Abandoning Objectives: Evolution Through the Search for Novelty Alone

Joel Lehman jlehman@eecs.ucf.edu
School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, Florida 32816, USA

Kenneth O. Stanley kstanley@eecs.ucf.edu
School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, Florida 32816, USA

Abstract

In evolutionary computation, the fitness function normally measures progress toward an objective in the search space, effectively acting as an objective function. Through deception, such objective functions may actually prevent the objective from being reached. While methods exist to mitigate deception, they leave the underlying pathology untreated: Objective functions themselves may actively misdirect search toward dead ends. This paper proposes an approach to circumventing deception that also yields a new perspective on open-ended evolution. Instead of either explicitly seeking an objective or modeling natural evolution to capture open-endedness, the idea is to simply search for behavioral novelty. Even in an objective-based problem, such novelty search ignores the objective. Because many points in the search space collapse to a single behavior, the search for novelty is often feasible. Furthermore, because there are only so many simple behaviors, the search for novelty leads to increasing complexity. By decoupling open-ended search from artificial life worlds, the search for novelty is applicable to real world problems. Counterintuitively, in the maze navigation and biped walking tasks in this paper, novelty search significantly outperforms objective-based search, suggesting the strange conclusion that some problems are best solved by methods that ignore the objective. The main lesson is the inherent limitation of the objective-based paradigm and the unexploited opportunity to guide search through other means.

Keywords

Evolutionary algorithms, deception, novelty search, open-ended evolution, neuroevolution.

1 Introduction

It is tempting to believe that measuring progress with respect to an objective, which is the customary role of the fitness function in evolutionary computation (EC; De Jong, 2006; Fogel et al., 1966; Holland, 1975; Mitchell, 1997), lights a path to the objective through the search space. Yet as results in this paper will remind us, increasing fitness does not always reveal the best path through the search space. In this sense, one sobering message of this paper is the inherent limitation of objective-based search. Yet at the same time, this message is tempered by the caveat that search need not always be guided by explicit objectives. In fact, a key insight of this paper is that sometimes, opening up

© 2011 by the Massachusetts Institute of Technology

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

the search, more in the spirit of artificial life than traditional optimization, can yield the surprising and paradoxical outcome that the more open-ended approach more effectively solves the problem than explicitly trying to solve it. There is no silver bullet to be found in these results; rather, the benefit is to widen our perspective on search, both to be more sober and at the same time more open-minded about the potential breadth of our toolbox.

The concept of the objective function, which rewards moving closer to the goal, is ubiquitous in machine learning (Mitchell, 1997). The overriding intuition behind the idea of the objective function, which is widely accepted, is that the best way to improve performance is to reward improving performance with respect to the objective. In EC, this objective measure of performance is called the fitness function (De Jong, 2006; Holland, 1975; Mitchell, 1997), which is a metaphor for the pressure to adapt in nature.

Yet although they are pervasive, objective functions often suffer from the pathology of local optima, that is, dead ends in the search space with respect to increasing the value of the objective function. In this way, landscapes induced by objective (e.g., fitness) functions are often deceptive (Goldberg, 1987; Liepins and Vose, 1990; Pelikan et al., 2001). Although researchers are still working to characterize the various reasons that search methods may fail to reach the objective (Davidor, 1990; Goldberg, 1987; Grefenstette, 1992; Hordijk, 1995; Jones and Forrest, 1995; Mahfoud, 1995; Mitchell et al., 1992; Weinberger, 1990), as a rule of thumb, the more ambitious the goal, the more difficult it may be to articulate an appropriate objective function and the more likely it is that search can be deceived by local optima (Ficici and Pollack, 1998; Zaera et al., 1996). The problem is that the objective function does not necessarily reward the stepping stones in the search space that ultimately lead to the objective. For example, consider fingers stuck within a Chinese finger trap. While the goal is to free one's fingers, performing the most direct action of pulling them apart yields no progress. Rather, the necessary precursor to solving the trap is to push one's fingers together, which seems to entrap them more severely. In this way, the trap is deceptive because one must seemingly move further from the goal to ever have the hope of reaching it.

Because of deception, ambitious objectives in EC are often carefully sculpted through a curriculum of graded tasks, each chosen delicately to build upon the prior (Elman, 1991; Gomez and Miikkulainen, 1997; Van de Panne and Lamouret, 1995). Yet such incremental evolution is difficult and ad hoc, requiring intimate domain knowledge and careful oversight. Incremental evolution, along with other methods employed in EC to deal with deception (Ficici and Pollack, 1998; Knowles et al., 2001; Mahfoud, 1995; Pelikan et al., 2001; Stewart, 2001), do not fix the underlying pathology of local optima; if local optima are pervasive, search will still likely be deceived. Thus, deception may be an unavoidable consequence of certain objective functions irrespective of the underlying search algorithm. Paradoxically, in these cases, pursuing the objective may prevent the objective from being reached.

In contrast to the focus on objective optimization in machine learning and EC, researchers in artificial life often study systems without explicit objectives, such as in open-ended evolution (Channon, 2001; Maley, 1999; Standish, 2003). An ambitious goal of this research is to reproduce the unbounded innovation of natural evolution. A typical approach is to create a complex artificial world in which there is no final objective other than survival and replication (Adami et al., 2000; Channon, 2001; Yaeger, 1994). Such models follow the assumption that biologically inspired evolution can support an open-ended dynamic that leads to unbounded increasing complexity (Bedau, 1998; Channon, 2001; Maley, 1999).

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

However, a growing yet controversial view in biology is that the drive toward complexity in natural evolution is a passive force, that is, not driven primarily by selection (Gould, 1996; Lynch, 2007b; McShea, 1991; Miconi, 2007). In fact, in this view, the path toward complexity in natural evolution can sometimes be inhibited by selection pressure. If selection pressure is too high, then any deviation from a locally optimal behavior will be filtered out by selection. Thus, in this view, instead of being a by-product of selection, perhaps the accumulation of complexity is better explained by a different characteristic of natural evolution and open-ended evolutionary systems in general: They continually produce novel forms (Standish, 2003).

This perspective leads to a key idea in this paper that approaches the problems in both EC and artificial life in a new way. Instead of modeling natural evolution with the hope that novel individuals will be continually discovered, it is possible to create an open-ended dynamic by simply searching directly for novelty. Thus, this paper offers a comprehensive introduction to the novelty search algorithm, that was first described in Lehman and Stanley (2008). The main idea is to search with no objective other than continually finding novel behaviors in the search space. By defining novelty in this domain-independent way, novelty search can be applied to real world problems, bridging the gap between open-ended evolution and the practical application of EC. Interestingly, because there are only so many ways to behave, some of which must be more complex than others (Gould, 1996), the passive force in nature that leads to increasing complexity is potentially accelerated by searching for behavioral novelty.

Paradoxically, the search for novelty often evolves objectively superior behavior to evolution that is actually driven by the objective, as demonstrated by experiments in this paper in both a deceptive two-dimensional robot maze navigation task and a challenging three-dimensional biped locomotion domain, as well as in experiments in several other independent works (Lehman and Stanley, 2010a; Mouret, 2009; Risi et al., 2009). Counterintuitively, in this paper, novelty search, which ignores the objective, evolves successful maze navigators that reach the objective in significantly fewer evaluations than the objective-based method. This result is further investigated under several different conditions, including expanding the dimensionality of the behavior space, and is found to be robust. In the biped domain, novelty search evolves controllers that walk significantly further than those evolved by the the objective-based method. These results challenge the premise that the objective is always the proper impetus for search.

The conclusion is that by abstracting the process through which natural evolution discovers novelty, it is possible to derive an open-ended search algorithm that operates without pressure toward the ultimate objective. Novelty search is immune to the problems of deception and local optima inherent in objective optimization because it entirely ignores the objective, suggesting the counterintuitive conclusion that ignoring the objective in this way may often benefit the search for the objective. While novelty search is not a panacea, the more salient point is that objective-based search, which is ubiquitous in EC, clearly does not always work well. The implication is that while it seems natural to blame the search algorithm when search fails to reach the objective, the problem may ultimately lie in the pursuit of the objective itself.

2 Background

This section reviews deception in EC, complexity in natural evolution, open-endedness in EC, and the neuroevolution method used in the experiments.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

2.1 Deception in Evolutionary Computation

The study of deception is part of a larger study by EC researchers into what may cause an evolutionary algorithm (EA) to fail and how to remedy such failures. For the purpose of this paper, it is instructive to study the role of the objective (fitness) function in such failures and remedies.

2.1.1 Deception and Problem Difficulty

The original definition of deception by Goldberg (1987) is based on the building blocks hypothesis, in which small genetic building blocks may be integrated by crossover to form larger blocks (Holland, 1975). In the original conception, a problem is deceptive if lower-order building blocks, when combined, do not lead to a global optimum. A variety of work has further refined this definition and investigated performance on deceptive problems (Goldberg, 1987; Liepins and Vose, 1990; Pelikan et al., 2001). Some researchers have argued that the importance of deception may be over-emphasized (Grefenstette, 1992; Mitchell et al., 1992), while Whitley (1991) concluded that the only challenging problems are those with some degree of deception. Such disagreements are natural because no measure of problem difficulty can be perfect; in general it is impossible to know the outcome of an algorithm on a particular set of data without actually running it (Rice, 1953). Hence, many different metrics of EA problem hardness have been explored (Davidor, 1990; Hordijk, 1995; Jones and Forrest, 1995; Kauffman, 1993; Manderick et al., 1991; Weinberger, 1990).

Some alternative measures of problem difficulty attempt to model or quantify the ruggedness of the fitness landscape, motivated by the intuition that optimizing more rugged landscapes is more difficult (Davidor, 1990; Hordijk, 1995; Kauffman, 1993; Manderick et al., 1991; Weinberger, 1990). Such approaches are often based on the concepts of *correlation*, the degree to which the fitness of individuals are well correlated to their neighbors in the search space, or *epistasis*, the degree of interaction among genes' effects. A low degree of correlation or a high degrees of epistasis may indicate a rough landscape that may be difficult to optimize (Davidor, 1990; Hordijk, 1995; Kauffman, 1993; Manderick et al., 1991; Weinberger, 1990). Importantly, because the fitness landscape is induced by the objective function, the problem of ruggedness, presupposing reasonable settings for the EA, can be attributed to the objective function itself.

Other researchers suggest that ruggedness is overemphasized and that neutral fitness plateaus (neutral networks) are key influences on evolutionary dynamics (Barnett, 2001; Harvey and Thompson, 1996; Stewart, 2001), a view which has some support in biology (Kimura, 1983). However, neutral networks actually suggest a deficiency in the objective function. In a neutral network, the map defined by the objective function is ambiguous with respect to which way search should proceed.

Another approach, by Jones and Forrest (1995), suggests that measuring the degree to which the heuristic of fitness relates to the real distance to the goal is a good measure of difficulty. This perspective perhaps most clearly demonstrates that problem difficultly may often result from an uninformative objective function. When the heuristic has a weak basis in reality, there is little reason to expect the search to perform well. However, as demonstrated by the experiments in this paper, sometimes pursuing what appears to be a reasonable objective produces an unreasonable objective function.

While in general the exact properties of a problem that make it difficult for EAs are still a subject of research, in this paper the term "deception" will refer to an intuitive

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

definition of problem hardness. A deceptive problem is one in which a reasonable EA will not reach the desired objective in a reasonable amount of time. That is, by exploiting objective fitness in a deceptive problem, a population's trajectory is unlikely to uncover a path through the search space that ultimately leads to the objective. It is important to note that this definition of deception is different from the traditional definition (Goldberg, 1987) and is not meant to trivialize the impact or difficulty of choosing the correct representation (Rothlauf and Goldberg, 2002), parameters (De Jong, 2006), or search operators (Yao, 1993), all of which affect the performance of the EA and the structure of the fitness landscape. Rather, the intuitive approach helps to isolate the general problem with particular objective functions because the word "deception" itself reflects a fault in the objective function. A deceptive objective function will deceive search by actively pointing the wrong way.

2.1.2 Mitigating Deception

A common approach to preventing premature convergence to local optima in EC is by employing a diversity maintenance technique (Goldberg and Richardson, 1987; Hornby, 2006; Hu et al., 2005; Hutter and Legg, 2006; Mahfoud, 1995; Stanley and Miikkulainen, 2002). Many of these methods are inspired by speciation or niching in natural evolution, wherein competition may be mostly restricted to occur within the same species or niche, instead of encompassing the entire population. For example, fitness sharing (Goldberg and Richardson, 1987) enforces competition between similar solutions so that there is pressure to find solutions in distant parts of the search space. Similarly, hierarchical fair competition (Hu et al., 2005) enforces competition among individuals with similar fitness scores, and the age-layered population structure (ALPS; Hornby, 2006) approach enforces competition among genomes of different genetic ages. The fitness uniform selection scheme (FUSS; Hutter and Legg, 2006) removes the direct pressure to increase fitness entirely. In FUSS, an individual is not rewarded for higher fitness, but for a unique fitness value. This approach can be viewed as a search for novel fitness scores, which is related to the approach in this paper. Although these methods encourage exploration, if local optima are pervasive or genotypic difference is not well-correlated with phenotypic/behavioral difference, these methods may still be deceived.

Other methods for avoiding deception tackle the problem of ruggedness by rearranging the genome so that crossover respects genes whose effects are linked (Goldberg et al., 1989), or by building models of interactions among genes (Pelikan et al., 2001). When successful, the effect is to smooth a rugged fitness landscape by deriving additional information from an imperfect objective function. However, given a sufficiently uninformative objective function, the advantage of such modeling is impaired.

Still other methods seek to accelerate search through neutral networks (Barnett, 2001; Stewart, 2001). While these methods may decrease the amount of meandering that occurs in neutral networks, if an unlikely series of specific mutations is needed to exit the neutral network, then search may still be stalled for a long time.

Some researchers incrementally evolve solutions by sequentially applying carefully crafted objective functions to avoid local optima (e.g., Elman, 1991; Gomez and Miikkulainen, 1997; Van de Panne and Lamouret, 1995). These efforts demonstrate that to avoid deception it may be necessary to identify and analyze the stepping stones that ultimately lead to the objective so that a training program of multiple objective functions and switching criteria can be engineered. However, for ambitious objectives,

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

these stepping stones may be difficult or impossible to determine a priori. Additionally, the requirement of such intimate domain knowledge conflicts with the aspiration of machine learning.

In addition to single-objective optimization, there also exist evolutionary methods that aim to optimize several objectives at once: multi-objective evolutionary algorithms (MOEAs; Veldhuizen and Lamont, 2000). These MOEAs are not immune to the problem of deception (Deb, 1999), and adding objectives does not always make a problem easier (Brockhoff et al., 2007), but the idea is that perhaps deception is less likely when optimizing multiple objectives because if a local optimum has been reached in one objective, then sometimes progress can be made with respect to an alternate objective (Knowles et al., 2001). In this spirit, some researchers have experimented with multi-objectivization, that is, extending single-objective problems into multi-objective problems to avoid deception (Knowles et al., 2001). Decomposing a single-objective problem into a multi-objective problem can either make it easier or harder (Handl et al., 2008b), and it is necessary to verify that the single-objective optima are multi-objective optima in the transformed multi-objective problem (Knowles et al., 2001). There have been several successful applications of multi-objectivization (De Jong and Pollack, 2003; Greiner et al., 2007; Handl et al., 2008a; Knowles et al., 2001), but as in other reviewed methods, the fundamental pathology of deception remains a concern.

Yet another avenue of research in EC related to deception is coevolution. Coevolutionary methods in EC attempt to overcome the limitations of a static fitness function by making interactions among individuals contribute toward fitness. The hope is that such an interaction will spark an evolutionary arms race that will continually create a gradient for better performance (Cliff and Miller, 1995). There have been several impressive successes (Chellapilla and Fogel, 1999; Hillis, 1991; Sims, 1994), but a common problem is that in practice such arms races may converge to mediocre stable states (Ficici and Pollack, 1998; Pollack et al., 1996; Watson and Pollack, 2001), cycle among various behaviors without further progress (Cartlidge and Bullock, 2004b; Cliff and Miller, 1995; Watson and Pollack, 2001), or one species may so far out-adapt another that they are evolutionarily disengaged (Cartlidge and Bullock, 2004a, b; Watson and Pollack, 2001). The difficulty for practitioners in coevolution, much like the difficulty of crafting an effective fitness function in standard EC, is to construct an environment that provides sustained learnability in which the gradient of improvement is always present (Ficici and Pollack, 1998).

Finally, outside of EC there is also general interest in machine learning in avoiding local optima. For example, simulated annealing probabilistically accepts deleterious changes with respect to the objective function (Kirkpatrick et al., 1983), and tabu search avoids re-searching areas of the search space (Glover, 1989). However, if local optima are pervasive, then these methods too can fail to reach the global optimum.

While methods that mitigate deception may work for some problems, ultimately such methods do not address the underlying pathology, that the gradient of the objective function may be misleading or uninformative. Instead, current methods that deal with deception attempt to glean as much information as possible from an imperfect objective function or encourage exploration in the search space. Given a sufficiently uninformative objective function, it is an open question whether any method relying solely on the objective function will be effective. Thus, an interesting yet sobering conclusion is that some objectives may be unreachable in a reasonable amount of time by direct objective-based search alone. Furthermore, as task complexity increases, it is more difficult to successfully craft an appropriate objective function (Ficici and Pollack,

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

1998; Zaera et al., 1996). Thus, as experiments in EC become more ambitious, deception may be a limiting factor for success.

The next section discusses complexity in natural evolution, which, in contrast to traditional EAs, is a search process without any final objective.

2.2 Complexity in Natural Evolution

Natural evolution fascinates practitioners of search because of its profuse creativity, lack of volitional guidance, and perhaps above all, its apparent drive toward complexity.

A subject of longstanding debate is the arrow of complexity (Bedau, 1998; Miconi, 2007), that is, the idea that evolutionary lineages sometimes tend toward increasing complexity. What about evolutionary search in nature causes complexity to increase? This question is important because the most difficult problems in search, for example, an intelligent autonomous robot, may require discovering a prohibitive level of solution complexity.

The topic of complexity in natural evolution is much in contention across biology, artificial life, and evolutionary computation (Arthur, 1994; Gould, 1996; Lynch, 2007b; McShea, 1991; Miconi, 2007; Nehaniv and Rhodes, 1999; Stanley and Miikkulainen, 2003). One important question is whether there is selective pressure toward complexity in evolution. A potentially heretical view that is gaining attention is that progress toward higher forms is not mainly a direct consequence of selection pressure, but rather an inevitable passive by-product of random perturbations (Gould, 1996; Lynch, 2007b; Miconi, 2007). Researchers such as Miconi (2007) in artificial life, Sigmund (1993) in evolutionary game theory, and Gould (1996), McShea (1991), and Lynch (2007a,b) in biology argue that natural selection does not always explain increases in evolutionary complexity. In fact, some argue that to the extent that fitness (i.e., in nature, the ability to survive and reproduce) determines the direction of evolution, it can be deleterious to increasing complexity (Lynch, 2007b; Miconi, 2007; Sigmund, 1993). In other words, rather than laying a path toward the next major innovation, fitness (like the objective function in machine learning) in effect prunes that very path away.

In particular, Miconi (2007) illustrates this point of view, noting that selection pressure is in fact a restriction on both the scope and direction of search, allowing exploration only in the neighborhood of individuals with high fitness. Similarly, Sigmund (1993, p. 85) describes how high levels of selection pressure in nature can oppose innovation because sometimes ultimately improving requires a series of deleterious intermediate steps. Gould (1996), a biologist, goes even further, and argues that an increasing upper bound in complexity is not a product of selection, but simply the result of a drift in complexity space limited by a hard lower bound (the minimal complexity needed for a single cell to reproduce). Lynch (2007b), another biologist, argues that selection pressure in general does not explain innovation, and that nonadaptive processes (i.e., processes not driven by selection) are often undeservedly ignored. The role of adaptation is also relegated to a lower level in Kimura's influential neutral theory of molecular evolution (Kimura, 1983). Finally, Huyen (1995) suggests that neutral mutations in nature allow a nearly limitless indirect exploration of phenotype space, a process that this paper seeks to directly accelerate.

These arguments lead to the main idea in this paper that abandoning the search for the objective and instead simply searching explicitly for novel behaviors may itself exemplify a powerful search algorithm.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

2.3 Open-Ended Evolutionary Computation

The open-ended evolution community in artificial life aims to produce simulated worlds that allow a similar degree of unconstrained exploration as earth. While a precise definition of open-ended evolution is not uniformly accepted (Maley, 1999; Standish, 2003), the general intuition is that it should continually create individuals "of a greater complexity and diversity than the initial individuals of the system" (Maley, 1999). Tierra (Ray, 1992), PolyWorld (Yaeger, 1994), and Geb (Channon, 2001) are typical attempts to attain such a dynamic. In such systems, there is no objective beyond that of survival and reproduction. The motivation behind this approach is that as evolution explores an unbounded range of life forms, complexity will inevitably increase (Channon, 2001; Miconi, 2007).

Tierra (Ray, 1992) is an open-ended evolutionary model in which self-replicating digital programs compete for resources on a virtual computer. In several runs of the system, cycles of parasitism and corresponding immunity arose, demonstrating a basic coevolutionary arms race. However, Tierra and other similar systems (Adami et al., 2000; Taylor and Hallam, 1997) all inevitably struggle to continually produce novelty (Channon and Damper, 2000; Kampis and Gulyás, 2008; Standish, 2003).

Polyword (Yaeger, 1994) is also a simulation of an ecological system of competing agents, but gives these basic agents embodiment in a two-dimensional world. The only goal for the agents is survival. Interestingly, emergent behaviors such as flocking and foraging, although not directly specified by the system, result from the interactions and evolution of agents. However, as in the Tierra-like systems, eventually innovation appears to slow (Channon and Damper, 2000). Geb (Channon, 2001), which addresses criticisms of Polyworld, follows a similar philosophy.

Bedau and Packard (1991) and Bedau et al. (1998) have contributed to formalizing the notion of unbounded open-ended dynamics by deriving a test that classifies evolutionary systems into categories of open-endedness. Geb and others are distinguished by passing this test (Channon, 2001; Maley, 1999), but the results nevertheless do not appear to achieve the levels of diversity or complexity seen in natural evolution. This apparent deficiency raises the question of what element is missing from current models (Maley, 1999; Standish, 2003), with the common conclusion that more detailed, lifelike domains must be constructed (Kampis and Gulyás, 2008; Maley, 1999; Yaeger, 1994).

However, this paper presents a more general approach to open-ended evolution that is motivated well by the following insight from Standish (2003): "The issue of open-ended evolution can be summed up by asking under what conditions will an evolutionary system continue to produce novel forms" (p. 1). Thus, instead of modeling natural selection, the idea in this paper is that it may be more efficient to search directly for novel behaviors. It is important to acknowledge that this view of open-endedness contrasts with the more commonly accepted notion of prolonged production of adaptive traits (Bedau and Packard, 1991; Bedau et al., 1998). Nevertheless, the simpler view of open-endedness merits consideration on the chance that a dynamic that appears adaptive might be possible to capture in spirit with a simpler process. Another benefit of this approach is that it decouples the concept of open-ended search from artificial life worlds, and can thus be applied to any domain, including real-world tasks.

The experiments in this paper combine this approach to open-ended evolution with the NEAT method, which is explained in the next section.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

2.4 NeuroEvolution of Augmenting Topologies (NEAT)

Because in this paper behaviors are evolved that are controlled by artificial neural networks (ANNs), a neuroevolution (NE) method is required. The NEAT method serves this purpose well because it is both widely applied (Aaltonen et al., 2009; Allen and Faloutsos, 2009; Stanley et al., 2005; Stanley and Miikkulainen, 2002, 2004; Whiteson and Stone, 2006) and well understood. However, the aim is not to validate the capabilities of NEAT yet again. In fact, in some of the experiments, NEAT performs poorly. Rather, the interesting insight is that the same algorithm can appear ineffective in an objective-based context yet excel when the search is open-ended. Thus, as a common approach to NE, NEAT is a natural conduit to reaching this conclusion.

The NEAT method was originally developed to evolve ANNs to solve difficult control and sequential decision tasks (Stanley et al., 2005; Stanley and Miikkulainen, 2002, 2004). Evolved ANNs control agents that select actions based on their sensory inputs. Like the SAGA method (Harvey, 1993) introduced before it, NEAT begins evolution with a population of small, simple networks and complexifies the network topology into diverse species over generations, leading to the potential for increasingly sophisticated behavior; while complexifying the structure of an ANN does not always increase the Kolmogorov complexity of the behavior of the ANN, it does increase the upper bound of possible behavioral complexity by adding free parameters. Thus, in NEAT, simpler behaviors must be encountered before more complex behaviors. A similar process of gradually adding new genes has been confirmed in natural evolution (Martin, 1999; Watson et al., 1987), and fits well with the idea of open-ended evolution.

However, a key feature distinguishing NEAT from prior work in complexification is its unique approach to maintaining a healthy diversity of complexifying structures simultaneously. Complete descriptions of the NEAT method, including experiments confirming the contributions of its components, are available in Stanley et al. (2005), and Stanley and Miikkulainen (2002, 2004). Let us review the key ideas on which the basic NEAT method is based.

To keep track of which gene is which while new genes are added, a historical marking is uniquely assigned to each new structural component. During crossover, genes with the same historical markings are aligned, producing meaningful offspring efficiently. Speciation in NEAT protects new structural innovations by reducing competition among differing structures, thereby giving newer, more complex structures room to adjust. Networks are assigned to species based on the extent to which they share historical markings. Complexification, which resembles how genes are added over the course of natural evolution (Martin, 1999), is thus supported by both historical marking and speciation, allowing NEAT to establish high-level features early in evolution and then elaborate on them later. In effect, then, NEAT searches for a compact, appropriate network topology by incrementally complexifying existing structure.

It is important to note that a complexifying neuroevolutionary algorithm such as NEAT induces an approximate order over the complexity of behaviors discovered during search from simple to complex. An important difference between an ANN with five connections and one with five million connections is that the larger network, by virtue of having more free parameters, can exhibit more complex behaviors. Thus, as new nodes and connections are added over the course of evolution, the potential complexity of the behaviors representable by the network increases. Encountering simple behaviors first is significant because the *most* complex behaviors are often associated with irregularity and chaos. Thus a search heuristic that encounters them last makes sense.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

In the experiments in this paper, NEAT is combined with novelty search, which is explained next.

3 The Search for Novelty

Recall that the problem identified with the objective fitness function in EC is that it does not necessarily reward the intermediate stepping stones that lead to the objective. The more ambitious the objective, the harder it is to identify a priori these stepping stones.

The suggested approach is to identify novelty as a proxy for stepping stones; instead of searching for a final objective, the learning method is rewarded for finding any instance whose functionality is significantly different from what has been discovered before. This idea is related to the concept of curiosity and seeking novelty in reinforcement learning research and developmental robotics (Kaplan and Hafner, 2006; Oudeyer et al., 2005, 2007; Schmidhuber, 2003, 2006), though novelty search operates on an evolutionary timescale and is motivated by open-ended evolution. In novelty search, instead of measuring overall progress with a traditional objective function, evolution employs a measure of behavioral novelty called a novelty metric. In effect, a search guided by such a metric performs explicitly what natural evolution does passively, that is, gradually accumulating novel forms that ascend the complexity ladder.

For example, in a biped locomotion domain, initial attempts might simply fall down. The novelty metric would reward simply falling down in a different way, regardless of whether it is closer to the objective behavior or not. In contrast, an objective function may explicitly reward falling the farthest, which likely does not lead to the ultimate objective of walking, and thus exemplifies a deceptive local optimum. In contrast, in the search for novelty, a set of instances are maintained that represent the most novel discoveries. Further search then jumps off from these representative behaviors. After a few ways to fall are discovered, the only way to be rewarded is to find a behavior that does *not* fall right away. In this way, behavioral complexity rises from the bottom up. Eventually, to do something new, the biped would have to successfully walk for some distance even though it is not an objective.

At first glance, this approach may seem naive. What confidence can we have that a search process can solve a problem when the objective is not provided whatsoever? Where is the pressure to adapt? Yet its appeal is that it rejects the misleading intuition that objectives are an essential means to discovery. The idea that the objective may be the enemy of progress is a bitter pill to swallow, yet if the proper stepping stones do not lie conveniently along its gradient, then it provides little more than false security.

Still, what hope is there that novelty is any better when it contains no information about the direction of the solution? Is not the space of novel behaviors unboundedly vast, creating the potential for endless meandering? One might compare novelty search to exhaustive search. Of course a search that enumerates all possible solutions will eventually find the solution, but at enormous computational cost.

Yet there are good reasons to believe that novelty search is not like exhaustive search, and that in fact the number of novel behaviors is reasonable and limited in many practical domains. The main reason for optimism is that task domains on their own provide sufficient constraints on the kinds of behaviors that can exist or are meaningful, without the need for further constraint from an objective function.

For example, a biped robot can only enact so many behaviors related to locomotion; the robot is limited in its motions by physics and by its own morphology. Although the search space is effectively infinite if the evolutionary algorithm can add new genes (like

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

NEAT), the behavior space into which points in the search space collapse is limited. For example, after an evaluation, a biped robot finishes at a specific location. If the robot's behavior is characterized only by this ending location, all of the many ways to encode a policy that arrives at a particular point will collapse to the same behavior. In fact, the behavior space may often collapse into a manageable number of points, significantly differentiating novelty search from exhaustive enumeration.

Furthermore, novelty search can succeed where objective-based search fails by rewarding the stepping stones. That is, anything that is genuinely different is rewarded and promoted as a jumping off point for further evolution. While we cannot know which stepping stones are the right ones, if the primary pathology in objective-based search is that it cannot detect the stepping stones at all, then that pathology is remedied.

A natural question about novelty search is whether it follows any principle beyond naively enumerating all possible behaviors. The answer is that while it does attempt to find all possible behaviors over time, when combined with a complexifying algorithm such as NEAT, the order in which they are discovered is principled and not random. To understand why, recall that NEAT evolves increasingly complex ANNs. That way, the amount of nodes and connections and thus the maximal possible complexity of behaviors discovered by novelty search increases over time, ensuring that simple behaviors must be discovered before more complex behaviors. Furthermore, this ordering from simple to complex is generally beneficial because of Occam's razor. Thus there is an order in a complexifying search for novelty; it is just a different one than in fitness-based search. While fitness-based search generally orders the search from low to high fitness, a structurally-complexifying search for novelty generally orders it from low to high complexity, which is principled in a different way.

The next section introduces the novelty search algorithm by replacing the objective function with the novelty metric and formalizing the concept of novelty itself.

4 The Novelty Search Algorithm

Evolutionary algorithms such as NEAT are well suited to novelty search because the population of genomes that is central to such algorithms naturally covers a wide range of expanding behaviors. In fact, tracking novelty requires little change to any evolutionary algorithm aside from replacing the fitness function with a novelty metric.

The novelty metric measures how unique an individual's behavior is, creating a constant pressure to do something new. The key idea is that instead of rewarding performance based on an objective, novelty search rewards diverging from prior behaviors. Therefore, novelty needs to be measured. There are many potential ways to measure novelty by analyzing and quantifying behaviors to characterize their differences. Importantly, like the fitness function, this measure must be fitted to the domain.

The novelty of a newly generated individual is computed with respect to the behaviors of an archive of past individuals whose behaviors were highly novel when they originated; because novelty is measured relative to other individuals in evolution, it is driven by a coevolutionary dynamic. In addition, if the evolutionary algorithm is steady state (i.e., one individual is replaced at a time) then the current population can also supplement the archive by representing the most recently visited points.

The aim is to characterize how far away the new individual is from the rest of the population and its predecessors in behavior space, that is, the space of unique behaviors. A good metric should thus compute the sparseness at any point in the behavior space. Areas with denser clusters of visited points are less novel and therefore rewarded less.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

A simple measure of sparseness at a point is the average distance to the k-nearest neighbors of that point, where k is a fixed parameter that is determined experimentally. If the average distance to a given point's nearest neighbors is large, then it is in a sparse area; it is in a dense region if the average distance is small. The sparseness ρ at point x is given by

$$\rho(x) = \frac{1}{k} \sum_{i=0}^{k} \text{dist}(x, \mu_i), \tag{1}$$

where μ_i is the ith-nearest neighbor of x with respect to the distance metric dist, which is a domain-dependent measure of behavioral difference between two individuals in the search space. The nearest neighbors calculation must take into consideration individuals from the current population and from the permanent archive of novel individuals. Candidates from more sparse regions of this behavioral search space then receive higher novelty scores. It is important to note that this behavior space cannot be explored purposefully; it is not known a priori how to enter areas of low density just as it is not known a priori how to construct a solution close to the objective. Therefore, moving through the space of novel behaviors requires exploration.

If novelty is sufficiently high at the location of a new individual, that is, above some minimal threshold ρ_{min}, then the individual is entered into the permanent archive that characterizes the distribution of prior solutions in behavior space, similar to archive-based approaches in coevolution (De Jong, 2004). The current generation plus the archive give a comprehensive sample of where the search has been and where it currently is; that way, by attempting to maximize the novelty metric, the gradient of search is simply toward what is *new*, with no explicit objective. Although novelty search does not directly seek to reach the objective, to know when to stop the search, it is necessary to check whether each individual meets the goal criteria.

It is important to note that novelty search resembles prior diversity maintenance techniques (i.e., speciation) popular in EC (Goldberg and Richardson, 1987; Hornby, 2006; Hu et al., 2005; Mahfoud, 1995). These also in effect open up the search by reducing selection pressure, but the search is still ultimately guided by the fitness function. In contrast, novelty search takes the radical step of only rewarding behavioral diversity with no concept of fitness or a final objective, inoculating it to traditional deception.

It is also important to note that novelty search is not a random walk; rather, it explicitly maximizes novelty. Because novelty search includes an archive that accumulates a record of where search has been, backtracking, which can happen in a random walk, is mitigated in behavioral spaces of any dimensionality.

The novelty search approach in general allows any behavior characterization and any novelty metric. Although generally applicable, novelty search is best suited to domains with deceptive fitness landscapes, intuitive behavioral characterizations, and domain constraints on possible expressible behaviors.

Once objective-based fitness is replaced with novelty, the NEAT algorithm operates as normal, selecting the highest-scoring individuals to reproduce. Over generations, the population spreads out across the space of possible behaviors, continually ascending to new levels of complexity (i.e., by expanding the neural networks in NEAT) to create novel behaviors as the simpler variants are exhausted. The power of this process is demonstrated in this paper through two experiments with results that conflict with common intuitions, which are described next.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

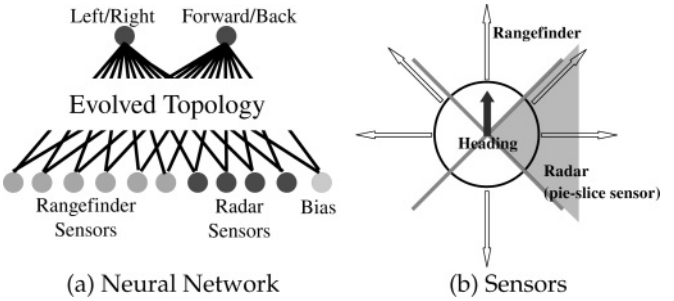

Figure 1: Maze navigating robot. (a) The artificial neural network that controls the maze navigating robot. (b) The layout of the sensors. Each arrow outside of the robot's body in (b) is a range finder sensor that indicates the distance to the closest obstacle in that direction. The robot has four pie-slice sensors that act as a compass toward the goal, activating when a line from the goal to the center of the robot falls within the pie slice. The solid arrow indicates the robot's heading.

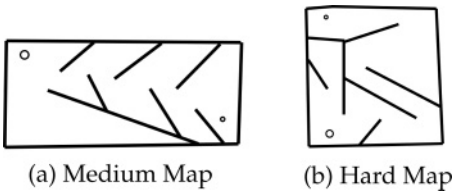

Figure 2: Maze navigation maps. In both maps, the large circle represents the starting position of the robot and the small circle represents the goal. Cul de sacs in both maps that lead toward the goal create the potential for deception.

5 Maze Experiment

An interesting domain for testing novelty search would have a deceptive fitness landscape. In such a domain, the search algorithm following the fitness gradient may perform worse than an algorithm following novelty gradients because novelty cannot be deceived with respect to the objective; it ignores objective fitness entirely. A compelling, easily visualized domain with this property is a two-dimensional maze navigation task. A reasonable fitness function for such a domain is how close the maze navigator is to the goal at the end of the evaluation. Thus, dead ends that lead close to the goal are local optima to which an objective-based algorithm may converge, which makes a good model for deceptive problems in general.

The maze domain works as follows. A robot controlled by an ANN must navigate from a starting point to an end point in a fixed time. The task is complicated by cul de sacs that prevent a direct route and that create local optima in the fitness landscape. The robot (see Figure 1) has six rangefinders that indicate the distance to the nearest obstacle and four pie-slice radar sensors that fire when the goal is within the pie slice. The robot's two effectors result in forces that turn and propel the robot. This setup is similar to the successful maze navigating robots in NERO (Stanley et al., 2005).

Two maps are designed to compare the performance of NEAT with fitness-based search and NEAT with novelty search. The first (see Figure 2(a)) has deceptive dead ends that lead the robot close to the goal. To achieve a higher fitness than the local optimum

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

provided by a dead end, the robot must travel part of the way through a more difficult path that requires a weaving motion. The second maze (see Figure 2(b)) provides a more deceptive fitness landscape that requires the search algorithm to explore areas of significantly lower fitness before finding the global optimum (which is a network that reaches the goal). Note that this task is not pathfinding. Rather, NEAT is searching for an ANN that itself can navigate the maze.

Fitness-based NEAT, which will be compared to novelty search, requires a fitness function to reward maze-navigating robots. Because the objective is to reach the goal, the fitness f is defined as the distance from the robot to the goal at the end of an evaluation: $f = b_f - d_g$, where b_f is a constant bias and d_g is the distance from the robot to the goal. Given a maze with no deceptive obstacles, this fitness function defines a monotonic gradient for search to follow. The constant b_f ensures all individuals will have positive fitness.

NEAT with novelty search, on the other hand, requires a novelty metric to distinguish between maze-navigating robots. Defining the novelty metric requires careful consideration because it biases the search in a fundamentally different way than the fitness function. The novelty metric determines the behavior-space through which search will proceed. It is important that the types of behaviors that one hopes to distinguish are recognized by the metric.

Thus, for the maze domain, the behavior of a navigator is defined as its ending position. The novelty metric is then the squared Euclidean distance between the ending positions of two individuals. For example, two robots stuck in the same corner appear similar, while a robot that simply sits at the start position looks very different from one that reaches the goal, though they are both equally viable to the novelty metric.

This novelty metric rewards the robot for ending in a place where none have ended before; the method of traversal is ignored. This measure reflects that what is important is reaching a certain location (i.e., the goal) rather than the method of locomotion. Thus, although the novelty metric has no knowledge of the final goal, a solution that reaches the goal can appear novel. In addition, the comparison between fitness-based and novelty-based search is fair because both scores are computed only based on the distance of the final position of the robot from other points. Furthermore, NEAT is given exactly the same settings in both (see Appendix A), so the only difference is the reward scheme.

Finally, to confirm that novelty search is indeed not anything like random search, NEAT is also tested with a random fitness assigned to every individual regardless of performance, which means that selection is random. If the maze is solved, the number of evaluations is recorded.

6 Maze Results

On both maps, a robot that finishes within five units of the goal counts as a solution. On the medium map, both fitness-based NEAT and NEAT with novelty search were able to evolve solutions in every run (see Figure 3(a)). Novelty search took on average 18,274 evaluations ($SD = 20{,}447$) to reach a solution, while fitness-based NEAT was three times slower, taking 56,334 evaluations ($SD = 48{,}705$), averaged over 40 runs. This difference is significant ($p < .001$; Student's t-test). NEAT with random selection performed much worse than the other two methods, finding successful navigators in only 21 out of 40 runs, which confirms the difference between novelty search and random search.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

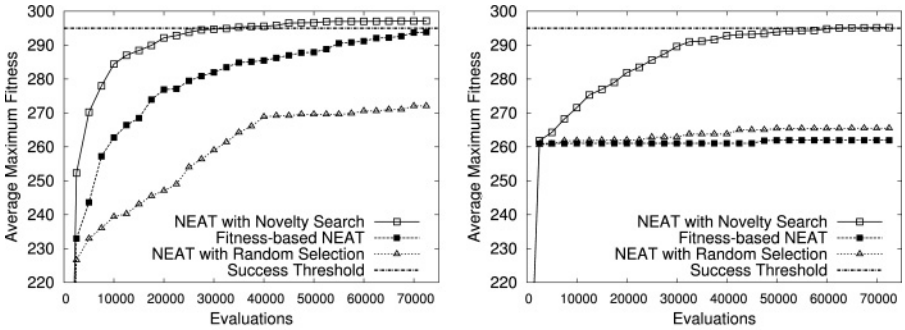

Figure 3: Comparing novelty search to fitness-based search. The change in fitness over time (i.e., number of evaluations) is shown for NEAT with novelty search, fitness-based NEAT, and NEAT with random selection on the medium (a) and hard (b) maps, both averaged over 40 runs of each approach. The horizontal line indicates at what fitness the maze is solved. The main result is that novelty search is significantly more effective. Only the first 75,000 evaluations (out of 250,000) are shown because the dynamics remain stable after that point.

Interestingly, the average genomic complexity of solutions evolved by fitness-based NEAT for the medium map (66.74 connections, $SD = 56.7$) was almost three times greater ($p < .05$; Student's t-test) than those evolved by NEAT with novelty search (24.6 connections, $SD = 4.59$), even though both share the same parameters.

On the hard map, fitness-based NEAT was only successful in three out of 40 runs, while NEAT with random selection fared marginally better, succeeding in four out of 40 runs, showing that deception in this map renders the gradient of fitness no more helpful than random search. However, novelty search was able to solve the same map in 39 out of 40 runs, in 35,109 evaluations ($SD = 30{,}236$) on average when successful, using 33.46 connections on average ($SD = 9.26$). Figure 3(b) shows this more dramatic divergence. Remarkably, because the second maze is so deceptive, the same NEAT algorithm can almost never solve it when solving the maze is made the explicit objective, yet solves it almost every time when finding novel behavior is the objective.

6.1 Typical Behavior

Figure 4 depicts behaviors (represented as the final point visited by an individual) discovered during typical runs of NEAT with novelty search and fitness-based NEAT on each map. Novelty search exhibits a more even distribution of points throughout both mazes. Fitness-based NEAT shows areas of density around local optima in the maze. The typical behavior of a successful robot on either maze was to directly traverse the maze for both methods.

The results so far raise a number of important questions and concerns, which the remainder of this section addresses.

6.2 Bounding the Size of the Archive in the Maze Domain

A possible concern about the computational effort required to search for novelty is that the archive of past behaviors may grow without limit as the search progresses.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

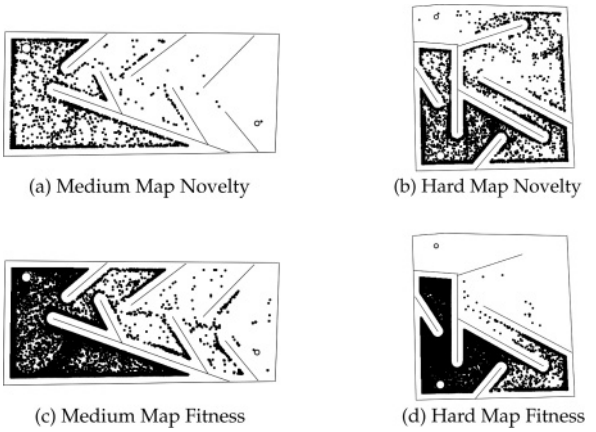

Figure 4: Final points visited over typical runs. Each maze depicts a typical run, stopping at either 250,000 evaluations or when a solution is found. Each point represents the end location of a robot evaluated during the run. Novelty search is more evenly distributed because it is not deceived.

As the size of the archive grows, the nearest-neighbor calculations that determine the novelty scores for individuals become more computationally demanding. Although in most complex domains the evaluation of individuals will likely be the computational bottleneck, it is true that the nearest-neighbor calculation increases the amount of computation beyond that required for a simple objective-based search method. Thus, it is interesting to consider ways in which the archive's size may be limited.

A simple approach is to change the structure of the archive from an ever-expanding list to a queue of limited size. When a new individual is added to the archive once the archive has reached its full size, the earliest entry is overwritten by the new individual instead of the new individual being appended to the end of the list. While this approach may lead to some backtracking through behavior space (which the archive is designed to avoid), the amount of backtracking may be limited because the dropped behaviors may no longer be reachable from the current population.

To explore the effects of limiting the archive size, 40 additional runs of novelty search were conducted in the hard maze with the size of the archive limited to the same size as the population (250). The hard maze was solved in all 40 runs, in 38,324 evaluations ($SD = 42{,}229$) on average, which is not significantly different from the original results of novelty search on the hard maze without a bounded archive. This result demonstrates that in some domains it is possible to limit the archive size, and thus the additional computational effort, without significantly decreasing the performance of novelty search.

6.3 Removing Walls in the Maze Domain

An important question for novelty search is whether it will suffer in relatively unconstrained domains. Domain constraints may provide pressure for individuals within a search for novelty to encode useful information about the domain because exploiting such information may be the easiest way for evolution to maximize novelty.

Both maps in Figure 2 are enclosed; that is, the space of possible behaviors is constrained by the outer walls of the map. It is interesting to consider an unenclosed

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

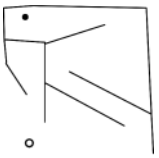

Figure 5: Unenclosed map. This unenclosed version of the hard map enlarges the behavior space because the robot can travel outside the maze without restriction.

map, in which the behavior space has no such limits. Figure 5 shows a variant of the hard map that has two walls removed; a navigator in such a map could travel into the vast empty area outside the map.

In such a map, the search for novelty may not efficiently explore the space of interesting behaviors that in some way relate to maze navigation. That is, it may be trivial to create variations of policies that appear novel by ending in significantly different areas outside the maze, yet encode no useful information about how to navigate within the maze. While in the constrained maze, learning about how to avoid walls may allow individuals to reach previously unexplored areas of the maze; in the unbounded maze, such learning may appear no more novel than a behavior that wanders off in a slightly different direction than previous robots.

To explore this issue, 100 runs of fitness-based NEAT and NEAT with novelty search were attempted in the unenclosed map (Figure 5). Interestingly, in this map, novelty search solved the maze only five times out of 100, which is not significantly better than fitness-based NEAT, which solved the maze two times out of 100. This result confirms the hypothesis that constraining the space of possible behaviors is important in some domains for novelty search to be efficient. However, fitness fares no better, highlighting that fitness-based search is not necessarily a viable alternative even when novelty search is not effective.

6.4 Conflation in the Maze Domain

Because the behavioral characterization in the maze experiment earlier consisted only of the location of the robot at the end of an evaluation, all navigators that end at the same point are conflated as equivalent even if they take entirely different routes. Thus they appear identical to the novelty search algorithm. Intuitively, a natural assumption is that in the maze domain, this conflation makes the search for novelty more efficient. By conflating individuals that end at the same location, the behavior space is greatly reduced and many possible uninteresting meandering behaviors that end at the same location are not rewarded for being different.

An important implication of this assumption is that lengthening the behavioral characterization might render the search for novelty intractable because the number of possible behaviors grows exponentially with the dimensionality of the characterization. However, as explained in Section 3, the hypothesis in this paper is that intuitive assumptions about the effect of the size of the behavior space do not take into account the effect of complexification and therefore wrongly predict that novelty search should fail in high-dimensional behavior spaces. Conversely, if the behavioral characterization does not include enough information, search might also be hindered. Interestingly,

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

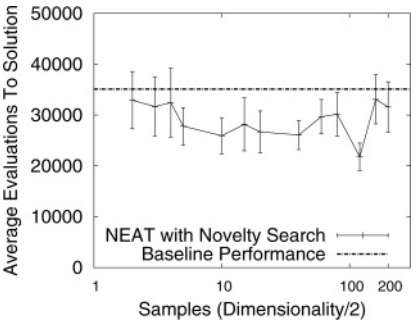

Figure 6: Increasing the dimensionality of the behavioral characterization. The average number of evaluations taken to solve the maze is shown for different numbers of samples of the navigator's location taken during an evaluation of NEAT with novelty search on the hard maze. For comparison, the baseline performance is shown as a separate line. Each point represents the average of 40 runs. The main result is that even with a behavioral characterization with 200 samples, the performance of novelty search is largely unaffected.

behavior can also be characterized by the fitness measure, which conflates all behaviors that would receive the same fitness score; if fitness as a behavioral characterization were always effective, then it would be trivial to apply novelty search to any domain that employs an objective function.

Three additional experiments in the maze domain are thus conducted to explore the effect of behavioral conflation by increasing the behavioral characterization's dimensionality, decreasing the amount of information in the behavioral characterization, and characterizing behavior by the fitness measure. The remaining experiments in this section all return to the regular bounded hard maze.

6.4.1 Lengthening the Behavioral Characterization

To increase the dimensionality of the behavioral characterization in the maze domain, the location of the navigator is sampled multiple times during an evaluation. For example, if the navigator's location is sampled 200 times during an evaluation, the dimensionality of the behavioral characterization becomes 400, which entails a vast space of behaviors. The question is whether an algorithm that simply searches for behavioral novelty can hope to find the objective behavior in such a large behavior space.

To answer this question, additional runs of novelty search in the (bounded) hard maze were conducted with behavioral characterizations of various sizes. In each such characterization, k equally-spaced samples of the navigator's (x, y) position are taken during an evaluation. Each (x_i, y_i) pair corresponds to the ith sample of the navigator's position in the maze. The resulting behavioral characterization is then the vector that concatenates all of these pairs, where $1 \leq i \leq k$, to form $(x_1, y_1, x_2, y_2, \ldots, x_k, y_k)$. The novelty metric remains the sum of the squared distances between the controllers' behavioral characterization vectors.

Figure 6 shows the results of novelty search with different characterization dimensionalities. Contrary to intuition, even in the 400-dimensional behavior space, novelty search still consistently finds controllers that reach the goal in the hard maze without

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

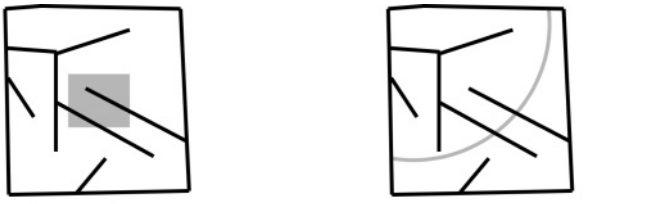

Figure 7: Comparing conflation across different behavioral characterizations. (a) Typical conflations of behavior are indicated in gray for reducing the amount of information in the behavioral characterization and (b) characterizing behavior by fitness. In (a) only geographically similar behaviors are conflated by a rectangle that is a part of a regular grid, whereas in (b), behaviors that end in very different locations (and are likely specified by very different policies) are conflated by the circle centered on the goal with radius equal to a particular fitness value.

any significant decrease in performance. Novelty search is largely unaffected by the vastness of the 400-dimensional behavioral characterization and is not lost exploring endless variants of uninteresting behaviors. The underlying rough order of behaviors imposed by a complexifying algorithm such as NEAT, from simple to more complex, may be beneficial to novelty search; as the behavioral characterization becomes higher-dimensional, conflation is reduced, but chaotic behaviors with no guiding principle still generally require a high level of structural complexity to be expressed in an ANN controller. That is, such unprincipled policies will be encountered later in search than simpler policies that exploit regularities in the environment. Thus, this experiment supports the hypothesis that high-dimensional behavior characterization is not a sufficient basis for predicting that novelty search should fail.

6.4.2 Reducing the Precision of the Behavioral Characterization

This section addresses what happens if conflation in the maze domain is increased such that some navigators that end in different locations within the maze nonetheless receive identical behavioral characterizations. Such conflation would effectively reduce the size of the behavior space and provide an upper bound on the possible size of the archive (see Section 6.2). The idea is to study the effect of reduced information on novelty search's ability to find the necessary stepping stones that lead to the ultimate objective.

To investigate this issue, the hard maze is discretized into a two-dimensional grid such that any two navigators that end in the same grid square are assigned the same behavioral characterization (see Figure 7(a)). In effect, as the grid becomes coarser, the resolution of the behavioral measure decreases. For example, if the grid is two by two, then only four distinct behaviors are recognized by the characterization. In this context, the original behavioral characterization for the maze domain had effective resolution of two single-precision floating point numbers (24 bits each) and thus could be approximately represented by a 2^{24} by 2^{24} grid. Conversely, the most trivial yet interesting grid that can be constructed is a two by two grid. To obtain a representative sampling of this wide range of grid sizes, 24 different behavioral characterizations are constructed. Of these, the ith characterization is the center of the final grid square

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

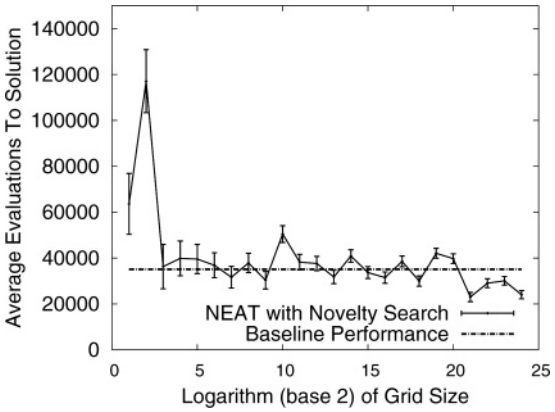

Figure 8: Reducing the amount of information in the behavioral characterization. The average number of evaluations taken to solve the hard maze is shown for behavioral characterizations that conflate behaviors that end in the same square of a virtual grid overlaid on the maze. Each point represents the average of 40 runs. The dotted line indicates the baseline performance. The main result is that even with a characterization that only recognizes four behaviors, novelty search still can consistently solve the same maze that fitness-based search nearly always fails to solve.

encountered by the navigator on a 2^i by 2^i grid overlaid on the hard maze. For each of these 24 characterizations, 40 additional runs of novelty search on the hard maze were conducted. If the maze is not solved within 500,000 evaluations, novelty search is restarted and the evaluations continue to accumulate until the maze is solved.

The results (see Figure 8) indicate that for all but the smallest grids (two by two and four by four), the performance of novelty search is largely unaffected. Only four runs out of 960 total runs failed to solve the hard maze in under 500,000 evaluations and required a restart. The worst-performing grid (four by four) recognized only 16 behaviors and required on average 117,107 evaluations to solve, which is about three times slower than the original results, but still significantly better than fitness-based search, which usually does not solve the hard maze at all.

This result is counterintuitive because it might be expected that discretizing the behavioral characterization would create plateaus in behavior space that would reduce novelty search in those areas to random search. However, even in these cases, as soon as the next plateau is discovered, novelty search then favors exploration there. Furthermore, the behaviors are conflated in a sensible way. Although nearby behaviors appear identical, distant behaviors still appear distant. Thus, novelty search still works well with large amounts of such behavioral conflation because the stepping stones are still respected, although at a lower resolution. This result suggests again that the archive size can be limited without significant loss of performance.

The next section explores characterizing an individual's behavior by its fitness value, which sometimes results in distant behaviors being conflated with each other.

6.4.3 Characterizing Behavior as the Fitness Measure

In traditional EAs, fitness in effect summarizes behavior with a single number, which conflates behavior in a fundamentally different way from the previous sections. In

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

the maze domain, fitness measures how close to the goal location the navigator ends. Thus, from a fitness standpoint, a navigator that ends at a specific point in the maze is conflated with all navigators that end on the circle of points that are at the same distance from the goal (see Figure 7(b)). Therefore, some behaviors are conflated that end at points separated by a large distance. In this way, the stepping stones may not be respected; if dead-end behaviors are discovered that mask distant stepping stones, the search for novelty may fail.

To test this hypothesis, a series of runs with fitness as the behavioral characterization were conducted in the maze domain. This kind of novelty search is very similar to FUSS (Hutter and Legg, 2006). The results validate the hypothesis that this type of conflation can be disruptive to novelty search. The maze was solved in only 11 out of 40 runs, though it is still better than the fitness-based NEAT performance, which only solved the hard maze in four out of 40 runs ($p < .05$; one-tailed Fisher's exact test). This result suggests that characterizing behavior by the fitness measure may be a simple way to improve performance over an objective-based method in a deceptive domain, although carefully constructing a more appropriate behavioral characterization may yield better results.

The conclusion is that behavioral conflation is complex, and cannot be characterized simply by the idea that more is worse, and less is better. Instead, it is more informative to think of conflation in terms of stepping stones. Conflation is harmful to the search for novelty if behaviors are conflated in a way that interferes with the discovery of stepping stones. Such harmful conflation can arise when characterizing the behavior of a navigator in terms of its fitness. Conversely, if some dimensions of behavior are completely orthogonal to the objective, it is likely better to conflate all such orthogonal dimensions of behavior together rather than explore them.

With the knowledge that novelty search can succeed with complex, high-dimensional behavior characterizations, the next section applies novelty search to the more challenging real-world domain of biped locomotion.

7 Biped Experiment

While the maze domain illustrates that the search for novelty can circumvent deception, a reasonable question is whether novelty search is simply well suited to the maze domain and therefore whether it can be applied successfully to more difficult domains.

Thus, there is a need to test novelty search on a well known problem of considerable difficulty. The intuition behind such an attempt is that novelty search may succeed in such a domain because problem difficulty is generally correlated with deceptiveness. That is, what makes a particular problem intractable to objective search is often that the gradient defined by the objective function nearly always leads to local optima. Because novelty search ignores the objective, it may be able to succeed in difficult domains simply by searching for novelty. The challenge domain in this paper is biped locomotion, a difficult control task that is popular within machine learning (Hein et al., 2007; McHale and Husbands, 2004; Reil and Husbands, 2002; Van de Panne and Lamouret, 1995). The key issue is whether the EA can overcome the deception inherent in the domain. Furthermore, the problem confronting novelty search is that the space of behaviors is far greater than in the maze domain and the solution, whatever gait is chosen, significantly more brittle from the need for balance and oscillation (Katić and Vukobratović, 2003).

In this domain, a three-dimensional biped robot in a realistic physics simulation is controlled by a type of ANN called a continuous time recurrent neural network

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

Figure 9: Biped robot. This visualization shows the biped robot controlled by an evolved ANN in the experiments in this paper.

(CTRNN) that is able to express the nonlinear dynamics found in natural gaits and is common in other biped experiments (McHale and Husbands, 2004; Reil and Husbands, 2002). The objective is to walk as far as possible within a given time limit. The task is difficult because it requires coordination, balance, and the discovery of oscillatory patterns. Initial random controllers for biped robots tend to provide a bad gradient for search because they and all of their immediate perturbations tend to simply fall (Van de Panne and Lamouret, 1995). Thus, even if oscillatory behavior is discovered (which could be useful), it is penalized and therefore ignored for falling down.

To mitigate deception, most methods that evolve controllers for biped robots implement domain-specific biases like enforcing symmetry (Allen and Faloutsos, 2009; Hein et al., 2007; Ok and Kim, 2005; Paul, 2003) or oscillation (Hein et al., 2007; Ishiguro et al., 2003), simplifying the task (Benbrahim, 1996; McHale and Husbands, 2004), or initially making the task easier (Allen and Faloutsos, 2009; Ishiguro et al., 2003; Van de Panne and Lamouret, 1995). These biases mitigate the deceptiveness of walking by restricting the search and introducing domain knowledge. However, they also impose a priori expectations on the creative evolutionary process. For example, imposing a constant oscillatory period may make the first few steps awkward and the restriction precludes considering alternatives. Thus, to make the problem as difficult and general as possible so that novelty search is forced to consider the entire space of possible behaviors, unlike typical approaches to biped walking, no domain-specific biases are implemented in the experiment in this paper.

The biped domain works as follows. A biped robot is controlled by an ANN for a fixed duration (15 simulated seconds). The evaluation is terminated if the robot falls or after the allocated time expires. The objective is that the robot must travel the greatest possible distance from the starting location.

The ANN that controls the biped has only two inputs, which are contact sensors that signal for each foot whether it is touching the ground. The sparsity of input makes the problem more difficult because the ANN has no information on the orientation of the robot or on the current angles of its joints.

The biped robot (see Figure 9) has a total of 6 DOF: two degrees in each hip joint (pitch and roll) and one degree in each knee joint (pitch). Simple sphere-shaped feet make ankle joints unnecessary in this model, although the lack of typical feet or a torso (which could provide a counterbalance) require knees to bend backward (as in birds) to balance, adding to the challenge of the model.

The ANN outputs movement requests for each DOF in the model, that is, for each independent axis of rotation for all joints in the model. The outputs are scaled to match the angular range of the corresponding DOF, which is interpreted as the angle that the neural controller is requesting. The difference between the requested angle and

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

the current orientation of the DOF denotes the disparity between the state the neural net is requesting and the current state of the model. A proportional controller applies torque to reduce this disparity. In other words, the ANN directs the low-level controllers toward a particular state. This model and method of control are similar to those in Reil and Husbands (2002).

This experiment again compares fitness-based NEAT to NEAT with novelty search. It is important to note that NEAT was extended for these experiments to evolve CTRNNs (which means that it also evolves the time constant assigned to each node). Parameter settings for NEAT in this experiment (which were the same for fitness and novelty) and the parameters of the biped physics simulation are given in Appendix A. A natural fitness function is the squared distance traveled from the starting location. This distance is measured by recording the location of the center of mass of the biped robot before and after evaluation, then calculating the distance between the two points. Distance traveled is a standard measure among evolutionary biped locomotion experiments (Hein et al., 2007; McHale and Husbands, 2004; Reil and Husbands, 2002; Van de Panne and Lamouret, 1995), and matches the intuitive notion of learning to walk with increasing stability and efficiency.

In contrast, NEAT with novelty search requires a behavioral characterization to distinguish biped gaits. The behavioral characterization in this domain is the offset of the biped's center of mass sampled at 1 s intervals during the evaluation:

$$x'_k = \text{sign}(x_k - x_0) * (x_k - x_0)^2, \tag{2}$$

$$y'_k = \text{sign}(y_k - y_0) * (y_k - y_0)^2, \tag{3}$$

where x_0 and y_0 correspond to the initial planar center of mass (i.e., ignoring the vertical z component) of the biped robot, and x_k and y_k correspond to the center of gravity sample taken after the kth second of simulation. The magnitude of the offsets is squared just as the fitness measure is to make the comparison uniform. If the robot falls, the center of gravity for all remaining samples is set to the robot's center of gravity when it fell. The behavioral characterization of a particular controller is then the vector that concatenates all pairs (x'_i, y'_i), where $1 \leq i \leq m$ and m is the final sample taken, to form $(x'_1, y'_1, x'_2, y'_2, \ldots, x'_m, y'_m)$.

The novelty metric for two different controllers is the same as in the maze domain, that is, the sum of the squared distances between the controllers' behavioral characterization vectors. Unlike in the maze domain, temporal sampling is necessary because the temporal pattern is fundamental to walking. This additional information allows the novelty metric to differentiate two gaits that end up at the same location by different means. However, most importantly, note that novelty search is ignorant of the objective of walking a long distance. In fact, to novelty search, walking clumsily and falling down can be rewarded just as much as walking well.

The comparison between novelty search and fitness-based search is fair because both the fitness function and novelty metric are calculated based on the robot's position. Although the novelty metric observes additional temporal information to differentiate biped gaits, the fitness function could not integrate this extra information without introducing a bias toward a certain style of walking. In contrast, novelty search integrates this information while remaining entirely agnostic about what is good or bad.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

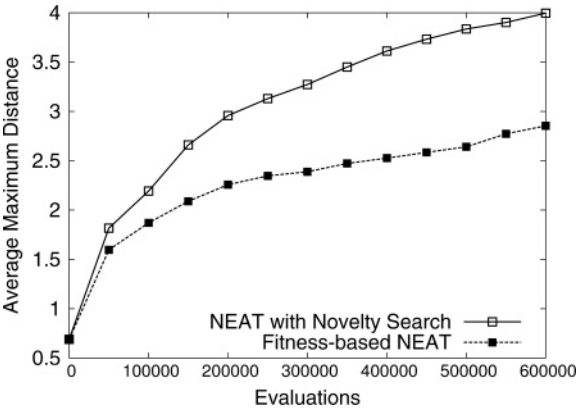

Figure 10: Comparing novelty search to fitness-based search in biped walking. The change in distance traveled by the best biped over time (i.e., number of evaluations) is shown for NEAT with novelty search and NEAT with fitness-based search on the biped task. Results are averaged over 50 runs. The main result is that, as in the maze domain, novelty search is significantly more effective.

8 Biped Results

Over 50 runs, novelty search evolved controllers that traveled on average 4.04 m ($SD = 2.57$) in the allotted 15 s while solutions evolved by fitness-based search traveled 2.88 m ($SD = 1.04$) on average. The difference in performance is significant ($p < .01$; Student's t-test). Furthermore, even in the beginning of a run, when one might expect a greedy search to temporarily show an advantage, novelty search still performs better (see Figure 10).

More dramatically, the best gait discovered by novelty search traveled 13.7 m, while the best gait discovered by fitness-based search traveled only 6.8 m. In fact, this latter solution was among only three gaits (in 50) discovered by fitness-based search that traveled a distance over 4 m (which was average for novelty search). Videos of these best evolved gaits are available at: http://eplex.cs.ucf.edu/noveltysearch/.

Qualitatively, the solutions evolved by both methods were different. A large proportion of runs of novelty search discovered oscillatory gaits (80%), while more than 40% of runs of fitness-based search converged to nonoscillatory gaits (i.e., less than four steps) corresponding to deceptive local optima of the objective function.

As in the maze domain, the average genomic complexity of champions from each run evolved by fitness-based NEAT (272.90 connections, $SD = 178.97$) was significantly larger ($p < .001$; Student's t-test) than those evolved by NEAT with novelty search (87.52 connections, $SD = 48.22$), even though both share the same parameters.

9 Discussion

Novelty search highlights the limitations of objectives, suggests a domain-independent concept of open-endedness, influences our interpretation of natural evolution, and foreshadows the potential for an artificial arrow of complexity. This section first discusses each of these implications in depth, and finally concludes by reviewing further results on novelty search that provide tentative evidence for its generality.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

9.1 Limitations of Objectives

Novelty search casts the performance of evolutionary algorithms in a new perspective. Based on the performance of fitness-based NEAT on the second maze and the biped, the usual conclusion would be that NEAT is ineffective for solving these problems. Yet NEAT with novelty search, which changes the reward function while preserving the rest of the algorithm, shows that the pathology is not in NEAT but rather in the pursuit of the objective itself. It is also notable that the problem is not simply diversity maintenance: NEAT itself already employs the diversity maintenance technique called explicit fitness sharing (Goldberg and Richardson, 1987), yet it is still fundamentally deceived when seeking higher fitness.

Novelty search also faces potential limitations. For example, because it ignores the objective, there is no bias toward optimization once a solution is found. An optimized solution may be produced by novelty search only if an individual can appear novel by reaching such a performance. However, it may be more efficient to take the most promising results from novelty search and further optimize them based on an objective function. This idea exploits the strengths of both approaches. Novelty search effectively finds approximate solutions, while objective optimization is good for tuning approximate solutions. One potential approach is to make novelty an objective and fitness another objective in a multi-objective formulation of the problem (Mouret, 2009). That is, it is possible to concurrently reward both fitness and novelty.

Although counterintuitive, the idea that the search for novelty can outperform the search for the objective introduces critical insight, to wit, objective fitness by necessity instantiates an imposing landscape of peaks and valleys. For complex problems it may be impossible to define an objective function through which these peaks and valleys create a direct route through the search space. Yet in novelty search, the rugged landscape evaporates into an intricate web of paths leading from one idea to another; the concepts of higher and lower ground are replaced by an agnostic landscape that points only along the gradient of novelty. What were once barriers become smooth links among nodes in a large lattice.

The problem with the objective is that it fails to identify the stepping stones. The more ambitious and complex the problem, the more difficult it is to formalize an objective that rewards the stepping stones along the way. Yet it is exactly those stepping stones that ultimately must be identified and rewarded if search is to find its way up the ladder of complexity (Miconi, 2007). Novelty search is designed to build gradients that lead to stepping stones. By abandoning the objective, all the steps along the way come into greater focus. While the trade-off is a more expansive search, it is better to search far and wide and eventually reach a summit than to search narrowly and single-mindedly yet never come close.

Of course, there are likely domains for which the representation is not suited to discovering the needed behavior or in which the space of behaviors is too vast for novelty search to reliably discover the objective, such as in the unenclosed maze in Section 6.3. Another example is the problem of pattern classification, which lacks the inherent domain restrictions of control tasks (e.g., simulated physics restricts the possible behaviors of the biped and walls restrict the movement of the maze navigator) and thus may not be directly amenable to a search for novelty without introducing some artificial constraints. Additionally, there are some domains in which it may be difficult to define behavior or a novelty metric; however, interestingly, Gomez (2009) defines a universal behavioral distance metric based on algorithmic information theory, and

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

research in adaptive curiosity in developmental robotics may also provide insight into defining appropriate characterizations of behavior (Kaplan and Hafner, 2006; Oudeyer et al., 2005, 2007). Characterizing when and for what reason novelty search fails is an important future research direction. Yet its performance has proven robust since it was first introduced in a conference paper in 2008 (Lehman and Stanley, 2008, 2010a; Mouret, 2009; Risi et al., 2009; Soltoggio and Jones, 2009).

Finally, because the results in this paper challenge common intuition, it is important to interpret them carefully. It would be a mistake to read them as a typical comparison between competing methodologies. In fact, we would be remiss if the reader infers a message that novelty search is "better" than objective-based search. Rather, a deeper interpretation is that search is about much more than objectives alone and ultimately can be guided by a diverse range of available information, much of it orthogonal to any explicit objective. Yet the field of EC, and even machine learning, has historically focused almost exclusively on this single guiding principle (i.e., objectives) among the many other kinds of searches that are possible, and hence the field is beginning to encounter the limits of what the objective-based paradigm has to offer. The results in this paper serve to confirm that indeed the objective is not the only impetus for search that can lead to interesting results. Thus this study is only beginning to scratch the surface of such alternative paradigms, which include natural evolution itself (which has no final objective in the search space). The remainder of this section ponders such connections and the possibilities that lie ahead.

9.2 Domain-Independent Open-Endedness

Novelty search suggests a perspective on open-endedness that is fitness-agnostic. Rather than viewing open-ended evolution as an adaptive competition, it can be viewed simply as a passive drift through the lattice of novelty. While this perspective bypasses a long-standing notion of adaptive innovation in open-ended evolution (Bedau and Packard, 1991; Bedau et al., 1998; Maley, 1999), it offers a complementary view that is recommended by its intuitive simplicity. Open-endedness can be defined simply as the continual production of novelty.

The benefit of this view is that it means that we can now endow any domain with this kind of open-endedness. No longer are we restricted to complex artificial life worlds in our pursuit of open-ended discovery. As long as novelty can be defined (which will not always be easy), it can be sought explicitly in every domain from simple XOR to the most complex artificial world, putting many practical problems in machine learning within its reach.

9.3 Novelty Search and Natural Evolution

Does the success of the search for novelty offer any insight into natural evolution? It is important to note that there is no evidence that natural evolution is explicitly a search for novelty. However, as Lynch (2007b) and Miconi (2007) suggest, it is often when the reigns of selection pressure are lifted that evolution innovates most prolifically. Novelty search can be viewed as an accelerated version of this passive effect in natural evolution; unlike in nature, it explicitly rewards drifting away in the phenotype/behavior space, thereby pushing the innovating process ahead.

Interestingly, there are several mechanisms in nature to encourage novelty. Novelty is preserved in nature as long as a novel individual meets minimal selection criteria.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

It is also encouraged through niching. By finding a new way to live, an organism may avoid competition and exploit untapped resources (Kampis and Gulyás, 2008). Moreover, there is evidence of active novelty search in natural evolution as well. Negative frequency dependent selection can encourage novelty (Endler and Greenwood, 1988) and lead to mate choice biased toward rare phenotypes (Hughes et al., 1999; Sigmund, 1993), which is the reward scheme implemented in novelty search. Thus it is not unreasonable to view natural evolution in part as a passive kind of novelty search subject to minimal criteria, which is an alternative to the more common interpretation of evolution as an adaptive competition.

This view of natural evolution leads to a unifying insight. To continue their lineages, all organisms ultimately must reproduce though they may live in vastly different ways. Therefore, perhaps natural evolution can be succinctly abstracted as an algorithm that finds many different ways to express the same functionality. Interestingly, in such an abstraction, the minimal criterion need not always be reproduction. In fact, it may be possible to construct minimal criteria in a practical domain such that the search for novelty is made more effective. Such minimal criteria, although they add restrictions to the search space, may provide a principled way of reducing the behavior space in tasks in which there are no inherent domain restrictions and a search for novelty alone may be infeasible. Thus, this perspective motivates future investigation into the search for novelty subject to minimal criteria. Recently, Lehman and Stanley (2010b) have begun investigating this research direction.

9.4 The Arrow of Complexity

Novelty search also provides a new hope for creating an artificial arrow of complexity. This follows the thought that once all the simple ways to live have been exhausted, the only way to do anything different is to become more complex (Gould, 1996). In a passive way, this explanation accounts for the arrow of complexity in nature. In novelty search, the principle should also hold true when coupled with a complexifying algorithm such as NEAT. Additionally, because novelty search actively encourages new behaviors, the arrow of complexity may be accelerated. In fact, the result that the ANN controllers from the maze domain and the biped walking task discovered by NEAT with novelty search contain about three times fewer connections than those discovered by fitness-based NEAT suggests that novelty search climbs the ladder of complexity more efficiently.

Although complexity may increase boundlessly within novelty search as the lower rungs of the ladder of complexity are filled, complexity is only ultimately interesting to humans with respect to functionality. For example, the search for novelty in the maze domain can continue long after the goal has been reached, generating new policies that yield continually novel behaviors by ending in slightly different areas of the maze. Each such policy's function can be viewed as reaching the point in the maze at which it ended. Because any infinitesimal coordinate in a maze can be seen as a functional task that a navigator can achieve, this view implies that open-ended evolution and complexity increase could potentially continue indefinitely. However, to a human observer the result of one of these potentially complex policies is underwhelming. The maze navigator simply navigates the maze and ends in a novel location.

The implication is that functional complexity alone is not interesting. Investigating the gap between functional complexity and interesting functional complexity is a direction for future research.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

9.5 Further Results on Novelty Search

A natural question is whether novelty search can also be applied with algorithms other than NEAT. In fact, since novelty search was first introduced by Lehman and Stanley (2008), several works have built upon it (Lehman and Stanley, 2010a; Mouret, 2009; Risi et al., 2009), some of which do not use NEAT or even neuroevolution.

Recently, the results from the maze domain were replicated in Mouret (2009) and combined with a multi-objective EA that did not include a genetic diversity-maintenance technique as in NEAT and was initialized with random network topologies instead of minimal networks as in NEAT. Experiments also compared optimizing various combinations of objectives on the same task; optimizing novelty as the sole objective evolved solutions the fastest, while combining novelty with fitness would find more exact solutions. Also, measuring the novelty of a new individual by comparing its behavior to both the population and the archive performed better than measuring novelty alone with respect to the archive.

Two independent works have also demonstrated that behavioral novelty is useful in evolving adaptive neural networks (i.e., neural networks that learn during their lifetimes; Risi et al., 2009; Soltoggio and Jones, 2009). Another recent result successfully applied novelty search to a different representation: genetic programming (GP; Lehman and Stanley, 2010a); on both a variant of the maze domain and the standard GP artificial ant benchmark, novelty search outperformed objective-based search. Thus, evidence is accumulating that novelty search is a general technique. While it is not always going to work well, the results so far suggest that novelty search is a viable new tool in the toolbox of EC.

10 Conclusions

This paper is a comprehensive introduction to novelty search (first introduced in Lehman and Stanley, 2008), a domain-independent method of open-ended search that illustrates the limitations of objectives. Motivated both by the problem of deceptive gradients in objective-based search and the desire for a simple approach to open-ended evolution, novelty search ignores the objective and instead searches only for individuals with novel behaviors. Counterintuitively, experiments in both a deceptive navigation task and a difficult biped locomotion task showed that novelty search can significantly outperform objective-based search. The idea that search can be more effective without an objective challenges fundamental assumptions and common intuitions about why search works. It is also the first machine learning approach to take seriously the growing (yet controversial) view in biology that adaptive selection does not explain the arrow of complexity in nature (Gould, 1996; Lynch, 2007a,b). Novelty search asks what is left if the pressure to achieve the objective is abandoned. Thus it teaches an important lesson on the limits of objective-driven search.

In summary, almost like a riddle, novelty search suggests a surprising new perspective on achievement: To achieve your highest goals, you must be willing to abandon them.

References

Aaltonen, T., Adelman, J., Akimoto, T., Álvarez González, B., Amerio, S., Amidei, D., et al. (2009). Measurement of the top quark mass with dilepton events selected using neuroevolution at CDF. *Physical Review Letters D, 79* (7): 072005.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

Adami, C., Ofria, C., and Collier, T. C. (2000). Evolution of biological complexity. *Proceedings of the National Academy of Sciences, USA*, 97:4463–4468.

Allen, B., and Faloutsos, P. (2009). Complex networks of simple neurons for bipedal locomotion. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4457–4462.

Arthur, W. (1994). On the evolution of complexity. In G. Cownan, D. Pines, and D. Meltzer (Eds.), *Complexity: Metaphors, Models and Reality* (pp. 65–81). Reading, MA: Addison-Wesley.

Barnett, L. (2001). Netcrawling—Optimal evolutionary search with neutral networks. In *Proceedings of the 2001 IEEE International Conference on Evolutionary Computation*, pp. 30–37.

Bedau, M. (1998). Four puzzles about life. *Artificial Life*, 4:125–140.

Bedau, M. A., and Packard, N. H. (1991). Measurement of evolutionary activity, teleology, and life. In C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (Eds.), *Proceedings of Artificial Life II*, pp. 431–461.

Bedau, M. A., Snyder, E., and Packard, N. H. (1998). A classification of longterm evolutionary dynamics. In C. Adami, R. Belew, H. Kitano, and C. Taylor (Eds.), *Proceedings of Artificial Life VI*, pp. 228–237.

Benbrahim, H. (1996). Biped dynamic walking using reinforcement learning. PhD thesis, University of New Hampshire, Durham, New Hampshire.

Brockhoff, D., Friedrich, T., Hebbinghaus, N., Klein, C., Neumann, F., and Zitzler, E. (2007). Do additional objectives make a problem harder? In *GECCO '07: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, pp. 765–772.

Cartlidge, J., and Bullock, S. (2004a). Combating coevolutionary disengagement by reducing parasite virulence. *Evolutionary Computation*, 12(2):193–222.

Cartlidge, J., and Bullock, S. (2004b). Unpicking tartan CIAO plots: Understanding irregular coevolutionary cycling. *Adaptive Behavior*, 12(2):69–92.

Channon, A. (2001). Passing the Alife test: Activity statistics classify evolution in Geb as unbounded. In *Proceedings of the European Conference on Artificial Life (ECAL-2001)*, 41–47.

Channon, A. D., and Damper, R. I. (2000). Towards the evolutionary emergence of increasingly complex advantageous behaviours. *International Journal of Systems Science*, 31(7):843–860.

Chellapilla, K., and Fogel, D. B. (1999). Evolving neural networks to play checkers without relying on expert knowledge. *IEEE Transactions on Neural Networks*, 10(6):1382–1391.

Cliff, D., and Miller, G. (1995). Tracking the red queen: Measurements of adaptive progress in co-evolutionary simulations. In F. Morán, A. Moreno, J. J. Merelo, & P. Chacón (Eds.), Advances in artificial life: Proceedings of the third European conference on artificial life (pp. 200–218). Berlin: Springer-Verlag.

Davidor, Y. (1990). Epistasis variance: A viewpoint on GA-hardness. In *Proceedings of the First Workshop on Foundations of Genetic Algorithms*, pp. 23–35.

De Jong, E. D. (2004). The incremental Pareto-coevolution archive. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2004)*, 525–536.

De Jong, E. D., and Pollack, J. B. (2003). Multi-objective methods for tree size control. *Genetic Programming and Evolvable Machines*, 4(3):211–233.

De Jong, K. A. (2006). *Evolutionary computation: A unified approach*. Cambridge, MA: MIT Press.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

Deb, K. (1999). Multi-objective genetic algorithms: Problem difficulties and construction of test problems. *Evolutionary Computation*, 7:205–230.

Elman, J. L. (1991). Incremental learning, or the importance of starting small. Technical Report 9101, CRL, La Jolla, CA.

Endler, J. A., and Greenwood, J. J. D. (1988). Frequency-dependent predation, crypsis and aposematic coloration [and discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 319(1196):505–523.

Ficici, S., and Pollack, J. B. (1998). Challenges in coevolutionary learning: Arms-race dynamics, open-endedness, and mediocre stable states. In *Proceedings of the Sixth International Conference on Artificial Life*, pp. 238–247.

Fogel, L., Owens, A., and Walsh, M. (1966). *Artificial intelligence through simulated evolution*. New York: John Wiley.

Glover, F. (1989). Tabu search, part I. *ORSA Journal on Computing*, 1(3):190–206.

Goldberg, D. E. (1987). Simple genetic algorithms and the minimal deceptive problem. In L. D. Davis (Ed.), *Genetic Algorithms and Simulated Annealing, Research Notes in Artificial Intelligence* (Chap. 6, pp. 71–88).

Goldberg, D. E., Deb, K., and Korb, B. (1989). Messy genetic algorithms: Motivation, analysis, and first results. *Complex Systems*, (3):493–530.

Goldberg, D. E., and Richardson, J. (1987). Genetic algorithms with sharing for multimodal function optimization. In *Proceedings of the Second International Conference on Genetic Algorithms*, pp. 41–49.

Gomez, F. (2009). Sustaining diversity using behavioral information distance. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2009)*, pp. 113–120.

Gomez, F., and Miikkulainen, R. (1997). Incremental evolution of complex general behavior. *Adaptive Behavior*, 5:317–342.

Gould, S. J. (1996). *Full house: The spread of excellence from Plato to Darwin*. New York: Harmony Books.

Grefenstette, J. J. (1992). Deception considered harmful. In *Foundations of Genetic Algorithms 2* (pp. 75–91). San Mateo, CA: Morgan Kaufmann.

Greiner, D., Emerador, J. M., Winter, G., and Galvan, B. (2007). Improving computational mechanics optimum design using helper objectives: An application in frame bar structures. In *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization*, 575–589.

Handl, J., Lovell, S. C., and Knowles, J. (2008a). Investigations into the effect of multiobjectivization in protein structure prediction. In *Proceedings of the 10th International Conference on Parallel Problem Solving from Nature*, pp. 702–711.

Handl, J., Lovell, S. C., and Knowles, J. (2008b). Multiobjectivization by decomposition of scalar cost functions. In *Proceedings of the 10th International Conference on Parallel Problem Solving from Nature*, pp. 31–40.

Harvey, I. (1993). The artificial evolution of adaptive behavior. PhD thesis, School of Cognitive and Computing Sciences, University of Sussex, England.

Harvey, I., and Thompson, A. (1996). Through the labyrinth evolution finds a way: A silicon ridge. In *Proceedings of the First International Conference on Evolvable Systems: From Biology to Hardware*, pp. 406–422.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

Hein, D., Hild, M., and Berger, R. (2007). Evolution of biped walking using neural oscillators and physical simulation. In *RoboCup 2007: Proceedings of the International Symposium, Lecture Notes in Artificial Intelligence* (vol. 5001, pp. 433–440). Berlin: Springer.

Hillis, W. D. (1991). Co-evolving parasites improve simulated evolution as an optimization procedure. In J. D. Farmer, C. Langton, S. Rasmussen, and C. Taylor (Eds.), *Artificial Life II*.

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.

Hordijk, W. (1995). A measure of landscapes. *Evolutionary Computation*, 4:335–360.

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)*, pp. 815–822.

Hu, J., Goodman, E., Seo, K., Fan, Z., and Rosenberg, R. (2005). The hierarchical fair competition (HFC) framework for sustainable evolutionary algorithms. *Evolutionary Computation*, 13(2):241–277.

Hughes, K. A., Du, L., Rodd, F. H., and Reznick, D. N. (1999). Familiarity leads to female mate preference for novel males in the guppy, *poecilia reticulata*. *Animal Behavior*, 58(4):907–916.

Hutter, M., and Legg, S. (2006). Fitness uniform optimization. *IEEE Transactions on Evolutionary Computation*, 10:568–589.

Huyen, M. (1995). Exploring phenotype space through neutral evolution. Working Papers 95-10-100, Santa Fe Institute.

Ishiguro, A., Fujii, A., and Hotz, P. E. (2003). Neuromodulated control of bipedal locomotion using a polymorphic CPG circuit. *Adaptive Behavior*, 11(1):7–17.

Jones, T., and Forrest, S. (1995). Fitness distance correlation as a measure of problem difficulty for genetic algorithms. In *Proceedings of the Sixth International Conference on Genetic Algorithms*, pp. 184–192.

Kampis, G., and Gulyás, L. (2008). Full body: The importance of the phenotype in evolution. *Artificial Life*, 14(3):375–386.

Kaplan, F., and Hafner, V. (2006). Information-theoretic framework for unsupervised activity classification. *Advanced Robotics*, 20(10):1087–1103.

Katić, D., and Vukobratović, M. (2003). Survey of intelligent control techniques for humanoid robots. *Journal of Intelligent Robotics Systems*, 37(2):117–141.

Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. New York: Oxford University Press.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge, UK: Cambridge University Press.

Kirkpatrick, S., Gelatt, C. D., Jr., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220:671–680.

Knowles, J. D., Watson, R. A., and Corne, D. (2001). Reducing local optima in single-objective problems by multi-objectivization. In *EMO '01: Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization*, pp. 269–283.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

Lehman, J., and Stanley, K. O. (2008). Exploiting open-endedness to solve problems through the search for novelty. In *Proceedings of the Eleventh International Conference on Artificial Life (ALIFE XI)*, pp. 329–336.

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, pp. 837–844.

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, pp. 103–110.

Liepins, G. E., and Vose, M. D. (1990). Deceptiveness and genetic algorithm dynamics. In G. J. E. Rawlins (Ed.), *FOGA* (*Foundations of Genetic Algorithms*), pp. 36–50.

Lynch, M. (2007a). The evolution of genetic networks by non-adaptive processes. *Nature Reviews Genetics*, 8:803–813.

Lynch, M. (2007b). The frailty of adaptive hypotheses for the origins of organismal complexity. In *Proceedings of the National Academy of Science, USA*, 104: 8597–8604.

Mahfoud, S. W. (1995). Niching methods for genetic algorithms. PhD thesis, University of Illinois, Champaign, IL.

Maley, C. C. (1999). Four steps toward open-ended evolution. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-1999)*, pp. 1336–1343.

Manderick, B., de Weger, M. K., and Spiessens, P. (1991). The genetic algorithm and the structure of the fitness landscape. In R. K. Belew and L. B. Booker (Eds.), *ICGA* (*International Conference on Genetic Algorithms*), pp. 143–150.

Martin, A. P. (1999). Increasing genomic complexity by gene duplication and the origin of vertebrates. *The American Naturalist*, 154(2):111–128.

McHale, G., and Husbands, P. (2004). Gasnets and other evolvable neural networks applied to bipedal locomotion. In *From Animals to Animats 8, Proceedings of the Eighth International Conference on the Simulation of Adaptive Behavior*, 163–172.

McShea, D. W. (1991). Complexity and evolution: What everybody knows. *Biology and Philosophy*, 6(3):303–324.

Miconi, T. (2007). Evolution and complexity: The double-edged sword. *Artificial Life: Special Issue on the Evolution of Complexity*, 14(3): 325–344.

Mitchell, M., Forrest, S., and Holland, J. H. (1992). The royal road for genetic algorithms: Fitness landscapes and GA performance. In F. J. Varela and P. Bourgine (Eds.), *Proceedings of the First European Conference on Artificial Life*, 245–254.

Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill.

Mouret, J.-B. (2009). Novelty-based multiobjectivization. In *Proceedings of the Workshop on Exploring New Horizons in Evolutionary Design of Robots, 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Nehaniv, C. L., and Rhodes, J. L. (1999). On the manner in which biological complexity may grow. In *Mathematical and Computational Biology. Lectures on Mathematics in the Life Sciences*, vol. 26 (pp. 93–102). Washington, DC: American Mathematical Society.

Ok, S., and Kim, D. (2005). Evolution of the CPG with sensory feedback for bipedal locomotion. In *Advances in Natural Computation* (pp. 714–726). Berlin: Springer.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

Oudeyer, P., Kaplan, F., and Hafner, V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11(2):265–286.

Oudeyer, P., Kaplan, F., Hafner, V., and Whyte, A. (2005). The playground experiment: Task-independent development of a curious robot. In *Proceedings of the AAAI Spring Symposium on Developmental Robotics*, pp. 42–47.

Paul, C. (2003). Bilateral decoupling in the neural control of biped locomotion. In *Proceedings of the 2nd International Symposium on Adaptive Motion of Animals and Machines*.

Pelikan, M., Pelikan, M., Goldberg, D. E., and Goldberg, D. E. (2001). Escaping hierarchical traps with competent genetic algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, pp. 511–518.

Pollack, J., Blair, A. D., and Land, M. (1996). Coevolution of a backgammon player. In *Proceedings of Artificial Life V*, pp. 92–98.

Ray, T. (1992). Evolution, ecology and optimization of digital organisms. Technical Report Working paper 92-08-042, Santa Fe Institute.

Reil, T., and Husbands, P. (2002). Evolution of central pattern generators for bipedal walking in a real-time physics environment. *IEEE Transactions on Evolutionary Computation*, 6(2):159–168.

Rice, H. G. (1953). Classes of recursively enumerable sets and their decision problems. *Transactions of the American Mathematical Society*, 74(2):358–366.

Risi, S., Vanderbleek, S. D., Hughes, C. E., and Stanley, K. O. (2009). How novelty search escapes the deceptive trap of learning to learn. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2009)*, 153–160.

Rothlauf, F., and Goldberg, D. E. (2002). *Representations for genetic and evolutionary algorithms*. Berlin: Physica-Verlag.

Schmidhuber, J. (2003). Exploring the predictable. In A. Ghosh and S. Tsutsui (Eds.), *Advances in Evolutionary Computing: Theory and Applications* (pp. 579–612). Berlin: Springer.

Schmidhuber, J. (2006). Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connection Science*, 18(2):173–187.

Sigmund, K. (1993). *Games of life: Explorations in ecology, evolution and behaviour*. New York: Oxford University Press.

Sims, K. (1994). Evolving 3D morphology and behavior by competition. In R. A. Brooks and P. Maes (Eds.), *Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems (Artificial Life IV)*, pp. 28–39.

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)*, 169–176.

Standish, R. (2003). Open-ended artificial evolution. *International Journal of Computational Intelligence and Applications*, 3:167.

Stanley, K. O., Bryant, B. D., and Miikkulainen, R. (2005). Real-time neuroevolution in the NERO video game. *IEEE Transactions on Evolutionary Computation Special Issue on Evolutionary Computation and Games*, 9(6):653–668.

Stanley, K. O., and Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evolutionary Computation*, 10:99–127.

Stanley, K. O., and Miikkulainen, R. (2003). A taxonomy for artificial embryogeny. *Artificial Life*, 9(2):93–130.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

Stanley, K. O., and Miikkulainen, R. (2004). Competitive coevolution through evolutionary complexification. *Journal of Artificial Intelligence Research*, 21:63–100.

Stewart, T. C. (2001). Extrema selection: Accelerated evolution on neutral networks. In *Proceedings of the 2001 IEEE International Conference on Evolutionary Computation*, 25–29.

Taylor, T., and Hallam, J. (1997). Studying evolution with self-replicating computer programs. In *Fourth European Conference on Artificial Life*, pp. 550–559.

Van de Panne, M., and Lamouret, A. (1995). Guided optimization for balanced locomotion. In D. Terzopoulos and D. Thalmann (Eds.), *Sixth Eurographics Workshop on Animation and Simulation*, pp. 165–177.

Veldhuizen, D. A. V., and Lamont, G. B. (2000). Multiobjective evolutionary algorithms: Analyzing the state-of-the-art. *Evolutionary Computation*, 8(2):125–147.

Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A., and Weiner, A. M. (1987). *Molecular biology of the gene, 4th ed.* Menlo Park, CA: Benjamin Cummings.

Watson, R., and Pollack, J. (2001). Coevolutionary dynamics in a minimal substrate. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2001)*, 702–709.

Weinberger, E. (1990). Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biological Cybernetics*, 63(5):325–336.

Whiteson, S., and Stone, P. (2006). Evolutionary function approximation for reinforcement learning. *Journal of Machine Learning Research*, 7:877–917.

Whitley, L. D. (1991). Fundamental principles of deception in genetic search. In *Foundations of genetic algorithms* (pp. 221–241). San Mateo, CA: Morgan Kaufmann.

Yaeger, L. (1994). Computational genetics, physiology, metabolism, neural systems, learning, vision and behavior or polyworld: Life in a new context. In C. G. Langton (Ed.), *Proceedings of Artificial Life III, Volume XVII*, pp. 263–298.

Yao, X. (1993). An empirical study of genetic operators in genetic algorithms. *Microprocessing and Microprogramming*, 38(1–5):707–714.

Zaera, N., Cliff, D., and Bruten, J. (1996). (Not) evolving collective behaviours in synthetic fish. In *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, 635–644.

Appendix A Experimental Parameters

In this paper, novelty search was implemented as an extension of NEAT, which when run with a traditional objective function also served as the control algorithm. Thus, to facilitate a fair comparison, NEAT with novelty search and objective-based NEAT have identical settings for all NEAT-specific parameters. Because the biped experiment is more challenging and operated with CTRNNs, the NEAT parameters required some adjustment, but were once again identical for both NEAT with novelty search and objective-based NEAT. The software package used in these experiments, Novelty Search C++, which is an extended version of the real-time NEAT (rtNEAT) software package, is available at: http://eplex.cs.ucf.edu.

Table A1 shows the parameters used in both experiments. NEAT has been found to be robust to moderate variations in parameters (Stanley et al., 2005; Stanley and Miikkulainen, 2002, 2004). Table A2 gives the physical parameters of the biped model.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021

Table A1: NEAT parameter settings. Archive threshold and *k*-nearest neighbor parameters apply only to NEAT with novelty search, while time constant and bias parameters apply only to the biped experiment.

Parameter	**Maze experiment**	**Biped experiment**
Population Size	250	500
c_1	1.0	1.0
c_2	1.0	1.0
c_3	3.0	3.0
C_t	Variable	Variable
Probability of adding link	0.1	0.06
Probability of adding node	0.005	0.005
Probability of mutating time constant	*N/A*	0.3
Probability of mutating bias	*N/A*	0.3
Initial archive threshold	6.0	1.0
k-nearest neighbors	15	15

Table A2: Biped simulation parameter settings. Parameters are given for the physical simulation of the biped robot, implemented using the freely available Open Dynamics Engine library (see http://www.ode.org). Maximum torque is the most torque a proportional controller can apply. The proportional constant is multiplied by the disparity between the actual angle of a joint and the angle that the ANN requests to derive the torque to be applied at that joint.

Parameter	**Value**
Foot radius	0.17 m
Foot density	1.0 kg/m^3
Torso radius	0.1 m
Torso length	0.33 m
Torso density	1.0 kg/m^3
Leg segment radius	0.2 m
Leg segment length	0.33 m
Leg segment density	1.0 kg/m^3
Maximum torque	5.0 N·M
Proportional constant	9.0

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/EVCO_a_00025 by guest on 26 September 2021