Discovering Lunar Lander Behaviors with Quality Diversity Algorithms

Melissa Lorenzo, mendezlorenzo.melissa@gmail.com
ICAROS Lab
John H. Francis Polytechnic Senior High School, Class of 2023
USC Viterbi Department of Computer Science, SHINE 2022

USC Viterbi School of Engineering

SHINE Summer High School Intensive in Next-Generation Engineering

USC-Meta Center for Research and Education in AI and Learning

Introduction

Quality Diversity (QD) finds a diverse set of high performing solutions. In this project we apply QD to reinforcement learning (RL) so the agent (lunar lander) acts in an environment and maximizes the reward it gets (Figure 1). QD shows that there is not always a single solution for a problem but various unique solutions that all meet the same goal. We use the OpenAI Gym lunar lander example to give a visual representation of the different solutions.

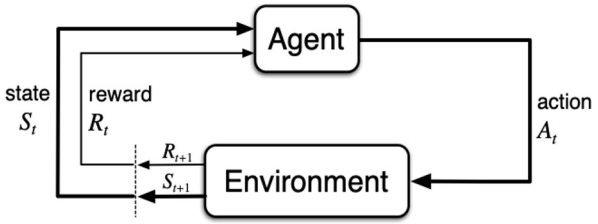
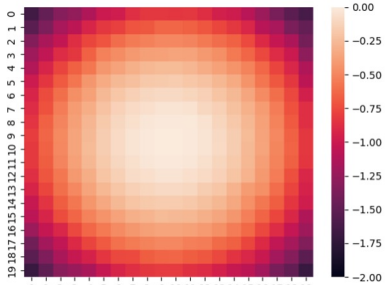


Figure 1.This chart shows the agent-environment loop. The agent, the learner, in this case is the lunar lander and behavior characteristics (BCs) change based on the agent's interactions with the environment. As the agent does the action the environment changes then it tells the agent its new state and it gets a reward, then this loop continues until the agent maximizes its reward.

Method

Multi-dimensional Archive of Phenotypic Elites (MAP-Elites) is an algorithm that searches for high performing solutions. What MAP-Elites does is it mutates a solution to find other solutions. Then it maps the solution on the heat map. It continues to do so until it fills up the heat map as much as it can. The map helps show the performance of the solution. The key on the side of the map is color coded to help understand which are the best solutions, with the lightest color being the best and the darker color the worst. This helps with QD because it finds various solutions with different behavior characteristics. In this specific project we use CMA-ME which is like MAP-Elites but runs faster than MAP-Elites making it more efficient to use.

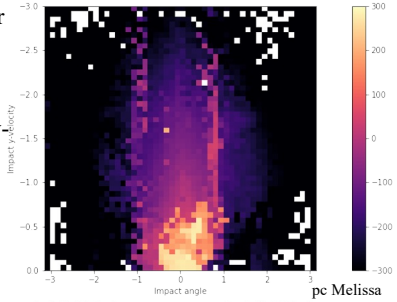
Work



pc Melissa

My mentor helped me first implement MAP-Elites to the Sphere function for optimization which created this heat map. This algorithm helped me get a better understanding as to how MAP-Elites works and how to use numpy, a Python library.

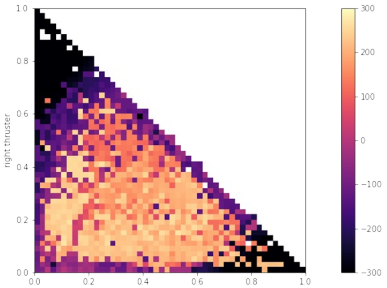
In this version of the lunar lander the behavior characteristics are the impact angle (x-axis) and the impact velocity (y-axis). The landers behaviors vary from floating above the target zone before landing (Figure 2) to swinging down towards its target zone (Figure 3).



pc Melissa

Figure 2. (-0.2,-0.1)

Figure 3. (0.8,-.15)



pc Melissa

Figure 4. (0.2,0.5)

Figure 5. (0.4,0.35)

In this version of the lunar lander the behavior characteristics are the number of left (x-axis) and right (y-axis) engine activations. Although the BCs for this lander aren't as different as the impact angle and impact velocity one there is still one that hovers longer over the target zone before landing (Figure 4) while the other swings down towards the target zone.

How This Relates to Your STEM Coursework

The SHINE experience has not only showed me what research really is but has helped me further enhance my programming skills. Personally, I want to go into computer science and this program helped me learn things about programming that I would not have found on my own right away.

Advice for Future SHINE Students

For future SHINE students I would suggest not underestimating the amount of time we have in this program. Yes, it is 7 weeks, but those 7 weeks go by quickly, so I suggest not second guessing asking for help when you're stuck. Our mentors are here to help us so if you're stuck don't be afraid to ask. But most importantly enjoy your time in SHINE.

Acknowledgement

I would like to thank Professor Nikolaidis for allowing me to work in the ICAROS Lab this summer and my PhD mentor, Bryon Tjanaka, for guiding me through my work. I would also like to thank Dr. Mills for starting SHINE and allowing us all do research this summer, and the USC-Meta Center for Research and Education in AI and Learning for the scholarship which allowed me to join the program.

References

- B.Tjanaka, S.Sommerer, N.Klapsis, M.C. Fontaine & S.Nikolaidis (2021). "Using CMA-ME to Land a Lunar Lander Like a Space Shuttle" https://docs.pyribs.org/en/stable/tutorials/lunar_lander.html
- J.Mouret & J.Clune (2015). "Illuminating search spaces by mapping elites" https://arxiv.org/pdf/1504.04909.pdf
- B.Tjanaka, J.Togelius, M.C. Fontaine & S.Nikolaidis "Approximating Gradients for Differentiable Quality Diversity in Reinforcement Learning" https://dqd-rl.github.io/
- R.S. Sutton and A.G. Barto "Reinforcement Learning An Introduction"
- M.C. Fontaine, J.Togelius, S.Nikolaidis & A.K.Hoover (2020). "Covariance matrix adaptation for the rapid illumination of behavior space" https://dl.acm.org/doi/10.1145/3377930.3390232