QWOP Bot Progress Report

Written by Matthew Oros, Sonny Smith, and Michael Terekhov
Baldwin Wallace University
275 Eastland Road
Berea, Ohio 44017

Abstract

Abstract Here

Content Here

Problem Description

QWOP is a simple 2-dimensional game with the goal being to make the main character run to the right as far as possible without falling. The challenge of the game is its physics simulation. The character is controll using four keyboard keys: Q, W, O, and P. Q and W control the thighs [TODO: Check if this is the right way around] and O and P control the knees by either rotating the joints one way or the other. As soon as an upper-body limb contacts the flat ground, the game is over and is restart. The score of the game is based on the distance traveled by the character.

The game is intentionally difficult as the physics simulation makes movements hard to predict and require a quick reaction time as well as understanding of how the combinations of key pressed translate to a desired movement on screen.

Proposed Implementation

The proposal of the simulation is an agent that is able to learn on its own to control the character to play the game of QWOP. Two main implementations that will be considered and attempted will be both based on neural networks [TODO: Make sure we are still using neural networks for Deep Q]. The first method is a genetic algorithm approach to learning while the second approach will be a deep Q method of learning.

Consideration of Backpropagation

A traditional method training a neural network is to use a method called backpropagation. Backpropagation is done by providing the network with input data and labelled output data where the term labelled refers to the desired output of the given input. The input is fed forward through the network in its current state and its output is compared to the labelled output. The weights are then adjusted from back to front based on how incorrect the network's output is.

The issue with this approach for this domain of problem is that there is not a viable method of obtaining such labelled data in this problem. The game of QWOP is difficult for humans and such the goal of the simulation is for the learning process to generate a solution from its fitness feedback and environment rather than by example from human playthrough. Thus the two methods that were chosen incorporate some form of random exploration of the state space to converge on a desired solution.