Evolutionary Behavioral Dynamics in Ecological Grid-Worlds

Richard W. Dionne

Department of Mathematics, Dartmouth College

Abstract

This project combines a grid-world approach—using discretized models of decision problems to improve clarity and tractibility—with replicator dynamics in an ecological simulation to study the behavioral dynamics of simple ecological agents. A small Python program allows for simulations of large numbers of competing agents which approximate standard fixed-population replicator dynamics on large time scales. Though shortcomings in parameter tuning, experimental design, and agent complexity limit the current effectiveness of the program as an evolutionary simulator, its display of concrete dynamics and capacity for extension make it a useful learning tool nonetheless.

Introduction

Background

Grid-worlds are a common approach in Computer Science and Artificial Intelligence research which use simplifying assumptions to tractably simulate detailed dynamics and observe the interactions and learning processees of digital agents under highly controlled conditions. In AI Safety, there has been recent work to benchmark known machine learning systems using grid-worlds designed to simulate the sorts of safety tests engineers might perform on a real autonomous device.[1] This project expands on prior research by developing a grid-world environment with ecological dynamics to simulate the decision space of agents competing for scarce natural resources.

Motivation

Ecological models compound the complexity of standard grid-worlds by introducing multiple competing agents and studying population dynamics over multiple agent durations rather than focusing on a single defined task. Thus, machine learning techniques can be effectively benchmarked according to their abilities to develop strategies for maintaining high populations and resisting invasion by competing strategies. By applying evolutionary game theory to these models, we can better predict how strategies may emerge by examining the dynamics of the game itself, allowing us to continue expanding our understanding of these complex decision spaces.

Methods

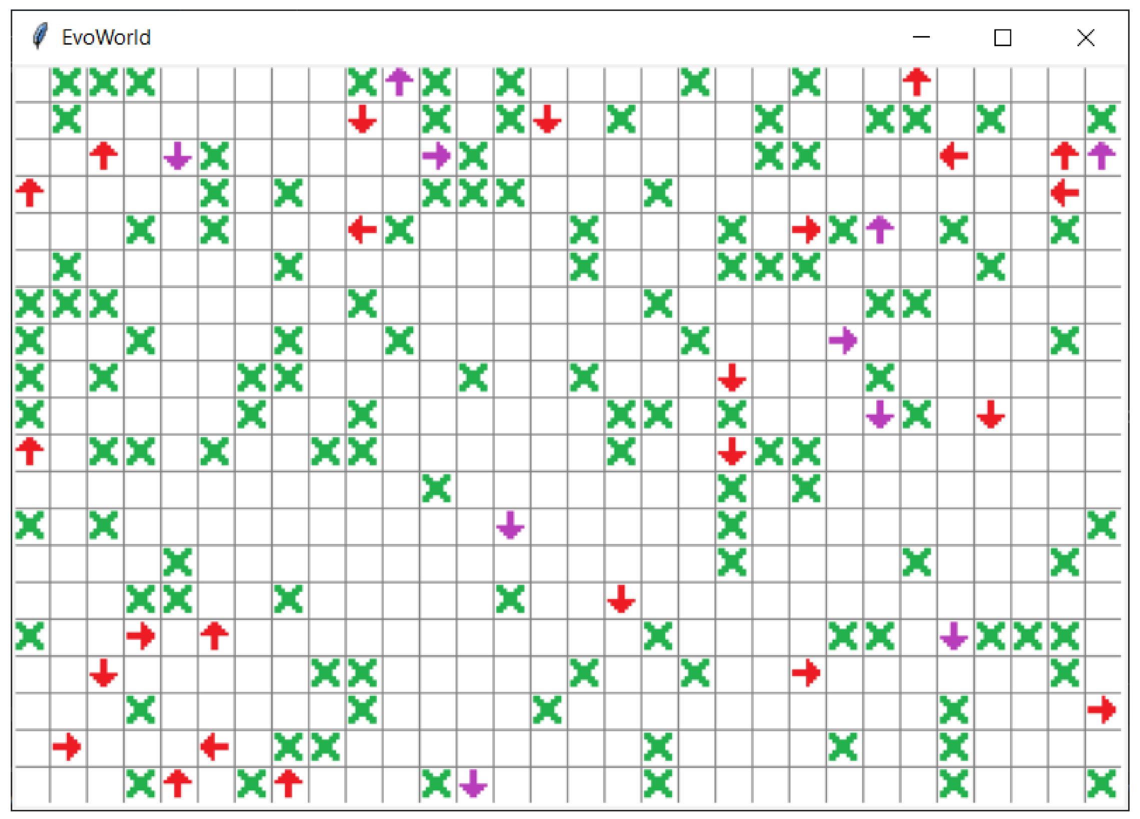

Figure 1: An example iteration at $t = 0$. Green crosses are plants, red arrows are main population animals, and purple arrows are mutant strategy animals. Plants grow at random locations over time and eventually wilt if not eaten, while animals can only reproduce by cloning after gathering sufficient calories, and can die by starvation or old age.

My project is implemented in roughly 700 lines of python split into three files (model logic: 500 lines, graphical display: 100 lines, testing apparatus: 100 lines). The eponymous grid-world is an $M \times N$ toroidal manifold with taxicab topology embedded in the plane, populated with plants which appear exogenously and provide food for animals, which must collect enough food to reproduce by cloning in order to grow their population. To simulate evolutionary pressures, each iteration of the model features a main population with strategy A and a competing, initially smaller population of mutants with related strategy A'. Population ratios after t time steps can then be used to approximate the probability of fixation of A', which would become the main population in the next iteration if fixation is likely.

Model

Modeling Limitations

Analyzing the direct dynamics of this stochastic discrete spatial game with many agents competing over time is beyond the scope of this project, so I will make a number of simplifying assumptions. First, since plants grow exogenously at a fixed rate and each animal requires a certain amount of food during its life to reproduce, this environment naturally has a maximum carrying capacity which depends on the balance of food demand and plant growth as well as agent strategy. As the size of the field (and therefore the maximum food supply at any given time) increases to arbitrarily large M and N, we can approximate strategic dynamics using fixed population replicator dynamics.

An Approximate Model

Let A represent the resident population strategy, and $B = A'$ represent a competing mutant strategy. Then x_A and x_B are the portions of the population following each respective strategy, and

$$\dot{x}_A = x_A\left[f_A(\vec{x}) - \phi\right]$$
$$\dot{x}_B = x_A\left[f_B(\vec{x}) - \phi\right]$$

Where ϕ is the overall population fitness. From this model, we see that for $B = A'$ to grow into a dominant strategy, it must have some net fitness advantage such that $\dot{x}_B > 0$ for small x_B. Though these dynamics alone are not sufficient to prove that A and B cannot have an equilibrium population balance, for the purpose of exploration let us make the further simplifying assumption that a mutant strategy which is sufficiently successful will eventually fixate in the population.

Consider agents with the following four choices for each action:

{Move Forwards, Turn Left, Turn Right, Do Nothing}

Probabilistic agents choose their action at random with a fixed probability distribution, which with three degrees of freedom lies on or inside a 3-simplex. Heuristically, one might expect the evolutionarily stable strategies to favor forward movement (in order to gain energy), a bias towards one direction to mitigate against turns canceling one another out at the cost of time and energy, and a general bias away from inaction.

Results

Exploring Decision Space

To test these assumptions, I explored the decision space by tracing strategy and fitness development from four key points: the uniform distribution, and heavy biases towards moving, staying still, and turning one direction (by symmetry, the two directions are equivalent).

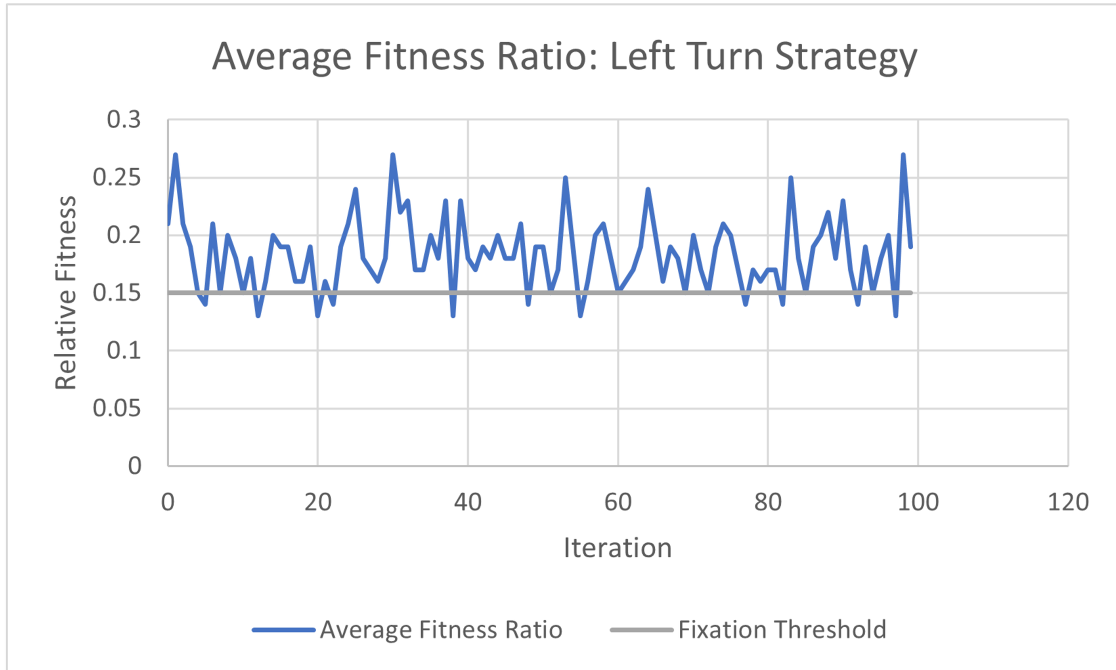

Figure 2: Relative fitness of the **Turn Left** strategy over time, as approximated by the ratio of mutant births to main population births after 1000 time steps, averaged over 10 samples per iteration. Tendency to turn left never dropped below its starting value.

The figure representing the trajectory of the strategy beginning with a heavy bias towards turning in one direction is representative of my finding across starting strategies. Because most mutant strategies were successful enough to replace the residents, in general the path through decision space is a random walk without obvious direction. These results suggest that some aspect of the training protocol is deficient.

Discussion

Model Parameters

One possible confounding factor could be that the underlying model parameters are not properly chosen to incentivize active ecological behavior. Parameters for agent and plant lifespans, energy costs and rewards, and the metabolic rate of agents could all be adjusted and the values for the prior experiment were chosen effectively arbitrarily based on observations during development. Further development of the balance of these parameters could yield steeper evolutionary gradients.

Training Parameters

Beyond the model itself, many aspects of the training protocol could be refined or improved. Larger field sizes and longer simulations would improve measurement of fixation probability, and smaller maximum shifts, which were abandoned during development due to slow convergence, might allow for more precise navigation of the decision space.

Agent Limitations

Finally, it is worth considering that due to the highly limited nature of this class of probabilistic agent, there may not be a sufficient gradient between these strategies to drive the agent toward optimality in the light of all the stochastic error brought upon by small sample sizes. Initial plans for this project included significantly more capable agents with vision and primitive neural network brains, but limited computing resources and development time required a reduction in scope.

Conclusion

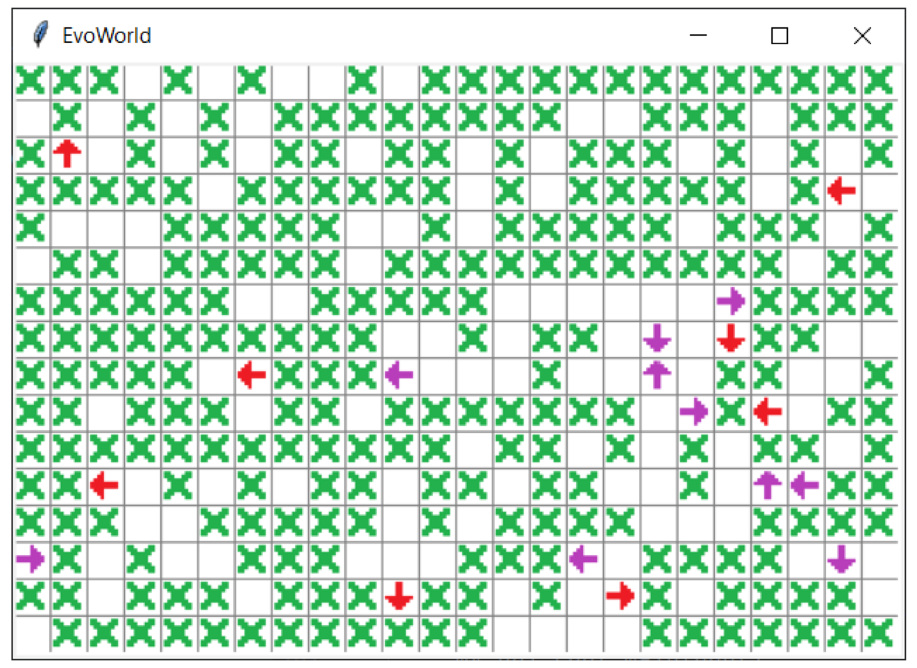

Figure 3: In this example configuration with a higher plant density, the purple strategy, in this case more likely to move forwards, is becoming dominant. This observational use of the program is helpful for intuition.

Uses and Limitations

This project has many obvious limitations, many of which reflect the short time in which it was developed and the limited complexity of the underlying dynamics. The food source is entirely exogenous, and the agents follow very simply defined strategies. On the other hand, the focus on concrete dynamics on a grid allows the user to directly apprehend what the agents are doing and develop a sense for their behavioral patterns. In this way, the program used to perform this research can be a helpful intuition pump, as well as a basis for future explorations.

Next Steps

Time was a major limiting factor for this project, because machine learning is unpredictable. Though agents with proper sensory aparati and simulated brains were not practical to design or train in this limited period, they would be of great interest as future additions. By adding new ways for agents to sense and interact, I hope to observe stronger evidence for evolutionary pressure in this environment and am excited to learn more about how populations are able to thrive.

Acknowledgments

I would like to deeply thank Feng Fu and Mark Lovett for their support and guidance on this project. I would also like to acknowledge the effectiveness of ChatGPT, which saved many hours of scouring the web for code snippets and basic syntax reminders. All written work is my own, as are all errors.

References

[1] Jan Leike, Miljan Martic, Victoria Krakovna, Pedro A. Ortega, Tom Everitt, Andrew Lefrancq, Laurent Orseau, and Shane Legg, *AI safety gridworlds*, CoRR **abs/1711.09883** (2017).