The Effect of Sub Sampling on the Preservation of Specialists

Daniel Junghans[1], Austin Ferguson[2 3 4], Alex Lalejini[2 3 4], Charles Ofria[2 3 4]

Business Department[1], Michigan State University, East Lansing, MI 48824

Computer Science and Engineering[2], Ecology, Evolutionary Biology, and Behavior[3], BEACON Center[4], Michigan State University, East Lansing, MI 48824

Abstract

Genetic algorithms are tools that use the principles of natural evolution to solve computational or engineering problems. Genetic algorithms start by generating populations of random candidate solutions, evaluating each of them, and selecting the “best” to move on to the next generation. When running genetic algorithms, different selection schemes exist to determine which parents should be selected for the next generation of candidate solutions. These selection schemes often determine which candidate solution is best by administering several tests to each candidate solution. These tests are referred to as test cases, and each selection scheme uses these tests to determine which candidate solution should be selected. Something that effects the quality of these evolutionary outcomes is the preservation of candidate solutions that are exceptional at solving only a few test cases. These candidate solutions are often called specialists, and many selection schemes have trouble preserving them because they don’t solve many test cases. A selection scheme called Lexicase can preserve specialists at a higher rate by continuously shifting which test cases are most important. Different sub sample variants of Lexicase (Down-Sampled Lexicase and Cohort Lexicase) have emerged in an effort to make Lexicase selection faster. In this work, I examined Lexicase and its randomly sub sampled variants by focusing on the differences in selection probabilities for specialists. The selection probabilities for selecting a specialist was calculated for each data set, and a Kruskal-Wallis test was performed to measure significance. After performing statistical analysis, significant differences in the probability of selecting a specialist between Lexicase and its randomly sub sampled variants were revealed.

I. Introduction

Evolution in nature has demonstrated its ability to produce amazing and diverse solutions to all sorts of problems. Genetic algorithms seek to answer complex computational and engineering problems by implementing evolution. These algorithms start with a population of possible solutions to a problem and then seek to pick the best performing possible solutions. There are many different selection schemes that can be implemented, and each one behaves differently. The most direct selection method is called "Elite selection" and it will always choose the highest fitness individual in the population. The problem with Elite selection is that one false step can take the population to a local "peak", where the solution is better than those around it, but not perfect. Most alternatives put more randomness into the choice to keep some variation in the population (e.g., Roulette selection and Tournament selection), or even actively encouraging diverse solutions (e.g., Lexicase selection and Fitness Sharing.) These differences in sampling methods can have a profound effect on whether a solution is found.

Due to the small number of theorists in the field of Evolutionary Computation, there has been limited research into how sub sampling affects the preservation of specialists and the differences in specialist preservation between selection schemes. There have, however, been other similar efforts in this field to determine the differences of selection schemes by measuring their overall performance, without digging into the dynamics of the system to understand why performance was affected in a particular way. For example, Goldberg and Deb have studied the differences in performance of several selection schemes. Their work has revealed that there are significant differences in the performance of each selection scheme when comparing the solutions, growth ratio estimates, and takeover time produced by the scheme (Goldberg & Deb, 1991). These research efforts are useful for building a better understanding of selection schemes, but my research builds off this kind of work in order to build a better understanding of the effect of random sub sampling on specialist preservation.

In the following sections of this paper, I will first provide more information on the prior work using the selection schemes that we are analyzing and will discuss the details of how we conducted our experiments and analyses in Section II. Results from the research will then be found in Section III and will be followed by an analysis of those results in Section IV. Lastly, I will describe the broader impact of this work in the Conclusions section as well as discuss future work in Section V.

II. Methods

When beginning this research, I chose several popular selection schemes to use in this study. Two common selection schemes (Roulette selection and Tournament selection) were used as well as three versions of lexicase selection (Lexicase, Down-Sampled Lexicase, and Cohort Lexicase). The randomly sub sampled variants of Lexicase were chosen in order to see if there was any significant difference in the preservation of specialists. The two common selection schemes used in this experiment were added as controls.

1. Roulette Selection
2. Tournament Selection (Tournament Sizes Used: 2,4, and 7)
3. Lexicase
4. Down Sampled Lexicase (20% sub sampling and 50% sub sampling)
5. Cohort Lexicase (20% sub sampling and 50% sub sampling)

Roulette Selection

In Roulette selection candidate solutions are selected with a probability directly tied to the number of test cases they solve. Every single candidate solution can be represented as portion of a roulette wheel. The more test cases a candidate solution solves the bigger its portion is on the roulette wheel. This selection scheme spins an imaginary roulette wheel and the resulting candidate is selected to move onto the next generation of candidate solutions (Razali & Geraghty, 2011).

Tournament Selection

Tournament selection aims to pick the best parent population by making individual possible solutions "compete" against one another. For every round of selection, this selection scheme holds a tournament with a variable tournament size. The candidate solution that solves the most test cases in the tournament is deemed to have the highest fitness by this selection scheme. The winner is then put into a pool of other winners and will become a parent for the next generation of candidate solutions (Goldberg and Miller, 1995).

Lexicase Selection

Lexicase selection is a complex selection scheme that is able to preserve specialists at a higher rate than other selection schemes. Lexicase selection starts by randomly ordering the test cases and then throwing out all the candidate solutions that could not solve the first random test case. If more than one candidate solution remains, this process is repeated until there is one winner. If there are two or more candidate solutions with the same performance, Lexicase selection will randomly pick one of the identical solutions (Helmuth, Spector, & Matheson, 2014).

Down-Sampled Lexicase Selection

Down-Sampled Lexicase selection is a selection scheme that tries to increase the speed of Lexicase by utilizing random sub sampling. This selection scheme behaves the same way as Lexicase, but instead of using every test case in the selection process, it instead randomly picks a certain percentage of test cases. The percentage of test cases that are used can vary and depends on the genetic algorithm. The entire population is then subjected to Lexicase selection, but only using this reduced set of test cases (Hernandez, Lalejini, Dolson, & Ofria, 2019).

Cohort Lexicase Selection

Cohort Lexicase is another randomly sub sampled variant of Lexicase and it aims to decrease the population size and test case count for each round of selection. This selection scheme behaves the same way as lexicase, but instead of using the entire population and every test case at once, it instead randomly picks candidate solutions and test cases and puts them into smaller cohorts. The percentage of individuals and test cases in every cohort can vary and depends on the genetic algorithm. By placing the population and test cases into small cohorts, Cohort Lexicase selection is much quicker than Lexicase (Hernandez, Lalejini, Dolson, & Ofria, 2019).

This research began with creating 30 experimental conditions of different configuration variables. The configuration variables that were altered in creating these data sets included:

1. The size of the population in the data set (10, 20, 100 candidate solutions)
2. The probability of a baseline organism passing a non-focal test case (1.0, 0.9, 0.5, 0.1, and 0.01)
3. The number of test cases (10 and 20)

In order to determine how random sub sampling affects lexicase's ability to preserve specialist, 10% of each population was designated as a focal specialist that could solve exactly 1 focal test case that could not be completed by the rest of the population. The specialist was only able to solve the focal test case with a 100% pass rate. Every single experimental condition had 1 focal test case. After creating 30 experimental conditions, I generated 100 data sets for each experimental condition for every selection scheme used. Once these data sets were created, every single selection possibility was enumerated for all selection schemes except Tournament and the sub sampled Lexicase variants. In order to see how random sub sampling affects the probability of preserving specialists, the selection probabilities for specialists in every individual population was summed. After the total specialist selection probability was calculated, the data was graphed, and a Kruskal-Wallis test was performed on Lexicase and its sub sample variants in order to determine if there was a significant difference in the probability of selecting a specialist between selection schemes.

IV. Results

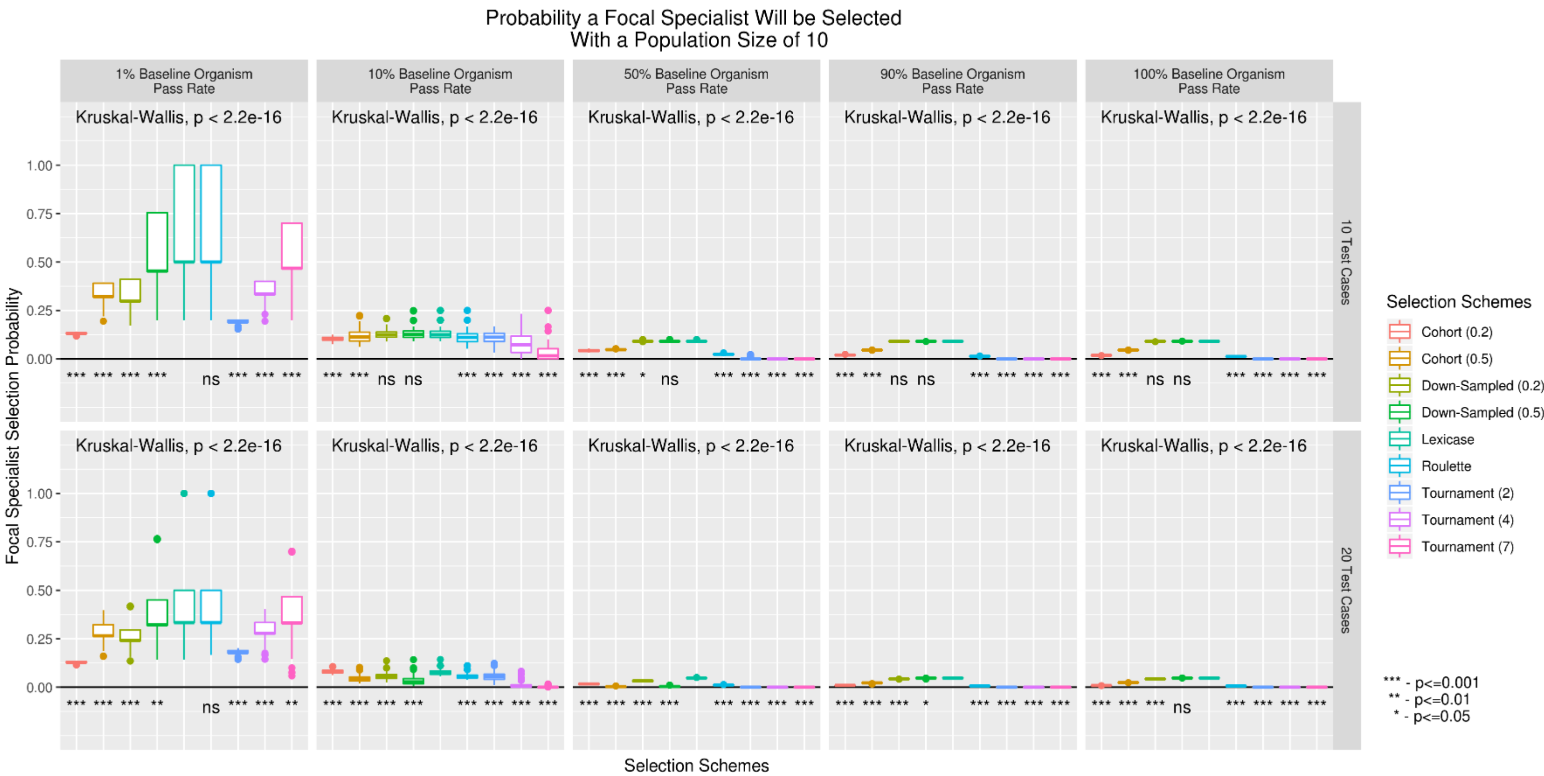

Figure 1. Shows the spread of focal specialist probability for a population size of 10. Kruskal-Wallis shown for Lexicase variants. Asterisks denote Mann-Whitney tests against Lexicase.

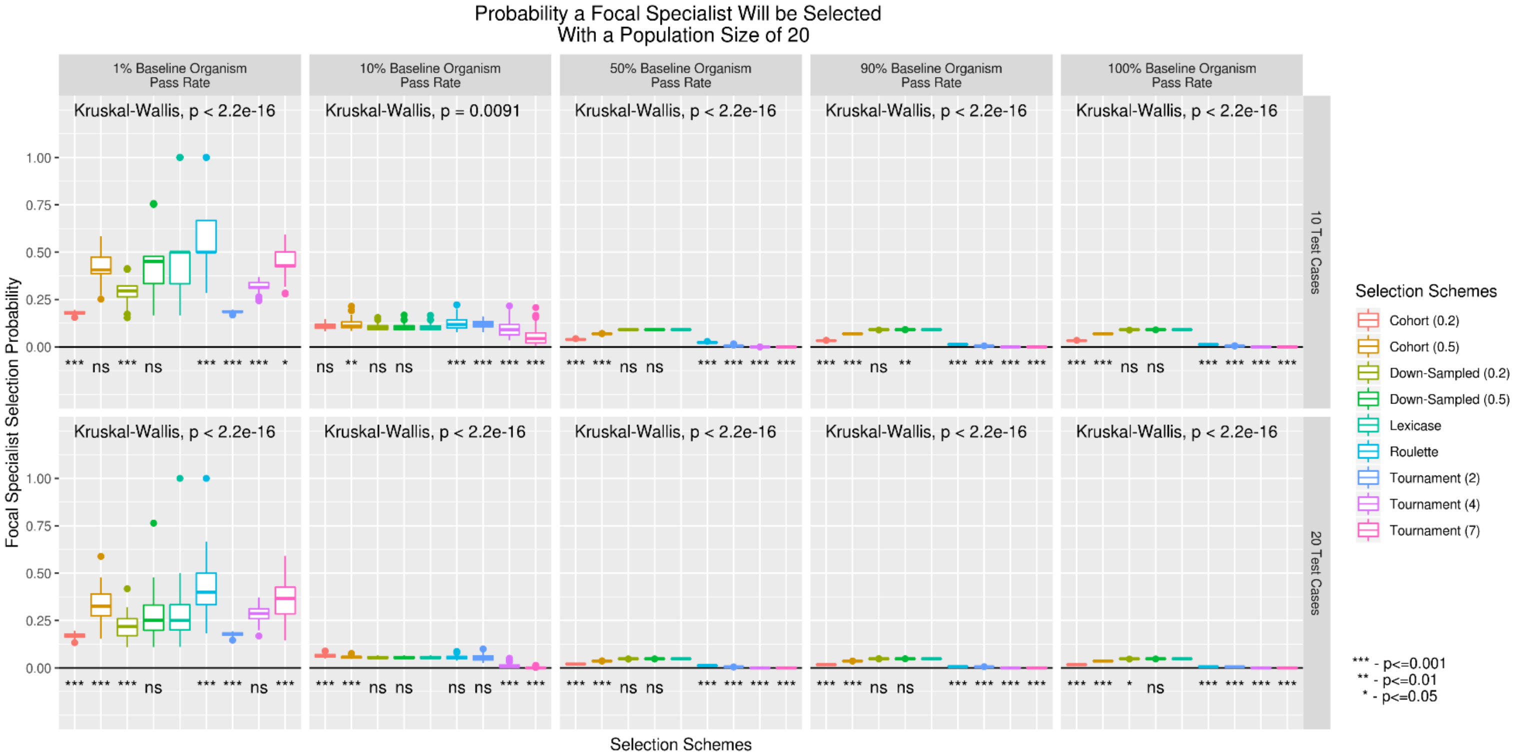

Figure 2. Shows the spread of focal specialist probability for a population size of 20. Kruskal-Wallis shown for Lexicase variants. Asterisks denote Mann-Whitney tests against Lexicase.

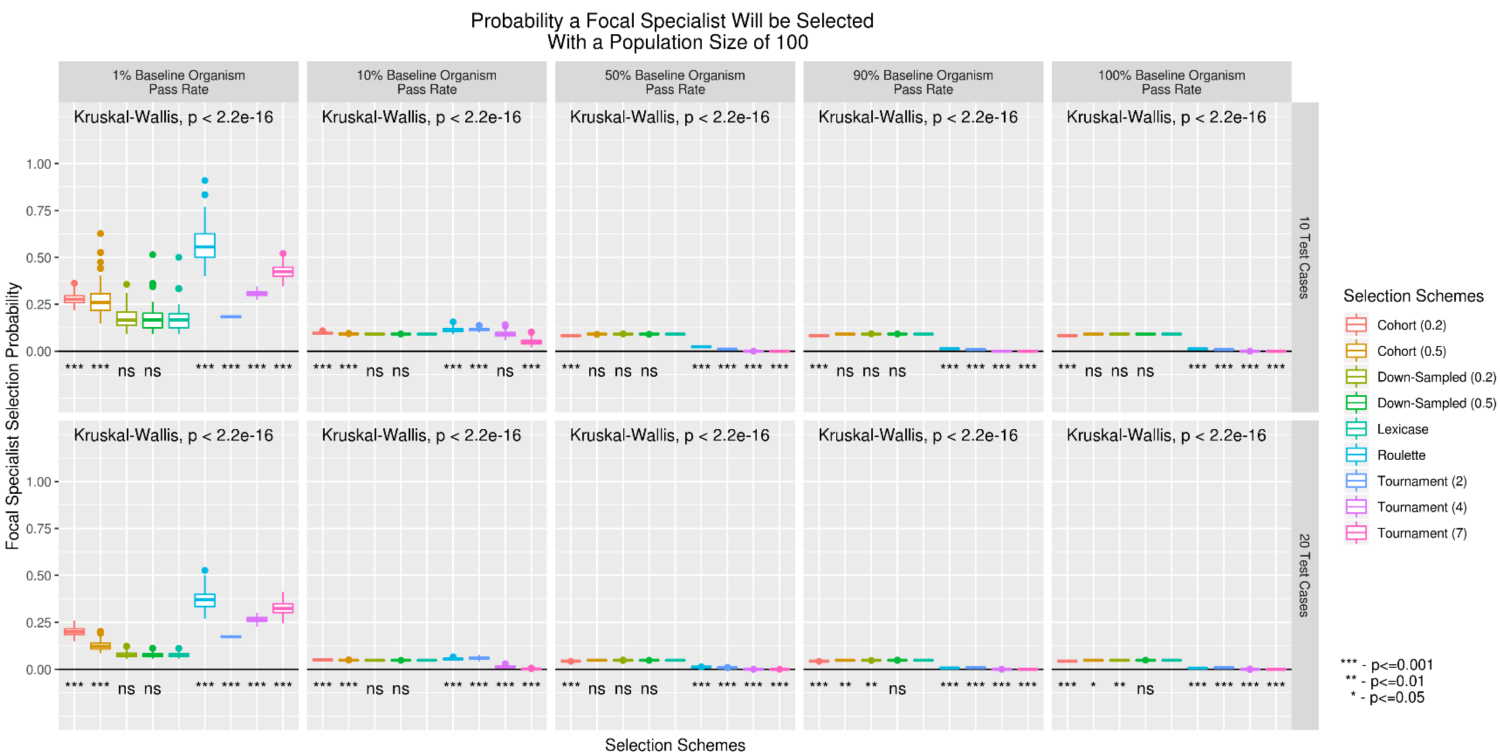

Figure 3. Shows the spread of focal specialist probability for a population size of 100. Kruskal-Wallis shown for Lexicase variants. Asterisks denote Mann-Whitney tests against Lexicase.

V. Discussion

After performing a Kruskal-Wallis test, a significant difference was discovered between the selection probabilities for specialists between Lexicase and Cohort (20% randomly sub sampled) in 29 of the 30 experimental conditions. In the 29 experimental conditions that had a statistically significant difference between Cohort (20% randomly sub sampled) and Lexicase, the p-value was below 0.001. A significant difference was present between the probabilities for specialist selection between Lexicase and Cohort (50% randomly sub sampled) in 25 of the 30 experimental conditions. In the 25 experimental conditions that had a statistically significant difference present, 22 conditions had a p-value below 0.001. In Down-Sampled Lexicase (20% randomly sub sampled) there was only a significant difference in selection probabilities when being compared to Lexicase for 12 of the 30 experimental conditions. Of the 12 conditions with statistical significance, 8 conditions had a p-value below 0.001. There was a significant difference between the selection probabilities for specialists between Lexicase and Down-Sampled Lexicase (50% randomly sub sampled) in 6 of the 30 experimental conditions. In the 6 conditions containing a statistically significant difference for the selection probabilities for specialists, 3 conditions had a p-value below 0.001.

VI. Conclusions and Future Work

With this information there will be more of a focus on Down-Sampled Lexicase due to its ability to maintain a high selection probability for specialists despite using random sub sampling. This experiment has shed light on how random sub sampling affects Lexicase selection in terms of specialist preservation, and it will lead to a more in-depth investigation into the differences between Lexicase and its variants. Finally, the tool used to determine selection probabilities will be used more in the future experiments dealing with selection probabilities.

VII. References

Goldberg, D. E., & Deb, K. (1991). A Comparative Analysis of Selection Schemes Used in Genetic Algorithms. In Foundations of Genetic Algorithms (Vol. 1, pp. 69–93). https://doi.org/10.1016/B978-0-08-050684-5.50008-2

Helmuth, T., Spector, L., & Matheson, J. (2015). Solving Uncompromising Problems with Lexicase Selection. IEEE Transactions on Evolutionary Computation, 19(5), 630–643. https://doi.org/10.1109/TEVC.2014.2362729

Blickle, T., & Thiele, L. (1996). A Comparison of Selection Schemes Used in Evolutionary Algorithms. Evolutionary Computation, 4(4), 361–394. https://doi.org/10.1162/evco.1996.4.4.361

Miller, B. L., & Goldberg, D. E. (1995). Genetic algorithms, tournament selection, and the effects of noise. *Complex systems*, *9*(3), 193-212.

Helmuth, T., Spector, L., & Matheson, J. (2014). Solving uncompromising problems with lexicase selection. *IEEE Transactions on Evolutionary Computation*, *19*(5), 630-643.

Razali, N. M., & Geraghty, J. (2011, July). Genetic algorithm performance with different selection strategies in solving TSP. In *Proceedings of the world congress on engineering* (Vol. 2, No. 1, pp. 1-6). Hong Kong: International Association of Engineers.

Hernandez, J. G., Lalejini, A., Dolson, E., & Ofria, C. (2019, July). Random subsampling improves performance in lexicase selection. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (pp. 2028-2031). ACM.