CS 747: Programming Assignment 1

Vinayak K (150050098)

19 Aug 2018

Code for bandit agent can be found in **agent.py**.
It consists of a base class **BanditAgent** and subclasses derived from this for each of the algorithms. (**EpsilonGreedyAgent**, **UCBAgent**, etc.)

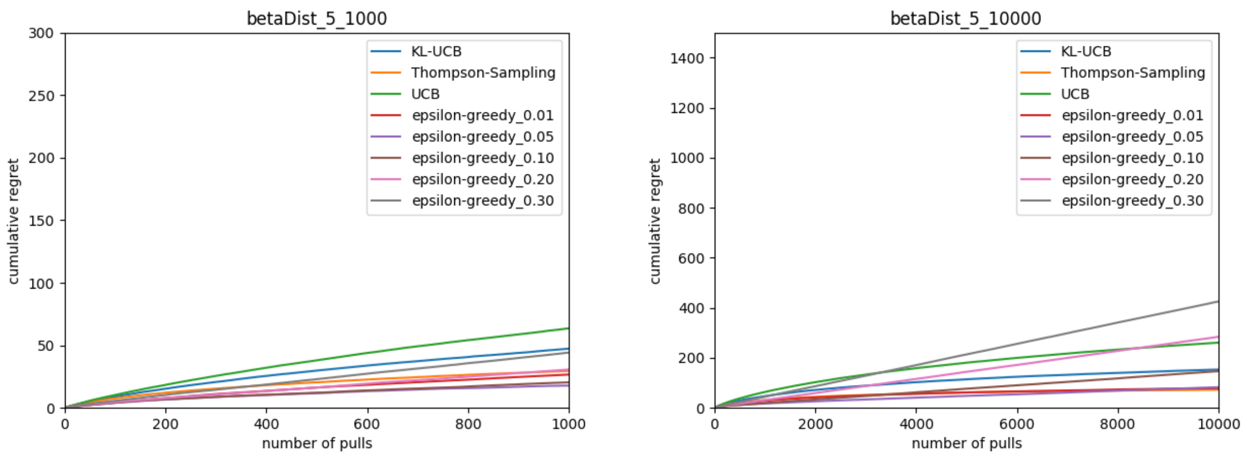
Following modification was made to **Thompson-Sampling** method for general distributions.

* Sample arm (like in normal Thompson-Sampling) and get corresponding reward.
* Do a Bernoulli trial with this reward as success probability.
* increment #success if you get a 1, else increment #failure.

`<>/report/experiment_logs/*` : contains rewards and cumulative regrets corresponding to each experiment.

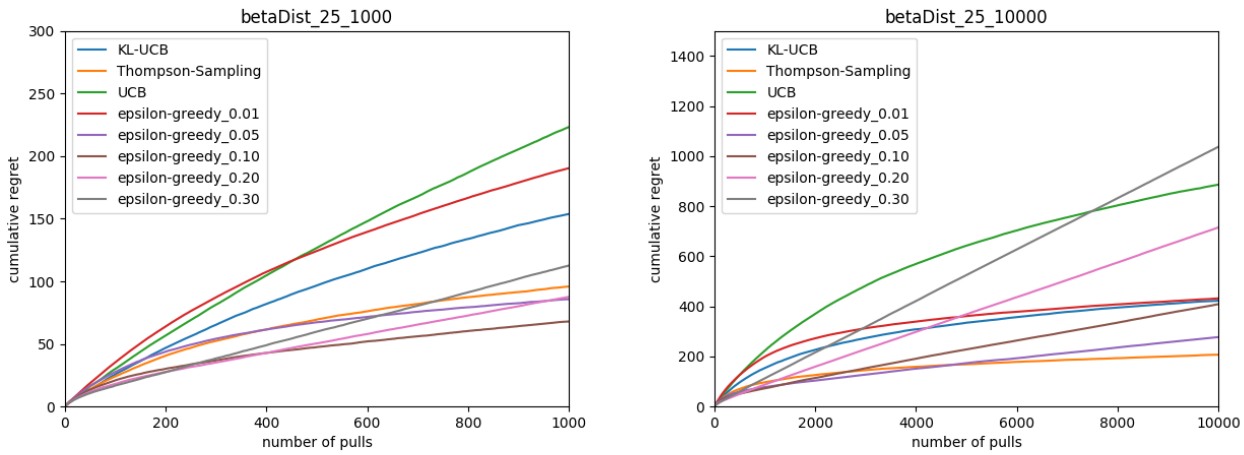
From the experiments, it seems like epsilon-greedy with appropriate ϵ value performs better than other algorithms when the horizon in small. But as the number of pulls increases Thompson-Sampling performs best in almost all cases.
For arms with beta distribution, performance of UCB algorithm does not seem to be very good. KL-UCB performs slightly better. But epsilon-greedy with $\epsilon = 0.05$ performed better than these two. Thompson sampling performed better than other algorithms.
For bernoulli distribution, KL-UCB is the second best algorithm in minimizing regret with thompson sampling minimizing the regret most again.
In case of histogram-based bandits, all the algorithms seem to perform poorly with cumulative regret almost being linearly dependent on the horizon.
In epsilon-greedy algorithm, choice of ϵ greatly impacts the performance. In many experiments with $\epsilon = 0.05$ performance was comparable to UCB.

Plots



(a) horizon: 1000 (b) horizon: 10000

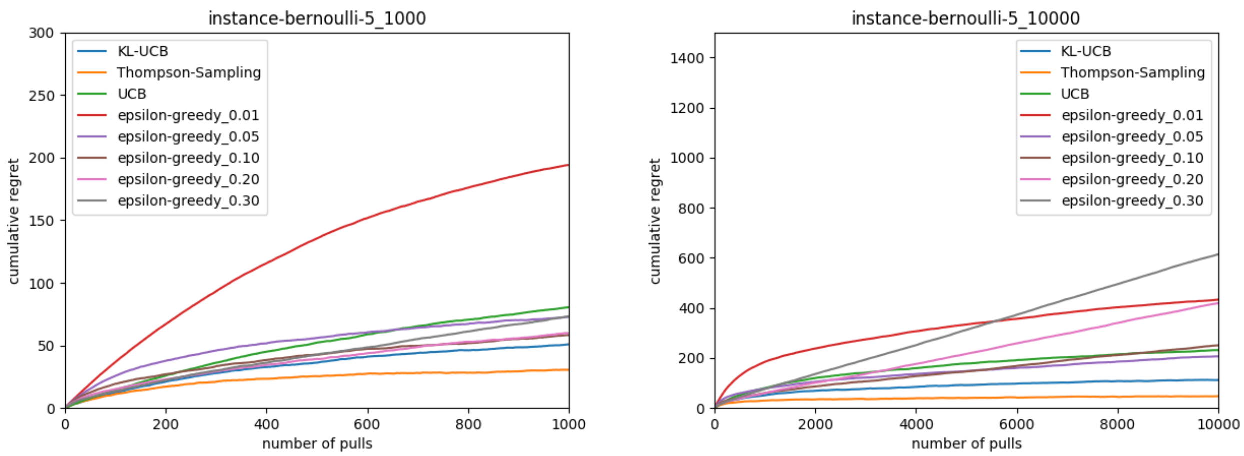
Figure 1: 5 armed bandit, Beta distributed



(a) horizon: 1000

(b) horizon: 10000

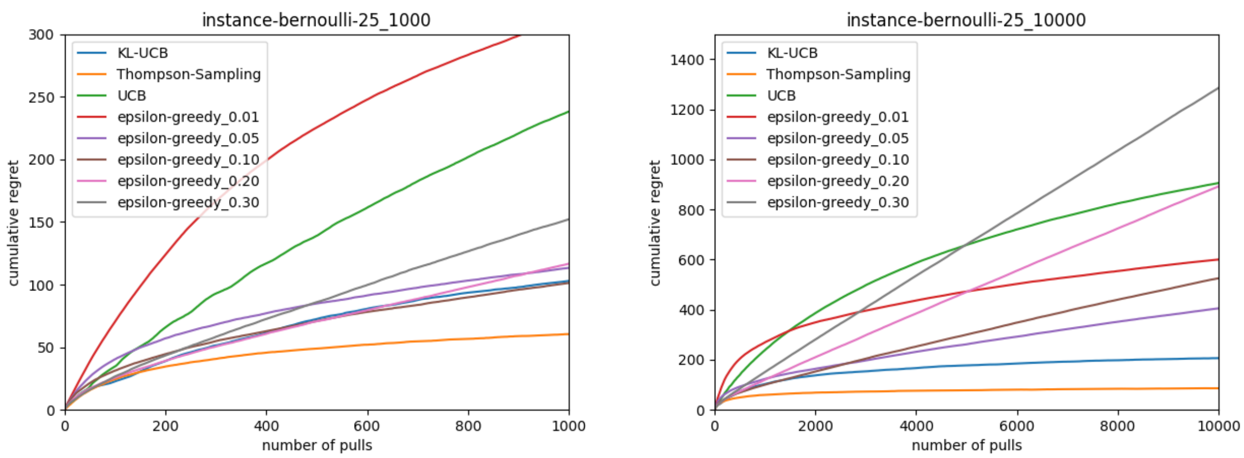
Figure 2: 25 armed bandit, Beta distributed



(a) horizon: 1000

(b) horizon: 10000

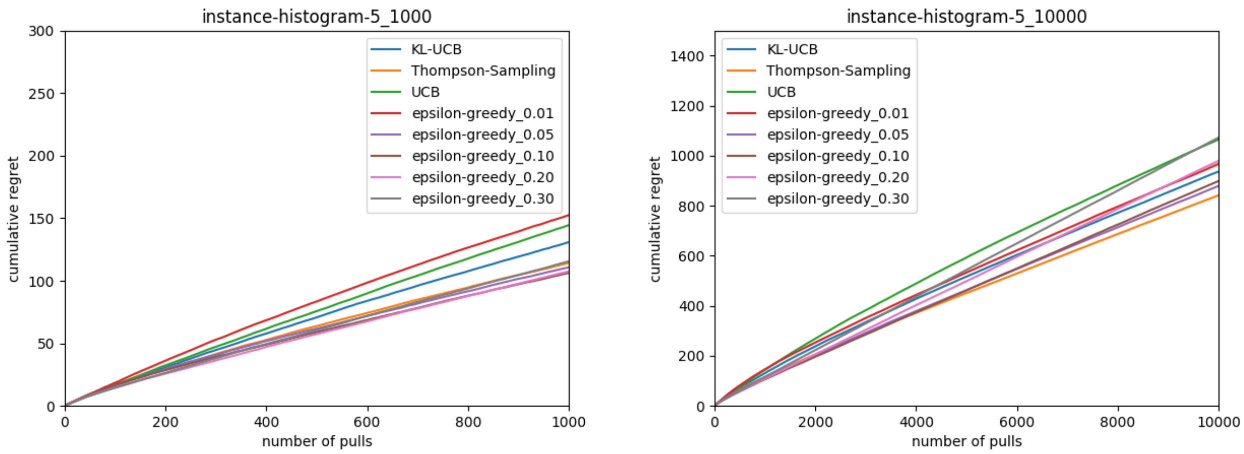
Figure 3: 5 armed bandit, Bernoulli



(a) horizon: 1000

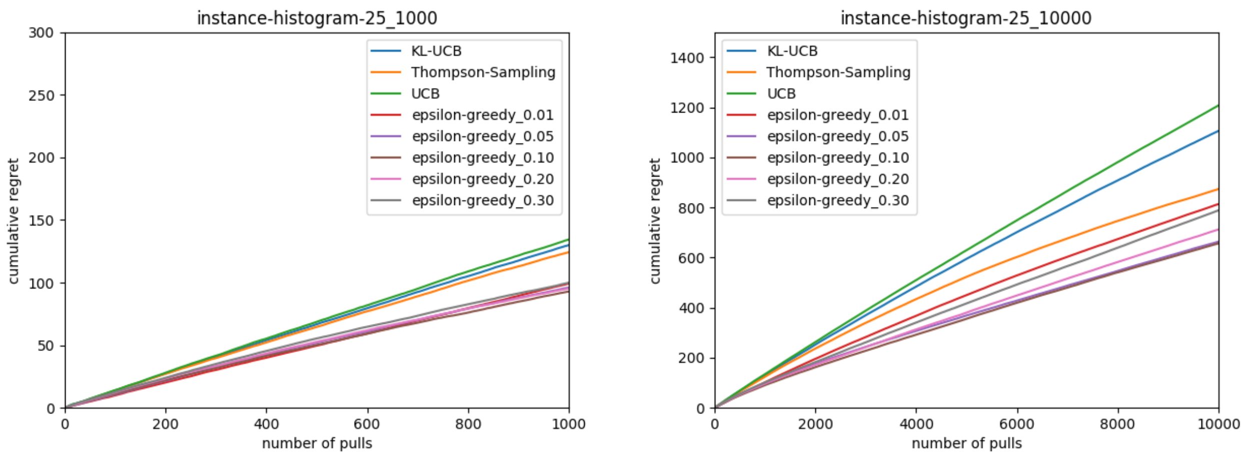
(b) horizon: 10000

Figure 4: 25 armed bandit, Bernoulli



(a) horizon: 1000 (b) horizon: 10000

Figure 5: 5 armed bandit, Histogram



(a) horizon: 1000 (b) horizon: 10000

Figure 6: 25 armed bandit, Histogram