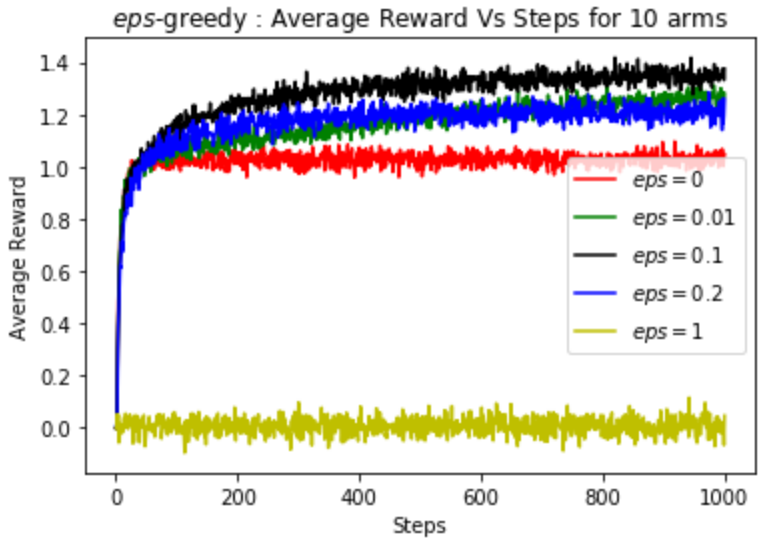
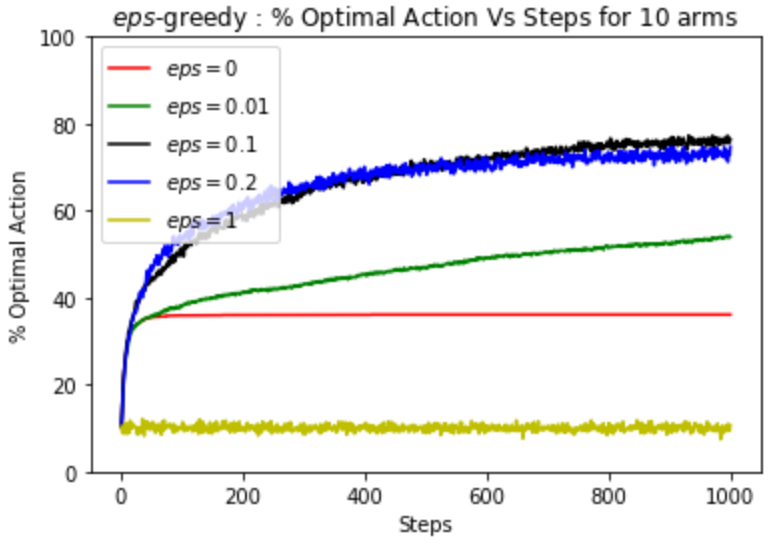
RL Assignment 1

Multi Arm Bandits:

Eps-Greedy Algorithm:

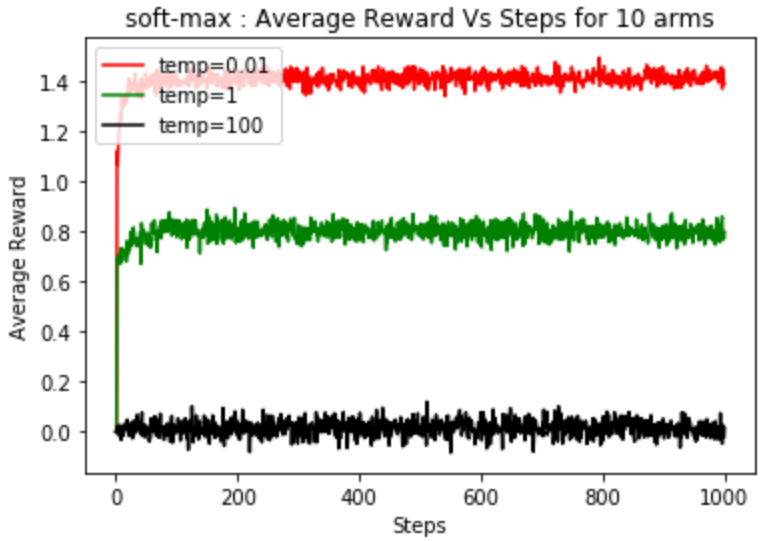


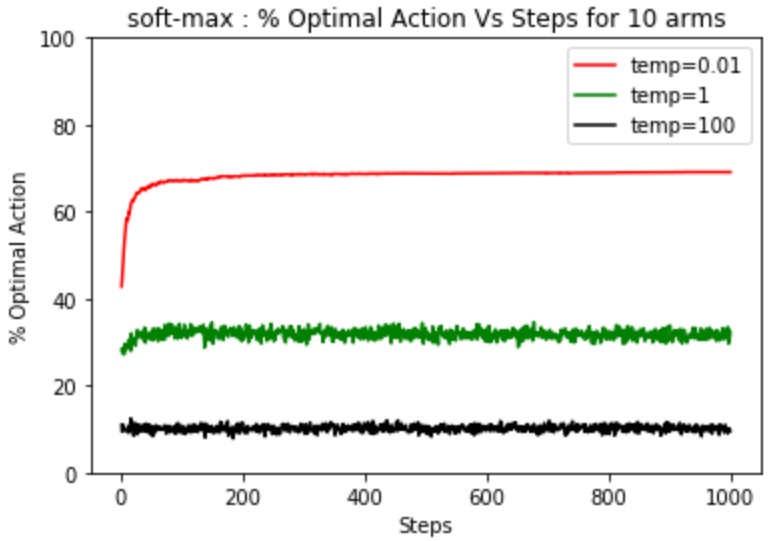


Inference

1. Best results are obtained for eps = 0.1 for both the optimal action and avg. reward.
2. Poor rewards is for eps = 1 i.e. every pull is random
3. Complexity is O (total pulls * K) K = number of arms

Soft-Max Algorithm:

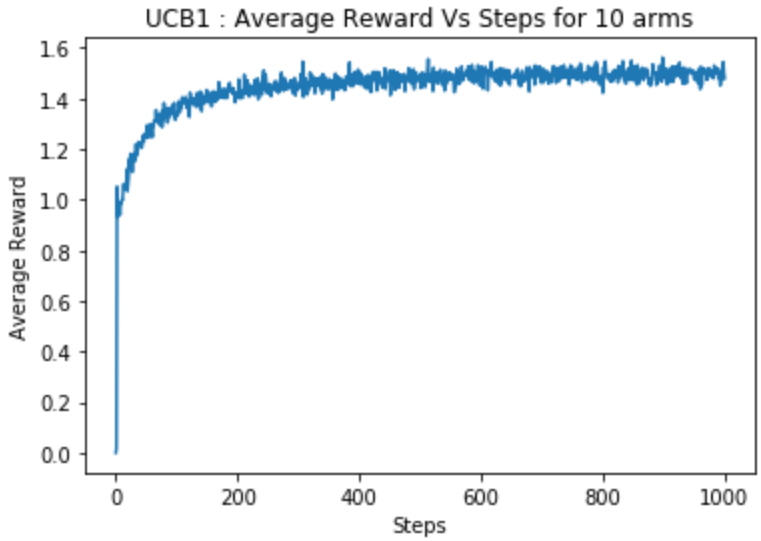


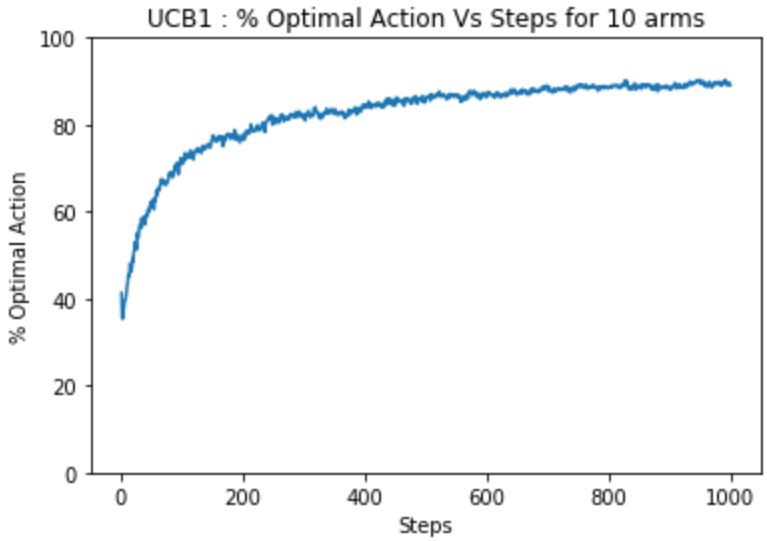


Inference :

1. Best results are obtained at temperature = 0.01
2. This temperature (beta) magnifies the term exp(Q(a)/(beta)) and thereby gives good prob. for arms with high reward.
3. Opposite is true for temp = 100
4. Avg reward converges quickly owing to assigning probabilities for each arm instead of selecting argmax one
5. Complexity is same O(num pulls * K)

UCB1 Algorithm:

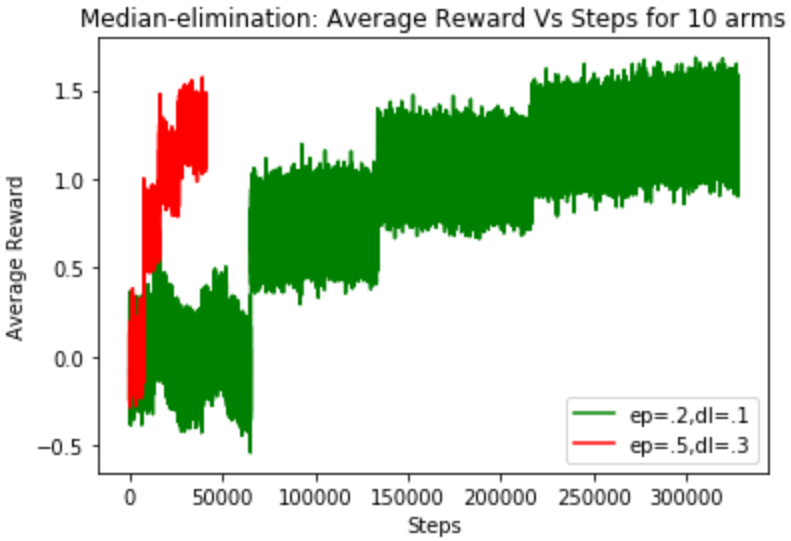


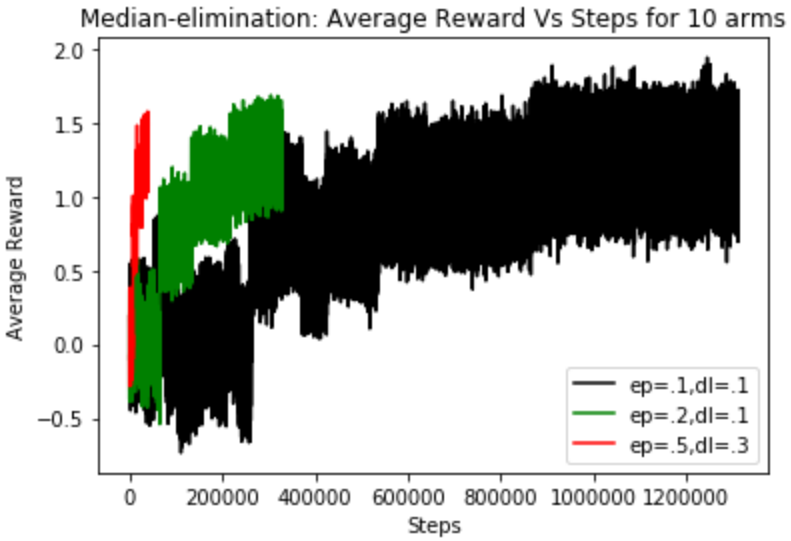


Inference:

1. This algorithm is computationally easy although complexity is same as above algorithms
2. Some complex forms of this include exploration parameter (c).
3. Complexity is same O(num pulls * K)

Median- Elimination Algorithm :

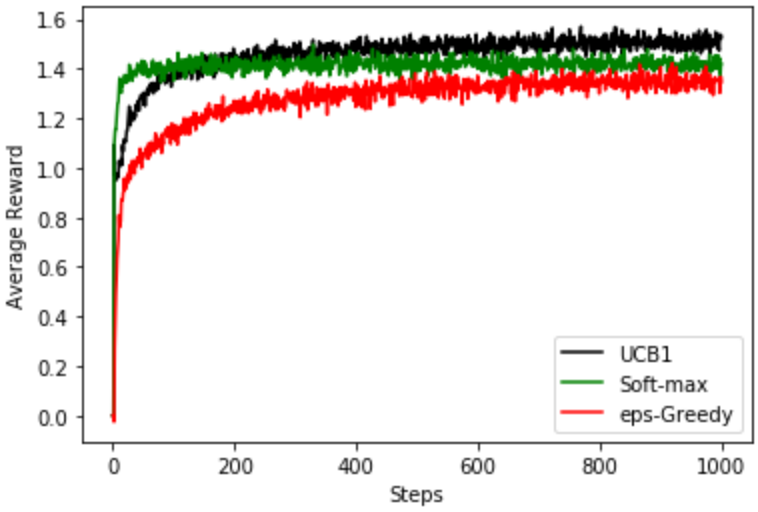




Inference:

1. The high variance corresponds to continuously picking same arm large number of times
2. The computational costs exponentially increase with decrement of eps,del
3. Red is e = 0.5,d = 0.3 green is e = 0.2,d = 0.1
4. Complexity = (K/(eps^2))*log(1/del)
5. As k = 10 it involves 4 runs for each bandit
6. Total steps are 41285, 286712, 1310966 respectively
7. Comparing avg reward, median-elimination gives best results.
8. Rate determining steps is sampling having above complexity.
9. We can simplify by summing over l and replace by function of eps

Comparison plots:



General Inference:

1. Intuitively soft-max is reaching convergence in lesser steps and giving higher reward.
2. UCB is taking more steps compared to Soft-max or ep-Greedy but providing better avg reward
3. 1000 arms bandit problem goes in around 10 runs for each bandit in median - elimination
4. The complexity of all other algorithms increase whereas median - elimination does not affect much