Laith Asabeh

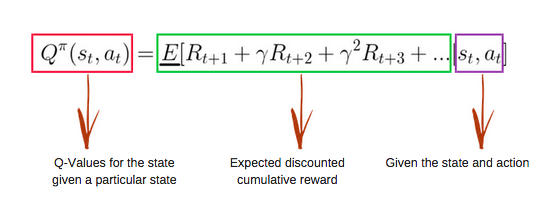
Yaman Salman

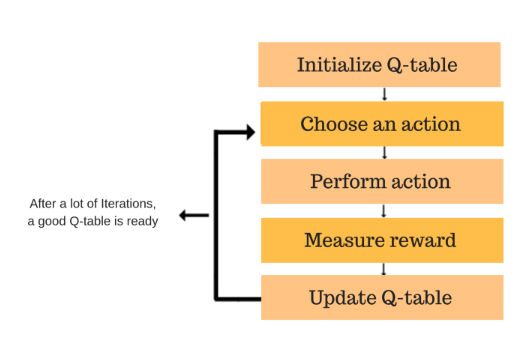
**Evolutionary Algorithms – (Progress Report II)**

**Q-Learning Algorithm**

Implementation of Q-Learning Algorithm was done onto our Tic Tac Toe game.

The algorithm was studied carefully and some values were tweaked and fixed inside the code so we could understand how it works. The Q-function uses the Bellman equation and takes two inputs: state (s) and action (a).



Using the above function, we get the values of Q for the cells in the table. When we start, all the values in the Q-table are zeros. There is an iterative process of updating the values. As we start to explore the environment, the Q-function gives us better and better approximations by continuously updating the Q-values in the table. After every move, the above equation is re-evaluated and the Q-Table is updated based on the newest reward.

**Expected Results**

After implementing the Q-Learning algorithm with almost every variable tweaked to perform the best onto the game, the simulation is run 20,000 times. Each time the equation is re-evaluated and the Q-Table is updated. This means after 20,000 runs, the equation will reach an almost perfect state of decisions, which is impossible to beat.