neural networks for facilitating animal-like intelligence in life of any form, e.g. are neural networks a key feature of all possible evolved intelligences, or are they merely an artifact of the particular evolutionary trajectory taken by Earth’s biology?

The problem is that appealing only to an ANN model to explore a question that encompasses both ANN models and other alternatives ignores the existence of the other possibilities: ANNs cannot address whether substrates for intelligence *other than* neural networks might be equally or more effective. More broadly, it is challenging for abstract models based only on biological examples of a particular phenomenon to convincingly characterize an entire class of phenomenon in general. The conclusion is that the possible divergence between existing biological examples and an artificial reimplementation is limited if the reimplementation is based only on increasing levels of abstraction.

Given the ubiquity of alife models that directly abstract biology at a high level, limitations of such an approach are problematic for alife’s quest to answer questions about life independently of any particular form. For example, alife experiments exploring general conditions for the evolution of important classes of behaviors (e.g. cooperation, altruism, sexual reproduction) often take direct inspiration from terrestrial examples or from existing understanding inferred from biology. Because so much depends on abstracting at a high level from existing example, there is no guarantee of separating the essential features from the incidental. Thus there is a need for divergent models (and measures of divergence) based on principles other than levels of abstraction alone.

In particular, one way to isolate an abstraction of a phenomenon from its biological inspiration is to explicitly reimplement it in a maximally divergent way. That is, a dif-

ferent measure of distance between two implementations of an underlying abstraction is how few mechanisms they share in common. That is, a helicopter achieves flight by mechanisms much divergent from a bird: Fast-spinning rotors have little in common with flapping wings. In contrast, an ornithopter is more directly motivated by bird wings and implements many of the same mechanisms, even if they are implemented in a different material or lack feathers. In this way, what is meant by radical reimplementation is an implementation guided by an abstraction of a biological phenomenon that is nevertheless realized by as few relevant shared mechanisms as possible with existing biological examples.

Note that the claim here is not that radical reimplementation is the *only* sound approach to alife experimentation; instead, the idea is that radical reimplementation provides a useful philosophical tool for isolating an abstract mechanism from its particular biological instantiation, thereby avoiding the influence of many potential confounding factors.

4 Radical Reimplementation

This section describes the radical reimplementation approach in more detail and then illustrates its utility through case studies of its application in practice.

A central analogy in this paper connects thinking about flight specifically to biological principles in general. The insight is that inferring too directly from nature's examples can lead to deception. That is, many features that superficially appear essential to life may reflect only one possible implementation among many of a *deeper* but less obvious concept. Problematically, because this deeper concept may be counterintuitive, imagining any such alternative implementations may be difficult as well. Thus a tempting assumption is that ubiquitous and obvious features observed in nature are *universal*, i.e. characteristic of all possible life in general. While such assumptions require verification, it is difficult to explore their validity when there is only one example (i.e. life on Earth) from which to infer. Even where Earth provides multiple instances of a biological phenomenon, these instances all are drawn from the same biased sampling method (i.e. natural evolution). Such bias may often cause evolution to converge to similar implementations of a phenomena and prevent it from reaching others; for example, evolution is unlikely ever to invent a large-scale jet engine. Thus, just as a helicopter is more informative than a bat in broadening inferences drawn about the nature of flight, the approach highlighted in this paper advocates creating maximally divergent examples of biological principles to facilitate stronger inference. In this way, radical reimplementation can act as a philosophical tool for investigating whether rarely questioned biological features are masking more fundamental explanations.

While such investigation may seem like a promising application of alife in general, an issue with many alife models is that because they are so directly inspired by mechanisms in nature they may not illuminate the general phenomena they attempt to explore. For example, inventing a mechanical bird with flapping wings demonstrates less understanding of flight than does inventing a helicopter; the reason is that successful synthesis through a more general mechanism demonstrates greater mastery than

does simply applying the most immediately evident concrete mechanism in a new context. In other words, evidence for understanding a general principle is better provided by a *radical reimplementation* than by a rote recreation. Although artificial life models are rarely rote recreations, such models often focus on biological plausibility and transplanting existing biological understanding into a new (often computational) medium. The danger of such transplantation is that features incidental to biology may be included; in this way alife can sometimes run the risk of assuming what it sets out to prove.

Thus another benefit of viewing results through the lens of radical reimplementation is that such an approach requires stating explicitly what is being reimplemented, which better frames how the results should be interpreted. For example, an experiment focusing on development might explore a radical reimplementation of a *developmental process*, and embed that radical reimplementation within a more conventional *evolutionary process*. Thus while the results of such an experiment can potentially shed light on the generalities of development, they may not reflect much about generalities of evolutionary systems. In other words, an experimenter may radically reimplement only what is necessary to explore the biological hypothesis being explored. In this way, this approach can clarify for which aspects of universal biology an experiment may provide evidence.

Given the potential benefits of radical reimplementation, an important practical question is how to create such a model (figure 1). The first step is to examine a biological principle of interest and to abstract it at a high level, that is, to formulate a hypothesis on the root mechanism behind a particular biological feature. From such a hypothesis, the next step is to synthesize a concrete implementation of the core mechanism that deviates significantly from natural examples. In other words, the idea is to return from the highly abstract to the concrete, but to land in as different a place as possible from the starting point. In particular, the reimplementation should share in common as few concrete mechanisms as possible with its biological inspiration. For example, if inspired by birds to form the abstract hypothesis that lift is central to flight, the radical reimplementer might create a helicopter that produces lift by blades that rotate at high speed about a central mast. If such a helicopter can be made to fly, then by virtue of its deviation from biology it can provide evidence for the underlying hypothesis about lift. Additionally, evidence for the abstraction’s generality can be established by the vast diversity of aircraft that an understanding of lift facilitates (e.g. propeller planes, jumbo jets, and helicopters). In this way, a successful abstraction both unifies understanding of all examples of a phenomenon and facilitates different applications and implementations.

The promise of this approach is that it can facilitate probing deep explanations of biological principles. The next sections illustrate the potential of radical reimplementation by reviewing examples of alife research through its lens.

4.1 Questioning the Necessity of Representations for Intelligent Behavior

Intelligent behavior is characteristic of animal life on Earth. Although its realization is not fully understood, such biological intelligence is largely enabled by brains composed

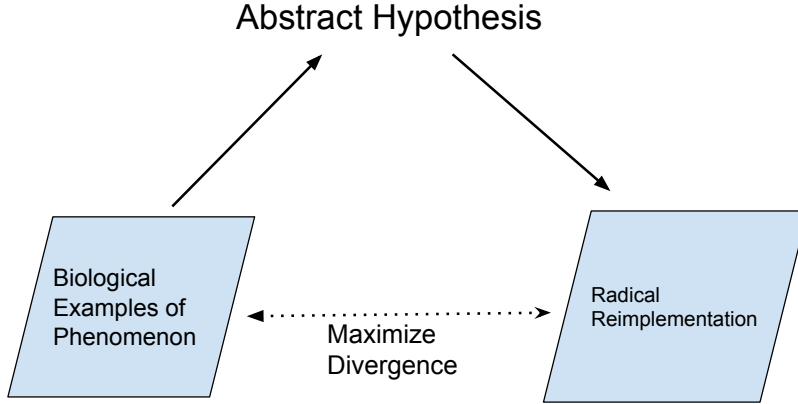


Figure 1: The radical reimplementation approach. To radically reimplement a particular biological phenomenon requires first creating an abstract hypothesis about the root mechanism enabling that phenomenon. Then, the abstract hypothesis is reimplemented in such a way to maximally diverge from existing biological example.

of many interconnected neurons. While the functionality of individual neurons is becoming more well-understood, the abstract principles guiding how large networks of such neurons enable intelligence is still in part elusive. Furthermore, beyond the singular biological anecdote provided by neural networks, an open question is whether there exists more general abstract principles for constructing intelligence.

Answering this question would lead to a better understanding of the nature of intelligence, and could also enable engineering intelligent machines that might benefit humanity and industry. As a result, many fields, including cognitive science, alife, and artificial intelligence (AI) explore how to understand and recreate intelligence computationally. In particular, to facilitate such intelligent machines it may be important to understand what kinds of general computational architectures can enable intelligent behavior.

That is, there are many different ways to approach modeling intelligence. One possibility is that an essential component of intelligent behavior is building an internal model of the world and thereby explicitly *representing* world knowledge. In other words, it is possible that increasingly intelligent behavior requires explicit mental representations of the world and its state, which can be leveraged through planning to achieve desired goals. If this view is true, any productive AI approach would need to provide mechanisms to implement or realize such internal models.

Indeed, guided by intuitions about human thought, early AI researchers focused largely on AI algorithms and architectures built on centralized control, planning, and symbolic manipulation of knowledge [18]. Yet an interesting but counter-intuitive hypothesis is that intelligent behavior is possible *without* representation of the world and its state. Evidence for such a hypothesis is given by the complex behavior of insects,

which is realized through relatively simple neural networks likely unable to implement rich models of the world. However, the principles connecting biological neural networks to intelligent behavior are not conclusively understood, which complicates exploring the hypothesis directly in biology.

Rather than analyze nature, an alternative approach is to create artificial examples through engineering. In particular, if an engineered implementation achieves intelligent behavior without including the capacity for world modeling or knowledge representation, it provides a counter-example to the hypothesis that representations are strictly necessary for intelligence. An engineering approach that can be viewed in such a light is the subsumption architecture [16, 14], a methodology that prescribes decomposing complex behavior into layers of simple computational behaviors that *subsume* each other. As an example of how radical reimplementation can frame experimental results to yield potential insight into universal biology, the next section reviews results from research on the subsumption architecture through the lens of radical reimplementation.

4.1.1 Subsumption Architecture

It is possible that in nature, the evolution of intelligence proceeded through building networks of behaviors that are each simple in construction. In this view, higher-level behaviors are enabled through leveraging and subsuming lower-level behaviors. Such a possibility is supported by the Turing-complete substrate for complex computation provided by layered recurrent neural networks [92]. However, the computational abilities of such neural networks also encompass building world model and representing knowledge. Thus a more convincing existence proof for intelligence without representation is possible through programs implementing intelligent behavior but *without any possibility* for representing the world. In this way, it is possible for a radical reimplementation of intelligence to illuminate the underlying hypothesis that such world models are not necessary for realizing at least some level of intelligent behavior.

One such reimplementation is provided by the subsumption architecture [16, 14], which controls robots through human-engineered layers of augmented finite state machines (AFSMs). Each AFSM is composed of a small number of states that the machines can transition between, one or two internal registers, and an optional timer. Additionally, the AFSMs are able to perform simple computations such as vector sums. The machines run asynchronously and can communicate with other AFSMs by sending fixed-length messages (often only one bit long).

Each of the AFSMs is designed to implement a particular behavior, and is connected by messages to other relevant behaviors in a fixed-topology network. Transitions between states in each AFSM are triggered by timers or by messages received from other AFSMs. Importantly, the system is designed such that no particular AFSM can be aware of the global state of the system, nor can any particular AFSM send messages that control all other AFSMs. Furthermore, because the computational abilities of each AFSM is limited, no AFSM on its own can implement a world model or retain much information about the state of the world. Because of such intentional limitations, any intelligent behavior resulting from the subsumption architecture cannot be the product of an internal world model or explicit representation

of information about the world. Note that while superficially similar to ANNs (e.g. AFSMs are also computationally simple and interconnected), neurons in ANNs are homogeneous and are designed to provide a substrate for distributed computations. In contrast, each AFSM is hand-engineered to implement a particular behavior, and interconnections between AFSMs thereby do not permit a universal mechanism for distributed computation, but instead only inhibit or influence other simple human-engineered behaviors.

Despite their lack of central control, models of the world, or explicitly represented knowledge, implementations of the subsumption architecture have successfully realized diverse autonomous behaviors in complex dynamic environments [14, 15, 25, 26]. In addition, subsumption architecture has proved productive in many extensions, such as when combined with evolutionary algorithms [105, 49], reinforcement learning [67], logical inference [2], or when applied to atypical domains such as generating musical accompaniment [17]. Finally, as evidence of its robustness even in practical commercial applications, the subsumption architecture also enables autonomous consumer robots such as the Roomba [51] that automatically vacuum rooms. These results serve to provide indirect evidence for the hypothesis underlying the radical reimplementation: Some aspects of autonomous intelligent behavior can be constructed without internal models or explicit knowledge representation. In this way, this example demonstrates the potential for viewing results through the lens of radical reimplementation.

4.2 Questioning Genetic Regulatory Networks in Biology

Life on Earth depends upon regulatory networks realized through nucleic acids and proteins. That is, cells in nature function and self-reproduce through genetic programs implemented through gene regulatory networks (GRNs; [28]). An interesting biological question is whether evolutionary regularities inferred from life on Earth are specific to such self-reproducing networks, or whether they reliably generalize across evolving self-reproducing entities embedded in other substrates.

Such a question is difficult to answer by extrapolation from existing biological examples because there is no substantial variety in the root mechanisms by which life on Earth functions and reproduces: All known life depends upon GRNs based on nucleic acids. In this way, inferring principles about the necessity of GRNs from biology is subject to the $N = 1$ problem (recall that this problem implies that inferences made from a sample size of one are often misleading [101, 53]). Thus an intriguing possibility is that this problem can be circumvented by radically reimplementing self-reproduction through alife such that it holistically differs from biological self-reproduction on Earth. Interestingly, many alife implementations of hereditary encodings may not equally aid in isolating the generality of existing evolutionary laws. For example, many approaches in alife and evolutionary computation abstract away self-reproduction entirely, relying upon an explicit evolutionary algorithm to copy and breed organisms externally from the organisms themselves. While useful for many other purposes, such an abstraction prevents studying how evolutionary regularities may emerge from the dynamics of self-reproduction. Thus, what is needed to explore the hypothesis that evolutionary laws do not depend upon GRN-based

self-reproduction is an alife approach that reimplements self-reproducing entities in an alternative substrate.

4.2.1 Avida

An interesting possibility for an alife radical reimplementation of self-reproduction is self-copying computer programs. That is, given a sufficiently expressive instruction set, a computer program can be written that replicates itself through copying the data that composes it. Indeed, such a replication mechanism enables destructive computer viruses, although they typically do not evolve like biological viruses. However, it is also possible to create computing environments in which self-reproducing digital organisms evolve [86, 61], driven by pressures to replicate faster [86] or to successfully perform desired computations [61]. Because of their significant divergence in function and means of self-reproduction, such digital organisms provide a suitable means to radically reimplement the hypothesis explored in this section.

In particular, the Avida alife platform [61] is the most widely-applied evolutionary model for digital organisms. In Avida, each digital organism is located in one cell of a two-dimensional grid, and is composed of a genetically-represented computer program and a simple simulated CPU. Evolution is seeded with a simple self-replicating computer program that is written in a turing-complete assembly language with 26 instructions. Importantly, the assembly language is designed such that mutations of functioning programs often remain viable. The CPU has access to its internal memory (in which the program is stored), several registers for holding intermediate results, and input and output buffers for interacting with the digital organism’s environment. Self-reproduction is enabled in Avida through individual instructions that allocate additional program memory, that copy one instruction in memory to another memory location, and that divide an organism into two separate organisms. Finally, digital organisms can reproduce more quickly by gaining additional CPU time. In particular, experiments in Avida are designed such that whenever desired computations are performed by an organism, its execution speed increases. Taken together, the features of Avida instantiate the potential for open-ended evolution of digital organisms.

In this way, it is possible to view experiments in the Avida alife platform [61] as radical reimplementations of self-reproductive entities. That is, in Avida, the behavior and replication of organisms is based on a digital computer instead of on GRNs. Despite such divergence, and with few domain-specific modifications, experiments in Avida have led to demonstrations of how complex features can evolve [61], how sleep can be adaptive [7], and how motility can evolve [40]. Furthermore, Avida has practical applications in software development [72], asynchronous sensor communication [72], and energy management [72]. In this way, the diversity and quality of results from Avida gives indirect evidence that many evolutionary patterns observed in natural biology may generalize to classes of self-reproductive entities much different from biological GRNs.

4.3 Questioning Development

Development is among the most prominent features of complex biological organisms. An important biological question is, what is the most fundamental principle behind development that we can identify to explain its central role in the realization of phenotypic form and shape? More specifically, what features of biological development enable compactly representing highly complex, functional, and evolvable organisms? That is, the vast complexity of the 100 trillion cells that compose an adult human are organized over time through biological development and are efficiently encoded by less than 50,000 genes [102]. Furthermore, such development has been much modulated by natural evolution to enable the vast diversity of complex multicellular life on Earth. Importantly, identifying the most general laws underlying development might facilitate human engineering of similarly complex artifacts, which would have profound implications for society.

However, when contemplating development, the danger is that superficial aspects of the process might seemingly provide plausible answers to philosophical questions about the deepest principles governing it. On its surface, development generally begins with a single cell that then repeatedly splits and differentiates into the multicellular adult form, e.g. a human adult develops over time from its zygote origins. Notably, during development each cell exploits chemical gradients and local signals from its neighbors to determine its identity and further differentiate. From this traditional perspective, the most intuitive inference is that *temporal unfolding* and *local interaction* may be the key principles that *explain* the expressive power of developmental processes. In fact, the process of development as growth over time is so familiar in biology that even questioning the necessity of such ubiquitous features may seem at first entirely pointless or misguided. Yet it is interesting to consider whether development in the traditional biological sense might be only a particular implementation detail that masks a deeper principle.

Problematically, investigating this possibility is challenging through inference from biology because of biology’s singularity: Biological life has mostly converged to a singular overarching means of development and cellular replication. Thus inferring principles about development’s necessity or purpose from biology is subject to the $N = 1$ problem, which can be potentially be circumvented again through radical reimplementation. Interestingly, most existing alife approaches to development may not equally aid in isolating its general principles. For example, while much work in alife has explored alternative developmental schemes [99, 12, 41, 33], nearly all of these schemes operate fundamentally in the spirit of natural development, i.e. one cell-like entity gives rise to many more through an iterative process of splitting and differentiation, or at least growth [99, 64, 93, 41, 33, 12]. While such approaches are thought-provoking and can illuminate some features of development, because of their *similarity* to terrestrial biology it is not always clear what insights they bring to *isolate* the deeper principles behind development.

For this reason it may be more informative to investigate models through the radical reimplementation approach to alife. Thus, if an interesting hypothesis can be derived for the root mechanism of development, it can be empirically investigated by reimplementing it in a way purposefully divergent from biological development.

In particular, if the results of reimplementations exhibit similar characteristics to the biological outcomes, the abstracted root mechanism gains credibility.

One such high-level hypothesis for the root mechanism of development is that it is a means for expressing an organism as a *function of geometry*. In other words, the fully-developed form of an organism can be understood as the product of an abstract function that maps from points in three-dimensional space to the type of cell (if any) that should occupy each point in the space. The main motivation for this more broad abstraction is that the most important aspect of development is the final pattern that it generates, and that development over time is only one way to realize such patterns. The implication is that there may be many such ways and that focusing solely on how it is accomplished in terrestrial biology may be misleading.

While this perspective departs from the more common (and intuitive) view of development as a temporal unfolding process that relies upon local interaction, there are reasons to pursue such an abstraction. For example, that such a functional relationship is possible at all is supported by the universal approximation theorem [27], which establishes that a series of enough nested functions can approximate *any* pattern (including biological patterns such as the phenotypic form of an organism). However, though they thus can clearly approximate phenotypic form, whether nested functions of geometry are a *productive* way to view developmental patterns is a separate question. Interestingly, it has been demonstrated that many heavily conserved genes active in development work to establish chemical gradients that act as nested *coordinate frames* that provide context to a particular cell about its role within the organism [84, 74]. This observation supports the idea that establishing nested patterns is an important function of development.

Because nested coordinate frames appear to be an important feature of development, an interesting hypothesis is that the formation of such nested frames may be the main abstract mechanism of development. In other words, development in nature may be only one way among many to approximate an abstract mapping from geometry to form through a series of nested functions, and such mapping may be what enables the compact representation and significant evolvability of multicellular organisms in nature. While this hypothesis might appear tenuous in the context of biology alone, it is possible to support it through empirical investigation, i.e. through a radical reimplementations of the hypothesized abstraction of development as a mapping from geometry to form.

4.3.1 Compositional Pattern Producing Networks

One way to reimplement this abstraction that differs significantly from development in nature is *explicitly* to represent development as a function of geometry. This approach is embodied by *compositional pattern producing networks* (CPPNs; [97]). The CPPN is a variant of a traditional ANN that composes a set of functions together in potentially complex nested ways to produce a pattern when queried over some input geometry (e.g. a two-dimensional coordinate frame). The basic idea is to represent a phenotype as a function over a scalar field. In this way, CPPNs can represent complex patterns from nested coordinate frames in a form much divergent from biological development.

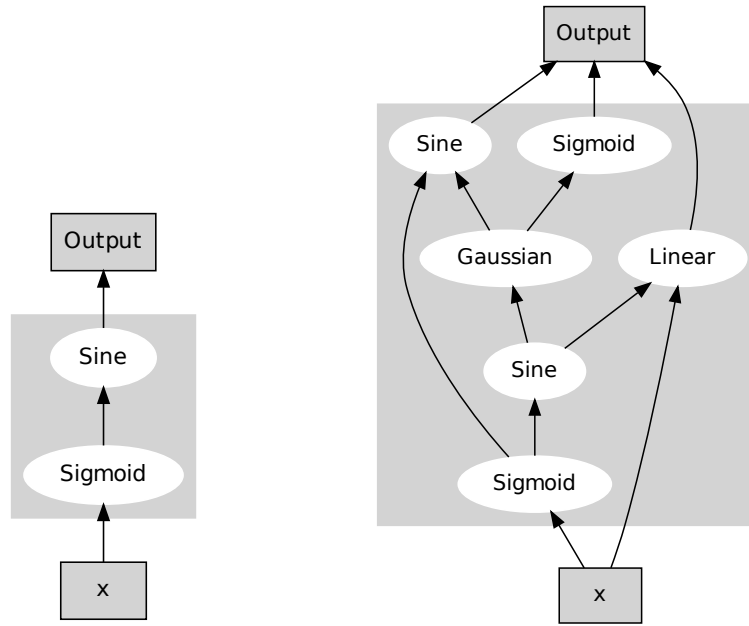
In particular, the main idea behind using CPPNs to investigate this abstraction is to apply computationally simulated evolution or human-guided breeding to structures (e.g. pictures, songs, or ANNs) represented indirectly by nested functions of global coordinates. These evolved functions are queried at each coordinate within a regular grid to produce a final structure *without* local interaction or iterative temporal unfolding. Thus the motivation is that the underlying hypothesis about development’s root mechanism is supported if over many domains CPPNs facilitate computational evolution of artifacts with biologically-similar regularities.

More concretely, a CPPN is based on the idea that a structure can be represented as a function of the geometric space it inhabits. For example, consider a function of one dimension; a three-dimensional printer moving along a line can apply such a function to realize a linear structure. That is, when evaluated with a scalar input corresponding to the printer’s current offset along its linear trajectory, the function’s output modulates whether or not the three-dimensional printer lays down material. In this way, a mathematical function can direct the realization of a physical structure. Note that this approach generalizes to larger-dimensional spaces and non-physical structures (e.g. temporal structures like music).

A CPPN is an approach to representing such functions. In particular, it represents a function as a collection of simple functions composed together in potentially complex nested relationships. Signature regularities often observed in organisms created by natural evolution such as symmetry, repetition, repetition with variation, and imperfect symmetry can be captured by a compact set of common functions that form the vocabulary for compositions. For example, because a Gaussian function is symmetric about the origin, it can create symmetric patterns in a CPPN. Similarly, a sine function can induce repetition. Nesting these functions can generate more complicated patterns; for example, a sigmoid function can modulate repetition when composed with a sine wave. For more complicated compositions, it is convenient to represent the composed functions visually as a network, where connections represent composition relationships, as shown in figure 2. In this way, CPPNs can facilitate representing complex structures that share familiar properties with the forms of natural organisms (e.g. symmetry, repetition, and repetition with variation).

Note that although the CPPN is a network, it does actually abstract a subtle form of temporal computation. Because the nodes of the CPPN are activated from bottom to top, their order of activation (and the functions they compute in that sequence) can be viewed as a kind of temporal progression. In this sense, they do in effect encompass a kind of temporal process. However, this form of temporality deviates importantly from that in natural development because it is not implemented through a progressive unfolding, and of course there is no local interaction between different parts of the phenotype. Thus although CPPNs do capture an essential aspect of temporal development, they deviate significantly from the approach seen in nature.

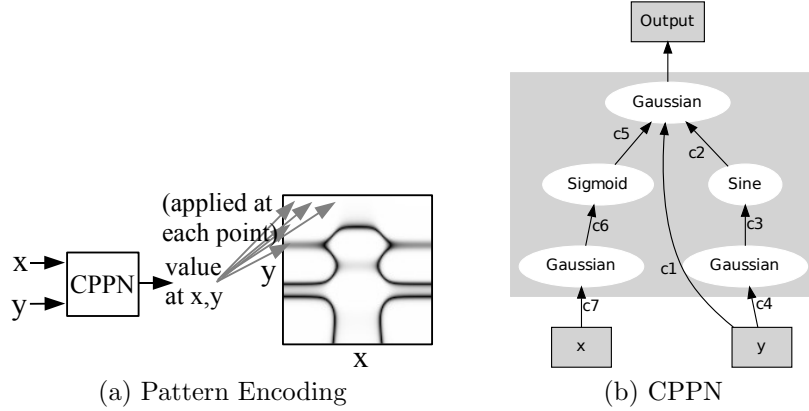
The same general process of composing functions works also in higher dimensional spaces. For example, figure 3 shows how a two-dimensional picture can be represented as the output of a CPPN that takes an x, y coordinate pair as input. Similar applications of CPPNs can facilitate representing songs [44], three-dimensional physical structures [21], or ANNs [98]. While there are many ways to represent functions, the main advantage of CPPNs is that they provide a means for increasingly complex



(a) Simple network

(b) More complex network

Figure 2: Network-based notation. Networks are one possible way to visualize compositions of functions. The network shown in (a) depicts the composition $\sin(\sigma(x))$, while the network shown in (b) depicts the function $\sin\left(\sigma(x) + G(\sin(\sigma(x)))\right) + \sigma\left(G(\sin(\sigma(x)))\right) + \sin(\sigma(x)) + x$. Note that $G(x)$ is the Gaussian function and $\sigma(x)$ is the Sigmoid function.



$$f(x, y) = G\left(c_1 y + c_2 \sin\left(c_3 G(c_4 y)\right) + c_5 \sigma\left(c_6 G(c_7 x)\right)\right)$$

(c) Formula

Figure 3: How a CPPN represents a picture. (a) The CPPN takes arguments x and y , which are coordinates in a two-dimensional space. When all the coordinates are drawn with an intensity corresponding to the output of the CPPN, the result is a spatial pattern. (b) Internally, the CPPN is a graph that determines which functions are connected. The connections are weighted such that the output of a function is multiplied by the weight of its outgoing connection. The CPPN in (b) actually produces the pattern in (a). (c) The weighted-graph structure in (b) can also be expressed as a mathematical formula. The constants c_n in the formula represent the weights for each connection, $G(x)$ is the Gaussian function, and $\sigma(x)$ is the sigmoid function.

structures to be evolved or bred through human guidance through a computational process.

That is, through an evolutionary algorithm or interactive evolutionary process (where humans perform selection in an evolutionary algorithm), it is possible to evolve complex CPPNs with desirable characteristics. Random mutations to CPPNs can change the weights of connections, or can alternatively *add* new nodes or connections to the network, which effectively add new compositions to the function the CPPN as a whole represents. In this way, over generations of breeding a CPPN can complexify over time to represent arbitrarily complex patterns.

In addition to their favorable theoretical properties, CPPNs have yielded interesting practical results in a wide variety of different applications. They have facilitated an online picture-breeding service [91], an online three-dimensional shape-breeding service [21], an interactive dance evolution program [32], an interactive music evolution program [44], a computer game based on evolved particle weapons [42], formed

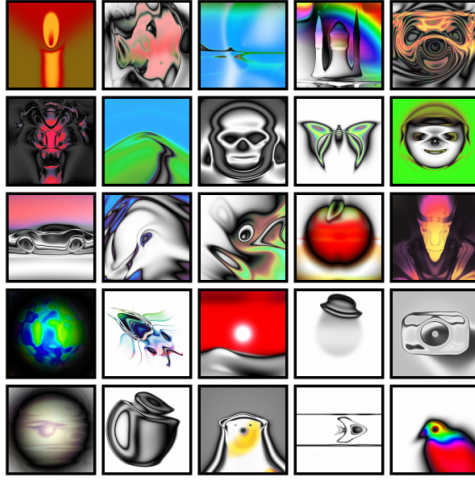
the foundation for a practical method of evolving complex neural networks [98, 24], and have been applied to evolving players for board games like checkers and Go [35, 34], virtual creatures in artificial life simulations [3, 103], and controllers for both simulated and real robots [54, 20]. Interestingly, CPPN-represented shapes evolved through the web service Endless Forms [21] can be realized physically through a three-dimensional printer. In this way, combining CPPNs with such printers can be seen as an alternative *physical* mechanism for realizing complex patterns; unlike biological development, such a mechanism does not rely on growth over time or local interaction. Examples of the complex phenotypic structures reminiscent of nature that are encoded by CPPNs are shown in figure 4.

Thus the conclusion from their diverse applications is that CPPNs appear to be generally well-suited for representing complex structures with natural regularities. In this way credence is given to the original hypothesis that inspired the radical reimplementa-tion: Development may be more at heart about realizing nested functions of geometry than temporal unfolding or local interaction. In other words, temporal unfolding and local interaction may merely be the particular tools that nature exploits to realize a functional relationship between geometry and form; yet CPPNs demonstrate that these particular tools are not strictly necessary to evolve complex forms. The practical implication is that engineering approaches can exploit abstractions of development like CPPNs (similarly to how helicopters exploit the principles of flight) to enable computationally representing and evolving complex geometrical structures. At a deeper level, conceiving development as a set of functions of nested coordinate frames broadens our understanding of the phenomenon and hints at a fundamental mechanism independent of physics and time.

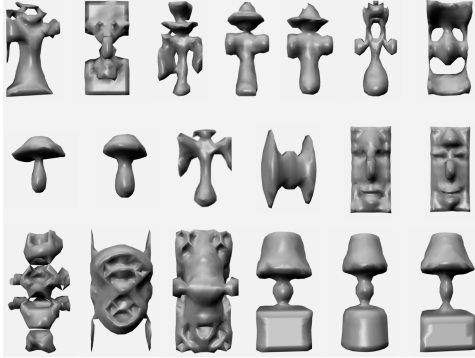
4.4 Questioning Adaptive Explanations

Natural evolution is a profoundly creative process responsible for crafting the diverse forms that populate Earth. While much is understood about natural evolution, there remain some questions that are still debated among biologists. For example, such questions include the origin of life [80], the maintenance of sexual reproduction [100], and the relative importance of evolutionary forces to evolution’s creativity [81, 83]. In particular, this section focuses on investigating the importance of natural selection relative to non-adaptive evolutionary forces like genetic drift or exaptation. Such investigation is conducted through an evolutionary algorithm that radically reimplements evolution as a search without any pressure to adapt.

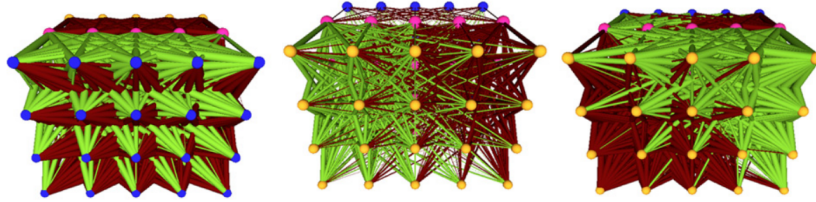
The issue of adaptation’s relative importance has been thoroughly debated although it remains unresolved [81]. Some researchers argue that natural selection is paramount [29, 94, 30], while others stress the role of non-adaptive forces [39, 38, 73, 66]. For example, Dawkins [29] lays out a selection-centric argument for complex macro-level adaptations. In contrast, Gould [39] attributes more of evolution’s creativity to historicity, the accumulation of variation, and the constraints imposed by a genome’s architecture (e.g. the body plan). While there are theoretical and experimental approaches to investigating this issue [81, 77], this section focuses on how radical reimplementa-tion can act as a new tool to explore this and other issues related to evolutionary pressures in a principled way.



(a) Picbreeder



(b) Endless Forms



(c) HyperNEAT

Figure 4: Variety of CPPN-represented structures with natural properties. A selection of applications of CPPNs to representing complex phenotypic structures is shown. In particular, (a) shows a gallery of pictures evolved by users on the Picbreeder (<http://picbreeder.org>) online service [91], (b) shows some of the products of an online service called Endless Forms [21] where three-dimensional shapes are evolved (reproduced from <http://endlessforms.com/press/>), and (c) shows examples of a complex ANNs evolved by the CPPN-based HyperNEAT method (reproduced from Clune et al. [24]).

The motivation for such new means of exploration is again that natural evolution as a whole is subject to the $N = 1$ problem: There is only one example of life evolving through natural evolution that we are aware of, which complicates inferring statistically-valid principles about evolution in all its possible instantiations. Furthermore, natural evolution itself is only one example of an evolutionary process, and may be biased in significant ways; this potential bias also serves to motivate exploring evolutionary processes divergent from nature.

A related problem for studying the relative importance of different evolutionary forces is that the effects and dynamics of natural evolution span both the entire planet and billions of years. Thus comprehensive and direct study of previous major evolutionary events is challenging because such events occurred over huge swaths of geography and time, and the exact temporal evolutionary forces expressed leave little trace. The vast scale of evolution similarly complicates recreating such events through directed evolution in laboratories. Even the longest experimental evolution studies are only around 50,000 generations long and culminate in relatively simple adaptations [107, 8]. For these reasons, experimental methods facilitating indirect investigation, like radical reimplementations, may be important tools for isolating the effect and importance of particular evolutionary forces.

The main idea is that evolution can be abstracted in different ways and reimplemented to stress a particular abstraction in a manner explicitly unlike nature. For example, a coarse way of abstracting evolution (in a selection-centric interpretation) is to view evolution as an *optimizer*, driving relentlessly towards higher fitness. An alternate abstraction that is non-adaptive is instead to conceive natural evolution as a process driven to continually create novelty without any direct pressure towards adaptation. Of course, both of these abstractions focus on only one aspect of natural evolution, but radical reimplementations allow for exploring the impact of those aspects in isolation.

Interestingly, the abstraction of evolution as an optimizer is prevalent in evolutionary computation (EC; [9, 43]). In most evolutionary algorithms (EAs), biological fitness is abstracted as a *fitness function*. Individuals from a population of computational genomes are selected for their ability to maximize the fitness function and are mated algorithmically to form the next generation. Because the fitness function is central in such an abstraction, these models can be seen as selection-centric radical reimplementations of natural evolution. That is, there is little room for other evolutionary forces in such EAs because of tiny population sizes (from a biological perspective) and exaggerated mutation rates. In other words, traditional EAs abstract natural evolution’s key driving force as optimization and reimplement such optimization much differently from natural selection. Unlike real organisms, such computational genomes do not reproduce inherently (i.e. they do not encode how to reproduce themselves and must be copied by the underlying evolutionary algorithm) and thus do not have a traditional biological fitness, but instead are rated by their ability to achieve a preconceived objective by an engineered fitness function. However, although they have proven successful in many cases, such objective-driven EAs have never generated artifacts exhibiting complexity on the level of biological organisms. Furthermore, the high-level properties of such EAs often contrast starkly with those of natural evolution. For example, EAs are nearly always *convergent*

[89, 59, 96], i.e. over time diversity in the population tends to be extinguished. In fact, a wide range of techniques have been investigated to actively maintain diversity in EAs [89, 59, 68, 46]. In contrast, a characteristic feature of natural evolution is its open-ended march towards greater diversity. These problems (i.e. failure to evolve biological levels of complexity and the tendency to converge) put into question the central abstraction of evolution as an optimizer [59], which abstracts to some degree the adaptationist view of natural evolution.

However, it is important to note that such a negative result of radical reimplementa- tion is not as informative as a positive result. It could be that a particular radical reimplementa- tion is misguided, and that another more well-designed imple- mentation of the guiding abstraction would yield better results. For example, the kinds of representations applied in most EAs may not compare favorably to biolog- ical DNA, begging the question of whether the negative result is due either to an ineffective abstraction of evolution or to an ineffective abstraction of biological repre- sentation. On the other hand, a positive result for a radical reimplementa- tion more assuredly provides evidence for its underlying abstraction, because it is improbable that reimplementing an incorrect abstraction would yield desirable results. For exam- ple, demonstrating in reality a new form of flight provides evidence for the principles the engineer applied; and because a new form of flight is so difficult to achieve, it would be unlikely that relying on a fundamentally incorrect abstraction would prove productive merely by chance.

A more positive example of a radical reimplementa- tion of natural evolution is based on an abstraction other than optimization: evolution as a generator of novelty. This optimization-free abstraction is plausible because a signature of evolution is its tendency to diverge and fill reachable niches. Interestingly, such an abstraction can be radically reimplemented as an algorithm that searches only for novelty. This idea motivates *novelty search*, an EA that focuses on novelty instead of natural selection [56, 59].

4.4.1 Novelty Search

Novelty search [56, 59] is a non-adaptive¹ EA that is not driven towards any particular form over another. Instead it diverges, finding forms different from what has been encountered by the search in the past. While in nature the accumulation of novelty is mainly passive (although frequency-dependent selection may sometimes directly encourage novelty [4]), novelty search *explicitly* seeks it. In this way, novelty search allows investigating the properties of a search without the pressure for organisms to adapt to their environment.

In particular, novelty search replaces the fitness function that characterizes the optimization-based EA abstraction of natural evolution with a *novelty metric*, which is a user-defined measure of distance between evolved artifacts in a particular domain.

¹Note that in this context, non-adaptive is meant to indicate that novelty search is not driven by an abstraction of natural evolution including natural selection (i.e. the idea that organisms are selected for increasing fit to their niche). That is, novelty search is non-adaptive in that it is driven by an abstraction of natural evolution based only on the accumulation of novelty, which is generally a non-adaptive process in natural evolution.

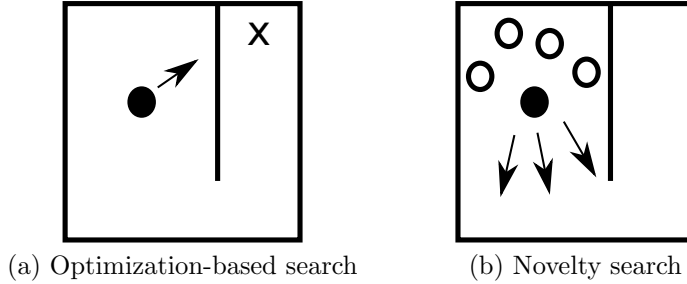


Figure 5: Contrast between optimization-based search and novelty search. (a) Traditional EAs guided by a fitness function tend to converge towards a prescribed objective. (b) Novelty search is instead driven to diverge from previously encountered artifacts.

In this way, novelty search can be driven to find only what is different from what it has previously encountered in a user-defined space of artifacts. Figure 5 illustrates the main difference between novelty search and a traditional objective-based EA. Importantly, although novelty might seem like an uninformative gradient of information to follow, often performing something novel requires exploiting regularities in a domain. For example, for a robot to behave in a novel way may require learning about its environment, e.g. how to avoid walls or navigate a corridor. In this way, a search for novelty can lead to functional and interesting results; a similar human drive may underlie artistic and literary creation [69].

What is particularly interesting about novelty search is that it offers the potential to investigate whether the *appearance* of adaptation can often deceptively result from non-adaptive search processes. This issue is biologically relevant because a common critique of “adaptationists” is that they are too quick to infer adaptive explanations before considering non-adaptive alternatives. The general problem is that while it is easy to form adaptive intuitions (e.g. feature X may have evolved to enhance fitness by means of explanation Y), our intuitions are less developed when understanding non-adaptive processes. In other words, understanding a direct search process is familiar because it is something we all employ in day-to-day life (e.g. a search for our car keys, or navigating from our house to work and back again). Because novelty search realizes a search without an overarching goal (just as natural evolution is not driven overall towards any one specific type of organism), it can act as a tool for understanding the potential of non-adaptive processes. In effect, novelty search allows *isolating* the potential of a raw search for novelty entirely separated from any optimization pressure, which would be a challenging experiment to perform with real organisms.

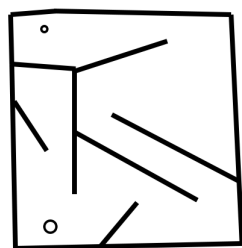
In this way, novelty search can both test the abstraction of evolution as a search for novelty and potentially provide evidence and insight in the adaptationist debate. To review such evidence, in EC there have been a series of experiments that compare novelty search with a more traditional objective-driven algorithm [60, 59, 57, 87]. The

basic idea is to investigate which type of search can better evolve an ANN that can control a robot to perform a target task (e.g. navigating a maze or walking bipedally). That is, an ANN is connected to a simulated robot, receiving sensory information and outputting motor commands; fitness is measured based on how well the robot controlled by the neural network performs, whereas novelty is measured by how different the robot’s behavior is from previous robots. That way, the structure and connectivity of the neural networks are evolved by feedback from either optimization pressure or the novelty of the controlled robot’s behavior. Two example problems to which novelty search has been applied are shown in figure 6; while the expectation might be that optimizing directly for the target task would be more successful, in practice novelty search has often performed better [60, 59, 57, 87], more effectively evolving controllers for maze-navigating robots [60, 59, 57], artificial ants [57], bipedal robots [59], and robots that learn from experience [87]. Thus while bipedal walking may normally be viewed as an adaptation, the explanation for its discovery by novelty search [59] cannot be adaptationist because novelty search does not explicitly favor walking over anything else. The reason that novelty search nevertheless does well is that a search driven to optimize fitness may converge on a fitness peak from which there is no path to the highest-fitness goal behavior. In other words, the stepping stones to the target behavior may not increase fitness themselves. While this phenomenon is known in biology (e.g. rugged fitness landscapes are known to be difficult to optimize [48, 22], and overall significant limitations of selection have been shown [5, 88]) it is informative to see how pervasive it is even in a wide range of simple problems, and how it can be overcome by searching only for novelty.

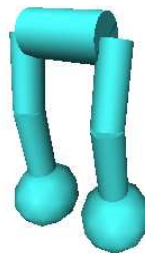
The generality of results from novelty search [60, 59, 57, 87, 58, 37, 31, 76, 50, 95, 55] is provocative and may hint at the importance of non-adaptive forces to the creativity of natural evolution. At minimum it provides a concrete example of how non-adaptive forces can consistently produce artifacts with the superficial appearance of adaptation. While natural selection may be a powerful honing force, it may not be well-suited for creativity, just as successful brainstorming sessions often result from a relaxed, non-judgmental atmosphere [82]. In other words, while adaptive arms races between organisms may optimize them relative to one another (e.g. gazelles and their cheetah predators may adapt to become faster), such arms races do not generally result in diverse novel species (e.g. naked mole rats). In this way novelty search offers a controversial perspective on the relative importance of natural selection: Perhaps evolution at heart is more of a novelty accumulator than an optimizer, even if on surface the mechanisms by which novelty search and natural evolution accumulate novelty are divergent.

5 Discussion

One motivation for the radical reimplement approach is to stress the broad possibilities for alife implementations. In particular, the universality of alife results may often be unnecessarily limited by taking direct inspiration from natural evolution. In other words, when alife acts to export current biological understanding into a new medium, it may shed light only on how current understanding can generalize



(a) Maze Navigation



(b) Biped Locomotion

Figure 6: Example applications of novelty search. (a) In the maze navigation domain, a maze-navigating robot begins at the location within the maze marked by the larger circle. A robot that travels from the starting location to the ending location (marked by the smaller circle) is considered a solution. Although novelty search does not directly seek to solve the problem, applying it to discover robots that end in novel locations within this deceptive maze proves much more effective than directly rewarding robots that get closer to the ending location: Novelty search nearly always evolves a successful navigator whereas a more traditional EA nearly always fails. (b) In the biped locomotion domain, successful evolution will yield a robot with a stable gait. Results in this domain illustrate that searching for maximally novel types of falling and walking behaviors is more effective than directly searching for the farthest-traveling gait.

over various mediums. Yet by including current biological understanding, such approaches may not help in achieving one of alife’s central goals: to illuminate the deeper underlying principles of *universal* biology. In this way, a contribution of radical reimplementations is to introduce a philosophical framework for isolating particular hypotheses about life from potential confounding details.

Additionally, it is often difficult for alife researchers to convince biologists of the relevance of their models. Part of the difficulty is that biology generally studies life as it is, centered on the concrete. In contrast, alife is centered on the abstract and aims to investigate life in all possible incarnations. As the examples of radical reimplementations reviewed here illustrate, focusing on the abstract and divergent possibilities for life can yield informative and thought-provoking insights. And while the philosophy of biology and theoretical biology also deal in abstractions, usually the abstract models are designed for biological purposes and are inspired by life as it is. Thus this paper helps demonstrate that models that are deliberately biologically *implausible* may still be of interest to biology.

The key point is that it is possible for abstractions, even those not specifically designed to illustrate biological principles, to be biologically relevant. Some alife models, like Avida, specifically aim at and achieve such relevance. But it may be exactly because the subsumption architecture, CPPNs, and novelty search were designed as practical means to extend the state of the art in their respective fields, and not to speak to biology, that they *are* biologically relevant. That is, the *practicality*

of these technologies in many different contexts provides significant evidence for their driving abstractions, and the divergence of the implementations of these abstractions from nature *isolates* the concept being tested.

In this way it is possible to interpret abstractions from non-biological fields as providing evidence for biological theories. The subsumption architecture is not a direct model of biological brains, Avida is not a model of terrestrial self-reproduction, CPPNs are not a model of animal development, nor is novelty search a plausible biological model of natural evolution. Yet the success of CPPNs in representing complex phenotypes in a wide range of applications provides evidence for abstracting development as a product of nested functions of geometry. In a similar way, novelty search provides a means to investigate the power of searches without pressure to achieve anything in particular. Because abstractions are general, it is not surprising that philosophical insight can bleed from one field to another. However, the merit of radical reimplementations is that it suggests a principled approach to such philosophical overlap.

Interestingly, much progress in science results from overturning assumptions widely accepted as fact. In this spirit an unconventional tool to investigate such assumptions about biology may be important, especially when it is difficult to investigate them in a more direct manner. What biological principles widely assumed essential are in fact only incidental? What might separating these two classes teach us about life?

6 Conclusion

This paper defined the radical reimplementations approach to artificial life. By coarsely abstracting a biological phenomenon and reimplementing it in a way maximally divergent from biological example, it is possible to investigate the validity of the abstraction. In this way, abstractions about biology that are divergently implemented in EC models or alife simulations can potentially be relevant to biology. The strength of this approach is that there is no need to defend the reimplementations themselves as biologically plausible or accurate; what is being tested is the underlying abstract *hypothesis*. In conclusion, radical reimplementations is a new philosophical tool for investigating and isolating fundamental biological principles, thereby allowing researchers to ask and investigate provocative questions.

Acknowledgements

This research was supported by DARPA and ARO through DARPA grant N11AP20003 (Computer Science Study Group Phase 3), and US Army Research Office grant Award No. W911NF-11-1-0489. This paper does not necessarily reflect the position or policy of the government, and no official endorsement should be inferred.

References

- [1] Abbott, R., Forrest, S., and Pienta, K. (2006). Simulating the hallmarks of cancer. *Artificial life*, 12(4):617–634.
- [2] Amir, E. and Maynard-Zhang, P. (2004). Logic-based subsumption architecture. *Artificial Intelligence*, 153(1):167–237.
- [3] Auerbach, J. and Bongard, J. (2010). Dynamic resolution in the co-evolution of morphology and control. In *Artificial Life XII: Proceedings of the Twelfth International Conference on the Simulation and Synthesis of Living Systems*.
- [4] Ayala, F. and Campbell, C. (1974). Frequency-dependent selection. *Annual Review of Ecology and Systematics*, 5:115–138.
- [5] Barton, N., Partridge, L., et al. (2000). Limits to natural selection. *BioEssays*, 22(12):1075–1084.
- [6] Beatty, J. (2006). Replaying life’s tape. *The Journal of philosophy*, 103(7):336–362.
- [7] Beckmann, B. E., McKinley, P. K., and Ofria, C. (2007). Evolution of an adaptive sleep response in digital organisms. In *Advances in Artificial Life*, pages 233–242. Springer.
- [8] Bell, G. (2013). Evolutionary rescue of a green alga kept in the dark. *Biology letters*, 9(1):20120823.
- [9] Bethke, A. D. (1981). *Genetic Algorithms as Function Optimizers*. Phd thesis, University of Michigan.
- [10] Blount, Z., Borland, C., and Lenski, R. (2008). Historical contingency and the evolution of a key innovation in an experimental population of escherichia coli. *Proceedings of the National Academy of Sciences*, 105(23):7899.
- [11] Boden, M. (1996). The philosophy of artificial life.
- [12] Bongard, J. and Pfeifer, R. (2001). Repeated structure and dissociation of genotypic and phenotypic complexity in artificial ontogeny. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 829–836.
- [13] Brockhurst, M., Colegrave, N., Hodgson, D., and Buckling, A. (2007). Niche occupation limits adaptive radiation in experimental microcosms. *PLoS One*, 2(2):e193.
- [14] Brooks, R. (1986). A robust layered control system for a mobile robot. *Robotics and Automation, IEEE Journal of*, 2(1):14–23.
- [15] Brooks, R. A. (1989). A robot that walks; emergent behaviors from a carefully evolved network. *Neural computation*, 1(2):253–262.

- [16] Brooks, R. A. (1991). Intelligence without representation. *Artificial intelligence*, 47(1):139–159.
- [17] Bryson, J. J. (1992). The subsumption strategy development of a music modelling system.
- [18] Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. *AI Magazine*, 26(4):53.
- [19] Chow, S., Wilke, C., Ofria, C., Lenski, R., and Adami, C. (2004). Adaptive radiation from resource competition in digital organisms. *Science*, 305(5680):84–86.
- [20] Clune, J., Beckmann, B. E., Ofria, C., and Pennock, R. T. (2009). Evolving coordinated quadruped gaits with the hyperneat generative encoding. In *Evolutionary Computation, 2009. CEC’09. IEEE Congress on*, pages 2764–2771. IEEE.
- [21] Clune, J. and Lipson, H. (2011). Evolving three-dimensional objects with a generative encoding inspired by developmental biology. In *Proceedings of the European Conference on Artificial Life*.
- [22] Clune, J., Misevic, D., Ofria, C., Lenski, R. E., Elena, S. F., and Sanjuán, R. (2008). Natural selection fails to optimize mutation rates for long-term adaptation on rugged fitness landscapes. *PLoS Computational Biology*, 4(9):e1000187.
- [23] Clune, J., Pennock, R. T., Ofria, C., and Lenski, R. E. (2012). Ontogeny tends to recapitulate phylogeny in digital organisms. *The American Naturalist*, 180(3):E54–E63.
- [24] Clune, J., Stanley, K., Pennock, R., and Ofria, C. (2011). On the performance of indirect encoding across the continuum of regularity. *Evolutionary Computation, IEEE Transactions on*, 15(3):346–367.
- [25] Connell, J. (1987). Creature design with the subsumption architecture. In *IJCAI*, pages 1124–1126.
- [26] Connell, J. (1989). A colony architecture for an artificial creature.
- [27] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCSS)*, 2(4):303–314.
- [28] Davidson, E. H. (2010). *The regulatory genome: gene regulatory networks in development and evolution*. Academic Press.
- [29] Dawkins, R. (1997). *Climbing mount improbable*. WW Norton & Company.
- [30] Dennett, D. (1996). *Darwin’s dangerous idea: Evolution and the meanings of life*. Simon and Schuster.
- [31] Doucette, J. (2010). Novelty-based fitness measures in genetic programming. Master of science in computer science, Dalhousie University.

- [32] Dubbin, G. and Stanley, K. (2010). Learning to dance through interactive evolution. *Applications of Evolutionary Computation*, pages 331–340.
- [33] Fleischer, K. and Barr, A. (1994). A simulation testbed for the study of multicellular development: The multiple mechanisms of morphogenesis. In *Proc. of the Third Intl. Conf. on Artificial Life (ALIFE III)*.
- [34] Gauci, J. and Stanley, K. O. (2010a). Autonomous evolution of topographic regularities in artificial neural networks. *Neural computation*, 22(7):1860–1898.
- [35] Gauci, J. and Stanley, K. O. (2010b). Indirect encoding of neural networks for scalable go. In *Parallel Problem Solving from Nature, PPSN XI*, pages 354–363. Springer.
- [36] Goings, S., Clune, J., Ofria, C., and Pennock, R. T. (2004). Kin selection: The rise and fall of kin-cheaters. In *Proceedings of the Ninth International Conference on Artificial Life*, pages 303–308.
- [37] Goldsby, H. and Cheng, B. (2010). Automatically Discovering Properties that Specify the Latent Behavior of UML Models. In *Proceedings of MODELS 2010*.
- [38] Gould, S. and Lewontin, R. (1979). The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161):581–598.
- [39] Gould, S. J. (1996). *Full House: The Spread of Excellence from Plato to Darwin*. Harmony Books.
- [40] Grabowski, L. M., Elsberry, W. R., Ofria, C., and Pennock, R. T. (2008). On the evolution of motility and intelligent tactic response. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 209–216. ACM.
- [41] Gruau, F., Whitley, D., and Pyeatt, L. (1996). A comparison between cellular encoding and direct encoding for genetic neural networks. In *Proceedings of the First Annual Conference on Genetic Programming*, pages 81–89. MIT Press.
- [42] Hastings, E., Guha, R., and Stanley, K. (2009). Automatic content generation in the galactic arms race video game. *Computational Intelligence and AI in Games, IEEE Transactions on*, 1(4):245–263.
- [43] Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. University of Michigan Press, Ann Arbor, MI.
- [44] Hoover, A., Szerlip, P., Norton, M., Brindle, T., Merritt, Z., and Stanley, K. (2012). Generating a complete multipart musical composition from a single monophonic melody with functional scaffolding. In *International Conference on Computational Creativity*.
- [45] Horgan, J. (1995). From complexity to perplexity. *Scientific American*, 272(6):104–109.

- [46] Hornby, G. S. (2006). ALPS: the age-layered population structure for reducing the problem of premature convergence. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2006)*, pages 815–822, New York, NY, USA. ACM.
- [47] Jacob, F. et al. (1977). Evolution and tinkering. *Science*, 196(4295):1161–1166.
- [48] Kauffman, S. and Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of theoretical Biology*, 128(1):11–45.
- [49] Koza, J. R. (1992). Evolution of subsumption using genetic programming. In *Proceedings of the First European Conference on Artificial Life*, pages 110–119.
- [50] Krcak, P. (2010). Solving deceptive tasks in robot body-brain co-evolution by searching for behavioral novelty. In *ISDA*, pages 284–289. IEEE.
- [51] Kurt, T. E. (2006). *Hacking Roomba: ExtremeTech*, volume 48. John Wiley & Sons.
- [52] Langton, C. (1984). Self-reproduction in cellular automata. *Physica D: Nonlinear Phenomena*, 10(1):135–144.
- [53] Langton, C. (1995). *Artificial life: An overview*. MIT Press: Cambridge.
- [54] Lee, S., Yosinski, J., Glette, K., Lipson, H., and Clune, J. (2013). Evolving gaits for physical robots with the hyperneat generative encoding: the benefits of simulation. In *Applications of Evolutionary Computation*, pages 540–549. Springer.
- [55] Lehman, J. and Stanley, K. (2011a). Improving evolvability through novelty search and self-adaptation. In *Evolutionary Computation (CEC), 2011 IEEE Congress on*, pages 2693–2700. IEEE.
- [56] Lehman, J. and Stanley, K. O. (2008). Exploiting open-endedness to solve problems through the search for novelty. In *Proc. of the Eleventh Intl. Conf. on Artificial Life (ALIFE XI)*, Cambridge, MA. MIT Press.
- [57] Lehman, J. and Stanley, K. O. (2010a). Efficiently evolving programs through the search for novelty. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2010)*. ACM.
- [58] Lehman, J. and Stanley, K. O. (2010b). Revising the evolutionary computation abstraction: Minimal criteria novelty search. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2010)*. ACM.
- [59] Lehman, J. and Stanley, K. O. (2011b). Abandoning objectives: Evolution through the search for novelty alone. *Evol. Comp.*, 19(2):189–223.
- [60] Lehman, J. and Stanley, K. O. (2011c). Novelty search and the problem with objectives. In *Genetic Programming in Theory and Practice IX (GPTP 2011)*, chapter 3, pages 37–56. Springer.

- [61] Lenski, R. E., Ofria, C., Collier, T. C., and Adami, C. (1999). Genome complexity, robustness and genetic interactions in digital organisms. *Nature*, 400(6745):661–664.
- [62] Lenski, R. E., Ofria, C., Pennock, R. T., and Adami, C. (2003). The evolutionary origin of complex features. *Nature*, 423(6936):139–144.
- [63] Lewens, T. (2009). Seven types of adaptationism. *Biology and Philosophy*, 24(2):161–182.
- [64] Lindenmayer, A. (1968). Mathematical models for cellular interactions in development. *Journal of theoretical biology*, 18(3):280–299.
- [65] Lougheed, V. (1911). *Vehicles of the Air*.
- [66] Lynch, M. (2007). The frailty of adaptive hypotheses for the origins of organismal complexity. In *Proc Natl Acad Sci USA*, volume 104, pages 8597–8604.
- [67] Mahadevan, S. and Connell, J. (1992). Automatic programming of behavior-based robots using reinforcement learning. *Artificial intelligence*, 55(2):311–365.
- [68] Mahfoud, S. W. (1995). *Niching methods for genetic algorithms*. PhD thesis, Champaign, IL, USA.
- [69] Martindale, C. (1990). *The clockwork muse: The predictability of artistic change*. Basic Books.
- [70] Maynard Smith, J. (1995). Life at the edge of chaos. *New York Rev. Books*, March, 2:28–30.
- [71] McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133.
- [72] McKinley, P., Cheng, B. H., Ofria, C., Knoester, D., Beckmann, B., and Goldsby, H. (2008). Harnessing digital evolution. *Computer*, 41(1):54–63.
- [73] McShea, D. (1996). Perspective: Metazoan complexity and evolution: Is there a trend? *Evolution*, pages 477–492.
- [74] Meinhardt, H. and Meinhardt, H. (1982). *Models of biological pattern formation*, volume 6. Academic Press London.
- [75] Miller, G. (1995). Artificial life as theoretical biology: How to do real science with computer simulation. Cognitive Science Research Paper 378, School of Cognitive and Computing Sciences, University of Sussex.
- [76] Mouret, J.-B. (2009). Novelty-based multiobjectivization. In *Proc. of the Workshop on Exploring New Horizons in Evol. Design of Robots, 2009 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*.
- [77] Nei, M. (2005). Selectionism and neutralism in molecular evolution. *Molecular biology and evolution*, 22(12):2318–2342.

- [78] North, M., Collier, N., and Vos, J. (2006). Experiences creating three implementations of the repast agent modeling toolkit. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 16(1):1–25.
- [79] Ofria, C. and Wilke, C. O. (2004). Avida: A software platform for research in computational evolutionary biology. *Artificial life*, 10(2):191–229.
- [80] Orgel, L. (1998). The origin of life: a review of facts and speculations. *Trends in Biochemical Sciences*, 23(12):491–495.
- [81] Orzack, S. and Sober, E. (1994). Optimality models and the test of adaptationism. *American Naturalist*, pages 361–380.
- [82] Paulus, P., Kohn, N., and Arditti, L. (2011). Effects of quantity and quality instructions on brainstorming. *The Journal of Creative Behavior*, 45(1):38–46.
- [83] Pigliucci, M. and Kaplan, J. (2000). The fall and rise of Dr. Pangloss: Adaptationism and the spandrels paper 20 years later. *Trends in ecology & evolution*, 15(2):66–70.
- [84] Raff, R. (1996). *The shape of life: genes, development, and the evolution of animal form*. University of Chicago Press.
- [85] Ray, T. (1992). An approach to the synthesis of life. *Artificial life II*, 10:371–408.
- [86] Ray, T. S. (1994). Evolution, complexity, entropy and artificial reality. *Physica D: Nonlinear Phenomena*, 75(1):239–263.
- [87] Risi, S., Hughes, C., and Stanley, K. (2010). Evolving plastic neural networks with novelty search. *Adaptive Behavior*.
- [88] Salazar-Ciudad, I., Marín-Riera, M., et al. (2013). Adaptive dynamics under development-based genotype-phenotype maps. *Nature*.
- [89] Sareni, B. and Krahenbuhl, L. (1998). Fitness sharing and niching methods revisited. *Evolutionary Computation, IEEE Transactions on*, 2(3):97–106.
- [90] Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford University Press, USA.
- [91] Secretan, J., Beato, N., D’Ambrosio, D., Rodriguez, A., Campbell, A., Folsom-Kovarik, J., and Stanley, K. (2012). Picbreeder: A case study in collaborative evolutionary exploration of design space. *Evolutionary Computation*.
- [92] Siegelmann, H. T. and Sontag, E. D. (1995). On the computational power of neural nets. *Journal of computer and system sciences*, 50(1):132–150.
- [93] Sims, K. (1994). Evolving virtual creatures. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 15–22. ACM.
- [94] Smith, J. (1993). *Did Darwin get it right?: Essays on games, sex and evolution*. Penguin Books.

- [95] Soltoggio, A. and Jones, B. (2009). Novelty of behaviour as a basis for the neuro-evolution of operant reward learning. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2009)*, New York, NY. ACM.
- [96] Srinivas, M. and Patnaik, L. M. (1994). Genetic algorithms: A survey. *Computer*, 27(6):17–26.
- [97] Stanley, K. (2007). Compositional pattern producing networks: A novel abstraction of development. *Genetic Programming and Evolvable Machines*, 8(2):131–162.
- [98] Stanley, K., D’Ambrosio, D., and Gauci, J. (2009). A hypercube-based encoding for evolving large-scale neural networks. *Artificial Life*, 15(2):185–212.
- [99] Stanley, K. and Miikkulainen, R. (2003). A taxonomy for artificial embryogeny. *Artificial Life*, 9(2):93–130.
- [100] Stearns, S. (1990). The evolutionary maintenance of sexual reproduction: The solutions proposed for a longstanding problem. *Journal of Genetics*, 69:1–10.
- [101] Sterelny, K. (1997). Universal biology. *The British Journal for the Philosophy of Science*, 48(4):587–601.
- [102] Stix, G. (2006). Owning the stuff of life. *Scientific American*, 294(2):76–83.
- [103] Szerlip, P. and Stanley, K. (2013). Indirectly encoded sodarace for artificial life. In *Advances in Artificial Life, ECAL*, volume 12, pages 218–225.
- [104] Taylor, C. E. and Jefferson, D. (1994). Artificial life as a tool for biological inquiry. *Artificial Life*, 1(1-2):1–14.
- [105] Togelius, J. (2004). Evolution of a subsumption architecture neurocontroller. *Journal of Intelligent and Fuzzy Systems*, 15(1):15–20.
- [106] Wilke, C., Wang, J., Ofria, C., Lenski, R., and Adami, C. (2001). Evolution of digital organisms at high mutation rates leads to survival of the fittest. *Nature*, 412(6844):331–333.
- [107] Wisser, M. J., Ribbeck, N., and Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science*, 342(6164):1364–1367.