CS 238
Dellarontay Readus - dreadus@stanford.edu
Final Project Status Update

November 16, 2018

UPDATE

After finding that the original custom pip package for the super mario openai gym environment was too deprecated to use, I instead found another mario environment based on gym. One key difference between these two environments is that the newer mario environment doesn't give an easily managed state object to work with. Instead I'm given a 240 x 256 2-D array at every time step in the environment that will enable me to see the RGB pixel representation of the game. This state configuration is cumbersome in comparison; however, I will work through it by manipulating the raw rgb pixel data as is and computing the state space into a 16 x 13 grid representing the current screen of the game with the possible objects defined as {0: no object, 1:object, 2: enemy, 3:Mario}

For this update I initially planned to submit results from a simple Q-Learning algorithm trained to get Mario through the World-1-1 level. However issues with grappling with the state have caused me to experiment with the notion of other types of states (i.e. a 5x5 grid closest to mario or have the state be the time with a max of 500 seconds) and finally decide on using a grid of possible objects to define the state. My intuition is that working from the 16 x 3 grid of discretized state objects will enable me to perform better than if I were to search for specific RGB colors or patterns within the environment at every time step.

UPDATED TIMELINE

Now that I have a firm grasp on the state space that I can use, I can start making progress in actually attempting to train the dataset with different levels of bias and different types of algorithms like Q-Learning or SARSA. Currently I'm on track to enable some decision agent trained with Q-Learning to finish World-1-1 of Super Mario Bros by November 23rd. A more interesting algorithm and field of study I also want to implement for this assignment is NEAT(Neuroevolution of Advancing Topologies). I believe that it will perform better than a simple Q-Learning agent while requiring more time to create its underlying structure. I'll be working on it in parallel with Q-Learning and SARSA, and it should be completed by November 29th.