Exploring the Effects of Adding Cohorts Within Lexicase Selection

Jose Guadalupe Hernandez, Alexander Lalejini, Dr. Charles Ofria
Michigan State University, East Lansing, Michigan

Motivation

Lexicase selection has shown the ability to perform well on modal problems. It does so by evaluating organisms on individual test cases. It uses these test cases as a filter to find the best performing organism/organisms per individual test case. As the amount of test cases and population size increases, the number of evaluations required increases. Hence, looking into how to reduce these evaluations would be beneficial.

Research Question

What are the effects on fitness and solution counts from adding cohorts within lexicase selection?

Experiment Setup

- Population size: 1000.
- Number of generations: 10,000.
- Used a linear genetic programming representation.
- Problems:
  1. Number IO
  2. Median
  3. Sum of Squares
- Mutations:
  - Insertions, .009
  - Deletions, .009
  - Substitutions, .009
- Number of Cohorts, Cohort Size:
  - (20,50)
  - (10,100)
  - (5, 200)
  - (4, 250)
  - (2, 500)
  - (1, 1000)

Methodology

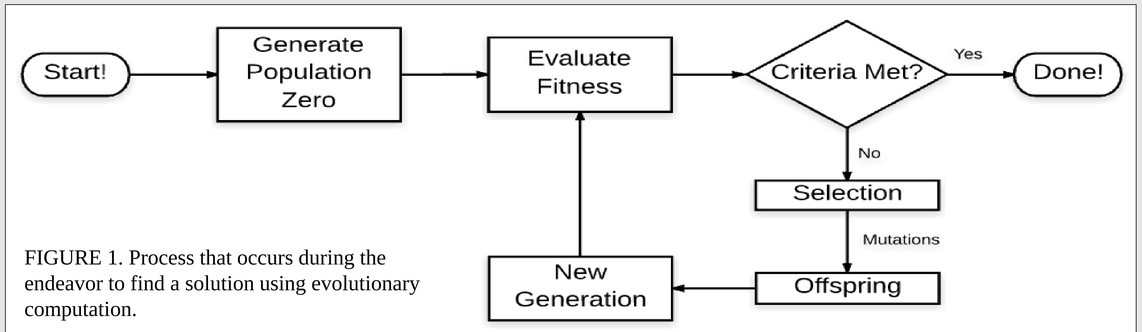


FIGURE 1. Process that occurs during the endeavor to find a solution using evolutionary computation.

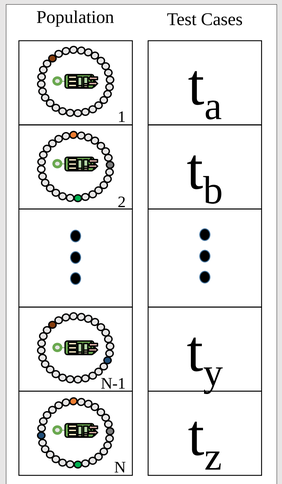
Lexicase Selection

To select a parent for use in a genetic operation:

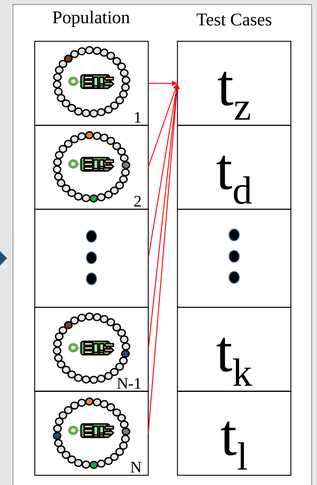
1) `Initialize`:
   a) Set `candidates` to be the entire population.
   b) Set `cases` to be a list of all of the test cases in random order.
2) `Loop`:
   a) Set `candidates` to be the subset of the current `candidates` that have exactly the best performance of any individual currently in `candidates` for the first case in `cases`.
   b) If `candidates` contains just a single individual then return it.
   c) If `cases` contains just a single test case then return a randomly selected individual from `candidates`.
   d) Otherwise remove the first case from `cases` and go to `Loop`.

FIGURE 2. Lexicase selection algorithm [1].

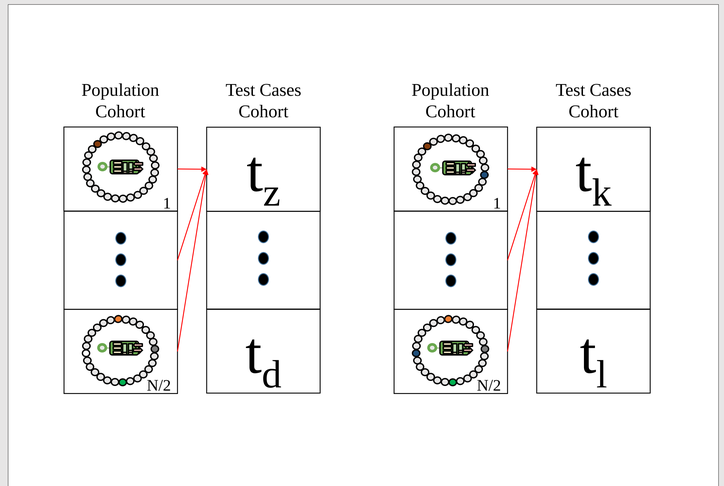
Lexicase & Cohort Lexicase Example



Initial population and test cases.



Lexicase selection example.

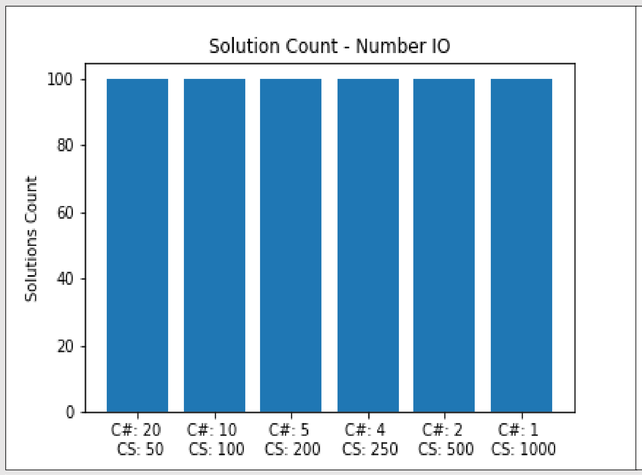


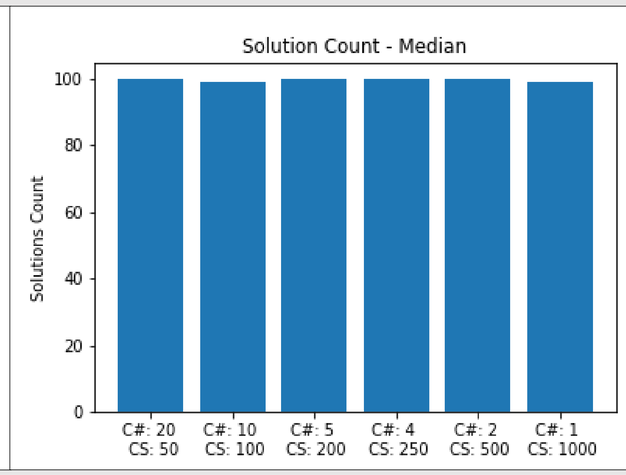
Cohort lexicase selection example.

Conclusion

- Average population fitness and solution count not heavily impacted by integrating cohorts within lexicase selection.
- Future work:
  - Increase amount of trials
  - Implement additional problems to solve
  - Track additional metrics such as diversity
  - Compare against other selection methods

Results





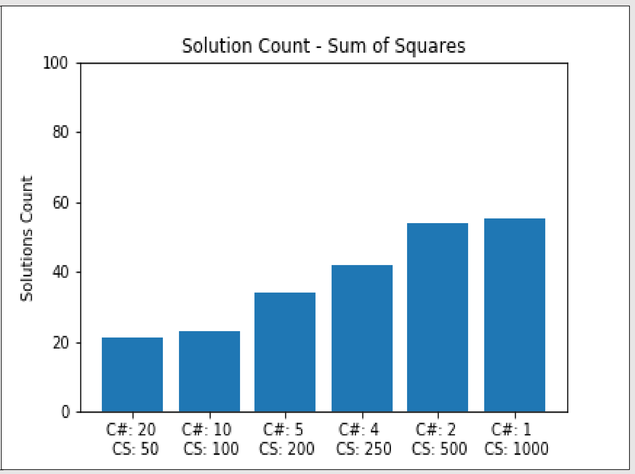


FIGURE 3. Solution counts for each problem. C# represents the number of cohorts and CS represents the cohort size.

References:
1. T. Helmuth, L. Spector, and J. Matheson. “Solving Uncompromising Problems With Lexicase Selection”. In: IEEE Transactions on Evolutionary Computation 19.5 (Oct. 2015), pp. 630– 643. issn: 1089-778X. doi: 10.1109/TEVC.2014.2362729.
2. Thomas Helmuth, Nicholas Mcphee, and Lee Spector. “The Impact of Hyperselection on Lexicase Selection”. In: (July 2016), pp. 717–724.
3. Lee Spector. “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. GECCO ’12. Philadelphia, Pennsylvania, USA: ACM, 2012, pp. 401–408. isbn: 978-1-4503-1178-6. doi: 10.1145/ 2330784.2330846. url: http://doi.acm.org/10.1145/2330784.2330846.

Acknowledgments: I would like to thank Alexander Lalejini, Dr. Charles Ofria, TRIO, SROP, College Assistant Migrant Program, and BEACON.