

Exploring Biological Intelligence through Artificial Intelligence and Radical Reimplementation

Joel Lehman

Department of Computer Sciences
The University of Texas at Austin, Austin, TX
joel@cs.utexas.edu

Kenneth O. Stanley

Department of Electrical Engineering and Computer Science
The University of Central Florida, Orlando, FL
kstanley@eecs.ucf.edu

Abstract

An important goal in artificial intelligence and biology is to uncover general principles that underlie intelligence. While artificial intelligence algorithms need not relate to biology, they might provide a synthetic means to investigate biological intelligence in particular. Importantly, a more complete understanding of such biological intelligence would profoundly impact society. Thus, to explore biological hypotheses some AI researchers take direct inspiration from biology. However, nature's implementations of intelligence may present only one facet of its deeper principles, complicating the search for general hypotheses. This complication motivates the approach in this paper, called *radical reimplementation*, whereby biological insight can result from purposefully *unnatural* experiments. The main idea is that biological hypotheses about intelligence can be investigated by reimplementing their main principles intentionally to explicitly and maximally *diverge* from existing natural examples. If such a reimplementation successfully exhibits properties similar to those seen in biology it may better isolate the underlying hypothesis than an example implemented more directly in nature's spirit. Two examples of applying radical reimplementation are reviewed, yielding potential insights into biological intelligence despite including purposefully unnatural underlying mechanisms. In this way, radical reimplementation provides a principled methodology for intentionally artificial investigations to nonetheless achieve biological relevance.

Introduction

Artificial intelligence (AI) researchers seek to develop intelligent software and machines. Interestingly, such AI research need not relate to the specific biological implementation of human intelligence. Thus through more broadly exploring intelligence, AI research may enable discovering the general laws underlying intelligence and intelligence-generating processes. However, understanding biological intelligence in particular remains important and could profoundly impact humanity's self-perception. Furthermore, any general theory of intelligence should explain both biological and artificial instances. For these reasons, researchers

both in biology and AI attempt to decipher the design principles underlying human intelligence.

Such efforts are motivated by how studying prior precedent can facilitate deeper understanding. As a result, observation, analysis, and inference from nature are common techniques in biological fields such as ethology and cognitive psychology (Bekoff and Allen 1997). In contrast, computational approaches often probe questions of biological intelligence through synthesis. For example, biologically-inspired AI abstractly emulates biological processes such as information processing in neural networks or natural evolution. The idea is that such biologically-inspired algorithms may help isolate the key elements of biological intelligence and its origins. That is, through simulations with varied features, it may be possible to separate the essential features of biological processes from those that are only incidental. However, a danger inherent in any biologically-inspired approach is that inferring generalities from a limited or biased distribution can often be deceptive.

For example, a digital multi-gigahertz computer is not found among the forms crafted by natural evolution, though such computation may have enhanced biological intelligence. Yet its absence is not surprising: Biological evolution is not an engineer, but a *tinkerer* constrained to making incremental tweaks to previous designs instead of holistic revisions (Jacob and others 1977). As a result, the attractors for biological design differ from those of more directed design methodologies such as human engineering. Additionally, historical contingency may lead natural evolution to converge to particular solutions to particular biological problems, leaving other plausible possibilities unrealized (Blount, Borland, and Lenski 2008). In total, the effect of these constraints (i.e. being restricted to tinkering and historical contingency) biases natural evolution's solutions to a systematically reduced subset of all that are possible.

Importantly, because of this systematic bias in how nature generates solutions to biologically-relevant problems, features that are only incidental to a deeper phenomenon may be amplified to become ubiquitous. The result is that inferences based on observing only biological examples may be misleading (Sterelny 1997). This is particularly true in fields like biologically-inspired AI where it is often assumed that particular mechanisms such as biological development or natural selection are essential rather than only particular

realizations of a deeper principle.

As an analogy, consider trying to infer the principles of flight (instead of intelligence) from natural examples. If a man isolated on an island observed only birds, he might conclude that flappable wings with feathers are essential to flight. However, after seeing a bat flying, that same man could 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; in other words, ornithopters are a powerful attractor when extrapolating flight from natural examples. If the man on the island, having already seen a bird and a bat, then observed a helicopter, he could correctly conclude that the action of flapping – and even wings in the traditional sense – are unnecessary for flight.

This example illustrates the value of divergent implementations for isolating and validating basic principles. Through its vast deviation from prior natural examples, the helicopter helps to illuminate that the general principles of flight run deeper than wings and feathers. Importantly, inferring the deeper principles of aerodynamics, which is easier given the helicopter, helps to unify understanding of both artificial and biological instances of flight. At the same time, the helicopter validates aerodynamics by demonstrating what engineering with such knowledge enables: agile powered flight for human transportation. In other words, the value in abstracting a phenomenon a particular way can in part be estimated by the divergent implementations it enables, i.e. what it makes possible, and the degree to which it elegantly unifies all examples of a phenomenon.

These insights motivate a principled approach to investigating biological principles such as intelligence through biologically-inspired AI 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 biological intelligence may also more clearly expose the core of the underlying phenomenon than would considering additional biological examples.

Thus AI, which seeks to explore the abstract possibilities for intelligence in all forms, can enable exploring radical reimplementations of biological intelligence. Note that AI researchers otherwise seeking to be relevant to biology face the significant challenge of demonstrating that their models relate convincingly to biological truth (Sterelny 1997). This challenge is especially relevant to researchers in biologically-inspired AI where a direct connection to biology is often expected or necessary. As a result, such researchers often must argue that the abstractions made by their models are principled, i.e. that the biological details filtered out through abstraction into a compact algorithm are non-essential.

Thus one way to view radical reimplementation is as a principled alternative to the tradition within biologically-inspired AI 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.

In this way, the advantage of the radical reimplementation approach is that it escapes the underlying restrictiveness of attempting to mimic nature directly. There is no attempt to claim that a model is intrinsically biologically plausible or even that it directly relates to biology; instead the idea is to provide radically novel examples of intelligence. In short, the 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.

In this paper the radical reimplementation approach is illustrated by reviewing two examples of existing research in the spirit of radical reimplementation that aim to shed light on the principles underlying biological intelligence and its origins.

Novelty Search: Questioning Optimization

Natural evolution is a profoundly creative process responsible for crafting the diverse animal intelligences found on Earth. While much is understood about natural evolution, one question still debated is the relative importance of evolutionary forces to evolution's creativity (Orzack and Sober 1994; Pigliucci and Kaplan 2000). In particular, this section focuses on investigating the importance of optimizing biological fitness relative to non-adaptive evolutionary forces like genetic drift or exaptation. The overall idea is that evolution's creativity is what facilitated the evolution of complex intelligence and thus better understanding such creativity may lead to algorithms able to realize similar intelligence. Furthermore, this biological question itself may be of interest to AI researchers given the ubiquity of optimization in AI and machine learning. That is, if optimization is not the key driver of evolution's creativity it may also point to potential advantages for driving AI search processes by other mechanisms.

The motivation for exploring this question through radical reimplementation is 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 (Sterelny 1997). Thus experimental methods facilitating indirect investigation, like radical reimplementation, may be important tools for isolating the effect and importance of particular evolutionary forces.

In particular, 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 non-adaptive abstraction 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 reimplementation allows for exploring the impact of those aspects in isolation.

Interestingly, the abstraction of evolution as an optimizer is prevalent in evolutionary computation (EC; Holland 1975). 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. In other words, traditional EAs abstract natural evolution’s key driving force as optimization and reimplement such optimization much differently from natural selection.

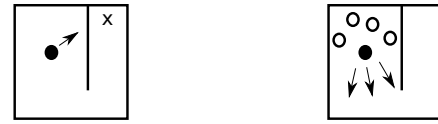
However, although they have proven successful in many cases, such objective-driven EAs have never generated artifacts exhibiting complexity on the level of biological organisms. Furthermore, the high-level properties of such EAs often contrast starkly with those of natural evolution. For example, EAs are nearly always *convergent*, while a characteristic feature of natural evolution is its open-ended march towards greater diversity. These problems put into question the central abstraction of evolution as an optimizer (Lehman and Stanley 2011a), which abstracts to some degree the adaptationist view of natural evolution.

However, such a negative result of radical reimplementation is not as informative as a positive result. It could be that a particular radical reimplementation is misguided, and that another more well-designed implementation of the guiding abstraction would yield better results. On the other hand, a positive result for a radical reimplementation more assuredly provides evidence for its underlying abstraction, because it is improbable that reimplementing an incorrect abstraction would yield desirable results. For example, 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 reimplementation 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. 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 (Lehman and Stanley 2011a).

Novelty search is a non-adaptive EA driven only to diverge, continually finding forms different from what has been encountered by the search in the past. While in nature the accumulation of novelty is mainly passive, 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



(a) Optimization-based search (b) Novelty search

Figure 1: 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.

natural evolution with a *novelty metric*, which is a user-defined measure of distance between evolved artifacts in a particular domain. 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 1 illustrates the main difference between novelty search and a traditional objective-based EA. While novelty might at first seem 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 how to avoid walls or navigate a corridor. In this way, a search for novelty can lead to functional and interesting results.

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 for the importance of fitness optimization in evolution. To review such evidence, in EC there have been a series of experiments that compare novelty search with a more traditional objective-driven algorithm (Lehman and Stanley 2011b; 2011a; 2010; Risi, Hughes, and Stanley 2010). The basic idea is to investigate which type of search can better evolve an artificial neural network that can control a robot to perform a target task (e.g. navigating a maze or walking bipedally). That is, an artificial neural network 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. While the expectation might be that optimizing directly for the target task would be more successful, in practice novelty search has often performed better (Lehman and Stanley 2011b; 2011a; 2010; Risi, Hughes, and Stanley 2010; Krcah 2010; Mouret 2009), more effectively evolving controllers for maze-navigating robots (Lehman and Stanley 2011b; 2011a; 2010; Mouret 2009), artificial ants (Lehman and Stanley 2010; 2011b), bipedal robots (Lehman and Stanley 2011a),

and robots that learn from experience (Risi, Hughes, and Stanley 2010). Thus while bipedal walking may normally be viewed as an adaptation, the explanation for its discovery by novelty search (Lehman and Stanley 2011a) 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, 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 (Lehman and Stanley 2011b; 2011a; 2010; Risi, Hughes, and Stanley 2010; Mouret 2009; Krcah 2010) is provocative and hints at the importance of non-adaptive forces to the creativity of natural evolution. It also provides a concrete example of how non-adaptive forces can 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. In this way novelty search offers a controversial perspective on the key mechanism behind evolutionary search in nature: Perhaps evolution at heart is more of a novelty accumulator than an optimizer, even if on surface their mechanisms are divergent. Interestingly, such results may also hint at the potential of AI search algorithms driven by gradients other than optimization towards an a priori objective.

HyperNEAT: Neural Connectivity as a Function of Geometry

Development is a prominent feature of biological organisms that enables the large-scale nervous systems underlying complex biological animal intelligence. Thus an important biological question is what features of biological development enable compactly representing highly complex, functional, and evolvable neural networks? For example, the vast complexity of the adult human nervous system is organized over time through biological development and is efficiently encoded by many fewer genes than there are neurons in the final structure (Stix 2006). Furthermore, such neural development has been much modulated by natural evolution to enable the wide diversity of complex animal lifestyles on Earth. Identifying the most general laws underlying neural development may facilitate human engineering of similarly complex artificial neural networks (ANNs), 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, with the nervous system forming as a coupled subcomponent of the overall developing organism. Notably, during development each cell exploits local signals from its neighbors to

determine its identity and further differentiate (Raff 1996). 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. Yet it is interesting to consider whether development in the traditional biological sense (and neural development in particular) might be only a particular implementation detail that masks a deeper principle.

Problematically, investigating this possibility through biological inference is challenging because life has largely converged to a single overarching means of development and cellular replication. Thus inferring principles about development's necessity or purpose from biology, like when considering natural evolution, is also subject to the $N = 1$ problem. However, an intriguing possibility is that this problem can be circumvented by radically reimplementing development such that it holistically differs from biological development on Earth. While much work has explored alternative developmental schemes (Stanley and Miikkulainen 2003), 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 (Stanley and Miikkulainen 2003). Such approaches are thought-provoking and can illuminate some features of development, but because of their *similarity* to terrestrial biology they may not well *isolate* the deeper principles behind neural development.

For this reason it may be more informative to investigate models through the radical reimplementation approach. 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.

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 patterns that it generates, and that development over time is only one way among many to realize such patterns.

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 (Cybenko 1989), which establishes that a series of enough nested functions can approximate *any* pattern. 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 (Raff 1996).

This observation supports the idea that establishing nested patterns are 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 biology. 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 organisms in nature, including the complex brains of animals. 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 reimplementing of the hypothesized abstraction of development as a mapping from geometry to form.

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; Stanley 2007). The CPPN is a variant of a traditional artificial neural network 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). In this way, CPPNs can represent complex patterns from nested coordinate frames in a form much divergent from biological development.

Relevant to the biological hypothesis explored in this section, an approach facilitated by this alternate conception of development, called HyperNEAT, can be seen as a radical reimplementing of neural development. In particular, HyperNEAT evolves complex ANNs that are represented by CPPNs. The main idea is that given an arbitrarily large neural network embedded in a geometry (i.e. each node is given a coordinate in space), a CPPN can represent the strengths of connections between two nodes as a function of the nodes' coordinates. For example, if the nodes were embedded in a two-dimensional geometry, then a CPPN with four-dimensional input could encode the connectivity among the nodes as a function of their positions. In this way, the final complex ANN is created *without* local interaction or iterative temporal unfolding. Thus the motivation is that the underlying hypothesis about neural development's root mechanism is supported if over many domains HyperNEAT facilitates computational evolution of functional ANNs with biologically-similar regularities.

Interestingly, HyperNEAT has proved successful over many domains, including many-joint robot arm control (Woolley and Stanley 2010), real-world Khepera robot control (Lehman et al. 2013), quadruped locomotion (Clune et al. 2011), checkers board evaluation (Gauci and Stanley 2010), and robocup soccer (Verbancsics and Stanley 2010). Figure 2 shows representative ANNs evolved with HyperNEAT that demonstrate regular connectivities reminiscent of biology. Thus the conclusion from its diverse applications is that HyperNEAT appears to be generally well-suited for representing complex ANNs with natural regularities.

In this way credence is given to the original hypothesis that inspired the radical reimplementing: Neural devel-

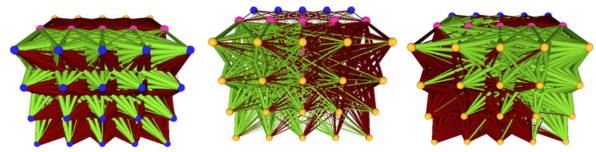


Figure 2: Examples of Evolved HyperNEAT ANNs. A selection of complex artificial neural networks evolved by the HyperNEAT method (reproduced from Clune et al. 2011). The ANNs exhibit key biological regularities like symmetry and modulated repetition.

opment may be more at heart about realizing connectivity patterns of nested geometrical functions than temporal unfolding or local interaction. If other mechanisms were also physically available to evolution, then if it were “run again” the form of development realized in this second realization might appear as radical to us as HyperNEAT appear when compared to biological development. Or perhaps more practically, 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 ANNs. At a deeper level, conceiving development as a set of functions of nested coordinate frames may broaden our understanding of the phenomenon and hint at a fundamental mechanism independent of physics and time.

Discussion

It is often difficult for AI researchers, even those directly inspired by biology, to convince biologists of the relevance of their models. Part of the difficulty is that biology generally studies intelligence as it is, centered on the concrete. In contrast, AI is centered on the abstract and aims to investigate intelligence in all possible incarnations. As the examples of radical reimplementing reviewed here illustrate, focusing on the abstract and divergent possibilities for life can yield informative and thought-provoking insights. Thus a contribution of this paper is to demonstrate that models that are deliberately biologically *implausible* may still have implications for understanding biological intelligence.

The key point is that it is possible for abstractions, even those not specifically designed to illustrate biological principles, to be biologically relevant. It may be exactly because novelty search and CPPNs were designed as practical means to extend the state of the art in EC, and not to speak to biology, that they *are* biologically relevant. That is, the *practicality* of novelty search and CPPNs 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. Traditional EAs or novelty search are not direct models of natural evolution, nor are CPPNs a direct model of development. Yet traditional EAs offer an indirect means of exploring the abstraction of evolution as an optimizer, and

novelty search provides a means to investigate the power of searches without pressure to achieve anything in particular. In a similar way, 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. Because abstractions are general, it is not surprising that philosophical insight can bleed from one field to another. However, the merit of radical reimplementation 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 a new tool to investigate such assumptions about biology may be important, especially when it is difficult to investigate them in a more direct manner. What principles underlying biological intelligence that are widely assumed essential are in fact only incidental? What might separating these two classes teach us about intelligence?

Conclusion

This paper defined the radical reimplementation approach and illustrated how it can facilitate AI research achieving biological relevance. By coarsely abstracting a biological phenomenon and reimplementing it in a way maximally divergent from biological example, the validity of the abstraction can be probed. In this way, abstractions about biology that are divergently implemented in AI simulations can potentially be relevant to biologists. The strength of this approach is that there is no need to defend the reimplementation itself as biologically plausible or accurate; what is being tested is the *abstraction*. In conclusion, radical reimplementation is a new philosophical tool for investigating and isolating fundamental biological principles like animal intelligence, thereby enabling AI researchers to ask provocative questions.

References

- Bekoff, M., and Allen, C. 1997. Cognitive ethology: Slayers, skeptics, and proponents. *Anthropomorphism, Anecdotes, and Animals: The Emperor's New Clothes?* 313–334.
- 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.
- 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.
- Cybenko, G. 1989. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCSS)* 2(4):303–314.
- Gauci, J., and Stanley, K. O. 2010. Autonomous Evolution of Topographic Regularities in Artificial Neural Networks. *Neural Computation Journal* 22(7):1860–1898.
- Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. Ann Arbor, MI: University of Michigan Press.
- Jacob, F., et al. 1977. Evolution and tinkering. *Science* 196(4295):1161–1166.
- Kraeh, P. 2010. Solving deceptive tasks in robot body-brain co-evolution by searching for behavioral novelty. In *ISDA*, 284–289. IEEE.
- Lehman, J., and Stanley, K. O. 2010. Efficiently evolving programs through the search for novelty. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2010)*. ACM.
- Lehman, J., and Stanley, K. O. 2011a. Abandoning objectives: Evolution through the search for novelty alone. *Evol. Comp.* 19(2):189–223.
- Lehman, J., and Stanley, K. O. 2011b. Novelty search and the problem with objectives. In *Genetic Programming in Theory and Practice IX (GPTP 2011)*. Springer. chapter 3, 37–56.
- Lehman, J.; Risi, S.; D'Ambrosio, D.; and Stanley, K. O. 2013. Encouraging reactivity to create robust machines. *Adaptive Behavior*.
- 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*.
- Orzack, S., and Sober, E. 1994. Optimality models and the test of adaptationism. *American Naturalist* 361–380.
- 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.
- Raff, R. 1996. *The shape of life: genes, development, and the evolution of animal form*. University of Chicago Press.
- Risi, S.; Hughes, C.; and Stanley, K. 2010. Evolving plastic neural networks with novelty search. *Adaptive Behavior*.
- Stanley, K., and Miikkulainen, R. 2003. A taxonomy for artificial embryogeny. *Artificial Life* 9(2):93–130.
- Stanley, K. 2007. Compositional pattern producing networks: A novel abstraction of development. *Genetic Programming and Evolvable Machines* 8(2):131–162.
- Sterelny, K. 1997. Universal biology. *The British Journal for the Philosophy of Science* 48(4):587–601.
- Stix, G. 2006. Owning the stuff of life. *Scientific American* 294(2):76–83.
- Verbancsics, P., and Stanley, K. O. 2010. Evolving static representations for task transfer. *Journal of Machine Learning Research* 1737–1763.
- Woolley, B. G., and Stanley, K. O. 2010. Evolving a single scalable controller for an octopus arm with a variable number of segments. In Schaefer, R.; Cotta, C.; Kołodziej, J.; and Rudolph, G., eds., *Proceedings of Parallel Problem Solving from Nature*, 270–279. Springer.