

Learning and Evolution in Simulated World

Cameron Rudd & Avinoam Henig

May 10, 2015

Abstract

Inspired by the work of Ackley and Littman [1], a two dimensional tile world inhabited by self directed creatures was implemented in python and run for variable lengths of time. The behavior of the creatures was then analyzed in order to determine whether learned strategies and evolution caused older, larger populations, and speciation. Various control factors were altered between runs in an attempt to encourage certain behaviors. Each creature contained two neural networks, an evaluation network and an action network. The action network determined which of several actions the creature would make on a given turn. This network was improved each time step using complementary reinforced back propagation. This was implemented using the evaluation network to determine whether the creature had made a good or bad decision. The creatures had the ability to self reproduce, reproduce with others, and die.

Background

In their 1991 paper *Interactions Between Learning and Evolution* [1] Ackley and Littman investigated whether a learning algorithm could learn “given only natural selection as feedback.” In the paper, Ackley and Littman outline evolutionary reinforcement learning (*ERL*), which combines traditional neural network learning with genetic evolutionary approaches. The approach uses an artificial ecosystem in which *ERL* driven agents interact with the ecosystem, trying to prolong their “life” and “survive”.

In Ackley and Littman’s world, called *AL*, uses a 100×100 two dimensional array of cells inhabited by various landscape features, adaptive lifeforms and non adaptive lifeforms. Different objects in the world have varying affects on the *ERL* agents. For example, a carnivore eat action hurts an *ERL* agent while an *ERL* agent eat action on a plant increases the health of the agent.

The *ERL* agents are able to “see” the closest world object that is within four cells of itself. The agents die when their health and energy is low and reproduce when it is high. Reproduction can be done solely by an agent, whereby it mutates its own genome, or with another nearby agent, whereby the two agent’s genomes are crossed over and then mutated.

Ackley and Littman’s work was primarily a proof of concept; they demonstrated that learning can occur when given only life and death as feedback. Furthermore, Ackley and Littman found that in their world, the *ERL* approach was found to more successful then either a pure evolutionary approach or a pure learning approach. Interestingly, however, Ackley and Littman found that when compared to agents that behaved randomly, the *ERL* agents only outperformed the random agents after 500,000 time steps.

Methods

We implemented our artificial ecosystem in python. The world, which wrapped around, contained ten thousand tiles arranged in a 100×100 two-dimensional array. Each tile could contain a tree, a plant, or a creature. A creature moving into a tile containing a plant resulted in the creature eating the plant and the creature’s health increasing. Attempting to move into a cell containing a tree negatively impacted the creature’s health. Were a creature to attempt to move into a tile occupied by another creature, a probabilistic “success” test was called. This test used the health values of the two creatures to determine whether the

move into the occupied tile was successful. When the move resulted in success, the moving creature consumed the other creature and gained health. When the move failed, the moving creature’s health decreased.

Creatures were equipped with evaluation and action neural networks. The action network took sensory input and output an action. The evaluation network sensory input from after the creature’s action. It then output a scalar corresponding to whether the action improved or worsened the state of the creature. Both networks took thirteen inputs, twelve corresponding to sensory input (sight); the thirteenth corresponded to the health of the creature. The networks were built using the CONX module. The action network was trained over the life of the creature using the complimentary back propagation algorithm described by Ackley and Littman in [1].

Each creature in the world possessed a genome containing the weights for its action and evaluation networks. Initially, like Ackley and Littman, the genome was fixed for the life of the creature. However later on we departed from Ackley and Littman and tested an epigenetic feature that allowed creatures to pass on the learned weights of their action network. This was done by updating the genome of the creature at reproduction.

A graphical representation of the world was built using Pyglet and a database containing all creature data was built using MondoDB. Additionally, a command line interface was developed using python’s multiprocessing module.

Results

Conclusions

Documentation

Bibliography

- [1] David Ackley and Michael Littman. Interactions between learning and evolution. *Artificial life II*, 10:487–509, 1991.