

Confronting the Challenge of Quality Diversity

In: *Proc. of the 17th Annual Conference on Genetic and Evolutionary Computation (GECCO 2015)*. New York, NY: ACM.

Justin K. Pugh, L. B. Soros, Paul A. Szerlip, and Kenneth O. Stanley
Department of EECS, Computer Science Division
University of Central Florida
Orlando, FL 32816 USA
{ jpugh, lsoros, pszerlip, kstanley } @eecs.ucf.edu

ABSTRACT

In contrast to the conventional role of evolution in evolutionary computation (EC) as an optimization algorithm, a new class of evolutionary algorithms has emerged in recent years that instead aim to accumulate as diverse a collection of discoveries as possible, yet where each variant in the collection is as fit as it can be. Often applied in both neuroevolution and morphological evolution, these new *quality diversity* (QD) algorithms are particularly well-suited to evolution's inherent strengths, thereby offering a promising niche for EC within the broader field of machine learning. However, because QD algorithms are so new, until now no comprehensive study has yet attempted to systematically elucidate their relative strengths and weaknesses under different conditions. Taking a first step in this direction, this paper introduces a new benchmark domain designed specifically to compare and contrast QD algorithms. It then shows how the degree of alignment between the measure of quality and the behavior characterization (which is an essential component of all QD algorithms to date) impacts the ultimate performance of different such algorithms. The hope is that this initial study will help to stimulate interest in QD and begin to unify the disparate ideas in the area.

Categories and Subject Descriptors:

I.2.6 [Artificial Intelligence]: Learning;

I.2.6 [Software Engineering]: Metrics—*complexity measures, performance measures*

General Terms: Algorithms

Keywords: novelty search; non-objective search; quality diversity; neuroevolution

1. INTRODUCTION

In evolutionary computation (EC), evolution has conventionally played the role of an optimization algorithm, wherein the performance of the algorithm is judged by how quickly and how close it comes to reaching an objective [6]. The idea that evolution can act as an efficient objective opti-

mizer is inspired by the feats of evolution in nature, where it appears to have optimized a myriad of complex yet highly functional forms. While some niche areas of EC deviate from this objective-driven paradigm (such as open-ended evolution in artificial life [2, 21, 23]), it has nevertheless broadly encompassed even newer research areas such as generative and developmental systems (GDS [25]) and neuroevolution [8].

One potential downside of this emphasis on objective optimization is that it has pitted EC as a direct competitor to virtually the entire field of machine learning, which similarly focuses largely on the problem of efficient optimization [1, 16]. While one consequence of this similar emphasis is a view among some in machine learning that EC is an inferior optimization algorithm, even if that view is arguably uninformed, it still leaves evolution in the position of seemingly “just another optimization algorithm.”

Yet evolution in nature is clearly not only about optimization. It is notable not for solving one particular problem, but *many* problems, and in uncountably diverse ways. In this sense, evolution is not a mere alternative to (for example) backpropagation [22] or support vector machines [3], which are not designed to generate a vast diversity of high-quality solutions. (As explained later, this kind of evolutionary diversity is also not the same conceptually as a multi-objective Pareto front [7] either.) This incongruity between EC in practice and its natural inspiration suggests a misalignment between the conventional application of EC and its full potential as a genuinely unique paradigm for search that can separate it from much of machine learning.

Hints of this broader interpretation of evolution within EC have begun to emerge with non-objective algorithms like novelty search (NS) [12, 14], which searches for novelty without an explicit objective. NS has in turn inspired a succession of variant algorithms that elaborate on the idea of a search through novelty alone [9, 13, 18, 19]. Yet even while this new generation of evolutionary algorithms begins to deviate from the traditional objective paradigm, such algorithms are still predominantly *evaluated* based on their ability to reach a particular objective [14, 19]. That is, while they may not function internally through the compass of the objective, *externally* they are often still treated merely as alternatives to objective-driven optimization algorithms.

However, some researchers have begun to take non-objective search in a different direction. They have instead embraced the potential for evolution to collect *quality diversity* (QD) [4, 5, 11, 17, 27], which means searching for as many possible variations (with respect to a chosen characterization)

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

GECCO '15, July 11–15, 2015, Madrid, Spain

© 2015 ACM. ISBN 978-1-4503-3472-3/15/07...\$15.00

DOI: <http://dx.doi.org/10.1145/2739480.2754664>

as possible, but where each such variation is as good as it can possibly be (with respect to a separate measure of quality). QD is particularly intriguing as an application of EC because it is explicitly *not* what most algorithms in machine learning are suited to achieving, yet nevertheless well-aligned with the achievements of evolution in nature. In this way QD is an opportunity for EC to shine as an essential option for the kinds of problems that require accumulating a multitude of diverse yet high-quality alternatives.

In this spirit, the aim of this paper is to begin to introduce a framework for understanding and relating QD algorithms. While the literature analyzing the nuances of algorithms in EC and elsewhere for the purpose of optimization is vast, almost no studies to date investigate the tradeoffs among different approaches to QD. This paper begins such an enterprise by introducing a domain designed explicitly to analyze approaches to QD in the context of neuroevolution, and then compares several leading algorithms in that new framework, including one new variant algorithm called *MAP-Elites+Novelty*, elaborated from the original MAP-Elites algorithm [5, 17]. The result is fresh insight into the consequences for different approaches to QD of varying the degree of alignment of the behavior characterization (which is essential to all such algorithms so far) with the notion of *quality* in the domain: some algorithms perform best when such alignment is high, while others excel without it. Because QD is such a promising new direction for EC with many possible future applications (as discussed in the next section), this initial investigation of algorithms in this area helps to lay the groundwork for an emerging new field.

2. BACKGROUND

Early hints of a science of QD within EC first appeared in the 1990s with work on multimodal function optimization (MMFO); a good review is provided by Mahfoud [15]. MMFO aims to discover as many high-scoring (though not necessarily global) optima within a fitness landscape as possible. While MMFO spawned a number of important diversity-preservation mechanisms (mainly focusing on genetic diversity), a shift more than ten years later towards methods focused on *behavioral diversity* [12, 14, 18, 19] revived interest in the potential for diversity to transform the application of EC. This new wave of algorithms, starting with novelty search [12], differed from the older work on MMFO in particular in its focus on discovering the full spectrum of possible behaviors within a particular domain, often represented by evolved neural networks. While much work in this area centered on the possibility that a side effect of seeking behavioral diversity is sometimes the discovery of very high performers, another view that emerged is that the diversity of discoveries *itself* is a desirable byproduct.

Unlike in multi-objective optimization [7], where the aim is to collect a Pareto front of performance trade-offs, these diversity-collecting algorithms focus instead on sampling the full behavior space as comprehensively as possible without concern for objective performance. In some cases a collection of diverse behaviors on its own offers direct practical value; for example the recent divergent discriminative feature accumulation (DDFA) algorithm [28] accumulates a vast collection of diverse feature detectors that are then aggregated into a neural network for later training in classification problems.

However, for many applications a collection of diverse

behaviors on its own is not enough because the ideal outcome would be that each variant is not only unique but also as high-performing as it can be. For example, merely discovering that self-ambulating pogo sticks exist in the morphology and behavior space of virtual creatures is not as good as discovering the *best possible* self-ambulating pogo sticks, in addition to the best possible behaviors of all the other viable morphologies that exist [11]. Note that this notion of QD is not the same as finding the highest-fitness subset of all conceivable creatures because even the very best pogo stick may have significantly *lower* fitness than e.g. quadrupeds and bipeds. However, the point is that we still want to find that pogo stick and see how well it can do. The diversity components in QD also still differ from the separate objectives in multi-objective optimization because individual diversity components include no intrinsic ranking – for example, all morphologies in the space of creatures are equally of interest as long as they can walk at all.

To empower evolution to collect QD, a new breed of algorithms have appeared in recent years that augment algorithms like novelty search to push each unique variant towards its best possible performance, opening up new applications. For example, Lehman and Stanley [11] introduced the novelty search with local competition (NSLC) algorithm to evolve a wide diversity of successful ambulating creatures in a single run. Similarly, Szerlip and Stanley [27] collect a diversity of high-performing behaviors in a Sodarace-inspired two-dimensional creature domain. In an intriguing foreshadowing of future applications, Cully and Mouret [4] show that QD once collected can be aggregated into a useful *behavioral repertoire*, in this case a collection of diverse robot motion skills.

Cully et al. [5] later introduced an alternative algorithm for collecting QD called MAP-Elites, which showcased yet another intriguing application of behavioral diversity: a collection of ambulation strategies becomes a source of means to alter ambulation in the event of damage. Yet another application of QD is demonstrated by Nguyen et al. [20], who generate with MAP-Elites a diverse collection of high-quality spoof images that fool deep networks. This breadth of early applications of such QD algorithms suggests an important new research area with many applications yet to be realized.

The aim of this paper is to begin to understand the factors that contribute to successfully collecting QD under different conditions. A common component of all modern QD algorithms is the *behavior characterization* (BC), which formalizes the behavior of a candidate individual in the population and is used to determine its eligibility for selection [5, 11]. However, different approaches to QD treat the BC in different ways, and the consequences of these choices are currently poorly understood. For example, while novelty favors individuals with the most novel BC, MAP-Elites discretizes the behavior space into unique bins from which (when occupied) selection is equally probable. This paper focuses in particular on the trade-offs created by such differing treatments and begins to explore the possibility of hybridizations that may combine the benefits of different approaches.

3. EXPERIMENT

To effectively tease out the differences between competing quality diversity (QD) algorithms, a test domain is needed with both (1) a concrete measure of quality, and (2) a diversity of legitimate solutions. No standard such benchmark

domain yet exists in the area of QD, and thus one of the contributions of this paper is to propose one suited specifically to the concerns of QD.

The first experiments with novelty search [12, 14] introduced a domain called *HardMaze* that has since become popular within the literature because of its explicit visual depiction of deception within a fitness landscape. The catch in *HardMaze* is that for a robot to navigate to the goal point of the maze it is necessary first to navigate *away* from the goal multiple times along the path. Additionally, there is an easily-reachable cul-de-sac that is close to the goal in terms of Euclidean distance but is actually farther away from the goal in terms of distance along the correct path. Importantly, searching for behavioral novelty alone (i.e. novelty search) in the *HardMaze* solves the maze much faster than a conventional objective-based search wherein fitness is the Euclidean distance to the goal point at the end of a trial (which, unlike novelty search, is easily drawn into the cul-de-sac). Thus, in *HardMaze*, Euclidean distance serves as a concrete measure of quality. However, *HardMaze* only contains *one* correct path through the maze (i.e. the diversity of legitimate solutions is limited), diminishing its utility for testing QD algorithms that seek many solutions.

Instead of *HardMaze*, in this paper, we augment the idea of a maze navigation domain that uses Euclidean distance as a concrete measure of quality to also encompass a diversity of possible solutions to the maze. This new domain, QD-Maze (figure 1), still contains several deceptive traps (albeit less difficult than those in *HardMaze*: all algorithms tend to find at least one solution) while also allowing many possible “correct paths” to the goal, thus allowing an explicit diversity of paths to the goal.

Aside from the change in map configuration, the setup for QD-Maze is identical to *HardMaze*. A robot agent is given a fixed time to move around the maze before the trial ends and the agent’s *quality* is scored according to the Euclidean distance between its final position and the goal point (the score is maximum if the goal point is reached). Reaching the goal point prematurely ends the trial. Agents are equipped with a set of six rangefinders, five of which are equally spaced across the front-facing 180 degrees and the sixth facing the rear. Additionally, four pie slice sensors (front, right, rear, and left -facing) detect the direction of the goal by activating the pie slice facing the goal with maximal value. In all experiments in this paper, agents are driven by evolved neural networks.

3.1 Behavior Characterizations (BCs)

The nature of the “diversity of solutions” available for discovery within the search depends heavily on how behaviors are characterized. In the original *HardMaze* experiments [12], agent behaviors are characterized according to their final position at the end of the trial (**EndpointBC**). This BC is notable for its strong alignment with the implicit objective of a maze. In other words, the neighborhood of behaviors around any particular trajectory endpoint generally share a similar level of quality, thereby aligning behavior with a notion of quality. Under the pressure of a strongly-aligned BC, searching for novelty *alone* may often be enough to find not only diversity, but also high-quality solutions to the domain, without even the need to explicitly optimize each behavioral neighborhood independently. To investigate the effect on different approaches to QD of varying the degree

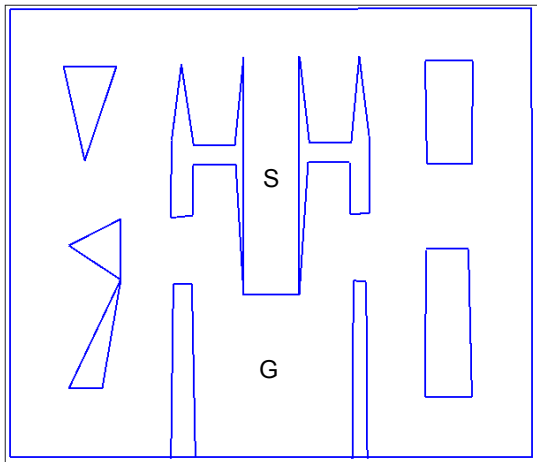


Figure 1: Quality diversity maze (QD-Maze). Individuals start at the point *S* and the highest quality solutions navigate to point *G*. While the maze is deceptive enough to challenge objective-based algorithms, it also contains a variety of solutions and thus serves as a benchmark for combining quality with diversity.

of the BC’s alignment with the objective, three alternative BCs are also tested, each less aligned than the previous: **FullTrajectoryBC**, **HalfTrajectoryBC**, and **DirectionBC**.

Under **FullTrajectoryBC**, the agent’s position is sampled at three points throughout its trial: one-third, two-thirds, and at the end of the trial, and all three (x, y) points are concatenated together to form the six-dimensional BC vector. This BC is less strongly aligned with the notion of quality than **EndpointBC** because only the final sampled point has a strong correlation to the final score within the domain. Even less aligned again is **HalfTrajectoryBC**, wherein the agent’s position is sampled at three points throughout *only the first half* of its trial. The second half of the trial leaves sufficient time for an agent to reach the goal from any location within the maze, rendering **HalfTrajectoryBC** even less aligned with the objective than **FullTrajectoryBC**. Highlighting this point, in theory, every behavior under **HalfTrajectoryBC** can potentially achieve a maximum quality score (i.e. ultimately solving the maze).

The final and least-aligned BC, **DirectionBC**, samples the direction the robot is facing on each tick (north, east, south, or west) and divides the duration of the trial into five equal-length time slices. The BC then consists of five dimensions, each corresponding to the direction that the agent was facing most often during each of the five time slices (north = 0.125, east = 0.375, south = 0.625, west = 0.875). This characterization is almost completely agnostic to the notion of quality because the goal point of the maze can be approached from any direction and the agent is free to wander around the maze and face any direction at any time. Under **DirectionBC**, every possible behavior in the behavior space can potentially achieve a maximum score on the objective of reaching the goal and there is little to no correlation between quality and behavioral neighborhoods.

3.2 QD Algorithms

A variety of algorithms can collect QD in a given domain, some designed explicitly to do so and some more as a side

effect. In this paper, we compare the performance of three existing algorithms (novelty search, novelty search with local competition, and MAP-Elites), the latter two designed explicitly for QD, and a new variant of MAP-Elites (MAP-Elites+Novelty). A conventional fitness-based search (which is not expected to excel at QD) also serves as a baseline for performance.

To normalize comparisons across the algorithms as much as possible, the SharpNEAT [10] distribution of the NEAT algorithm [24] serves as the underlying neuroevolution platform for all of them. Therefore, the neural networks are able to evolve increasing complexity, a central feature of NEAT. However, except in the pure fitness-driven baseline runs, to minimize potentially confounding factors, the conventional genetic diversity component of NEAT (also called *speciation*) is excluded. That way, the diversity facet of QD is exclusively focused on *behavior*, which is the focus of the majority of recent innovation in non-objective search algorithms and has exhibited significantly more impact on performance in such algorithms than genetic diversity [19]. Furthermore, because of preliminary experimental evidence that generational evolution can lead to instability or “jumping around the behavior space” in non-objective search populations, all algorithms are run in *steady state* mode, which means that a subset of the population (in this case 32 individuals evaluated in parallel) is replaced at a time instead of a whole generation (which is similar to the rtNEAT [26] variant of NEAT). In total five algorithms are compared:

The classic **novelty search (NS)** [12, 14] algorithm *only* searches for behavioral diversity with no notion of quality. It rewards behaviors that are maximally distant from previously discovered behaviors. However, if the BC is aligned with a notion of quality, then it is possible that NS without any special augmentation can itself collect QD effectively, a possibility thereby investigated in this experiment.

NS is augmented to explicitly seek QD in the **novelty search plus local competition (NSLC)** [11, 27] algorithm. In this variant, a multi-objective formulation implemented through NSGA-II [7] casts classic novelty as one objective and *local competition (LC)*, which means a quality rank relative only to those with a *similar* BC, as another. The LC component encourages unique behaviors to improve in quality within their local behavioral neighborhoods, thereby aligning the search explicitly with the notion of QD. The LC approach contrasts with the idea of combining novelty with a *global competition* objective (NSGC), which has been previously shown less effective than NSLC for discovering a diversity of solutions within a search space [11] because global competition explicitly prunes out diversity, and is therefore not considered in this paper.

Unlike NS and NSLC, the recent **MAP-Elites (ME)** algorithm [5, 17, 20] divides the behavior space of the BC into discrete bins. Each bin then remembers the fittest (elite) genome ever to occupy it, and only one genome occupies any given bin at one time. Thus the elites within each bin capture the notion of quality and the whole set of bins capture the notion of diversity. In the classic formulation of ME, selection is very simple: the chance of producing an offspring is equal for each filled bin. Thus search effort tends to be spread uniformly across the known behavior space (rather than explicitly biased by novelty). One technical limitation of ME is that the BC must be constrained to a low number of dimensions because the total number of bins

grows exponentially with BC dimensionality; for this reason and to keep the comparisons fair, none of the BCs in this paper includes more than six dimensions.

While the simplicity of selection in ME is appealing, its effect is that evolution is more likely to concentrate resources in regions of the behavior space already filled (i.e. because that is where most of the bins are occupied). This approach proved effective in ME applications so far, but it is easy to augment ME with pressure to concentrate effort on more novel regions by adding novelty pressure. In this **MAP-Elites+Novelty (ME-Nov)** variant tested for the first time here, the probability of selecting an occupied bin for reproduction is proportional to its novelty, creating a stronger pressure towards diversity than in the original ME. However, once all the bins are filled in ME-Nov, it collapses back to an approximation of regular ME. The question is whether the added initial pressure towards novelty might enhance ME’s drive towards QD. Also, while ME has effectively filled most bins in experiments published so far [5, 20], in future domains in which some bins require very advanced solutions and therefore are unlikely all to be filled, the additional novelty pressure in ME-Nov could prove instrumental in covering as much of the space as possible.

Finally, a regular **fitness-based search (fitness)** (implemented as NEAT [24]) simply searches for objectively superior performance (though with NEAT’s conventional genetic speciation fitness-sharing mechanism to give it at least some hope of maintaining diversity). This variant, which is susceptible to the trap of deception [14] and has no behavioral diversity mechanism, helps to highlight the need for specialized approaches to the problem of QD.

3.3 MAP-Elites as a QD assessment mechanism

To compare the ability of competing QD algorithms to collect quality diversity, a quantitative measure of quality diversity is needed. Interestingly, the MAP-Elites style of discretized behavior space can be collected for *all* variant algorithms, even those that are not MAP-Elites. As long as there are not too many dimensions in the BC (as is true in BCs in this paper), the best behavior found for every bin in the behavior space (regardless of through which method it was found) can simply be stored for later analysis.

For this purpose, for each BC, the behavior space is discretized into an appropriately small number of bins (i.e. small enough that MAP-Elites itself can be run efficiently and with a breeding pool of the same order of magnitude as the size of the breeding population in the NS, NSLC, and fitness-based algorithmic treatments): 900 bins for EndpointBC, 729 for FullTrajectoryBC and HalfTrajectoryBC, and 1,024 for DirectionBC. For each algorithmic treatment, all individuals are checked for eligibility in this MAP-Elites-style grid as they are evaluated. The performance (QD score) of each algorithm at any given point during the run is then measured as *the total fitness across all filled grid bins within the QD collection grid* (where higher fitness or quality means a final robot position closer to the goal). Throughout a run, algorithms therefore improve their QD score by either (1) discovering higher quality solutions for existing bins (increasing quality) or (2) discovering more bins (increasing diversity). To achieve the highest QD scores (which is possible in this domain), an algorithm must therefore excel with respect to both quality and diversity.

3.4 Experimental setup

Each of the methods (NS, NSLC, ME, ME-Nov, fitness) was combined with each of the behavior characterizations (EndpointBC, FullTrajectoryBC, HalfTrajectoryBC, DirectionBC) to comprehensively test the effects of both factors on QD. Except for the DirectionBC variants, each combination was evaluated over 20 runs capped at 250,000 evaluations. Because the DirectionBC variants converge to stable scores significantly more slowly (collecting QD across such a poorly-aligned BC is harder), they were capped at 750,000 evaluations. The population size was 500 for NS, NSLC, and fitness; NS and NSLC also keep an associated novelty archive of maximum size 2,500, after which the least novel is removed when a new entry is added. For MAP-Elites, as is standard the population at any given time was the size of the number of occupied bins. To enable parallel evaluation, steady state evolution is implemented in batches of 32 offspring at one time. Following the settings in the original novelty search experiments [14], the probabilities for mutating connections was 60%, for adding connections was 10%, and for adding neurons was 0.5%. Evolved networks are constrained to be strictly feedforward. All other settings follow standard SharpNEAT 1.0 defaults [10]. Source code for the entire experimental setup is available at <http://eplex.cs.ucf.edu/uncategorised/software#QD>

Performance for each method and BC combination was measured according to the total fitness across all filled grid bins (Section 3.3). Results were averaged across twenty runs for each of the 20 method-BC combinations (a total of 400 separate runs were performed). Additionally, pairwise comparisons were performed after every 160 evaluations to determine when exactly there are significant differences between methods. Tukey’s test (with $p < 0.05$) was used to establish significance instead of the Student’s t-test to account for the statistical problem of multiple pairwise comparisons.

4. RESULTS

Figures 2 through 5 show for each BC the total quality diversity (QD) discovered so far by each method over the course of 20 averaged runs for each method. Higher QD is better. Fitness runs, which are included only as a baseline, are unsurprisingly significantly below all other runs. Among the other four methods, a yellow bar at the top of each graph highlights the period during which significant differences are observed between the best and worst performing methods among those four. Note that in terms of both QD and significance testing, numerical comparisons are only meaningful between methods within the *same* BC because the total possible QD score depends on the BC. The main result is that the ranking of methods depends on the alignment of the BC with the notion of quality, suggesting that choosing a QD method for a particular problem can potentially become a principled decision in the future, though some methods are more generally robust across different BCs than others.

With the EndpointBC (figure 2), which is highly aligned with the notion of quality (i.e. closeness to the endpoint), the novelty search-based methods (NS and NSLC) significantly outperform all other methods for most of the run, with NS ending significantly (though only slightly) above NSLC. ME-Nov also significantly outperforms original ME, though is still significantly below NS and NSLC. (It is important to note that the differences between the methods

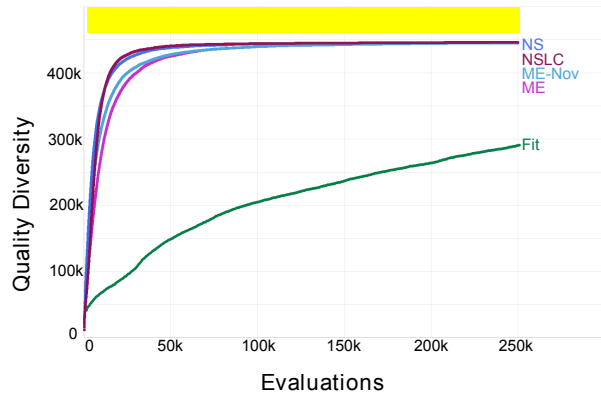


Figure 2: EndpointBC (very high alignment). In this performance comparison and others in this paper, the average QD (taken over 20 runs) for each variant method is shown across a run. The yellow strip at the top indicates the period during which there is a significant difference at least between the top and bottom method (excluding fitness, which is always significantly worse than all other methods). The method labels are color coded to match with their respective curves and are shown from top to bottom in their rank order during the period of significance. Note that because of the large QD scale, sometimes significant differences exist even when not visually apparent. For EndpointBC, NS performs best, followed by NSLC.

as they converge near the end of the run can be difficult to perceive in the figures, but the QD scale is actually vast, so significant difference actually do persist in some cases where it is difficult to perceive visually.) The fact that novelty alone (NS) does so well highlights the fact that when the BC is aligned with fitness, additional sophisticated machinery designed to simultaneously reward quality and diversity is potentially unnecessary (perhaps depending on the difficulty of the domain), and in such situations novelty alone becomes a powerful force for QD.

However, as the alignment of the BC with quality reduces, the story changes. In the FullTrajectoryBC (figure 3), which is still fairly aligned with the notion of quality (i.e. because trajectories that lead to the goal are high quality), NS, NSLC, and ME-Nov are tied (i.e. not significantly different) throughout most of the run. For this BC, only ME performs significantly below the other methods. Thus NS retains its utility as a top choice, but no longer holds an advantage over NSLC or ME-Nov. ME is not well suited to BCs with high alignment because it is agnostic with respect to novelty, leading to a relative delay compared to novelty-driven approaches (including ME-Nov) in exploiting promising areas of behavior space.

The HalfTrajectoryBC (figure 4) aligns with quality even less, which leads to an interesting phenomenon seen only with this BC: After a very brief period at the start of the run, all methods effectively perform the same (although ME is still a bit below, the difference is not significant). In effect, HalfTrajectoryBC hits just the right level of misalignment that the advantages and disadvantages of respective methods become a wash. Surprisingly, even NS alone remains competitive, despite its lack of an explicit quality component.

Finally, DirectionBC (figure 5), which tracks facing di-

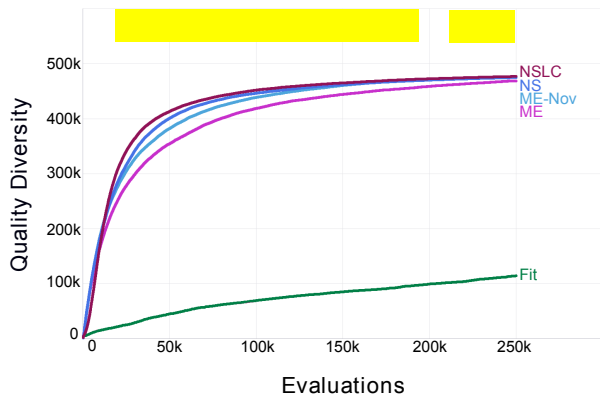


Figure 3: FullTrajectoryBC (high alignment). For this BC, which is slightly less aligned with quality than EndpointBC, NS, NSLC, and ME-Nov are effectively tied, with ME significantly behind.

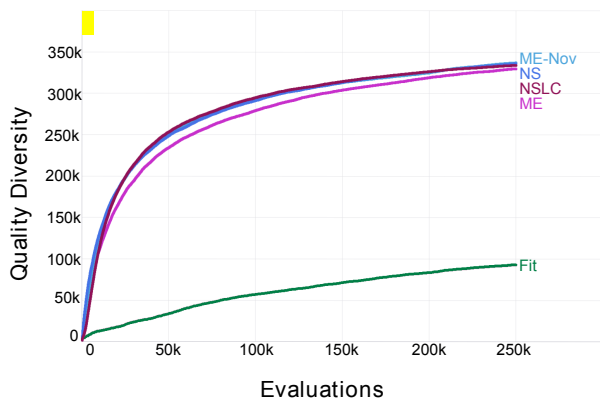


Figure 4: HalfTrajectoryBC (modest alignment). All the methods except fitness are tied for the vast majority of the run for this BC, which is even less aligned than FullTrajectoryBC.

rection only, aligns least with quality because it is largely detached from trajectory. For example, agents can arbitrarily spin around whenever they want without impacting their trajectory. Note that such low alignment is not unprecedented in serious applications. For example, an algorithm might search for diversity in terms of the height and mass of virtual creatures, which is hardly predictive of their quality as walkers [11]. The results with DirectionBC suggest that such a complete loss of alignment can significantly upend the utility of NS as a QD mechanism. In this scenario, NS drops to the bottom of the ranking, performing significantly below other approaches (each of which have quality-seeking components). Furthermore, while NSLC exceeds NS significantly (showing that the LC component of NSLC can indeed provide an advantage in QD), both ME and ME-Nov significantly outperform NSLC. In fact, ME alone effectively ties with ME-Nov, suggesting that in cases of high misalignment, the advantage of the novelty pressure is eclipsed by the advantage of a strong elitist mechanism that pushes towards higher quality.

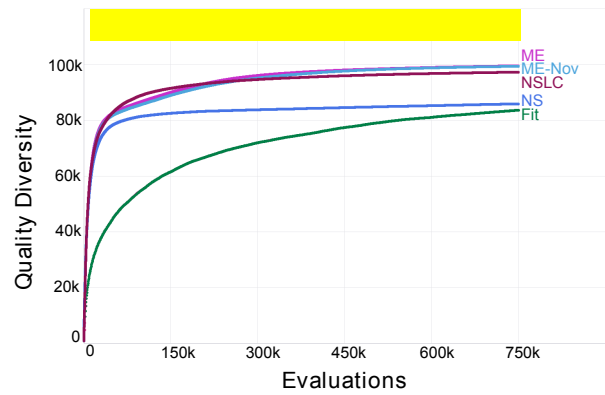


Figure 5: DirectionBC (low alignment). With an almost complete lack of alignment in this BC, ME and ME-Nov tie for first place, and NS trails far behind.

4.1 Visualizing Behavior Spaces

Recall that for all variant search algorithms, the highest-performing genomes are collected in a discretized behavior space “grid” (inspired by the MAP-Elites grid). Collecting elite behaviors in this way enables a convenient visualization of each algorithm’s ability to collect QD within the domain. The visualization technique, proposed by Cully et al. [5], uses two dimensions to display an entire behavior space grid by nesting two-dimensional grids within other two-dimensional grids, each two-dimensional grid representing two dimensions of the behavior space. As an example, figures 6 through 8 depict the final state of the QD collection grid at the end of one typical run of Fitness, ME, and NS in the FullTrajectoryBC. These visualizations reveal that: (1) Fitness (figure 6) and MAP-Elites (figure 7) exhibit a diminished ability to fill the grid for this BC (corresponding to a diminished pressure towards diversity) while the highest performing algorithm (on FullTrajectoryBC), NS (figure 8), is able to fill the grid almost entirely after 250,000 evaluations. The visualizations also expose (2) the relatively strong alignment of the behavior space with the fitness measure through the smooth gradient in the outer two dimensions from low fitness at the top of the grid towards high fitness at the bottom-middle of the grid.

This paper is also accompanied by a website where sample interactive behavior space visualizations (including the ability to browse through the discovered behaviors within each bin) are available for all 20 method-BC combinations (including the three shown here):

<http://eplex.cs.ucf.edu/QD/GECCO-15/compare.html>

5. DISCUSSION

The main insight to emerge from the experiment concerns the two key forces operating in QD algorithms: the push for novelty and the push for quality. If the BC, which supplies the main push for novelty, is sufficiently aligned with a notion of quality, then novelty-based approaches are most effectively empowered. In contrast, if the BC is orthogonal to a notion of quality, a strong push for quality (as in the elitism of ME) becomes the more instrumental factor in QD performance. The apparent ability under some BCs for NS *alone* to succeed, even without any explicit push for quality, and ME *alone* in other BCs to succeed without any explicit

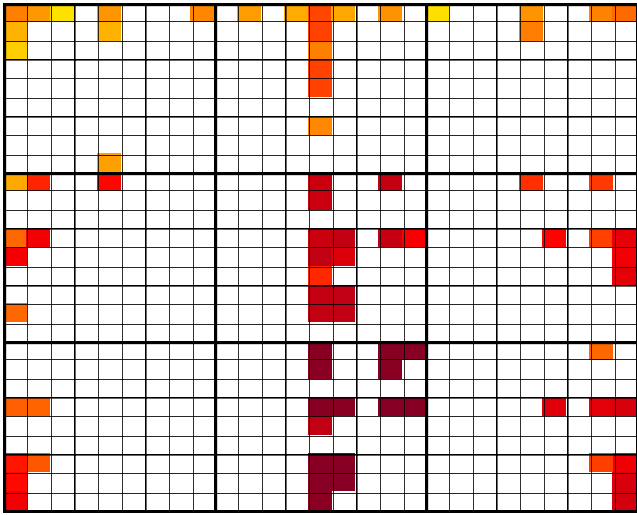


Figure 6: Fitness in the FullTrajectoryBC. In this example grid, the six-dimensional behavior space (FullTrajectoryBC) (discretized into three bins per dimension for a total of 729 bins) is visually depicted as a series of nested two-dimensional grids (each of which are 3×3). The color of each grid box corresponds to the quality of the solution found by the search algorithm after 250,000 evaluations: yellow corresponds to low quality, dark red to high quality, and white to unfilled bins. Fitness finds very few of the possible behaviors for this BC.

push for novelty, highlights just how profoundly this apparent principle applies. Yet it also appears that augmenting these most simple methods with a component to cover what they lack, e.g. augmenting NS with LC and ME with novelty, carries generally little risk. In fact, in some cases it improves their results, especially when they are being applied in their respectively less ideal alignment circumstances, hinting that aiming for “best of both worlds” is not unrealistic and may even be attainable for future QD algorithms yet unrealized.

However, the results in this initial study should be qualified by the observation that the QD-Maze domain is relatively easy; all methods except for fitness eventually find diverse solutions. Because more ambitious applications of QD in the future are likely significantly more challenging (imagine e.g. much more challenging parallel mazes), a natural question is whether the results here will extend to future experiments. While of course further empirical research is necessary and this study is only a first step, there is still good reason to believe that the results reported here are at least somewhat general: First, even in much harder domains, as the length of runs approaches infinity, all methods are likely eventually to find most QD. The experiment here is simply shorter because the domain is less difficult. Yet the stratification of different methods observed during the course of these short runs can conceivably emerge similarly in more challenging domains, though simply over longer timescales. Second, the fact that the ranking of methods in QD-Mazes matches well our expectations for each method (i.e. that those more reliant on a separate measure of quality do better when the BC is poorly aligned with fitness) further supports the credibility of the results.

Of course, there are some other considerations outside QD

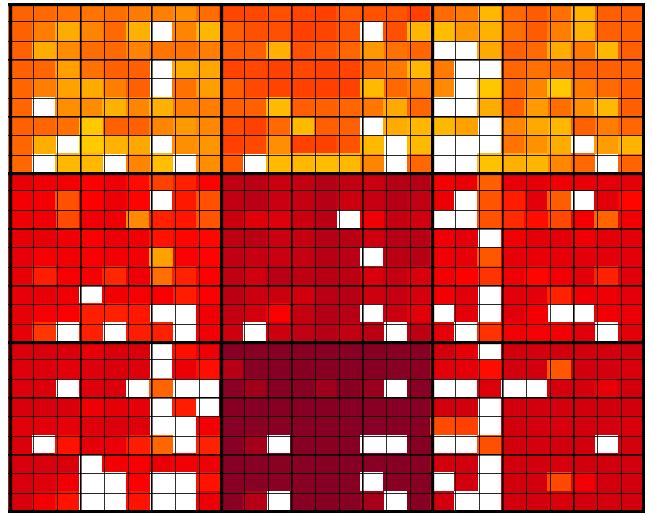


Figure 7: MAP-Elites in the FullTrajectoryBC. Compared to Fitness (figure 6), ME discovers far more QD under FullTrajectoryBC. Some bins remain unfilled (white), corresponding to the lack of pressure towards diversity within the ME algorithm.

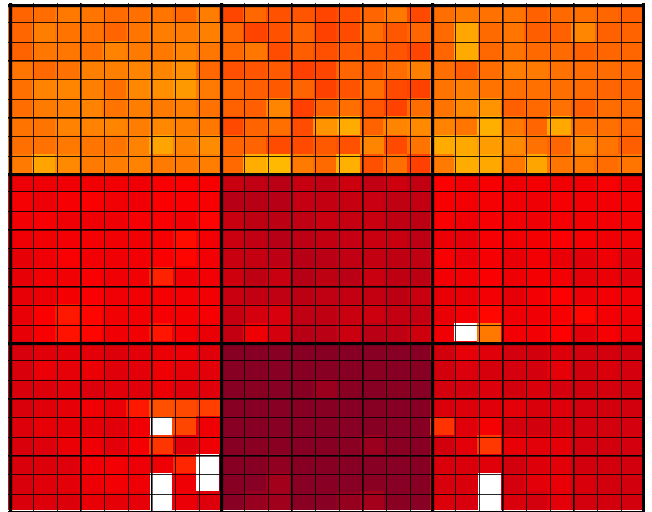


Figure 8: Novelty Search in the FullTrajectoryBC. Of the five compared algorithms, NS performs the best under the FullTrajectoryBC (featuring relatively high alignment between the BC and the objective) because it focuses exclusively on pursuing diversity. This conclusion is supported in the QD collection grid by almost all bins being filled (i.e. non-white).

performance alone that may arise in specific domains. For example, because it divides behavior space into discrete bins, ME (and ME-Nov) becomes increasingly difficult to apply to higher-dimensional BCs, an issue not present for NS or NSLC. Yet beyond its use as a *method*, the ME grid also turns out an excellent mechanism for collecting and measuring QD results, even for methods that are not ME. We anticipate that the results of future QD algorithms even unrelated to ME will ultimately often be visualized and quantified by placing them within in a classic ME grid.

6. CONCLUSION

The aim of this paper was to help to unify the emerging *quality diversity* (QD) research area by beginning to introduce benchmarks and principles for analyzing and contrasting methods in the area. For this purpose a new domain with many parallel solutions called the QD-Maze was introduced, and several QD methods (including a new variant called MAP-Elites+Novelty) were compared under different behavior characterizations in QD-Maze. The results begin to establish initial principles for understanding QD, namely that the alignment of the behavior characterization with the notion of quality significantly influences which QD methods will be most effective in a particular setup. This insight gives hope that an even broader understanding of QD can be reached in the future, thereby helping to expand the reach and impact of evolutionary computation.

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] C. M. Bishop. *Pattern Recognition and Machine Learning*, volume 1. Springer, New York, NY, 2006.
- [2] A. Channon. Improving and still passing the alife test: Component-normalised activity statistics classify evolution in geb as unbounded. *Artificial Life*, 8: 173–181, 2003.
- [3] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.
- [4] A. Cully and J.-B. Mouret. Behavioral repertoire learning in robotics. In *Proc. of the 15th Annual Conf. on Genetic and Evolutionary Computation (GECCO '13)*, pages 175–182, New York, NY, USA, 2013. ACM.
- [5] A. Cully, D. Tarapore, J. Clune, and J.-B. Mouret. Robots that can adapt like natural animals. *arXiv preprint*, abs/1407.3501, 2014.
- [6] K. A. De Jong. *Evolutionary Computation: A Unified Perspective*. MIT Press, Cambridge, MA, 2002.
- [7] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.
- [8] D. Floreano, P. Dürri, and C. Mattiussi. Neuroevolution: From architectures to learning. *Evolutionary Intelligence*, 1:47–62, 2008.
- [9] J. Gomes and A. L. Christensen. Generic behaviour similarity measures for evolutionary swarm robotics. In *Proceedings of the Fifteenth Annual Conference on Genetic and Evolutionary Computation (GECCO '13)*, pages 199–206, New York, NY, USA, 2013. ACM.
- [10] C. Green. SharpNEAT homepage, 2003–2006. URL <http://sharpneat.sourceforge.net/>.
- [11] J. Lehman and K. O. Stanley. Evolving a diversity of virtual creatures through novelty search and local competition. In *Proc. of the Thirteenth Annual Conf. on Genetic and Evolutionary Computation (GECCO '11)*, pages 211–218, Dublin, Ireland, 2011. ACM.
- [12] J. Lehman and K. O. Stanley. Exploiting open-endedness to solve problems through the search for novelty. In *Proc. of the 11th Intl. Conf. on Artificial Life (Alife XI)*, Cambridge, MA, 2008. MIT Press.
- [13] J. Lehman and K. O. Stanley. Revising the evolutionary computation abstraction: Minimal criteria novelty search. In *Proc. of the 12th Annual Conf. on Genetic and Evolutionary Computation (GECCO '10)*, pages 103–110, New York, NY, USA, 2010. ACM.
- [14] J. Lehman and K. O. Stanley. Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary Computation*, 19(2):189–223, 2011.
- [15] S. W. Mahfoud. *Niching Methods for Genetic Algorithms*. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, IL, May 1995.
- [16] T. M. Mitchell. *Machine Learning*. McGraw-Hill, New York, NY, 1997.
- [17] J.-B. Mouret and J. Clune. Illuminating search spaces by mapping elites. *arXiv preprint*, 2015.
- [18] J.-B. Mouret and S. Doncieux. Overcoming the bootstrap problem in evolutionary robotics using behavioral diversity. In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC-2009)*, pages 1161–1168. IEEE, 2009.
- [19] J.-B. Mouret and S. Doncieux. Encouraging behavioral diversity in evolutionary robotics: An empirical study. *Evolutionary Computation*, 20(1):91–133, 2012.
- [20] A. Nguyen, J. Yosinski, and J. Clune. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. *arXiv preprint*, abs/1412.1897, 2014.
- [21] C. Ofria and C. O. Wilke. Avida: A software platform for research in computational evolutionary biology. *Artificial Life*, 10(2):191–229, 2004.
- [22] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by error propagation. In *Parallel Distributed Processing, Volume 1*, pages 318–362. MIT Press, Cambridge, MA, 1986.
- [23] L. Spector, J. Klein, and M. Feinstein. Division blocks and the open-ended evolution of development, form, and behavior. In *Proceedings of the Ninth Annual Conference on Genetic and Evolutionary Computation (GECCO '07)*, pages 316–323. ACM, 2007.
- [24] K. O. Stanley and R. Miikkulainen. Evolving neural networks through augmenting topologies. *Evolutionary Computation*, 10:99–127, 2002.
- [25] K. O. Stanley and R. Miikkulainen. A taxonomy for artificial embryogeny. *Artificial Life*, 9(2):93–130, 2003.
- [26] K. O. Stanley, B. D. Bryant, and R. Miikkulainen. Real-time neuroevolution in the NERO video game. *IEEE Transactions on Evolutionary Computation Special Issue on Evolutionary Computation and Games*, 9(6):653–668, 2005.
- [27] P. Szerlip and K. O. Stanley. Indirectly encoded sodarace for artificial life. In *Proceedings of the European Conference on Artificial Life (ECAL-2013)*, volume 12, pages 218–225, 2013.
- [28] P. A. Szerlip, G. Morse, J. K. Pugh, and K. O. Stanley. Unsupervised feature learning through divergent discriminative feature accumulation. In *Proc. of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-2015)*, Menlo Park, CA, 2015. AAAI Press.