

When Specialists Transition to Generalists: Evolutionary Pressure in Lexicase Selection

Jared M. Moore¹ and Adam Stanton²

¹School of Computing and Information Systems,
Grand Valley State University, Allendale, MI, USA

²School of Computing and Mathematics,
Keele University, Keele, ST5 5BG, UK
jared.moore@gvsu.edu, a.stanton@keele.ac.uk

Abstract

Generalized behavior is a long standing goal for evolutionary robotics. Behaviors for a given task should be robust to perturbation and capable of operating across a variety of environments. We have previously shown that Lexicase selection evolves high-performing individuals in a semi-generalized wall crossing task—i.e., where the task is broadly the same, but there is variation between individual instances. Further work has identified effective parameter values for Lexicase selection in this domain but other factors affecting and explaining performance remain to be identified. In this paper, we expand our prior investigations, examining populations over evolutionary time exploring other factors that might lead to generalized behavior. Results show that genomic clusters do not correspond to performance, indicating that clusters of specialists do not form within the population. While early individuals gain a foothold in the selection process by specializing on a few wall heights, successful populations are ultimately pressured towards generalized behavior. Finally, we find that this transition from specialists to generalists also leads to an increase in tiebreaks, a mechanism in Lexicase, during selection providing a metric to assess the performance of individual replicates.

Introduction

Robotic systems should exhibit robust behaviors that are resilient to perturbation and operate effectively across a wide variety of conditions. However, realizing these goals is a continuing challenge from an algorithmic standpoint. In this paper, we expand our earlier investigation into Lexicase Selection in the context of evolutionary robotics (ER). Lexicase is a many-objective selection operator (Spector, 2012) which we applied as part of a generational genetic algorithm (GA) to an established wall-crossing task (Moore and Stanton, 2017). Figure 1 illustrates this generalized control problem where individuals are challenged to express a gait capable of reaching a target position by crossing a wall whose height varies on a per-trial basis. Our initial investigation demonstrated that a GA with Lexicase selection reliably discovered generalized controllers capable of crossing the majority of wall heights, significantly exceeding the performance of custom algorithms designed for the task (ibid.)

We then attempted to identify the underlying mechanisms responsible for the high performance of Lexicase in this

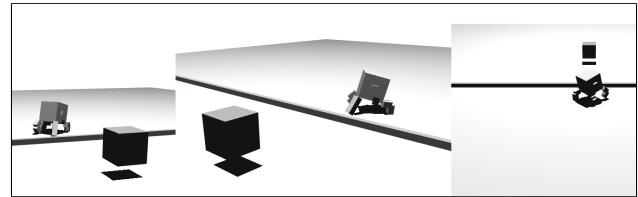


Figure 1: The wall crossing task examined in this study originally introduced in Stanton and Channon (2013). The animat must cross a wall of varying height to reach the target cube.

wall-crossing task (Moore and Stanton, 2018). Two key factors contributing to the success of Lexicase were identified. First, effective Lexicase parameter configurations experience an increasing frequency of tiebreaks. That is, if two or more individuals are tied after going through all objectives considered during the selection event, a random selection from the tied set of individuals is performed. Tiebreak events encourage an increase in population diversity because the choice is not based on performance. With a sufficient number of environments (≥ 5) considered during a selection event, the frequency of tiebreaks increases over time in effective replicates. The second factor is that high-performing individuals evolve in populations that settle into a specific range of population diversity. Still, as we demonstrated in Moore and Stanton (2019), tiebreaks appear to be one of the key factors of Lexicase selection. These experiments pushed the limits of Lexicase to understand when it breaks down. Lexicase exhibits a surprising resilience, maintaining high performance even when the range of wall heights considered in a selection event is heavily restricted to only a narrow band of adjacent wall heights. Under these constraints, tiebreaks continue to increase over time, but their correlation to performance breaks down as the search is biased towards small areas of the objective space.

The two prior investigations into Lexicase's effectiveness explored performance metrics aggregated across many replicates and their effect on population diversity. Another direction is to examine population genetics and the performance of ancestral populations across the 100 wall heights over the course of evolution identifying trends that might indi-

cate successful versus unsuccessful conditions. Here, we look for possible speciation in genome space. We hypothesize that since Lexicase evolves specialists (Pantridge et al., 2018), individuals that are high performing in one or a limited set of objectives, clusters will form in genome space exhibiting similar performance characteristics to each other but different from those of other clusters. Second, we evaluate performance of populations at different points in evolution to see how individuals, and the populations at large, transition from specialists to generalists.

In this paper, we expand our investigation further on data originally reported in Moore and Stanton (2018). We focus on exploring the assumption that initial populations contain specialists, individuals whose performance is poor on all but a few task configurations. These individuals may then begin to produce offspring that carry this genetic information forward, adding additional competencies resulting in a more general capability that comes to dominate the population. Alternatively, specialist genetic information may be ephemeral and persist in the population only until generalists evolve that can out-compete specialists. We examine ancestral populations over evolutionary time to determine what features arise that could indicate the procession of evolving species from sparse specialists to converged generalists. We look specifically at two areas. First, whether clusters of individuals (putative sub-species) in genotype space correspond with clusters in performance space (indicating specialists) and whether changes in these clusters over evolutionary time are indicative of the emergence of generalists in the population. Second, we consider whether the number of tiebreaks provide insight into the transition of species from being primarily specialists to primarily generalists.

Results of this paper are as follows. First, we find that individuals clustered by genomes do not exhibit similar performance, in contradiction of our initial hypothesis. There are also no apparent performance differences between clusters. Instead, populations transition from individual specialists towards generalized wall crossing behavior over the course of 5,000 generations. We also find additional evidence supporting earlier observations that a high number of tiebreaks is indicative of successful behavioral generalization.

Background and Related Work

Evolutionary robotics (Nolfi and Floreano, 2000) applies evolutionary principles to the optimization of robotic systems. ER applications include legged locomotion control (Baydin, 2012; Clune et al., 2009), morphological optimization (Paul and Bongard, 2001), and the transfer of controllers to reality (Ruud et al., 2016; Koos et al., 2010) among others. In many cases, ER approaches are moving beyond single objective fitness metrics to enhance the resilience, robustness, and generalizability of controllers (Pinville et al., 2011). For example, a legged animat might typically be assessed on locomotive performance by

using distance traveled as the primary metric, but secondary objectives such as efficiency can also be considered if multiple objectives are employed (Moore and McKinley, 2016). Other robust controllers consider evolving multiple behaviors for a single robotic platform (Doncieux and Mouret, 2013).

These pursuits of robustness (Lehman et al., 2013), resilience (Kriegman et al., 2019), behavioral diversity and generalized behavior thus involve multi-objective (Deb et al., 2002) or many-objective algorithms like Lexicase selection (Spector, 2012) that evaluate performance of individuals on more than one metric. Generalized behavior involves the expression of multiple robust behaviors, for example forward locomotion, turning, and reversing (Cully et al., 2015). In this paper, we continue to investigate the semi-generalized, many-objective task explored in Stanton and Channon (2013) and Moore and Stanton (2017, 2018, 2019) wherein individuals evolve to express a *robust* behavior capable of crossing a wall of varying heights with the same controller. We consider the behavior to be semi-generalized as individuals are evolved on different wall heights, each height constituting a unique objective. However, the overall goal for all objectives is similar, each being a variant of a general wall-crossing task, thus the composite problem requires generalization.

In our earliest work to address the generalized wall-crossing task, individuals were exposed to a variety of objectives over evolutionary time, according to a number of different presentation strategies (Stanton and Channon, 2013). It was found that commonly used strategies such as random presentation of the individual objectives are inferior to custom strategies that take into account linkage between wall heights. In these strategies, evolutionary pressure to adapt is thus maintained, whilst ensuring that species do not forget previously learned behaviors that address specific wall heights. However, these strategies all used single-objective selection mechanisms, albeit with the full spectrum of task parameterisations explored through environmental change over evolutionary time. In contrast, Lexicase accommodates a number of objectives simultaneously during a single selection event. The generalized wall-crossing task provides a problem which has been well explored in other work, and one that decomposes easily into the sub-components necessary to examine the application of Lexicase selection.

Methods

Quadrupedal Animat Figure 3 shows a wireframe of the animat including joint ranges of motion. The animat has 12 degrees of freedom (DOF) with 2-DOF hinges in the hip and 1-DOF hinges at the knee. The hips can sweep horizontally and extend the legs from parallel with the ground to vertical. When the legs are fully vertical, the maximum wall height is 67% of total leg length. In practice, the wall is typically higher relative to the animat as the legs are flexed during locomotion lowering the torso.

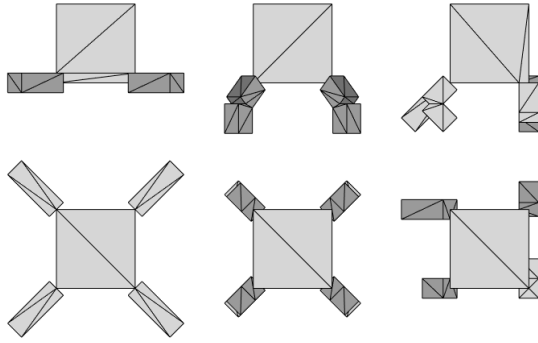


Figure 2: To-scale front and top schematics of animat morphology with legs at various positions within their range of motion. Figure adapted from Moore and Stanton (2019).

Animat’s Environment The quadrupedal animat walks on a flat high-friction surface with a target placed 10 body lengths in front of the animat’s starting position. A wall is placed between the animat and the target. Prior to a simulation, wall height is determined from 100 possibilities. Heights range from almost non-existent ($< 1\%$ of the animat’s extended leg) to a maximum height of 67% of the animat’s leg length. The Open Dynamics Engine (ODE) (Smith, 2013) handles physics in the simulation including: friction between the animat and the surface, operation of hinges connecting various components of the quadruped, and collisions between the various rigid bodies in the simulation.

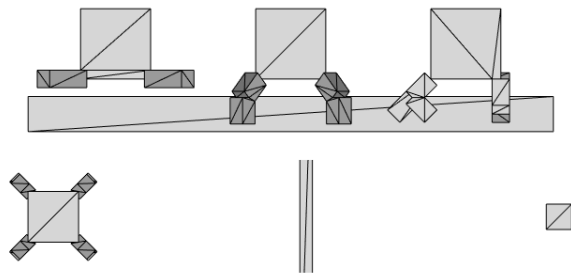


Figure 3: To-scale schematic of animat (left and top), target (right) and wall (centre and front), illustrating relative sizes and positions. Wall is shown at maximum height. Figure adapted from Moore and Stanton (2019).

Controller Controllers are Feed-Forward Artificial Neural Networks (ANN) with 16 inputs, 12 hidden nodes and 12 outputs. Table 1 lists the inputs providing the animat with: periodic signals promoting regular oscillating movement, the animat’s position relative to the target, and feedback on the current state of the 12 leg joints. Genomes consist of 336 evolved weights (16 inputs * 12 hidden nodes + 12 hidden nodes * 12 outputs).

1	Oscillator: $\sin(2\pi t)$
2	Oscillator: $\cos(2\pi t)$
3	Animat Balance Angle
4	Animat Orientation to Target
5–12	Hip Joint Angles
13–16	Knee Joint Angles

Table 1: ANN controller inputs.

Wall Crossing Task Each wall height is a unique objective. An individual animat’s distance from the target at the end of simulation is the fitness for one objective. Within a generation, individuals are simulated only on the environments drawn for the Lexicase selection step. In Moore and Stanton (2018), we found that 10 environments per generation is sufficient to select an individual to be a parent. The analysis conducted in this study therefore samples 10 environments during the Lexicase selection step. If there is a tie after 10 environments, a tiebreak occurs.

Evolutionary Algorithm with Lexicase Selection Evolutionary runs employ a generational genetic algorithm run for 5,000 generations with a population of 50 individuals. Offspring are created from two parents using one-point crossover with a gene mutation frequency of $\frac{2}{GenomeLength}$. Parents are selected each generation using Lexicase selection where performance in up to 10 environments is considered. An individual selection event samples a set of 5 potential parents, comparing them based on the shuffled order of the 10 selected environments. The best potential parent for the first environment is identified, and other potentials are considered “tied” if they are within 10% of performance. Any tied individuals are then compared on the second environment with a similar tie consideration. If two or more potential parents are tied in all 10 environments, a tiebreak occurs. A tiebreak is a random event where one of the remaining potential parents with similar performance across all 10 environments is selected. For a full description of the Lexicase algorithm employed in this paper we refer the reader to Moore and Stanton (2017).

Clustering In Genome Space Population genetics over evolution are considered to see if there are characteristics in the population that indicate the effectiveness of Lexicase. We evaluate genetics by clustering populations. Genomic and performance information across all 100 wall heights of the population is logged every 5 generations. Next, we perform a two-step dimensionality reduction of the genome space. First, a principal component analysis (PCA) capturing 95% of the variation in the replicate’s population. This typically reduces dimensionality from 336 genes down to 10-20 dimensions depending on the replicate. The PCA-reduced genome dataset is then run through the t-SNE algorithm (Maaten and Hinton, 2008) with parameters: 2 com-

ponents, n dimensions, and perplexity of 5 for the genome space, reducing dimensionality to 2 dimensions suitable for visualization. n dimensions indicate how many dimensions are the result of the PCA. Once the t-SNE projection is complete, we run the HDBScan clustering algorithm (Campello et al., 2013) with minimum cluster size of 4. This generates a clustering in genome space, identifying genomes that are similar to each other.

A second t-SNE is run on the performance of each individual with parameters: 2 components, 100 dimensions corresponding to the 100 wall heights, and perplexity of 10. We do not conduct a PCA on the performance space as 100 objectives are within the recommended parameter bounds of t-SNE.

Results and Discussion

The following analysis presents results from the highest and lowest performing replicates from the *Lexicase 1.1F 10E 1.0D 100B* treatment originally examined in Moore and Stanton (2019). Results from these two replicates are similar in the remaining replicates. Due to the nature of analyzing the population over time in this study, we do not present aggregated plots summarizing results across replicates. We observe similar results across the other treatments conducted in Moore and Stanton (2019) but due to space we present results only from a single treatment.

Cluster Analysis Figures 4a and 4b plot the populations at generation 5,000 in both genome and performance space. Our hypothesis is that if specialists are present in the population, clusters in genome space will map to clusters in performance space. The populations cluster in genome space readily, with outliers indicated by diamonds. However, genome based clusters do not map to performance space as seen in the right side of the figures. Instead, clustered individuals are evenly dispersed across the space with generally poor performance in Figure 4a and high performance in Figure 4b. Contrary to our hypothesis, it appears that the genomic clustering does not in fact map to performance space at this late generation.

Figures 5a and 5b show the performance of every individual in the generation 5,000 populations across the 100 wall heights. Clusters from Figures 4a and 4b are carried into these two plots. Individual clusters are separated by the white horizontal lines. The populations are primarily generalists with similar performance across environments. No single cluster exhibits a level of performance in any segment of wall heights that other clusters do not. If such a situation were to exist, with a cluster showing high performance in a specific range of wall heights, it would support the idea that clusters are specialized groups of individuals. In the lowest performing replicate, Figure 5a, all clusters evolve generalists for the lower half of wall heights. Whereas, in the highest performing replicate, Figure 5b, all clusters evolve individuals that express high generalizability typically capable of crossing all but the tallest of wall heights.

The final populations discussed above are the result of 5,000 generations of evolution. It may be that specialist clusters transition to generalists sometime in earlier generations. Figures 6a and 6b show the evolutionary history of the two replicates. Shapes indicate clusters formed from the t-SNE/HDBScan. Shapes are reused at different generations as no clusters persist over 1,200 generations. Each generation is thus a unique set of clusters. In the two figures, no single cluster exhibits large differences in performance compared to other clusters. Instead, each replicate depicts snapshots of the evolutionary trajectory that arrive at Figures 5a and 5b respectively. Examining the ancestral populations at these generations we notice similar performance to that shown in Figure 7. That is, within a cluster, there are a number of high performing individuals relative to the population but there are also a number of low performing individuals. We do not observe a single cluster dominating others in terms of performance across wall heights.

Although clustering doesn't indicate that groups of individuals are specializing on specific wall heights, it does appear that individual specialists emerge in the first few hundred generations of evolution. Figure 7 plots the ancestral population at generation 125 of Replicate 10. At this early point in the evolutionary process it is unlikely that every environment has been used for a Lexicase selection event. Indeed, there is a high prevalence of stuck on wall or lower performance across the wall heights. However, the figure also shows that some individuals exhibit effective performance on wall heights that are not adjacent to each other. These are likely the initial specialists that emerge in the early generations of Lexicase selection. (Bright colors in columns with primarily dark colors.) If one of these wall heights is an objective and the specialist is a potential parent, it has a higher likelihood of being selected. More individuals in future populations will then share the specialist's genetic material. With many individuals having the genetic material, they will likely also have high performance on the previous specialist's niche wall heights. In order to maintain competitiveness, future individuals will likely then be driven towards specializing on other wall heights. We hypothesize that this process continues over time, eventually resulting in specialists transitioning to generalists as more individuals in the population exhibit effective performance across wall heights. Thus, high performance sweeps through the population across environments, see Figure 5b.

Additionally it may be that some specialists exist in the population only until a generalist appears with capability that includes that exhibited by the specialist. In this view, specialists regularly appear in the population due to mutation, but do not persist beyond a few generations. This view does not require these independent specialists contribute their genetic material toward future generalists; they simply provide evolutionary pressure towards the general solution by competing with potential generalists in specific wall

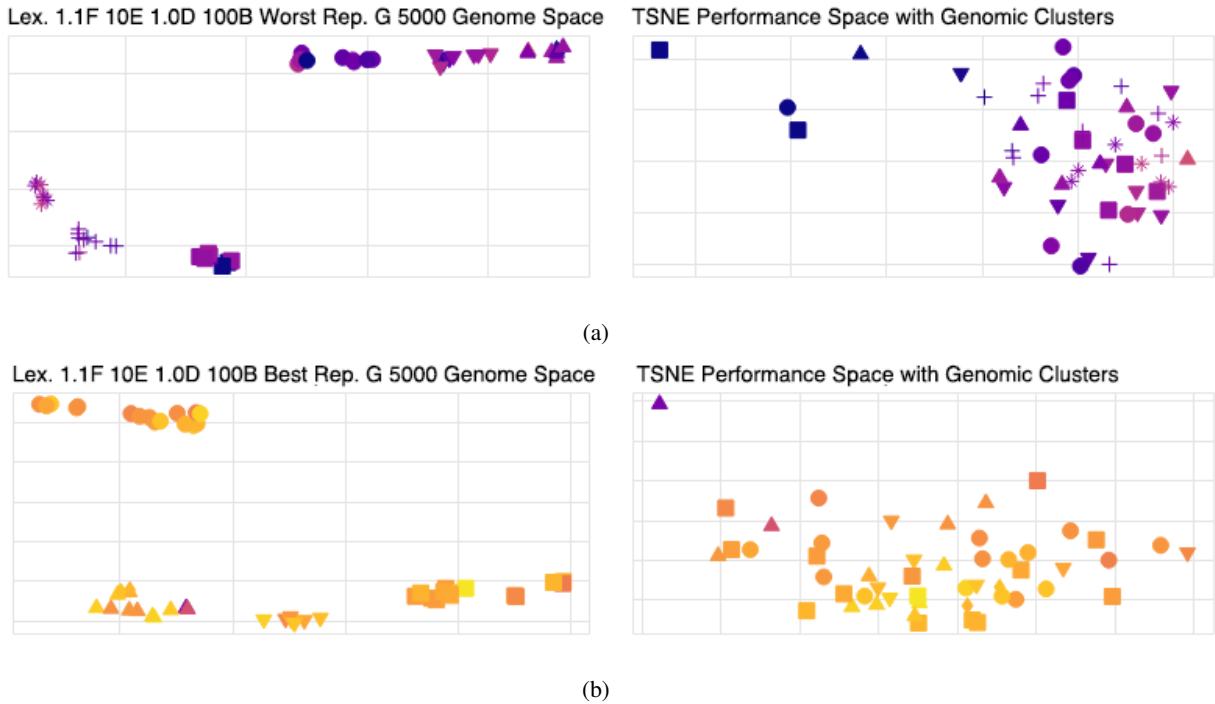


Figure 4: Final populations after 5000 generations of evolution. Left side depicts the t-SNE projection with HDBSCAN clustering algorithm applied in genome space. Colors indicate the average fitness of the individual as measured across the 100 objectives after evolution has concluded. Darker shades indicate poor performance. Shapes indicate the assigned cluster. Diamonds are outliers. Right side depicts t-SNE applied to performance space. Shapes are the clusters projected from genome space. (a) Worst performing replicate. (b) Best performing replicate.

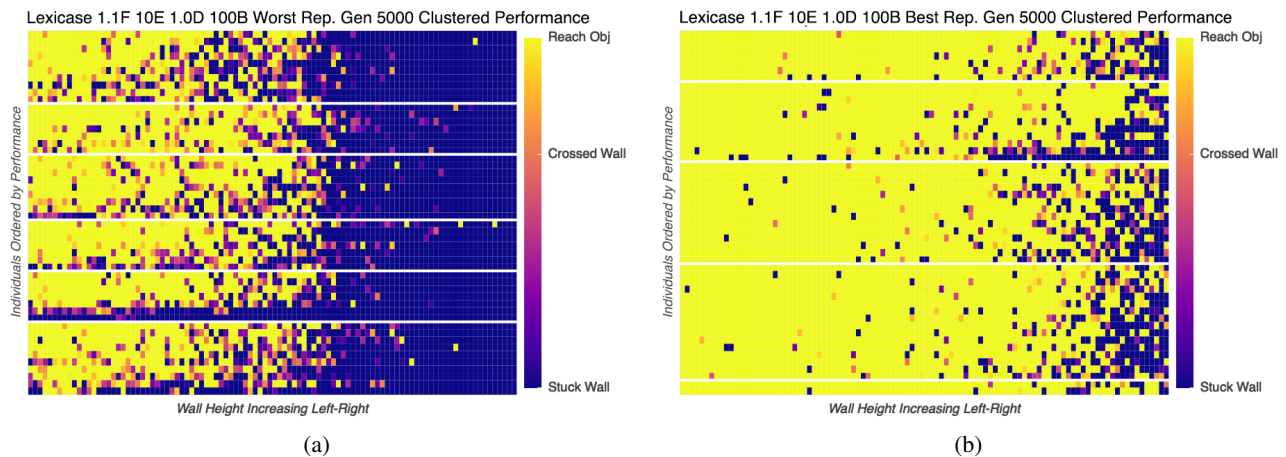


Figure 5: (a) A replicate with clusters in genome space but subpar performance. No distinct performance differences between clusters. Clusters separated by white horizontal lines. (b) A replicate with clusters formed in genome space yet high performance across the entire population.

heights.

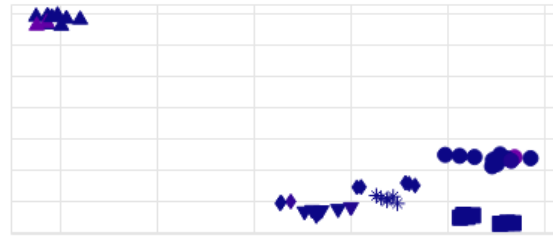
Tiebreak analysis In previous work (Moore and Stanton, 2018) we noted that tiebreaks are an indicator of success in the wall crossing domain. Here, we find additional relationships to tiebreaks in that when specialists are heavily present in the early populations, the number of tiebreaks are low. Figure 8 plots the rolling average of tiebreaks during

selection per generation. In very early generations ($< \text{Generation } 25$) the number of tiebreaks are high as randomly generated individuals all have similar poor performance. Selection thus cannot readily distinguish performance between individuals resulting in Lexicase exhausting the 10 environments used for selection. However, as specialists emerge and those environments are chosen as objectives in Lexicase selection, the number of tiebreaks drops to almost 0

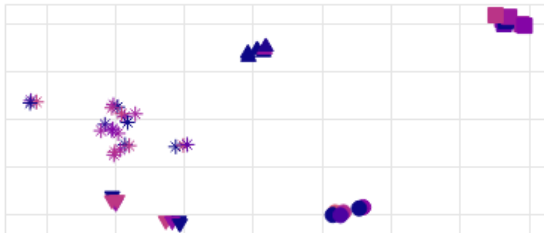
L 1.1F 10E 1.0D 100B Worst Rep. G 200 Genome Space



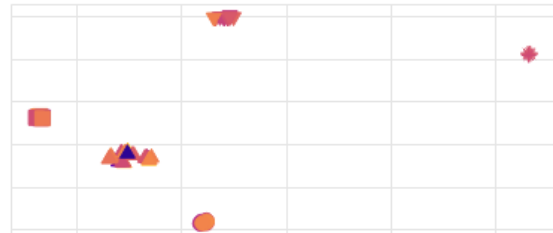
L 1.1F 10E 1.0D 100B Best Rep. G 200 Genome Space



L 1.1F 10E 1.0D 100B Worst Rep. G 1400 Genome Space



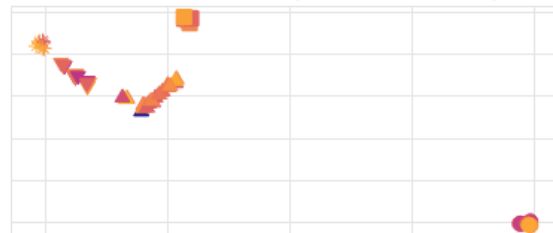
L 1.1F 10E 1.0D 100B Best Rep. G 1400 Genome Space



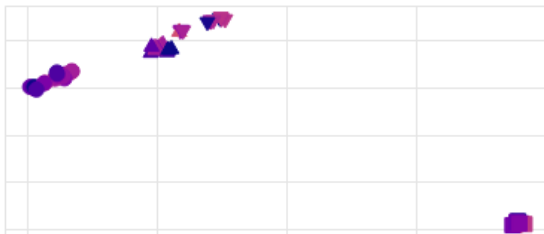
L 1.1F 10E 1.0D 100B Worst Rep. G 2600 Genome Space



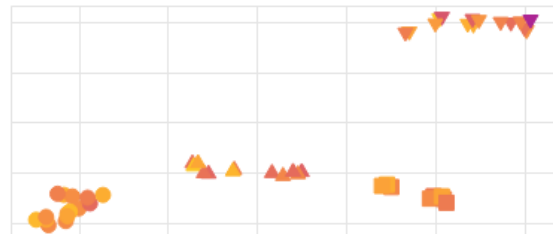
L 1.1F 10E 1.0D 100B Best Rep. G 2600 Genome Space



L 1.1F 10E 1.0D 100B Worst Rep. G 3800 Genome Space



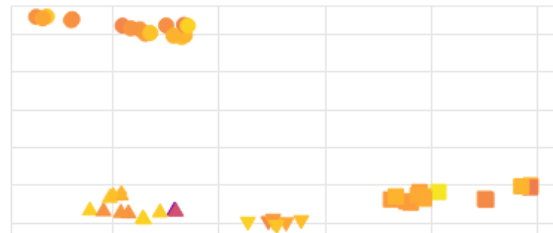
L 1.1F 10E 1.0D 100B Best Rep. G 3800 Genome Space



L 1.1F 10E 1.0D 100B Worst Rep. G 5000 Genome Space



L 1.1F 10E 1.0D 100B Best Rep. G 5000 Genome Space



(a)

(b)

Figure 6: (a) Genome clustering of the population every 1,200 generations for the worst performing replicate. Diamonds indicate outliers. Color shading indicates the average fitness of the individual across all 100 wall heights. (b) Genome clustering of the population every 1,200 generations for the best performing replicate. No apparent relationship between clusters and performance is evident in either replicate.



Figure 7: Early generations show that a few individuals have evolved effective wall crossing strategies for some, but not all wall heights. High performance is indicated by the bright shading in columns.

in both replicates for the first 800 generations. As the performance across the populations increases, the number of tiebreaks also increases. For the worst performing replicate, the number of tiebreaks remain low as the population fails to generalize performance across all environments primarily failing to evolve effective strategies for high wall heights. In this case, specialists are still present for the higher wall heights as the population has not transitioned to generalists as shown in the contrast between Figures 5a and 5b. We hypothesize that a high number of tiebreaks in a replicate might be an indicator that the population has transitioned from primarily specialists to generalists. We plan to test this hypothesis in future work.

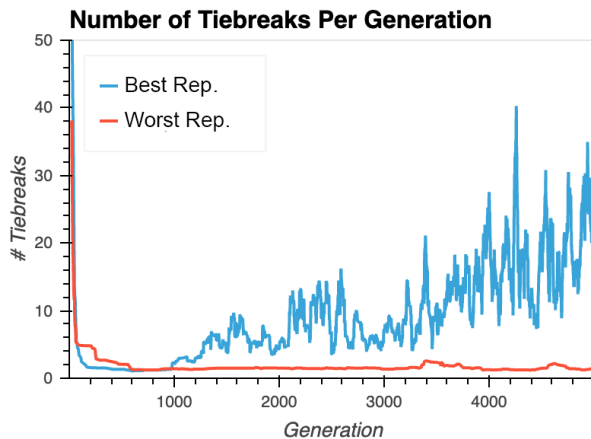


Figure 8: 25 generation rolling average of the number of tiebreaks for the two replicates over evolutionary time.

Conclusions and Future Work

We have previously shown that Lexicase selection evolves generalized performance for robot controllers in this wall crossing task. Prior results identified parameter configurations enabling Lexicase to perform effectively as a selection mechanism. Tiebreaks have been a key contributor allowing for small random selection events during evolution. Here, we expand on our earlier observations noting that specialists do not arise as a cluster of individuals in genetic space. Instead, a small number of individuals specialize on an environment and their genetic material is then spread throughout the population during the evolutionary process. When evaluating specific replicates, it appears that the transition from specialists to generalists might be indicated by the overall frequency of tiebreaks that occur during the selection step. As generalized behaviors emerge across the population, the number of tiebreaks increases. In addition, when a replicate does not evolve generalized performance, the number of tiebreaks is substantially lower than that of a successfully generalized replicate.

Our ongoing and future work continues altering the Lexicase selection algorithm to better understand factors that drive its behavior. First, we will examine the relationship between tiebreaks and behavioral generalization as discussed earlier. We also plan to examine strategies that influence how the population is evaluated on the various objectives to approach specialization versus generalization. Finally, we will examine task environments wherein animats are evolved to perform a variety of distinct tasks similar to Lexicase's original introduction of modal problems in Genetic Programming and see if our findings on the wall crossing task extend to other domains.

Acknowledgements

The authors would like to thank William La Cava, Thomas Helmuth, Edward Pantridge, Emily Dolson, and Lee Specator for their conversations about various aspects of Lexicase selection. We would also like to acknowledge Keele University and Grand Valley State University for supporting this research.

References

- Baydin, A. G. (2012). Evolution of central pattern generators for the control of a five-link planar bipedal walking mechanism. *Paladyn. Journal of Behavioral Robotics*, 3(1):45–53.
- Campello, R. J. G. B., Moulavi, D., and Sander, J. (2013). Density-based clustering based on hierarchical density estimates. In Pei, J., Tseng, V. S., Cao, L., Motoda, H., and Xu, G., editors, *Advances in Knowledge Discovery and Data Mining*, pages 160–172, Berlin, Heidelberg, Springer Berlin Heidelberg.

- Clune, J., Beckmann, B. E., Ofria, C., and Pennock, R. T. (2009). Evolving coordinated quadruped gaits with the HyperNEAT generative encoding. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 2764–2771, Trondheim, Norway. IEEE.
- Cully, A., Clune, J., Tarapore, D., and Mouret, J.-B. (2015). Robots that can adapt like animals. *Nature*, 521(7553):503–507.
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197.
- Doncieux, S. and Mouret, J. B. (2013). Behavioral diversity with multiple behavioral distances. In *Proceedings of the 2013 IEEE Congress on Evolutionary Computation*, pages 1427–1434, Cancun, Mexico. IEEE.
- 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., Walker, S., Shah, D. S., Kramer-Bottiglio, R., and Bongard, J. (2019). Automated shapeshifting for function recovery in damaged robots. In Antonio Bicchi, H. K.-G. and Hutchinson, S., editors, *In Proceedings of the Robotics: Science and Systems XV*, Freiburg im Breisgau, Germany.
- Lehman, J., Risi, S., D’Ambrosio, D., and Stanley, K. O. (2013). Encouraging reactivity to create robust machines. *Adaptive Behavior - Animals, Animats, Software Agents, Robots, Adaptive Systems*, 21(6):484–500.
- Maaten, L. v. d. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605.
- 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.
- Moore, J. M. and Stanton, A. (2019). The limits of lexicase selection in an evolutionary robotics task. In *The 2019 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 551–558, Newcastle-Upon-Tyne, UK. MIT Press, MIT Press.
- 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.
- Paul, C. and Bongard, J. C. (2001). The road less travelled: Morphology in the optimization of biped robot locomotion. In *Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 226 – 232, Maui, Hawaii, USA.
- Pinville, T., Koos, S., Mouret, J.-B., and Doncieux, S. (2011). How to promote generalisation in evolutionary robotics: The progab approach. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, pages 259–266, Dublin, Ireland. ACM.
- Ruud, E. L., Samuelson, 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. and Channon, A. (2013). Heterogeneous complexification strategies robustly outperform homogeneous strategies for incremental evolution. In *Proceedings of the 12th European Conference on Artificial Life*, pages 973–980, Taormina, Italy.