Name: Shaun Mathew

Assignment #1

**Question 2:**

We found that the inaction learning automaton (i.e. the one where beta is set to 0), creates quite polarizing action probabilities, meaning, a certain action will receive an incredibly high probability when compared to other actions. We found that having beta set to 0.1 (i.e. have a penalty for non-optimal actions), gave fewer polarizing probabilities. Here is a comparison of the results for beta = 0, alpha=0.4 and beta=0.1, alpha=0.4. We see that setting beta=0.1 improved the performance of the algorithm a lot and it performs better than the UCB algorithm, but it requires more tweaking.

Epsilon Probabilities: [0.5330997535581208, 0.9923466319781956, 0.3611753581835483, 0.14759482289331916, 0.445846164641964, 0.8429675948489133, 0.9088036350562749, 0.7461280795017567, 0.8041783971970722, 0.9677720230936874]

Optimal Action: 1

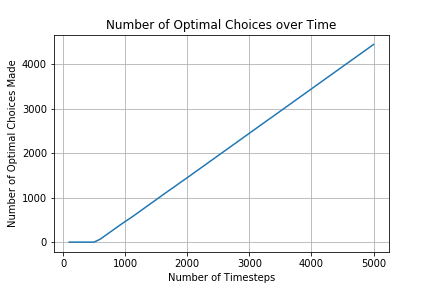
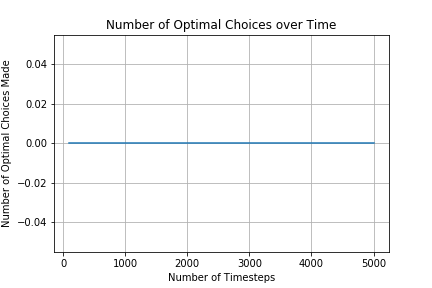
*Polarizing Probabilities with beta 0.0:*

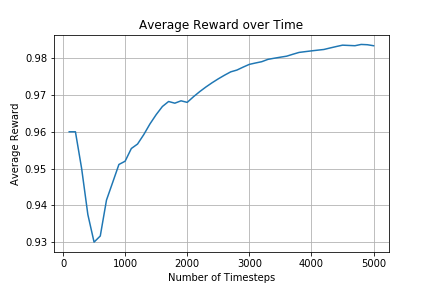
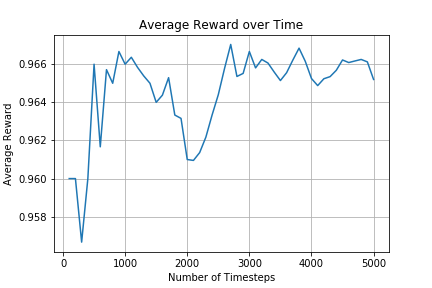
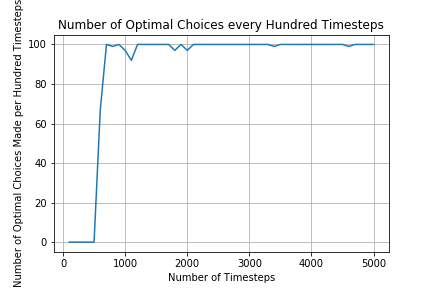
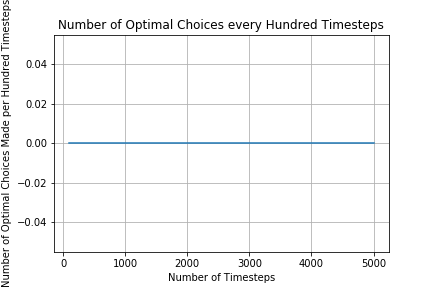
P\_t: [5e-324, 5e-324, 5e-324, 5e-324, 5e-324, 5e-324, 5e-324, 5e-324, 5e-324, 0.9999999999999999]

*Less Polarizing Probabilities with beta 0.1:*

P\_t: [2.5378293341678792e-11, 0.9999999997715954, 2.5378293341678792e-11, 2.5378293341678792e-11, 2.5378293341678792e-11, 2.5378293341678792e-11, 2.5378293341678792e-11, 2.5378293341678792e-11, 2.5378293341678792e-11, 2.5378293341678792e-11]

Note the first graph in every pair is the learning automata with beta 0 and the second is with beta 0.1

****

****