Searching for Quality Diversity when Diversity is Unaligned with Quality

In: Proceedings of the 14th International Conference on Parallel Problem Solving From Nature (PPSN 2016). New York, NY: Springer.

Justin K. Pugh, L. B. Soros, and Kenneth O. Stanley

University of Central Florida, Department of Computer Science,
4328 Scorpius Street, Orlando, FL 32816-2362, USA
{jpugh,lsoros,kstanley}@cs.ucf.edu

Abstract. Inspired by natural evolution's affinity for discovering a wide variety of successful organisms, a new evolutionary search paradigm has emerged wherein the goal is not to find the single best solution but rather to collect a diversity of unique phenotypes where each variant is as good as it can be. These *quality diversity* (QD) algorithms therefore must explore multiple promising niches simultaneously. A QD algorithm's diversity component, formalized by specifying a *behavior characterization* (BC), not only generates diversity but also promotes quality by helping to overcome deception in the fitness landscape. However, some BCs (particularly those that are *unaligned* with the notion of quality) do not adequately mitigate deception, rendering QD algorithms unable to discover the best-performing solutions on difficult problems. This paper introduces a solution that enables QD algorithms to pursue arbitrary notions of diversity without compromising their ability to solve hard problems: driving search with multiple BCs simultaneously.

Keywords: novelty search, non-objective search, quality diversity, behavioral diversity, neuroevolution

1 Introduction

Evolutionary computation (EC) has developed increasingly sophisticated search algorithms around the idea that increasing fitness is a powerful mechanism for optimization [2]. However, natural evolution is more than an optimizer. Unlike conventional optimization, nature has no single unifying target and often rewards being different in addition to being better. Indeed, natural evolution has discovered a vast diversity of organisms and ways of being, simultaneously solving an uncountable and ever-changing array of problems from sight to ambulation to cognition, not by finding a single "best" solution to each but instead by collecting a breadth of viable alternatives.

In a step away from EC's longstanding fixation on fitness for the purpose of optimization, a new algorithm called novelty search (NS) [8] was introduced, which searches only for diversity and is notably free from objective pressure. Ironically, novelty search and its variants [4, 13] were initially heralded themselves as powerful tools for optimization because their agnosticism to the objective sometimes allows them to bypass the problem of *deception* and thus succeed on tasks where traditional objective-based approaches fail. However, NS's ability to collect a wide breadth of phenotypes is largely unappreciated when applied as an optimization algorithm: any accumulated diversity in such an application is eventually discarded in favor of saving only the best-performing individual.

Taking NS in a different direction, a unique search paradigm has begun to emerge within EC wherein diversity itself is a desirable end product. New algorithms such as Novelty Search with Local Competition (NSLC) [9] and MAP-Elites [12] stand apart from the usual focus on optimization in that rather than simply trying to find the single best individual (or tradeoffs among a set of objective targets [3]), these algorithms are instead designed to find *quality diversity* – a maximally diverse collection of individuals in which each member is as high-performing as possible. For example, one classic application of QD is to collect as many successful ambulating virtual creature morphologies as possible [9, 18]. QD is distinct from other approaches designed to return multiple results (such as those that seek to return a handful of local optima) in that all parts of the diversity space are considered equally important and the goal is to sample the entire space, returning the best possible performance in each region (even lower-performing regions). Compared to simple optimization, QD represents a new style of search that more closely embodies the spirit of natural evolution and for which evolutionary treatments are uniquely well-suited due to their natural inclination for exploring many promising directions at the same time.

Applying QD algorithms such as NSLC or MAP-Elites requires both a notion of quality (a fitness function) and a notion of diversity, called the *behavior characterization* (BC), which defines the degree of difference between two individuals. So far, applications of QD have largely featured characterizations that are *unaligned* with quality, which means that where an individual is located in the diversity space has little bearing on its potential performance; examples include the time individual legs of a hexapod robot spend on the ground [1], the specific image class targeted by a generated image [14, 15], and the size and shape of ambulating stick-figure creatures [18]. We can see that this focus on characterizations that are orthogonal to quality is natural by examining our intuitive sense of QD in nature. Indeed, Earth has accumulated a diverse repertoire of organisms with respect to intuitive characterizations such as size, appearance, or locomotion strategy, but those characterizations themselves are not good predictors of a particular organism's reproductive capacity or cognitive function (intuitive measures of quality). In effect, the types of diversity that we consider to be interesting or salient are often unaligned with our notions of quality.

This observation is interesting because a recent study comparing state-of-the-art QD algorithms in a relatively easy maze domain called the "QD-Maze" (inspired by the "HardMaze" domain [8] that has become ubiquitous in studies involving NS) indicates that the degree of characterization-quality alignment has a significant impact on performance and furthermore suggests that unaligned characterizations may be sub-optimal for driving search [16]. If this hypothesis is true, then typical approaches to QD may break down on harder problems.

The goal of this paper is to specifically address the challenge of finding "unaligned QD" in the context of a difficult maze domain (such that finding solutions is non-trivial even for the most sophisticated approaches). Experimental evidence in this domain confirms that driving search with an unaligned BC indeed has catastrophic effects on the ability of QD algorithms to successfully collect QD. As a solution, this paper introduces the idea of driving search with multiple BCs *simultaneously.* The success of this new approach offers a promising strategy for applying QD algorithms even when there is an incongruity between the desired notion of diversity and the ideal characterization for driving search, thus opening the door to a wider breadth of potential domains in the future.

2 Domain: QD-Gauntlet

Because the QD-Maze domain of Pugh et al. [16] is relatively easy, its results in effect speak to search spaces with a variety of relatively simple solutions. Yet many spaces of interest in the future will likely require a significant degree of complexity to find the interesting needles in the haystack. Earth itself has this quality, where there are innumerable different species, yet each is highly complex in its own right. Thus the new *QD-Gauntlet* domain in this paper builds on the old QD-Maze by greatly increasing the complexity of the possible paths to solutions, but still providing the opportunity for variety among those paths. In particular, like the QD-Maze [16], the QD-Gauntlet is a maze domain featuring an egocentric robot with multiple viable paths to the goal. An egocentric maze is an appealing platform for studying QD algorithms because the results are easily visualized and egocentric mazes are well-studied in the context of novelty search, where results originally obtained on HardMaze [8] have been shown to generalize to a variety of other domains such as quadruped locomotion [11], game content generation [10], swarm robotics [4], and image classification [19]. Furthermore, while the *Euclidean distance to the goal* heuristic that traditionally drives search in maze domains is known to be deceptive (which is the primary source of difficulty), in this domain we can also compute the perfect solution paths, enabling concrete measurements of the progress towards solving the maze.

The new maze, QD-Gauntlet (Fig. 1), is significantly more complex than its predecessors. In QD-Gauntlet, there are four distinct corridors leading horizontally to a goal point on the right side of the maze. Each corridor (i.e. leg) of the maze is composed of four successive segments, where each segment is designed to be approximately equivalent to the size and complexity of the canonical HardMaze [8]. To ensure that the four legs are similarly difficult, the legs are near mirror images of each other with slight variations. Importantly, QD-Gauntlet contains several long, straight corridors that terminate in a dead-end close to the goal. These dead-end corridors serve to increase maze difficulty by deceiving quality-seeking (i.e. fitness-based) search mechanisms. To further increase maze difficulty, agents are given strict time constraints such that deviating too far from one of the optimal paths will cause the agent to run out of time before reaching the goal.

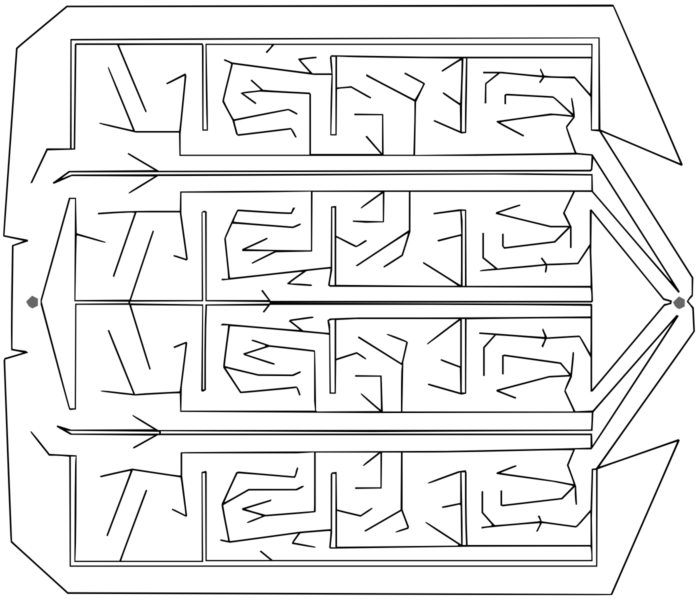

Fig. 1. QD-Gauntlet. An egocentric agent begins on the start point (left) and is presented with four viable paths to reach the goal point (right). The maze is riddled with dead-ends that serve to deceive objective-oriented search algorithms.

As in Lehman and Stanley [8], agents are driven by evolved neural networks and are equipped with a set of six wall-sensing rangefinders (five spanning the frontal 180 degrees and one facing the rear), four pie-slice sensors to sense the distance and relative direction of the goal point, and a single output to specify left-right turns. The challenge then is not to evolve a path, but rather to evolve a neural network that can correctly guide the agent through one of the long and deceptive corridors based on its sensory inputs. Furthermore, the hope is that QD algorithms can find *multiple* such solutions in the same run, corresponding to a variety of different driving strategies. Doing so would be a proxy for finding QD in any domain whose solutions are challenging and deceptive to reach.

3 Algorithms

This section describes the algorithms considered in this study, focusing first on variants of novelty search from the literature. Then, two additional algorithms are introduced to address the problem of finding QD when the desired notion of diversity is not aligned with quality. The hope is that the results from this study will also provide guidance for other QD algorithms in the future, such as MAP-Elites [1, 12].

Fitness. Included as a baseline to establish domain difficulty, a purely fitness-based search is implemented as standard generational NEAT [17] with a population size of 500, where fitness is the deceptive *Euclidean distance to the goal* heuristic (which also drives the quality portion of the QD algorithms that follow).

With the exception of Fitness, all of the other algorithms in this paper are implemented as steady state (only a small portion of the population is replaced at a time to avoid radical shifts in what is considered "novel" from one tick to the next: genomes are replaced in batches of 32 to allow moderate parallelism).

NS. While not technically a "quality diversity" algorithm because there is no quality component, the *novelty search* (NS) [8] algorithm forms the foundation of a number of other algorithms in this study and serves as another baseline for comparison, establishing what is possible without a drive towards quality. Novelty search works by rewarding *novelty* instead of fitness, where novelty measures how different an individual's behavior is from those that have been seen before. More formally, novelty is calculated by summing the distance to the k-nearest behaviors (in this paper, $k = 20$) from a set composed of the current population and an *archive* of past behaviors. The distance between two behaviors is simply the Euclidean distance between those behaviors when represented as a vector of numbers (called a *behavior characterization*). While there exist several different strategies for managing the archive [5], preliminary experiments indicated that a powerful strategy is to add all individuals to an archive with a maximum size that is enforced by deleting those with the lowest novelty (novelty is recomputed against the archive before each deletion). In all cases, NS is run with a population size of 500 and a maximum archive size of 2,500.

NSLC. *Novelty search with local competition* (NSLC) [9] combines the diversifying pressure of NS with a localized drive towards quality called *local competition* (LC), calculated as the proportion of 20 nearest behavioral neighbors with a lower fitness score. LC encourages increasing performance within local behavioral neighborhoods without suffering the deleterious effects of a global objective pressure. Novelty and LC are combined by Pareto ranking as in the NSGA-II multi-objective optimization algorithm [3].

3.1 Multi-BC QD Algorithms

In each of NS and NSLC, search is driven by some notion of behavioral diversity (i.e. a BC). Traditionally, the BC that drives search corresponds to the type of diversity that the researcher is interested in collecting (e.g. different types of robot morphologies or different walking gaits) and thus is typically unaligned with the notion of quality. Unaligned BCs are less capable of overcoming deception [7, 16] and on difficult tasks (such as QD-Gauntlet) may altogether fail to obtain high-performing solutions, creating a problem for researchers interested in finding unaligned QD. As a solution, we introduce the idea here of driving search with *multiple* BCs simultaneously and propose two possible methods for doing so. While each method can conceivably support three or more BCs, for simplicity the experiments that follow are restricted to only two BCs.

NS-NS. The basic NS algorithm can be extended to support multiple BCs simultaneously by combining their respective novelty scores in a multi-objective formulation (with NSGA-II [3]). In this algorithm, dubbed NS-NS, each BC maintains its own independent archive and individuals are evaluated against each archive in turn to calculate one novelty score per BC. There is only a single breeding population where the breeding potential for each member is decided by Pareto ranking according to novelty scores. The key idea is that in a two-BC formulation, one BC may be ideal for driving search while the other corresponds to the type of diversity the user is interested in collecting.

NS-NSLC. While NS-NS facilitates searching with multiple concepts of diversity, it lacks the explicit drive towards quality that is essential to QD algorithms. This omission is remedied in NS-NSLC by adding a local competition objective (in the same way as in NSLC) where behavioral neighbors are decided by the (unaligned) BC that corresponds to the user's desired notion of diversity.

4 Experiment

As discussed in Sect. 3, a common component of all QD algorithms is the BC, which formalizes the notion of diversity so that it can drive the search explicitly. In domains where there is really only one desired objective behavior, the BC serves only to drive search towards better solutions and thus a strongly-aligned BC is most appropriate (e.g. NS quickly solves the difficult HardMaze domain [8] because it circumvents the problem of deception by pursuing novel *endpoints*). However, when diversity itself is a desirable product of search, researchers must choose a BC that expresses the type of diversity they want to collect; often this choice results in a BC that is not well-aligned with the notion of quality, which recent research suggests may not be optimal for driving search towards better solutions [7, 16]. To investigate how the performance gap between aligned and unaligned BCs extends to hard problems, this paper compares the strongly-aligned EndpointBC from Lehman and Stanley [8] and the unaligned DirectionBC from Pugh et al. [16] on the much more challenging QD-Gauntlet (Fig. 1). However, congruent with the common practice of searching for unaligned QD, in this paper QD is always *collected* with respect to DirectionBC.

EndpointBC simply characterizes agent behavior by its (x, y) location at the end of its trial. This strongly-aligned BC is a powerful way to drive search on maze domains because it explores progressively more remote locations until the goal point is found. On the other end of the alignment spectrum, **DirectionBC** characterizes *how* the agent drives instead of where. DirectionBC consists of five values indicating whether the agent was most frequently facing north (0.125), east (0.375), south (0.625), or west (0.875) during each fifth of its trial. When driving search with this unaligned BC, it is possible to exhaust the entire behavior space without ever discovering high-performing solutions.

Each of the algorithms from Sect. 3 is implemented with each of DirectionBC and EndpointBC for a total of seven treatments: Fitness, NS_d, NS_e, NSLC_d, NSLC_e, $\mathrm{NS}_e\mathrm{NS}_d$, and $\mathrm{NS}_e\mathrm{NSLC}_d$. Each treatment is run 20 times on QD-Gauntlet, each for 1,000,000 evaluations (by which time all treatments reach a performance plateau). Networks are evolved with a modified version of SharpNEAT 1.0 [6] with mutation parameters validated by Pugh et al. [16]: 60% mutate connection, 10% add connection, 0.5% add neuron. Networks are feedforward and restricted to asexual reproduction; other settings follow SharpNEAT 1.0 defaults.

The performance of each treatment is evaluated according to the *QD-score* metric introduced by Pugh et al. [16] (and similar to the "global reliability" metric in Mouret and Clune [12]). Over the course of a run, a collection of individuals called the "QD grid" is gathered: the QD grid is managed such that each behavioral bin remembers the highest quality individual seen so far. While these bins are reminiscent of the bins in a MAP-Elites grid, the QD grid is completely external to the breeding population and thus does not interfere with or influence evolution. The QD-score is calculated as the total quality across all filled bins within the QD grid and reflects both how many distinct behaviors have been discovered and how good those behaviors are. Regardless of the BC driving search, QD-score for this study is always calculated with respect to DirectionBC (i.e. the QD grid represents a collection of different ways of driving). *Consequently, the EndpointBC-driven treatments in effect test whether QD with respect to one BC can be achieved passively by driving search with another BC altogether.*

While fitness for algorithms in this paper is the Euclidean distance to the goal, this heuristic is deceptive and does not accurately characterize how close collected behaviors are to actually solving the maze. Thus, for the purposes of evaluation, quality for individuals within the QD grid is instead represented by a *progress score* that respects that agents cannot drive through walls. Progress is defined as inversely proportional to the length of the shortest valid path between the agent's final location and the goal point of the maze. Importantly, this measure of quality (which draws a perfect, non-deceptive gradient over the drivable area of the maze) is not available to drive search and is only used for assessment.

5 Results

Figure 2 depicts the final QD-score (averaged over 20 runs) for each treatment after 1,000,000 evaluations. Unsurprisingly, Fitness performs significantly worse[1] than all other treatments ($p < 0.001$), underscoring the importance of specialized approaches to QD from outside the realm of conventional optimization. Notably, treatments driven by DirectionBC perform significantly worse than the EndpointBC and multi-BC treatments ($p < 0.001$), even though EndpointBC has nothing to do with the type of diversity being collected. Explaining this disparity, DirectionBC treatments fail to find even a single solution across all 40 runs, while each of the EndpointBC and multi-BC treatments consistently find multiple maze solutions per run – thus DirectionBC's low performance represents an inability to overcome deception in the QD-Gauntlet.

Among the best surveyed approaches, $\mathrm{NS}_e\mathrm{NSLC}_d$ and NSLC_e are not significantly different after the full 1,000,000 evaluations ($p = 0.163$). However, EndpointBC and multi-BC approaches demonstrate fundamentally different trends of QD-score over time (not shown): while multi-BC increases rapidly to

[1] Statistical significance is determined by an unpaired two-tailed Student's t-test.

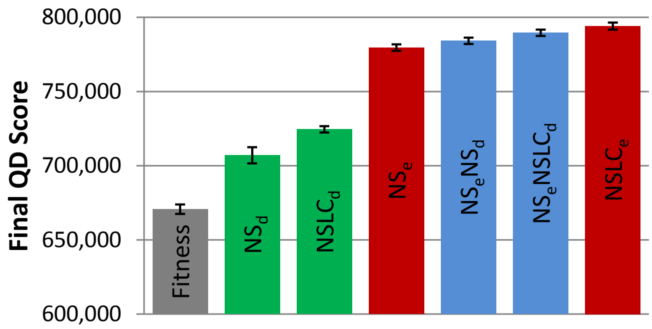

Fig. 2. Final QD-score. The final QD-score achieved by each treatment is shown (averaged over 20 runs). Error bars represent standard error. Bars are color-coded according to which BC drives search: DirectionBC-driven treatments (subscript d) are drawn in green, EndpointBC-driven (subscript e) in red, and multi-BC in blue.

an early plateau, EndpointBC takes much longer to reach similarly high scores. Correspondingly, *$NS_e NSLC_d$ scores significantly better than $NSLC_e$ ($p < 0.05$) until 224,000 evaluations*, after which point NSLC_e catches up.

6 Discussion and Conclusion

The QD-Gauntlet task, like the HardMaze [8] before it, contains a pronounced level of deception wherein the natural "Euclidean distance to the goal" fitness function does not always point the way to the goal but rather often leads search into an evolutionary dead end. In this task, as in any domain where deception is sufficiently present, strictly following the compass of increasing fitness (as in the Fitness treatment in this paper) is doomed to fail. While the idea of pursuing behavioral diversity instead of objective fitness (as in novelty search) is often recognized as a powerful way to confront such problems, it is clear that to be effective behavioral diversity cannot be applied naively.

Such a naive application of behavioral diversity lies in DirectionBC. Because it is unaligned with the concept of quality (driving closer to the goal), diversity with respect to DirectionBC can be achieved without making any progress towards conquering the problem of deception. Indeed, in the QD-Gauntlet, DirectionBC alone is incapable of unlocking the best-performing parts of the search space. Even if QD algorithms driven by such a BC succeed in generating a collection of locally optimal niches, if the best-performing behaviors cannot be found then the system is only operating at a fraction of its potential.

Experimental evidence on the QD-Gauntlet supports the idea that achieving QD with respect to quality-unaligned BCs in the traditional way is problematic: surprisingly, more QD is found with respect to DirectionBC by driving search with a completely *different* BC (EndpointBC – which is known to excel at solving maze tasks) than when search is driven by DirectionBC itself (Fig. 2). This result suggests that the ability of a BC to overcome deception can be just as instrumental in the search for QD as actually searching for diversity.

To allow QD algorithms to effectively bypass the problem of deception without sacrificing the desired notion of diversity, this paper introduces the idea of multi-BC QD algorithms that drive search with more than one BC at the same time: one that targets diversity and another that is better suited to driving search. Of the multi-BC variants surveyed, the best performance is achieved by $\mathrm{NS}_e\mathrm{NSLC}_d$, which is on par with NSLC_e for achieving the highest QD-score on the QD-Gauntlet (Fig. 2). Although EndpointBC alone in NSLC_e succeeds in collecting diversity with respect to DirectionBC in this study, it would be naive to assume that collecting diversity with respect to one BC can *always* succeed at covering diversity in arbitrary unrelated characterization spaces. Therefore, while NSLC with an aligned BC emerges highly successful in this study, for researchers aiming for long-term collection of unaligned QD, the multi-BC formulations offer an attractive alternative that allows explicitly searching for *both* unaligned and aligned diversity while potentially losing no significant performance.

While this paper has focused on QD variants of NS, in particular centered on NSLC, an interesting future direction will be to validate the advantage of multiple BCs on other QD algorithms such as MAP-Elites. In MAP-Elites, multiple BCs are possible because each BC can in principle control a separate grid, and all the grids can be run at the same time. In this way, it is possible to apply lessons from experiments here to a broad range of future QD algorithms.

Overall, this study reveals that the currently-accepted practice within QD of exclusively driving search with the desired notion of diversity (BC) can break down on hard problems and that an effective solution is to drive search with multiple BCs simultaneously. The insights gained here therefore expand the reach of quality diversity to more difficult problems and to arbitrary notions of diversity, bringing evolutionary computation one step closer to emulating the inventive power of natural evolution.

Acknowledgments. This work was supported by the National Science Foundation under grant no. IIS-1421925. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] A. Cully, J. Clune, D. Tarapore, and J.-B. Mouret. Robots that can adapt like animals. *Nature*, 521(7553):503–507, 2015.
[2] K. A. De Jong. *Evolutionary Computation: A Unified Perspective.* MIT Press, Cambridge, MA, 2002.
[3] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[4] J. Gomes, P. Urbano, and A. L. Christensen. Evolution of swarm robotics systems with novelty search. *Swarm Intelligence*, 7(2-3):115–144, 2013.

[5] J. Gomes, P. Mariano, and A. L. Christensen. Devising effective novelty search algorithms: A comprehensive empirical study. In *Proceedings of the 17th Annual Conference on Genetic and Evolutionary Computation (GECCO '15)*, pages 943–950. ACM, 2015.

[6] C. Green. SharpNEAT homepage, 2003–2006. URL `http://sharpneat.sourceforge.net/`.

[7] S. Kistemaker and S. Whiteson. Critical factors in the performance of novelty search. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation (GECCO '11)*, pages 965–972. ACM, 2011.

[8] J. Lehman and K. O. Stanley. Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary Computation*, 19(2):189–223, 2011.

[9] J. Lehman and K. O. Stanley. Evolving a diversity of virtual creatures through novelty search and local competition. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation (GECCO '11)*, pages 211–218. ACM, 2011.

[10] A. Liapis, G. N. Yannakakis, and J. Togelius. Sentient sketchbook: Computer-aided game level authoring. In *FDG*, pages 213–220, 2013.

[11] G. Morse, S. Risi, C. R. Snyder, and K. O. Stanley. Single-unit pattern generators for quadruped locomotion. In *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation (GECCO '13)*, pages 719–726. ACM, 2013.

[12] J.-B. Mouret and J. Clune. Illuminating search spaces by mapping elites. *arXiv preprint arXiv:1504.04909*, 2015.

[13] J.-B. Mouret and S. Doncieux. Encouraging behavioral diversity in evolutionary robotics: An empirical study. *Evolutionary Computation*, 20(1): 91–133, 2012.

[14] A. Nguyen, J. Yosinski, and J. Clune. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE, 2015.

[15] A. Nguyen, J. Yosinski, and J. Clune. Innovation engines: Automated creativity and improved stochastic optimization via deep learning. In *Proceedings of the 17th Annual Conference on Genetic and Evolutionary Computation (GECCO '15)*, New York, NY, USA, 2015. ACM.

[16] J. K. Pugh, L. B. Soros, P. A. Szerlip, and K. O. Stanley. Confronting the challenge of quality diversity. In *Proc. of the 17th Annual Conf. on Genetic and Evolutionary Computation*, New York, NY, USA, 2015. ACM.

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

[18] P. Szerlip and K. O. Stanley. Indirectly encoded sodarace for artificial life. In *Proceedings of the European Conference on Artificial Life (ECAL '13)*, volume 12, pages 218–225, 2013.

[19] P. A. Szerlip, G. Morse, J. K. Pugh, and K. O. Stanley. Unsupervised feature learning through divergent discriminative feature accumulation. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI '15)*, Menlo Park, CA, 2015. AAAI Press.