RL at Harvard

Encouraging Diversity in Reinforcement Learning with Lexicase Selection

Ryan Boldi[1], Charles Zhang[2] and Lee Spector[3,1]

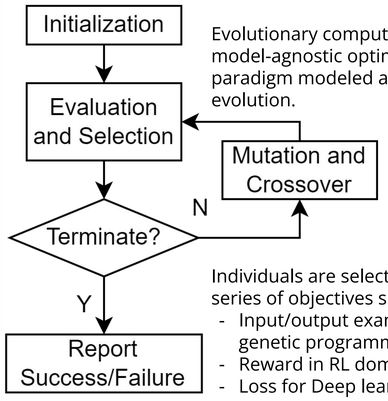
[1]University of Massachusetts Amherst, [2]Independent Researcher, [3]Amherst College

PUSH*lab
programs under selection + heredity

University of Massachusetts Amherst

Amherst College

Evolutionary Computation

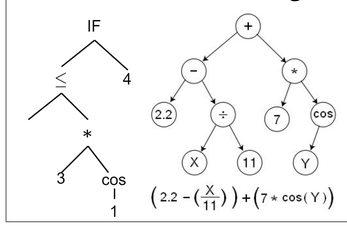


Evolutionary computation is a model-agnostic optimization paradigm modeled after biological evolution.

Individuals are selected based on a series of objectives such as:
- Input/output examples for genetic programming
- Reward in RL domains
- Loss for Deep learning

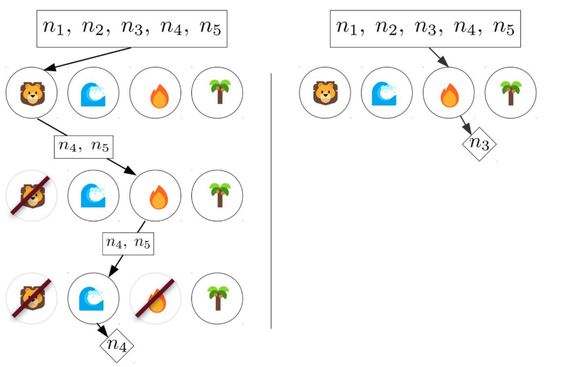
Lexicase selection is a selection strategy that has been shown to be efficacious at solving problems in the evolutionary computation literature, at least partially due to diversity maintenance. Here, we propose two use cases to potentially transfer these benefits to the reinforcement learning domain.

Genetic Programming



Genetic programming is a subset of evolutionary computation that deals with evolving symbolic expressions, such as computer programs.

Lexicase Selection



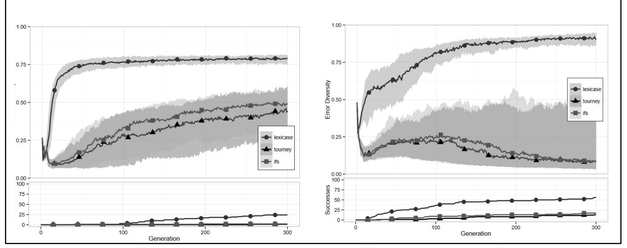
Lexicase Selection is a parent selection algorithm for evolutionary computation that selects individuals based on a random ordering of test cases.

Selecting based on specific test cases makes individuals that are good at unique subsets of tasks are more likely to survive. This has a clear evolutionary interpretation: Individuals that pass the random situations that occur during their lifetime survive and reproduce.

Performance on Program Synthesis Problems (# of successes)

Problem name	Lexicase	Tournament	IFS
Replace Space With Newline	57	13	17
Syllables	24	1	2
String Lengths Backwards	75	18	12
Negative To Zero	72	15	9
Double Letters	5	0	0
Scrabble Score	0	0	0
Checksum	0	0	0
Count Odds	4	0	0

Effect on Diversity



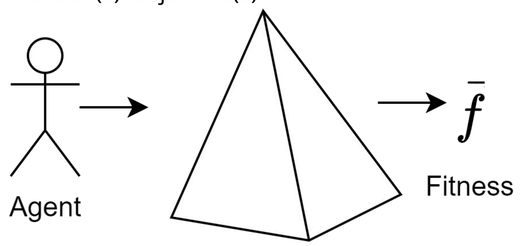
Lexicase Selection has been shown to significantly improve the success rate of genetic programming runs such as program synthesis. Relatedly, it has also been shown to significantly improve the diversity of populations over the course of evolution.

Lexicase Selection in Reinforcement Learning

Lexicase selection can be incorporated into RL on different levels. We outline two cases to encourage both action diversity and agent diversity

Agent Diversity

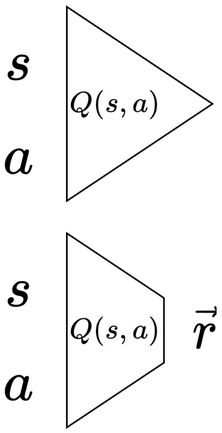
We can more generally apply evolutionary computation to RL training and use lexicase selection to select agents from a population based on their performance on the task(s)/objective(s).



This would help facilitate a diverse set of agents that accomplish the task(s) in different ways.

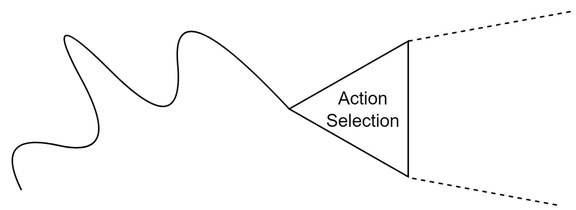
Importantly, lexicase selection lends itself to multi-agent environments, where maintaining a population does not add any additional computational overhead.

Action Diversity



Lexicase selection can be used to select the actions to take at each time step for a given agent. This would help facilitate exploration in state action space by allowing the agent to take more diverse actions

A simple example: Using the penultimate layer of a DQN Q function gives a vector-valued representation of the reward. Lexicase can be used on this vector to select the action that optimizes a particular value in the vector



References

Boldi, R, & Spector, L. *Can the Problem-Solving Benefits of Quality Diversity Be Obtained Without Explicit Diversity Maintenance?. arXiv preprint arXiv:2305.07767* (2023).

Helmuth, T., McPhee, N. F., & Spector, L. (2016). Lexicase Selection for Program Synthesis: A Diversity Analysis. In R. Riolo, W. P. Worzel, M. Kotanchek, & A. Kordon (Eds.), *Genetic Programming Theory and Practice XIII* (pp. 151–167). Springer International Publishing. https://doi.org/10.1007/978-3-319-34223-8_9

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

Wärnberg, E., & Kumar, A. (2023). Feasibility of dopamine as a vector-valued feedback signal in the basal ganglia. *Proceedings of the National Academy of Sciences*, *120*(32), e2221994120.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Spector, L. (2012). *Assessment of Problem Modality by Differential Performance of Lexicase Selection in Genetic Programming: A Preliminary Report*.

Spector, L., Ding, L., & Boldi, R. (2023). Particularity. *arXiv preprint arXiv:2306.06812*.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

This work was performed in part using high performance computing equipment obtained under National Science Foundation Grant No. 2117377. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.