Effective Diversity Maintenance in Deceptive Domains

Joel Lehman
University of Texas at Austin
1 Inner Campus Drive
Austin, TX 78712 USA
joel@cs.utexas.edu

Kenneth O. Stanley
University of Central Florida
Orlando, FL 32816-2362 USA
kstanley@eecs.ucf.edu

Risto Miikkulainen
University of Texas at Austin
1 Inner Campus Drive
Austin, TX 78712 USA
risto@cs.utexas.edu

ABSTRACT

Diversity maintenance techniques in evolutionary computation are designed to mitigate the problem of deceptive local optima by encouraging exploration. However, as problems become more difficult, the heuristic of fitness may become increasingly uninformative. Thus, simply encouraging genotypic diversity may fail to much increase the likelihood of evolving a solution. In such cases, diversity needs to be directed towards potentially useful structures. A representative example of such a search process is novelty search, which builds diversity by rewarding behavioral novelty. In this paper the effectiveness of fitness, novelty, and diversity maintenance objectives are compared in two evolutionary robotics domains. In a biped locomotion domain, genotypic diversity maintenance helps evolve biped control policies that travel farther before falling. However, the best method is to optimize a fitness objective and a behavioral novelty objective together. In the more deceptive maze navigation domain, diversity maintenance is ineffective while a novelty objective still increases performance. The conclusion is that while genotypic diversity maintenance works in well-posed domains, a method more directed by phenotypic information, like novelty search, is necessary for highly deceptive ones.

Categories and Subject Descriptors
1.2.6 [Artificial Intelligence]: Learning—connectionism and neural nets, concept learning

General Terms
Algorithms

Keywords
NEAT, Diversity Maintenance, Novelty Search, Deception

1. INTRODUCTION

Deception is a fundamental problem in evolutionary computation (EC) wherein optimizing a fitness function often leads an evolutionary algorithm (EA) to converge to a suboptimal solution [7, 16]. Thus, many researchers have proposed mechanisms for combating deception, often by encouraging or maintaining genotypic diversity during evolutionary search [7, 11, 12, 18]. The motivation is that by exploring a wider set of possible solutions it should be less likely for search to prematurely converge to degenerate locally optimal solutions.

However, recent research in evolutionary robotics suggests that deception may be a more serious problem than previously thought [14, 17]. While fitness heuristics (and objective-based optimization approaches in general) are founded on the assumption that steps to the solution will increasingly resemble it, in complex problems steps that ultimately lead towards the objective may yet poorly resemble it [17]. Thus with increasing problem difficulty, fitness functions may sometimes become perversely deceptive, i.e. nearly all of the steps in the sequence of evolutionary innovations necessary to solve a particular problem may actively be punished by optimizing resemblance to the desired outcome. For example, at nearly all points along the shortest path through a circuitous maze, traveling directly towards the goal will actually oppose progress.

Problematically, in such perversely deceptive problems the fitness function becomes a liability, and genotypic diversity maintenance alone may not significantly increase the effectiveness of search because the gradient of increasing fitness remains search’s primary compass. Indeed, it has been shown in many deceptive problems that driving search only by phenotypic diversity can be more effective than guiding the search for the objective by the objective (i.e. optimizing fitness) [13, 16, 17, 22]. Although there exist a variety of approaches for encouraging phenotypic diversity in evolutionary robotics [9, 16, 20, 27], novelty search is different in that it can be used alone to solve problems, i.e. without being combined with pressure to optimize objective fitness [13, 14, 22].

Motivated by such results, Lehman and Stanley [17] proposed that search processes driven primarily by phenotypic diversity (e.g. novelty search) may scale to solve difficult problems more readily than search processes guided primarily by the objective (e.g. traditional objective-based fitness functions). The reason is that such search does not depend upon the fitness function conveniently identifying the order of stepping stones that lead to the objective. Instead, it relies on the assumption that previously discovered innovations are stepping stones to further evolutionary innovations, which may potentially hold more broadly. In other words,
while genotypic diversity maintenance creates random raw material from which fitness selects the most promising ones, novelty search discovers structure in the domain, thereby making progress possible even when fitness is uninformative.

This paper evaluates this hypothesis further by testing various combinations of common evolutionary search objectives (e.g., novelty, fitness, and diversity maintenance objectives) empirically on problems of increasing difficulty. In this way, in addition to exploring how fitness-based search and novelty search scale with respect to problem difficulty, insight may be gained into what combinations of objectives might be most useful for practitioners to apply to different types of problems.

To facilitate such a study, two well-studied evolutionary robotics domains are adapted to include an axis of increasing difficulty: biped locomotion [16, 21, 28] and maze navigation [16, 17, 20]. Across both domains and over problem instances of varying difficulty, various combinations of fitness, novelty, and diversity maintenance objectives are employed to guide evolutionary search.

In the biped domain, diversity maintenance techniques in general increase performance, although the overall most effective approach is optimizing a fitness objective and a novelty objective simultaneously. The biped domain is an example of a challenging but not perversely deceptive problem, wherein diversity can allow search to cross locally optimal fitness peaks to improve solution quality.

In contrast, in maze navigation there is little benefit from any of the genotypic diversity maintenance objectives. However, search still benefits from the novelty objective. The reason is that many of the intermediate products that ultimately lead to successfully navigating a particular maze may be actively punished by the fitness function. Because mazes can be seen as an abstract model for search problems in general, this result suggests that increasingly ambitious problems may in general become perversely deceptive.

The conclusion for practitioners is that for difficult real-world problems in which deception is not perverse, combining fitness-based search with novelty search is an effective approach, as has indeed been demonstrated before [20]. However, for researchers scaling evolutionary algorithms to more deceptive problems, the message is different: The objective-based search paradigm may be poorly-suited, and instead, open-ended evolutionary approaches such as novelty search may be needed.

2. BACKGROUND

This section reviews the NEAT neuroevolution approach, multi-objective optimization, diversity maintenance techniques in EC, and novelty search.

2.1 NEAT

In experiments described in this paper, behaviors are evolved for robots that are controlled by artificial neural networks (ANNs). Thus a neuroevolution (NE) method is needed to underpin these experiments. The NEAT method is appropriate because it is widely applied [1, 3, 16, 22, 24–26] and well understood.

The NEAT method was originally developed to evolve ANNs to solve difficult control and sequential decision tasks [24–26]. Evolved ANNs control agents that select actions based on their sensory inputs. Like the SAGA method [10] 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 increasingly sophisticated behavior. A similar process of gradually adding new genes has been confirmed in natural evolution [19]. This section briefly reviews the NEAT method; for comprehensive introductions see e.g. [25, 26].

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 and network complexities, 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 [19], is thus supported by both historical markings and speciation, allowing NEAT to establish high-level features early in evolution and then later elaborate on them. Further, NEAT’s ability to evolve increasingly complex ANNs fits well with this paper’s motivation of solving increasingly difficult problems.

Note that as originally described, NEAT speciates the population to encourage genotypic diversity. Because this paper’s experiments explore the effects of diversity maintenance on search, NEAT is extended here to run without speciation, and speciation is replaced by other diversity maintenance techniques. Additionally, NEAT is also extended to simultaneously optimize two objectives through multi-objective optimization, which is reviewed next.

2.2 Multi-objective Optimization

Multi-objective optimization is a popular paradigm within EC that addresses how more than one objective can be optimized at the same time in a principled way [4]. It has previously been applied to neuroevolution with promising results [2, 20]. Some of the experiments in this paper apply an implementation of NSGA-II [5], a well-established Pareto-based multi-objective search algorithm, to optimize two objectives concurrently.

The concept of dominance is central to Pareto-based multi-objective search; the key insight is that when comparing the performance of two individuals over multiple objectives, if both individuals are better on different subsets of the objectives then there is no meaningful way to directly rank such individuals because neither entirely dominates the other. That is, ranking such mutually non-dominating individuals would require placing priority or weight on one objective at the cost of another; traditionally one individual dominates another only if it is no worse than the other over all objectives and better than the other individual on at least one objective.

In this way, the best individuals in a population are those that are not dominated by any others. Such best individuals form the non-dominated front, which defines a series of trade-offs in the objective space. That is, the non-dominated front contains individuals that specialize in various combinations of optimizing the set of all objectives. Some will maximize one at the expense of all the rest, while some may focus equally on all of the objectives. In this way, various trade-offs of competing objectives such as novelty, fitness, and measures of diversity can be explored during a single
evolutionary run. The idea is that such trade-offs allow for principled exploration of both objectives in cases where there is no clear and principled mechanism for combining two objectives together into one.

In the experiments in this paper, one of such simultaneously explored objectives maintains diversity in the population, which is reviewed next.

### 2.3 Diversity Maintenance in EC

Because naively optimizing an objective function often results in converging to a point that is optimal only locally [7, 16], practical EC approaches often employ diversity maintenance techniques that incentivize evolving individuals for reasons beyond simply increasing performance [7, 11, 12, 18]. A central motivating idea is that diversity maintenance may allow an EA to avoid being entirely deceived by overcommitting to a singular gradient of increasing fitness that actually leads away from the solution. That is, diversity maintenance encourages exploring multiple promising pathways through the search space simultaneously in hopes of thereby reaching higher fitness peaks.

In general, what differentiates most diversity maintenance techniques is the choice of what dimension to characterize diversity. For example, diversity can be encouraged in the space of genotypes [8, 18], the age of genotypes [11, 23], the fitness of genotypes [12], or the space of phenotypic behaviors [9, 16, 20, 27]. Specifically, this paper applies genotypic speciation, an age objective, and a behavioral novelty objective as representative diversity maintenance techniques in evolutionary robotics.

Speciation, as applied in the canonical NEAT approach [25], separates individuals into different species through applying a genotypic distance metric. The idea is to encourage exploring ANNs with different topologies and also to ensure that crossover is restricted to recombining sufficiently similar ANNs. In particular, NEAT’s speciation is based on explicit fitness sharing [8].

In contrast, the age-layered population structure (ALPS; [11]) characterizes genotypes more coarsely, by their age, i.e. how many generations have elapsed since they were randomly generated. In ALPS, on a regular basis new random genetic material is injected into the system, and genomes are segmented into layers based on their age such that older genomes do not compete with younger genomes. The idea is that similar to how fresh PhDs do not compete with senior professors in the job market, younger genomes may benefit from reduced competition to potentially converge to a different and perhaps more promising optimum than their older peers.

As originally described, ALPS consists of many separated layers with specific rules for transmission between them [11]; however, Schmidt and Lipson [23] showed that simply having an age objective in the context of multi-objective search often achieves similar performance. The idea is that if all else is equal then a search should prefer a younger genome over an older one to increase diversity. Thus for reasons of simplicity the experiments in this paper apply this concept of an age objective instead of the full ALPS setup.

It is important to note that speciation and ALPS encourage diversity without taking into account any aspect of phenotypic behavior. However, because many genotypes may map to the same phenotype and many genes may serve no functional purpose, distance metrics based only on genotypes may not always encourage sufficient functional diversity to escape deceptive local optima [16]. Likewise, simply protecting younger genomes may not help escaping from deceptive fitness gradients if they are pervasive. In other words, most diversity maintenance techniques do not drive towards inherently useful information and would not make sense as the sole guidance for search [14]. In contrast, searching for novel behaviors encourages different functionality and can solve problems even when it is not combined with optimizing objective-based fitness [16]. Thus it may sometimes be more effective to encourage behavioral novelty or diversity [9, 16, 20, 27] instead of diversity in the genotypic space.

This motivates the next section, which reviews the idea of searching for behavioral novelty. Although such novelty search can be applied to complement the search for higher fitness, in its original conception it acts as the primary guidance for evolutionary search.

### 2.4 Novelty Search

In contrast to most EAs, which tend to converge, novelty search is a divergent evolutionary technique. It is inspired by natural evolution’s drive to novelty, and directly rewards novel behavior instead of progress towards a fixed objective [15, 16]. This paper applies novelty search alone and in combination with other objectives to examine how the performance of such differing combinations of objectives scales to solve problems of increasing difficulty.

Tracking novelty requires little change to any evolutionary algorithm aside from replacing the fitness function with a novelty metric, which measures how different an individual is from other individuals, thereby creating a constant pressure to do something new. The key idea is that instead of rewarding performance on an objective, novelty search rewards diverging from prior behaviors. Therefore, novelty needs to be measured.

The novelty metric characterizes how far away the new individual is from the rest of the population and its predecessors in behavior space, i.e. 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.

A simple measure of sparseness at a point is the average distance to the k-nearest neighbors of that point. Intuitively, 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 $\rho$ at point $x$ is given by

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

where $\mu_i$ is the $i$th-nearest neighbor of $x$ with respect to the distance metric $\text{dist}$, which is a domain-dependent measure of behavioral difference between two individuals in the search space. Candidates from more sparse regions of the behavior space then receive higher novelty scores.

If novelty is sufficiently high at the location of a new individual, i.e. above some minimal threshold $\rho_{\text{min}}$, then the individual is entered into the permanent archive that characterizes the distribution of prior solutions in behavior space. 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 towards what is new, with no other explicit objective. However, even without an
explicit objective, novelty search is still driven by meaningful information; that is, behaving in a novel way often requires learning the structure of the domain.

Once objective-based fitness is replaced with novelty (or alternatively is augmented by novelty), the underlying evolutionary algorithm operates as normal, selecting the most novel individuals to reproduce. Over generations, the population spreads out across the space of possible behaviors.

3. APPROACH

The approach in this paper is to study empirically how the effectiveness of different combinations of common objectives for EAs scale with problems of increasing difficulty. Thus a means of generating problems of variable difficulty is needed. For this purpose, two well-explored evolutionary robotics domains are adapted: biped locomotion [16, 21, 28] and maze navigation [16, 17, 20].

3.1 Variable Difficulty in Biped Locomotion

The biped locomotion domain from Lehman and Stanley [16] is adapted for these experiments. In this domain, the goal is to evolve an ANN controller for a three-dimensional simulated biped robot (figure 1) that can locomote as far as possible within the given time limit, which is challenging because of the need for balance and oscillatory gaits. As in Lehman and Stanley [16], the fitness of a biped controller is evaluated as the squared distance the robot walks before it falls, while its behavior (for calculating novelty) is derived from sampling its center of gravity over time. That is, fitness-based search looks for walkers that travel farther while novelty search rewards different walking trajectories. Exact details of the simulated biped robot can be found in Lehman and Stanley [16]. Note that in all the biped experiments in this paper, evolution was run for 2,000 generations with a population size of 500; other parameters are the same as those used by Lehman and Stanley [16].

In its original formulation, the biped problem is deceptive because an attractive local optimum is to lunge forward, which maximizes distance traveled without having any relation to a successful oscillatory gait [16, 28]. To create easier variations of this problem with less deception, supportive forces are added that act to keep the robot upright, like the approach that van de Panne and Lamounet [28] called “the hand of God.” In particular, assuming the angular velocity of the robot’s torso is \( \omega \), the ideal up vector representing the desired orientation of the robot’s torso is \( \upsilon \), and the current imperfect up vector is \( \upsilon - \omega \), and a difficulty parameter ranging from 0.0 to 1.0 is \( d \), then a lifting force is added with intensity \( 20(1.0 - d) \) newtons, a damping torque is added equal to \( -0.3(1.0 - d) \upsilon \times \omega \) newton-meters, and a corrective force is added equal to \( 2.0(1.0 - d) \upsilon \times \omega \) newton-meters. The idea of these forces is to make balancing easier, in order to reduce the difficulty in evolving an effective cyclical gait. Thus the deceptiveness of the problem can be manipulated by varying the difficulty parameter \( d \), which controls the intensity of the corrective forces (e.g. decreasing \( d \) decreases the supportive forces and makes the problem more deceptive). For example, when \( d \) is set to its maximum (1.0), then no supportive forces are applied and the domain is at its most deceptive and is equivalent to how it is described in Lehman and Stanley [16].

The next section describes how a similar axis of difficulty is created in the maze navigation domain.

3.2 Variable Difficulty in Maze Navigation

The maze navigation domain is adapted from Lehman and Stanley [16], where simulated robots controlled by evolved ANNs must navigate in a maze from a starting point to a goal point within a fixed amount of time. Traversal of the maze is complicated by walls that require circumnavigation. The simulated robot is shown in figure 2. Note that the simulation is ego-centric and continuous (i.e. the robot has incomplete local information about its environment, and its sensors, effectors, position, and orientation are represented by continuous variables), which makes the problem more difficult. As in previous studies [16, 17, 20], the fitness function measures how close a maze navigator ends to the goal location, which is an intuitive measure of progress in navigation tasks; likewise, for novelty search the behavior of a navigator is defined as the Cartesian coordinate of the navigator at the end of an evaluation. In this way, fitness-based search seeks individuals that minimize distance to the goal while novelty search seeks individuals that end in different locations from those previously encountered during search.

To generate a wide variety of problems, a recursive backtracking algorithm was applied to generate random mazes [6]. The initial and goal points of the maze were also randomly chosen, which, when considered with the walls, specifies a complete problem instance. The difficulty of such a generated problem was estimated by the length of the optimal path from the starting point to the goal, guided by the intuition that longer paths will in general require more complex policies to navigate (this intuition is supported empirically by the results).

In particular, 15 random maze problems on a 4x4 grid were generated for each path length ranging from four to 14, yielding a total of 165 problem instances. Examples of an easy and a difficult maze problem are shown in figure 3. Evolution was run for a maximum of 1,000 generations with a population size of 250; other parameters are the same as those used by Lehman and Stanley [16].
of NEAT to simultaneously optimize the age objective and either a fitness or novelty objective, which are, respectively, the fitness with age and novelty with age methods. Note that this multi-objective version of NEAT operates without speciation, thereby isolating the effects of the age objective to encourage genotypic diversity.

Finally, the last method combines a fitness objective and a novelty objective together, similarly to the multi-objectivization approach by Mouret and Doncieux [20]. The same multi-objective version of NEAT as in the age-objective method is applied to optimize both a novelty and fitness objective simultaneously, resulting in the fitness and novelty method. The idea is that novelty search may drive towards a diverse set of behaviors and that when it uncovers an individual in the basin of attraction of a solution, optimizing the fitness objective may then efficiently encourage that solution to be evolved.

The next section describes the results of applying these various methods to problems of increasing difficulty in the two domains.

4. RESULTS

Figure 4 compares the average performance of various methods across instances of the biped problem of increasing difficulty. Note that performance is measured as the average distance traveled by run champions; such performance of methods generally declines as problem difficulty increases, supporting the way problem difficulty was implemented in this domain. A coarser view of the results is given by figure 5, which aggregates a method’s average performance across all instances of the biped problem.

Note that the overall best-performing method is the fitness and novelty method, reflected by its significantly higher aggregate performance when compared pairwise to the other methods (Student’s t-test; p < 0.05). At a finer granularity, it also generally performs better than the other methods in the 11 biped domain difficulty variants (in 54 out of 66 such pairwise comparisons; Student’s t-test; p < 0.05). Additionally, figures 4a and 4b show that combining either fitness or novelty with speciation provides a general benefit; for both fitness and novelty, the differences in performance between including and not including speciation are significant for each of the 11 biped difficulty levels (Student’s t-test; p < 0.05). In contrast, augmenting novelty or fitness with an age objective is significantly deleterious in this domain. It is possible that using the complete ALPS setup may provide different results; in such a case these results would suggest that the results of Schmidt and Lipson [23] do not completely generalize.

Figure 4c compares the most representative techniques, showing that novelty search with speciation tends to outperform fitness-based search with speciation, although the advantage decreases with increasing problem difficulty. In particular, of the 11 possible pairwise comparisons, novelty with speciation is significantly better in the seven intermediate difficulty settings from 0.3 to 0.9, is significantly worse when difficulty is set to 0.2, and does not differ significantly in the three remaining difficulties (0.0, 0.1, and 1.0). An interesting result is the lack of significant difference between novelty search with speciation and fitness-based search with speciation when the difficulty is set to 1.0; this particular setting (where there are no supportive forces) is equivalent to the original problem from Lehman and Stanley [16]. How-
ever, in that paper novelty search did significantly outperform fitness-based search. This discrepancy suggests that the generational novelty search applied in the experiments in this paper may provide a less informative gradient of novelty than the steady-state novelty search of Lehman and Stanley [16]. That is, gradients of novelty are more stable when replacing only one individual at a time instead of replacing the entire population. Generational search was used (as in some prior novelty search studies [17, 20]) because the NSGA-II search algorithm chosen for optimizing multiple objectives is generational, and standardizing on generational search allows for uniformity and easier interpretation of results.

Next, the results for the maze navigation experiment are shown in figures 6 and 7. As in the biped domain, performance of all methods tends to decrease as problem difficulty increases. Interestingly however, the maze results exhibit little benefit from either adding diversity maintenance to fitness-based search (figure 6a) or novelty search (figure 6b), or from optimizing fitness and novelty simultaneously (fig 6c). Instead, the main differentiator is simply whether or not a novelty objective is included: With a novelty objective, more of the difficult problem instances are solved.

Supporting this statement, figure 7 presents the total proportion of successful evolutionary runs from each of the methods. This quantity is significantly higher for methods that include a novelty objective compared to those that do not (Fisher’s exact test; $p < 0.001$). There is no significant difference among the methods including a novelty objective.

Furthermore, when looking more granularly at the performance of methods across the 11 individual problem difficulty levels shown in figure 6, there are no significant differences among methods including a novelty objective nor among methods not including it. However, in T3 out of 132 possible pairwise comparisons of performance on problems of a specific difficulty between methods including novelty and those not including it, those including novelty performed significantly better. Also, such significant differences only occurred when difficulty was seven or higher. In the remaining 59 pairwise comparisons (mainly comprising the easiest problem instances) there was no significant difference in performance between the novelty methods and those without novelty.

5. DISCUSSION

The results in this paper differ significantly over the two explored domains. In the biped domain, combining fitness and novelty is more effective than optimizing either one separately, and there is some benefit achieved from actively maintaining genotypic diversity. In contrast, in the maze navigation domain the most important factor is whether a behavioral novelty objective is included or not, and there is little benefit from the other objectives, whether fitness or genotypic diversity maintenance.

Important, these two different problems can be seen as representative of the distinct challenges facing practitioners and researchers. For practitioners seeking to exploit EC to solve challenging problems, the biped problem provides a good model; it is deceptive and difficult, but the stepping stones to a solution are known (oscillatory leg movements and balance) and are generally recognized by intuitive fitness functions (e.g., distance traveled before falling). This is a problem that objective-based search can reasonably solve,
6. CONCLUSION

This paper evaluated how genotypic diversity maintenance techniques combined with fitness-based search and novelty search scale to solve problems of increasing complexity. Traditional diversity maintenance techniques facilitate greater exploration around high-fitness areas by generating random diversity on a genetic level. In contrast, searching for novelty can drive towards useful stepping stones that reflect underlying structure in a domain; that is, incorporating information about the domain into the genome often facilitates generating novelty. Thus in difficult but only superficially deceptive problems, genotypic diversity maintenance is often effective: Random diversity can aid escaping from shallow local optima. However, with increased deception the gradients of
the fitness function become overwhelmingly misleading and genotypic diversity maintenance may not provide much benefit. Thus, abstractions of natural evolution that are based on phenotypic diversity, and especially those that discover fundamental structure in the domain, may become increasingly important to EC researchers hoping to scale EC to evolve biological levels of complexity.

References


