

# The Limits of Lexicase Selection in an Evolutionary Robotics Task

Jared M. Moore<sup>1</sup> and Adam Stanton<sup>2</sup>

<sup>1</sup>School of Computing and Information Systems,  
Grand Valley State University, Allendale, MI, USA

<sup>2</sup>School of Computer Science and Mathematics,  
Keele University, Keele, ST5 5BG, UK  
moorejar@gvsu.edu, a.stanton@keele.ac.uk

## Abstract

Agents exhibiting generalized control are capable of solving a theme of related tasks, rather than a specific instance. Here, generalized control pertains to the locomotive capacity of quadrupedal animats, evaluated when climbing over walls of varying height to reach a target. In prior work, we showed that Lexicase selection is more effective than other evolutionary algorithms for this wall crossing task. Generalized controllers capable of crossing the majority of wall heights are discovered, even though Lexicase selection does not sample all possible environments per generation. In this work, we further constrain environmental sampling during evolution, examining the resilience of Lexicase to the impoverished conditions. Through restricting the range of samples at given points in time as well as fixing environmental exposure over fractions of evolutionary time, we attempt to increase the ‘adjacency’ of environmental samples, and report on the response of the Lexicase algorithm to the pressure of this reduced environmental diversity. Results indicate that Lexicase is robust, producing viable agents even in considerably challenging conditions. We also see a positive correlation between the number of tiebreak events that occur and the success of individuals in a population, except in the most limiting conditions. We argue that the increased number of tiebreaks is a response to local maxima, and the increased diversity resulting from random selection at this point, is a key driver of the resilience of the Lexicase algorithm. We also show that in extreme cases, this relationship breaks down. We conclude that tiebreaking is an important control mechanism in Lexicase operation, and that the breakdown in performance observed in extreme conditions indicates an inability of the tiebreak mechanism to function effectively where population diversity is unable to reflect environmental diversity.

## Introduction

Generalized control remains a challenge in the field of robotics. Systems must be capable of addressing a task broadly, rather than solve only specific environmental configurations. In this paper, we use a wall crossing task, see Figures 1 and 2, wherein an agent is presented with a navigation target placed on the opposite side of a wall. It must then locomote over the obstacle and reach the target within a fixed amount of time. The wall height varies between environments, with 100 different gradations challenging a controller to learn the general task of ‘locomotion with wall crossing’.

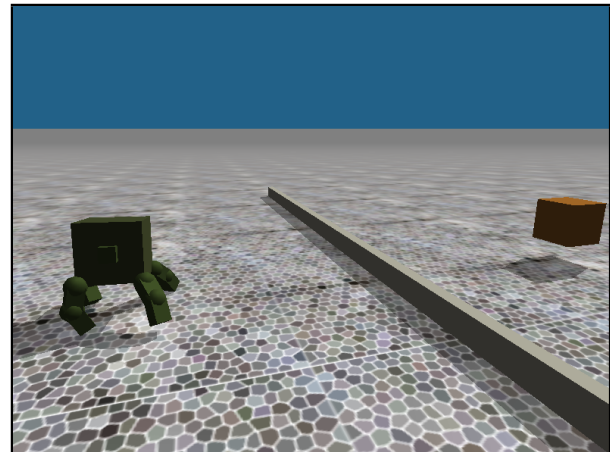


Figure 1: Neurocontrollers evolve to produce gaits that guide the quadrupedal animat (left) across a wall (center) and towards a target, represented by the box (right). Graphic originally presented in Moore and Stanton (2018)

In previous work (Moore and Stanton, 2018), we demonstrated evolved artificial neural network (ANN) controllers, generalized across different wall heights, using the Lexicase selection algorithm originally proposed by Spector (2012). We primarily focused on performance of the evolved agents, identifying a set of parameters for Lexicase that promote effective controllers. Even though there are 100 wall heights, simulating agents across 10 unique wall heights per generation was sufficient to evolve generalized behavior. From these results, we speculated that the advantages of Lexicase (aside from the exposure of species to a heterogeneous set of environments over evolutionary time) are due to automatic maintenance of population diversity facilitated by tiebreak events occurring as part of the selection process. During an individual selection event in Lexicase, a tiebreak occurs when two or more individuals have similar performance in all environments used in the selection event. A random selection is then performed choosing a parent from the set of tied individuals.

In this study, we attempt to elucidate characteristics of Lexicase that govern its behavior in extreme situations. This

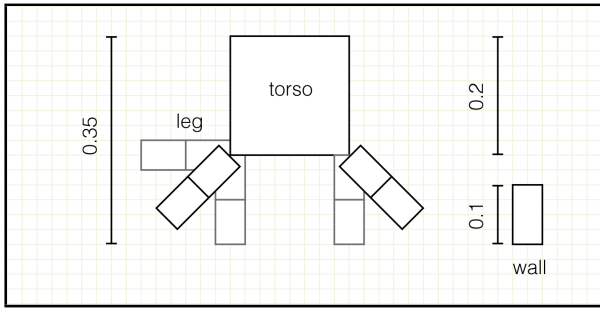


Figure 2: Schematic of robot and wall, illustrating relative sizes. Wall is shown at maximum height; robot leg shown at initial position (black) and at leg–torso joint extremes (gray).

allows us to identify areas where generalized task performance by evolved controllers begins to break down. Specifically, we use two restrictions to alter key aspects of the Lexicase algorithm. First, we constrain the sampling range of environments (i.e. wall heights) used to evaluate individuals within a generation. This increases *adjacency*, that is, the difference in wall height between environments chosen during selection. We hypothesize that this will hinder the generalizability of controllers as adjacency might also play a role in the success, or failure, of individual runs. Second, we limit how often over evolutionary time we reposition our sampling range, forcing populations to experience a limited subset of the 100 total wall heights for a number of consecutive generations.

Results show that in this task, even with fairly significant restrictions, Lexicase selection is still a resilient evolutionary algorithm. Most treatments are able to evolve at least some level of generalization with performance degrading as stronger restrictions are placed on Lexicase selection. There is a positive correlation in the number of tiebreaks and performance of the best individual for a given replicate. Although Lexicase does well to preserve mean population diversity, in this task, it does not appear to drive generalization performance across treatments.

## Background

Evolutionary robotics (ER) optimizes robotic systems by employing concepts from biological evolution in a digital system (Nolfi and Floreano, 2000). ER has been employed across robotic systems to optimize both body structure alone (Auerbach and Bongard, 2010; Cheney et al., 2013; Collins et al., 2018) or together with control (Jelisavcic et al., 2018; Kriegman et al., 2018). Evolved robots have been transferred to reality (Ruud et al., 2016) although the reality gap remains a persistent issue (Stanton, 2018; Koos et al., 2010; Jakobi, 1998). While enhancing performance of systems is often the primary goal, one outstanding area is quantifying the impact that different components of a robotic system (Powers et al., 2018), or the underlying evolutionary algorithm (Dolson and Ofria, 2018), have on evolved systems.

Lexicase selection was originally introduced for many-objective problems in genetic programming by Spector (2012) where it has been effective on a variety of problems (Helmuth et al., 2014). By replacing traditional selection methods in a genetic algorithm, Lexicase uses a variable number of objectives to perform selection between individuals. One of the potential strengths for Lexicase selection is its influence on population diversity, as proposed by Helmuth et al. (2016). In addition, Lexicase may also preserve specialists in the population as originally noted by Pantridge et al. (2018). An overview of the Lexicase selection algorithm and our specific modifications for this paper are described in the next section.

## Methods

The animat, simulation environment, and wall-crossing task are continuations of previous work originally reported with respect to Lexicase selection in Moore and Stanton (2017). Specific design parameters are detailed in that work and its references. Here, we report our animat’s genome parameters and simulation configuration for completeness.

**Animat** Figure 1 shows the quadrupedal animat in the wall crossing environment, and Figure 2 shows a schematic of the animat and obstacle to illustrate relative proportion. The torso is a cube with four legs, each placed at a corner. Hip joints can swing both horizontally and vertically. Knees move only on one axis allowing for flexion and extension. Table 1 shows the physical dimension, mass, and joint ranges for the animat.

Torso Dimension	$0.2 \times 0.2 \times 0.2$
Leg Component Dimension	$0.075 \times 0.05 \times 0.05$
Torso Mass	2.0
Leg Component Mass	0.5
Hip Vertical Axis	range $[-\frac{\pi}{4}, \frac{\pi}{4}]$
Hip Horizontal Axis	range $[0, \frac{\pi}{2}]$
Knee Horizontal Axis	range $[0, \frac{\pi}{2}]$
Max Torque	0.125

Table 1: Physical parameters of robot. Adapted from Moore and Stanton (2017, 2018).

Controllers are feed-forward ANNs with one output per leg motor. Each ANN output specifies a joint angle translated to motion using a Proportional Derivative (PD) controller (Reil and Husbands, 2002) which calculates a torque to apply to a joint for the given timestep based on Equation 1,

$$T = k_s \times (\theta_d - \theta) - k_d \dot{\theta} \quad (1)$$

where  $T$  is torque,  $k_s$  and  $k_d$  are spring and damper constants,  $\theta_d$  is the ANN output angle,  $\theta$  the present joint and  $\dot{\theta}$  the angular change from the last timestep.  $k_s = k_d = 0.5$ .

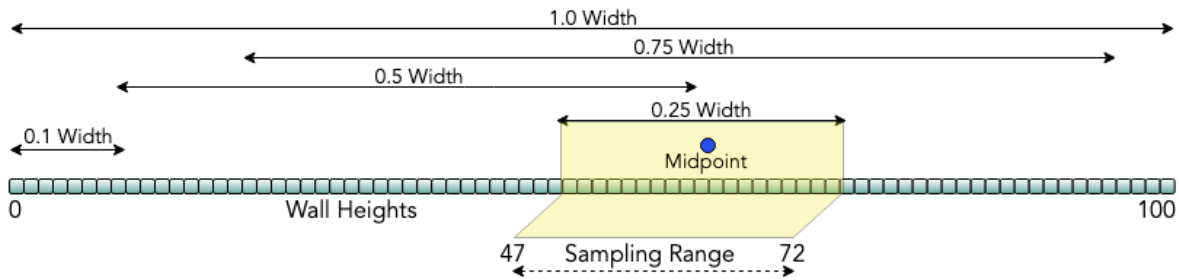


Figure 3: Five separate sampling widths are explored. The process of selecting environments proceeds as follows: 1. Pick a midpoint within the range of environments such that the maximum and minimum environments from the midpoints stay within the range of 0 to 100 exclusive. 2. The midpoint establishes the minimum and maximum wall heights that can be sampled for use in the Lexicase selection event.

**Simulation** The Open Dynamics Engine (ODE) (Smith, 2013) version 0.15.2 was used. ODE is a real-time rigid body physics engine handling the interaction between components of the animat, as well as interactions with the ground and wall.

**Wall Crossing Task** The goal for an individual is to reach a target placed on the opposite side of a single wall from the animat’s starting position. Animats must thus evolve a legged gait to move towards a target as well as an ability to cross a wall. Wall height is fixed within a simulation, but can be chosen from a range of 100 possible wall heights. The height is determined by the specific sampling determined by the Lexicase algorithm, described in the next section.

Controller effectiveness is characterized as one of five possible behaviors: (1) reached target, (2) crossed wall, (3) stuck on wall, (4) reached wall, and (5) did not reach wall.

**Evolutionary Algorithm and Lexicase Selection** Lexicase selection is applied to a subset of individuals sampled from the overall population between generations. To compare individuals, an objective is selected randomly within the subset ranked by performance. If two or more individuals are tied, selection moves on to the next objective only on the subset of tied individuals. If multiple individuals remain after all objectives have been used, a random selection is applied to the remaining individuals and recorded as a tiebreak event. For a full description of the Lexicase selection algorithm employed in this paper refer to Moore and Stanton (2017).

Originally proposed for GP, Lexicase selection only considered two individuals to be tied in an objective if they were exactly equal. Moore and McKinley (2016) introduced a *fuzz factor* to ease the consideration of ties to include individuals who are within a specified threshold of performance. La Cava et al. (2016) concurrently proposed  $\epsilon$ -Lexicase selection as an improvement for real-valued objectives. For continuity with our previous work, in this paper we use a *fuzz factor* of 10%, that is two individuals are tied if one individual is within 10% of the performance of the other. Ongoing work by others however should use the  $\epsilon$ -Lexicase terminology.

**Sampling Width** Figure 3 shows the various widths employed for sampling in this strategy. Treatments use widths of 0.1, 0.25, 0.5, 0.75, and 1.0. For example, 0.1 width allows for environments to be sampled within a range of 10 environments whereas 0.75 width allows sampling from 75 environments. The process of selecting environments within a given width proceeds as follows.

1. Select a midpoint within the range of 0 - 100 that will not cause the minimum or maximum of the width to exceed the range of environments.
2. Midpoint establishes the minimum and maximum.
3. Sample  $n$  environments uniformly from within the range according to the number of environments specified by the treatment. ( $n = 10$  in this paper.)

Figure 3 shows how a range could be specified for the 0.25 width. A midpoint of 60 establishes a range of 47-72 (floating point numbers are rounded down). A number of environments are then sampled from within this range establishing the objectives for Lexicase within that generation.

**Monte Carlo Simulation of Sampling** Figure 4 plots the expected sampling frequency of various widths and 10 environments for evaluation within a generation if the midpoint were randomly selected each generation. Each pair of width plots consists of 5000 distinct sampling events with 10 environments selected per sample. As the sampling width increases, the distance between selected wall heights increases presenting agents with a larger variety of wall heights. We hypothesize that this will help increase generalizability.

**Measuring Diversity** We quantify diversity by calculating the mean genotype for a population every generation with the following equation from (Moore and Stanton, 2018):

$$G^{gene} = \frac{1}{P} \sum_{i=1}^P I_i^{gene} \quad (2)$$

, where  $G$  is the mean genotype,  $P$  is the population size, and  $I_i^l$  is the  $l$ th gene of the  $i$ th individual). The gene-wise

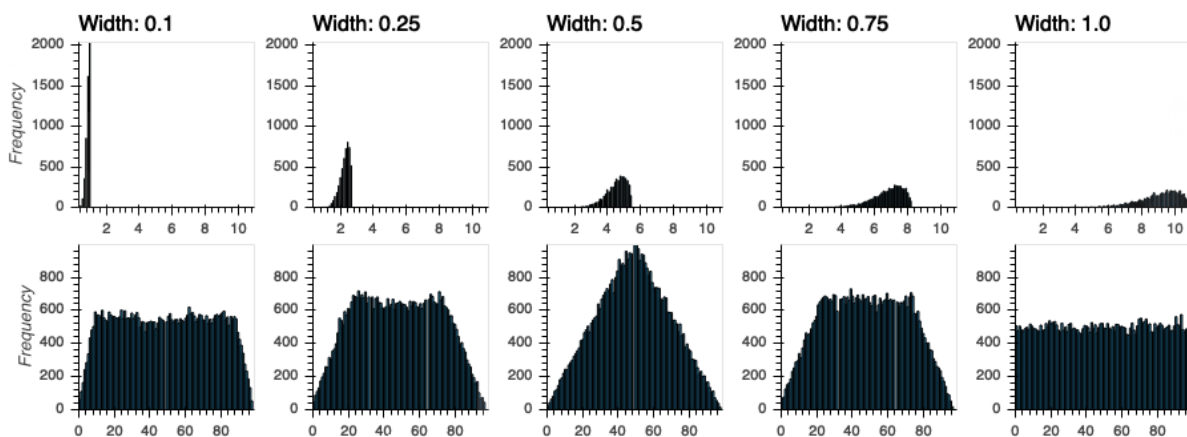


Figure 4: Simulation of 5000 sampling events per each of the sampling widths conducted in this study. Each sampling event draws 10 environments uniformly within a given sampling range discussed previously in Figure 3. Top row shows the average distance between wall heights (adjacency) in a sample. Bottom row shows the raw frequency that each wall height is sampled.

mean squared difference of each individual from the generations’s mean individual is calculated. This quantity is calculated for each individual, summed and averaged across the population. The value is then compressed to a single reportable measure in the following:

$$D = \frac{1}{P} \sum_{i=1}^P \left[ \frac{1}{L} \sum_{l=1}^L (I_i^l - G^l)^2 \right] \quad (3)$$

where  $D$  is the diversity value and  $L$  is genotype length).

**Treatments** In prior work (Moore and Stanton, 2018), we sampled between 1 and 20 environments per generation to use in the Lexicase selection process. Results indicated that the range of 5 to 10 environments are optimal for this task. Therefore, all treatments conducted in this study simulate agents in 10 environments per generation. This also cuts down on simulation time when evolving individuals as we only need to simulate 10 out of 100 possible environments to evaluate individuals in a selection event. Treatments consist of 20 replicate runs, each initialized with a random starting seed. Each replicate is evolved for 5,000 generations with a population size of 50 individuals.

Treatments are defined by two factors. First, there are five separate sampling widths (0.1, 0.25, 0.5, 0.75, and 1.0) described in the previous sections. Second, the midpoint of a sampling width is moved according to a specified braking interval. Possible intervals are 1, 2, 5, 10, 100, 500, 2500, and 5000 generations. For example, a braking interval of 5 generations (abbreviated as 5B) means that we select a midpoint that establishes the sampling range for the next 5 generations. A new set of environments are sampled every generation, however, braking effectively limits how often the midpoint moves around the wall height space. Some limiting of moving the midpoint might encourage a popu-

lation to specialize on a specific range of environments at the expense of not being exposed to other areas of the wall height space. At the conclusion of a braking interval, we then move the sampling range by selecting a new midpoint and evolving for another 5 generations. This process repeats until 5000 generations have elapsed. Once the evolutionary phase is complete, each individual in each population is evaluated in each environment (wall height), and these results are recorded to gauge the generalizing capability of that particular controller.

## Experiments and Results

The focus of this investigation is on aspects of Lexicase selection that enable generalized control. In prior work (Moore and Stanton, 2017, 2018) we have shown that Lexicase selection is an effective evolutionary algorithm for agents in this problem domain. Our results therefore are not centered on the question of “whether Lexicase can evolve effective individuals” but rather “which aspects of Lexicase selection might affect performance”.

**Generalization** We first examine the generalizability of the best individual per replicate per treatment across the 100 environments. Figure 5 plots the average performance of the best individual per treatment grouped by sampling width. Lower wall heights are not a challenge for any of the treatments as indicated by the brighter shading associated with wall crossing and reaching the target. Narrower widths and higher braking intervals also appear to have more individuals reach the objective attaining the highest fitness possible in these low wall heights compared to the 1.0 and 0.75 width treatments. We hypothesize that there may be an evolutionary pressure in these treatments to evolve highly effective behavior in the low wall heights to compensate for the generally poor performance in the higher walls.



## Average Performance of the Best Ind Per Replicate Per Treatment

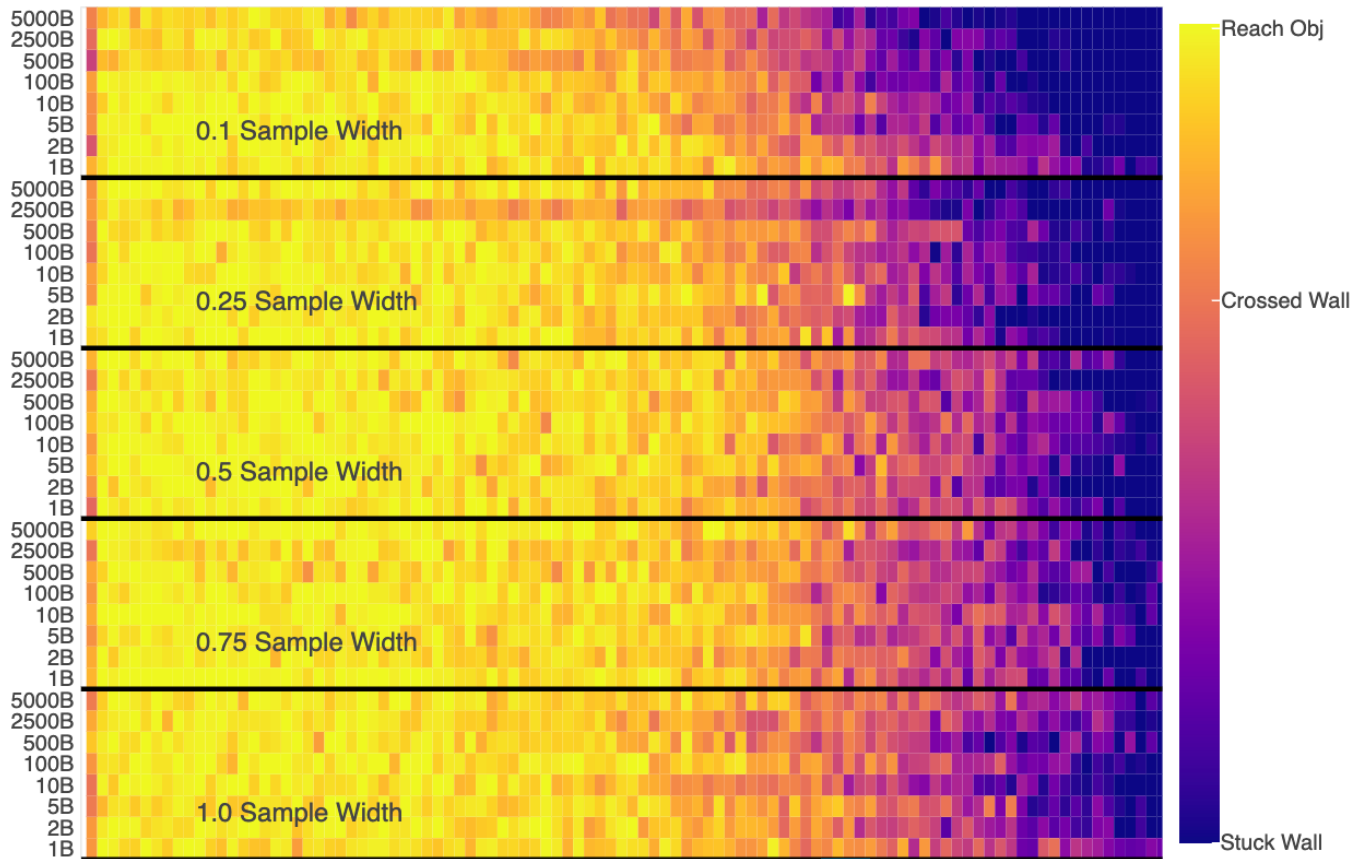


Figure 5: Average performance of the best individual per replicate by treatment during validation in each of the 100 environments. Treatments are grouped by the sampling width. Wall heights increase from left to right. Braking strategy increases from bottom to top in each width subgroup.

There is a decline in performance for wall heights above 50 with many treatments not able to cross the wall. Although the highest wall heights appear difficult for any treatment, the 1.0 sampling width group has the broadest generalization. As sampling width narrows, there is a subsequent decline as higher wall heights are more challenging. Physically these wall heights would be the most imposing for the animat, requiring gaits that lift the legs substantially to cross the wall.

In terms of sampling widths, 0.75 is not significantly different using a Wilcoxon rank-sum pairwise test compared to 1.0 counting the total number of successful wall crossings across wall heights for the best individual per replicate. As shown in Figure 4, sampling widths of 0.75 can still be expected to sample environments between 20 and 80 fairly uniformly. Generalizability declines in narrower widths (0.1, 0.25, 0.5), more drastically as the braking interval increases. The sum of the successful wall crossings for the best individual per replicate during validation for 0.25 and 0.1 sam-

pling width are significantly different than the 1.0, 0.75, and 0.5 sampling widths across braking intervals. The two factors of narrow sampling widths and long intervals evolving in the same range of environments prevent populations from evolving on a wide range of wall heights, apparently hindering generalizability.

**Tiebreaks** Figure 6 plots the number of tiebreaks in a replicate versus the count of wall heights with the target being reached for the best individual per replicate grouped by treatment. Within a specific braking regimen, treatments show a general increase in the number of tiebreaks as sampling width narrows. For the same sampling width this does not hold, as the number of tiebreaks does not vary due to the braking regimen. Observing the figure, it appears that there is generally a positive correlation between the number of tiebreaks and the total wall heights an individual solves for smaller braking values (1B, 2B, 5B, 10B, 100B) and wide sampling widths (1.0W, 0.75W, 0.5W). This begins to break-

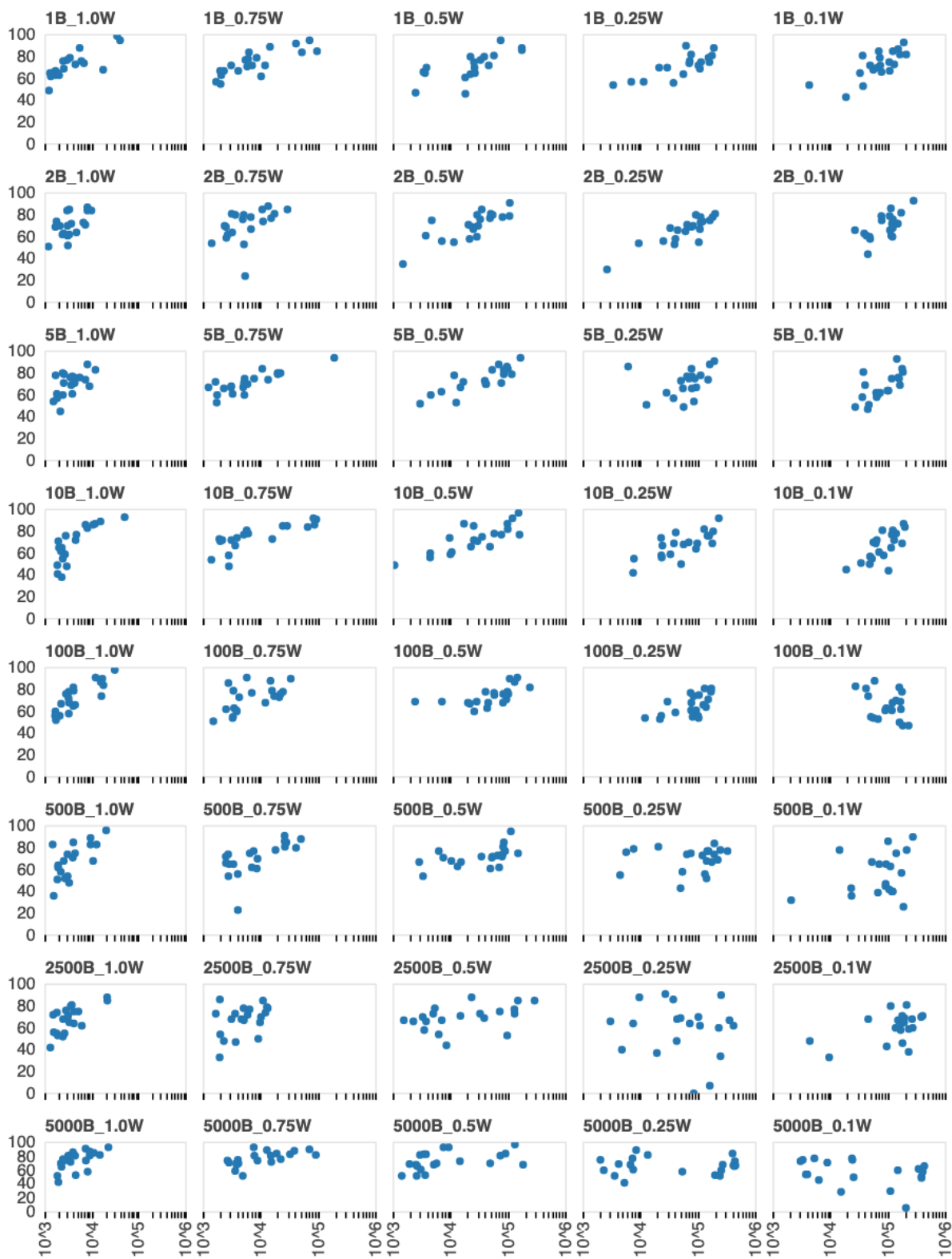


Figure 6: Number of tiebreaks in a replicate versus the number of environments where the best individual per replicate reached the objective, across treatments. Each sub-plot shows 20 replicates.

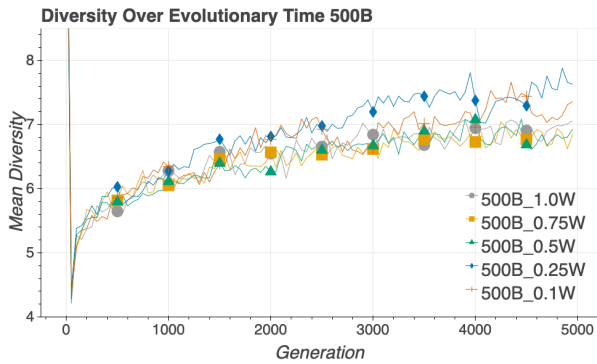


Figure 7: Mean population diversity across 20 replicates per treatment for the 500 braking treatments over evolutionary time.

down, becoming almost random, in the lower right plots as the combination of high braking and narrow width push the Lexicase algorithm to its limit. Here, the limited range of environments and lack of moving the sampling range restrict the total number of environments that a population is exposed to. Thus, individuals become specialized rather than evolving toward generalized performance.

**Diversity** Figure 7 plots the mean population diversity across replicates for the 500 braking interval treatments. Trends shown in this plot are similar for the other braking intervals. For widths 1.0, 0.75, and 0.5 there is no difference in mean population diversity. Mean population diversity is generally higher for widths 0.25 and 0.1. This trend holds for all braking intervals with a clear difference in mean diversity between (1.0, 0.75, 0.5 widths) and (0.25/, 0.1 widths) that increases as the braking interval lengthens. We surmise that this is due to the reduction in environments that individuals are exposed to over evolutionary time. We note that in evolutionary algorithms, an increase in diversity is unusual where individuals are unable to discover regions of higher fitness and that normally the converse is true.

In terms of mean population diversity predicting generalization performance of the best individual per replicate we do not observe any relationship. Instead, mean population diversity tends to fall in the range of 6-8, with low variation in small braking intervals or wide sampling widths. As the sampling widths narrow and braking intervals increase, mean population diversity tends to increase as well as the variation between replicates. However, no positive or negative correlation exists between mean population diversity and generalizability of the best individual per replicate. This clarifies one open question that we had raised in Moore and Stanton (2018), where we hypothesized that diversity might predict performance. We do note however, that for all treatments, mean population diversity still falls in the range of 6-8, consistent with earlier results.

## Conclusions and Future Work

In prior work, we identified ‘good’ parameter configurations that show Lexicase selection is the preferred algorithm for this task across a number of different algorithms investigated. Here, we identify specific factors and characteristics of the Lexicase algorithm that do, and do not, lead to generalized behavior on the wall crossing task. While there is a degradation in generalization performance as braking intervals increase and sampling widths narrow, Lexicase is still quite resilient. Even the lowest generalization treatments still on average are successful at crossing the wall in 60 environments.

Tiebreaks are a mechanism to add a random selection event during evolution. Tiebreak events occur due to multiple individuals being considered equal in the wall heights used during selection. When looking at generalization performance, there is a positive correlation between the number of tiebreaks over evolutionary time and generalization performance of the best individual per replicate until the restrictions placed on Lexicase become too strong. The correlation then disappears. As the sampling widths narrow, and braking increases, we see tiebreaks rise due to the higher adjacency of wall heights, see top of Figure 4. In these restricted treatments, populations are exposed to narrow ranges of wall heights which are likely highly related. From our results, it appears that some number of these random selection events are beneficial when Lexicase is relatively unhindered. Whereas, mean population diversity does not appear to be a predictor of performance.

In future work, we plan to continue pursuing questions regarding the underlying factors that make Lexicase selection effective. We plan to further expand the range of tasks an animat is evaluated in to include those that are not thematically related. That is, multiple tasks versus variations on a similar themed task.

## Acknowledgements

We would like to acknowledge the contributions of Thomas Helmuth, Bill La Cava, Edward Pantridge, and Lee Spector for their discussions during the formulation of this work. We would also like to thank Keele University and Grand Valley State University for their continued support of this research by providing computational resources and travel funding.

## References

- Auerbach, J. E. and Bongard, J. C. (2010). Dynamic resolution in the co-evolution of morphology and control. In *Proceedings of the Twelfth International Conference on Artificial Life*, pages 451–458, Odense, Denmark.
- Cheney, N., MacCurdy, R., Clune, J., and Lipson, H. (2013). Unshackling evolution: Evolving soft robots with multiple materials and a powerful generative encoding. In *Proceedings of the 15th Annual Conference on Genetic*

- and *Evolutionary Computation*, pages 167–174, Amsterdam, The Netherlands. ACM.
- Collins, J., Geles, W., Howard, D., and Maire, F. (2018). Towards the targeted environment-specific evolution of robot components. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 61–68, Kyoto, Japan. ACM.
- Dolson, E. and Ofria, C. (2018). Ecological theory provides insights about evolutionary computation. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '18, pages 105–106, Kyoto, Japan. ACM.
- Helmuth, T., McPhee, N. F., and Spector, L. (2016). The impact of hyperselection on lexicase selection. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '16, pages 717–724, Denver, Colorado, USA. ACM.
- Helmuth, T., Spector, L., and Matheson, J. (2014). Solving uncompromising problems with Lexicase selection. *IEEE Transactions on Evolutionary Computation*, PP(99):1–1.
- Jakobi, N. (1998). Running across the reality gap: Octopod locomotion evolved in a minimal simulation. In *Proceedings of the First European Workshop on Evolutionary Robotics*, pages 39–58, Paris, France. Springer-Verlag.
- Jelisavcic, M., Roijers, D. M., and Eiben, A. (2018). Analysing the relative importance of robot brains and bodies. In *Proceedings of the 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*, pages 327–334, Tokyo, Japan. MIT Press.
- Koos, S., Mouret, J. B., and Doncieux, S. (2010). Crossing the reality gap in evolutionary robotics by promoting transferable controllers. In *Proceedings of the 2010 ACM Genetic and Evolutionary Computation Conference*, pages 119–126, Portland, Oregon, USA. ACM.
- Kriegman, S., Cheney, N., Corucci, F., and Bongard, J. C. (2018). Interoceptive robustness through environment-mediated morphological development. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 109–116, Kyoto, Japan. ACM.
- La Cava, W., Spector, L., and Danai, K. (2016). Epsilon-lexicase selection for regression. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, pages 741–748, Denver, Colorado, USA. ACM.
- Moore, J. M. and McKinley, P. K. (2016). *A Comparison of Multiobjective Algorithms in Evolving Quadrupedal Gaits*, pages 157–169. Springer International Publishing, Aberystwyth, UK.
- Moore, J. M. and Stanton, A. (2017). Lexicase selection outperforms previous strategies for incremental evolution of virtual creature controllers. In *Proceedings of the 14th European Conference on Artificial Life*, pages 290–297, Lyon, France. MIT Press.
- Moore, J. M. and Stanton, A. (2018). Tiebreaks and diversity: Isolating effects in lexicase selection. In *Proceedings of the 16th International Conference on the Simulation and Synthesis of Living Systems*, pages 590–597, Tokyo, Japan. ACM.
- Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence and Technology of Self-Organizing Machines*. The MIT Press.
- Pantridge, E., Helmuth, T., McPhee, N. F., and Spector, L. (2018). Specialization and elitism in lexicase and tournament selection. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 1914–1917, Kyoto, Japan. ACM.
- Powers, J., Kriegman, S., and Bongard, J. (2018). The effects of morphology and fitness on catastrophic interference. In *Proceedings of the 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*, pages 606–613, Tokyo, Japan. MIT Press.
- Reil, T. and Husbands, P. (2002). Evolution of central pattern generators for bipedal walking in a real-time physics environment. *IEEE Transactions on Evolutionary Computation*, 6(2):159–168.
- Ruud, E. L., Samuelsen, E., and Glette, K. (2016). Memetic robot control evolution and adaption to reality. In *Proceedings of the 2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–7, Athens, Greece. IEEE.
- Smith, R. (2013). Open Dynamics Engine, <http://www.ode.org/>.
- Spector, L. (2012). Assessment of problem modality by differential performance of Lexicase selection in genetic programming: A preliminary report. In *Proceedings of the 14th Annual Conference Companion on Genetic and Evolutionary Computation*, pages 401–408, Philadelphia, Pennsylvania, USA. ACM.
- Stanton, A. (2018). Stochastic ontogenesis in evolutionary robotics. In *Artificial Life Conference Proceedings*, pages 214–221. MIT Press.