Assignment-1 Report

Harsh Siloiya - 160050011

The submission folder contains 3 python files -

1. init.py which is the script which starts the code execution
2. banditAlgorithms.py which contains the algorithms for solving multi-arm bandit problem
3. bandit.py which simulates the pulling of a Bernoulli bandit arm

Bandit.sh runs init.py taking instance and other parameters as an argument and calls run_function in banditAlgorithms.py to run the specific algorithm. The algorithms use a Bernoulli arm (reward $\in$ {0, 1}) to simulate a bandit. It returns 1 with probability μ and 0 with probability 1-μ (μ is the mean of the arm provided in the bandit instance).

For KL-UCB, 20 iterations have been assumed in order to find the maximum Q $\in$ [0, 1] such that the KL divergence is less than $log(t) + 3log(log(t))$.

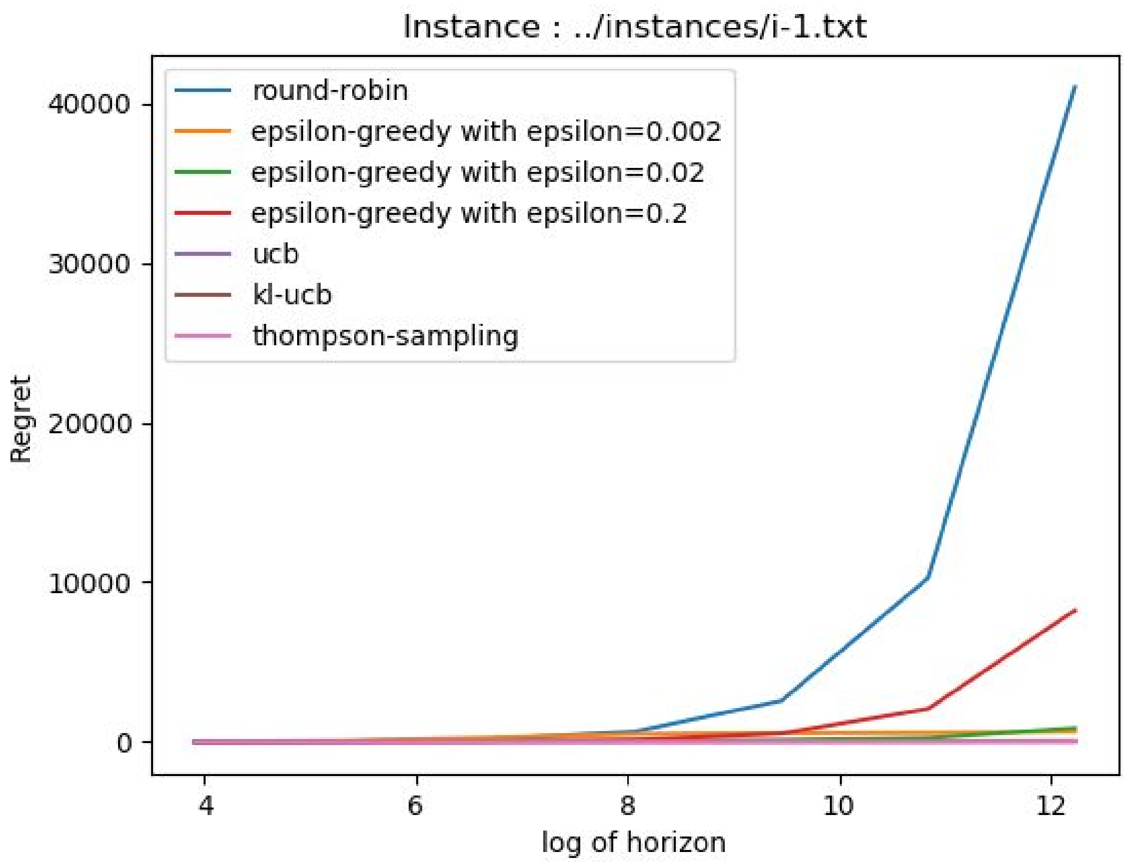
Necessary documentation has been provided in the code for the functions.

The graphs are attached in the report. The interpretation from the graphs is as follows -

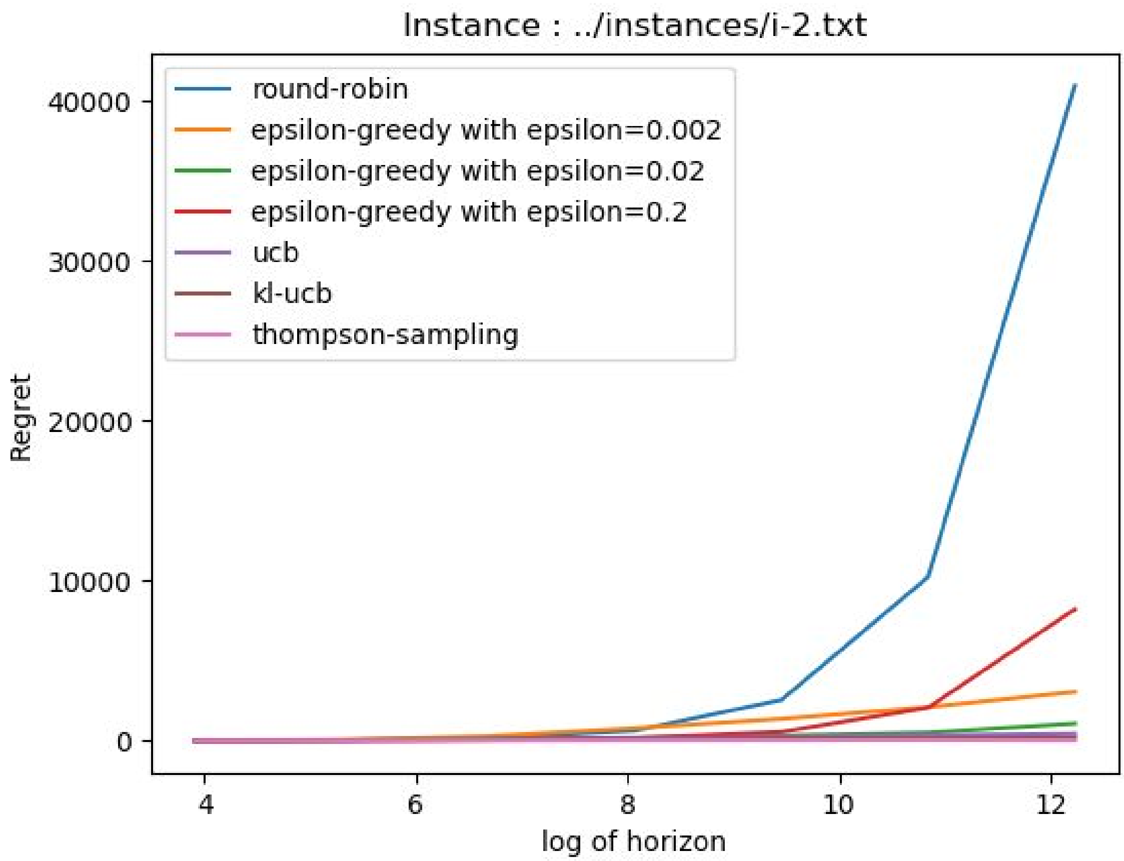
- Thompson Sampling seems to perform better in all 3 of the instances, giving the minimum regret for all horizons.
- Round-robin performs similar to Thompson sampling for low values of horizon, but the regret increases as the horizon value increases.
- ε-greedy seems to perform with intermediate results as that of Thompson Sampling and Round-robin. ε-greedy performs worse for large values of ε for all instances.
- For instance-1, ε-greedy with low ε values perform as good as UCB and KL-UCB.
- UCB and KL-UCB perform at par with Thompson sampling but with a little more regret.
- For instance-3, KL-UCB performs similar to that of ε-greedy with a large epsilon value, implying that KL-UCB is not optimal for that instance.

Graphs :-

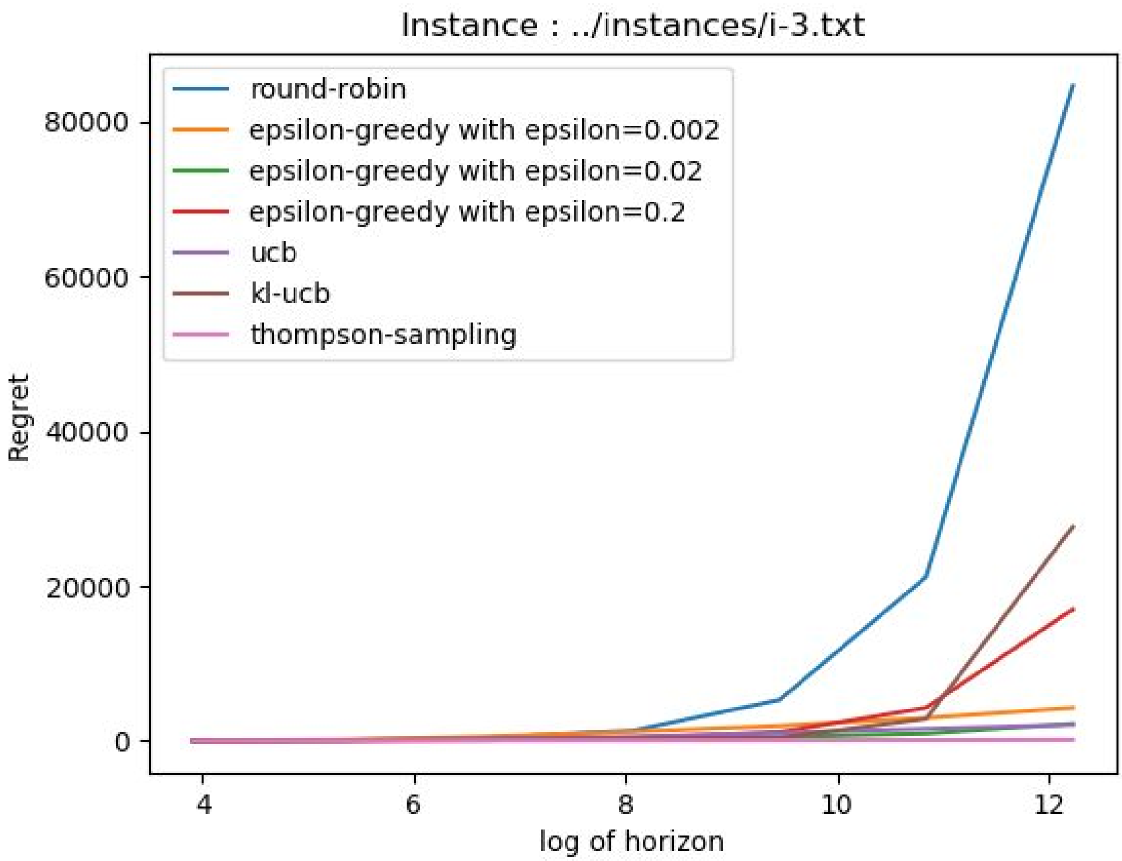
Instance-I



Instance-2


Instance : ../instances/i-2.txt
round-robin
epsilon-greedy with epsilon=0.002
epsilon-greedy with epsilon=0.02
epsilon-greedy with epsilon=0.2
ucb
kl-ucb
thompson-sampling
Regret
40000
30000
20000
10000
0
4
6
8
10
12
log of horizon

Instance-3



Instance : ../instances/i-3.txt
round-robin
epsilon-greedy with epsilon=0.002
epsilon-greedy with epsilon=0.02
epsilon-greedy with epsilon=0.2
ucb
kl-ucb
thompson-sampling
80000
60000
40000
20000
0
Regret
4
6
8
10
12
log of horizon