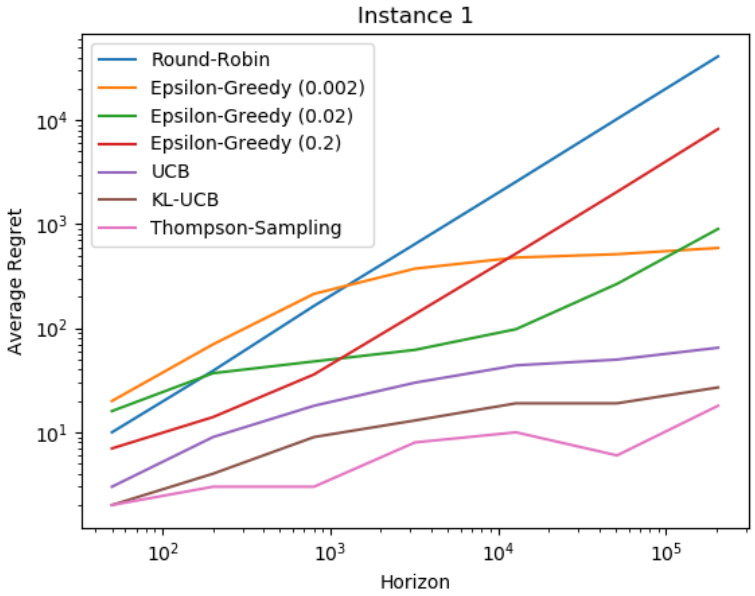
CS 747: Foundations of Intelligent and Learning Agents - Assignment 1

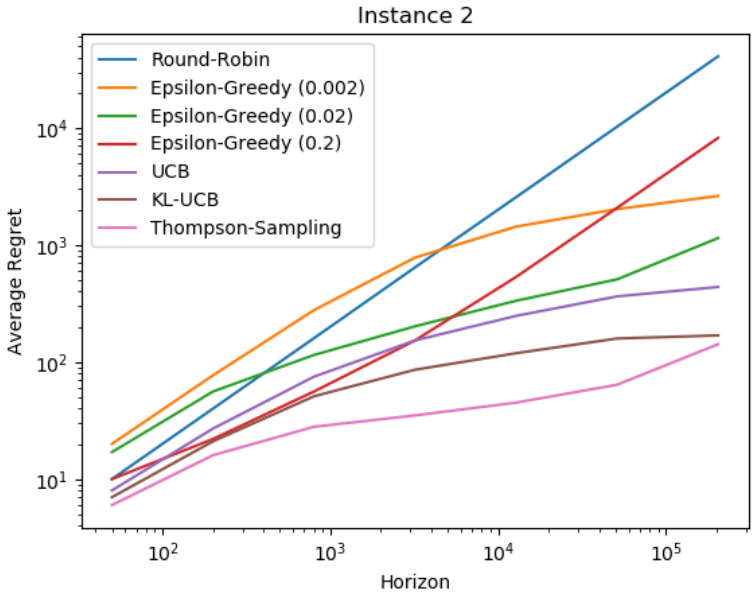
Kumar Ashutosh, 16D070043

September 3, 2019

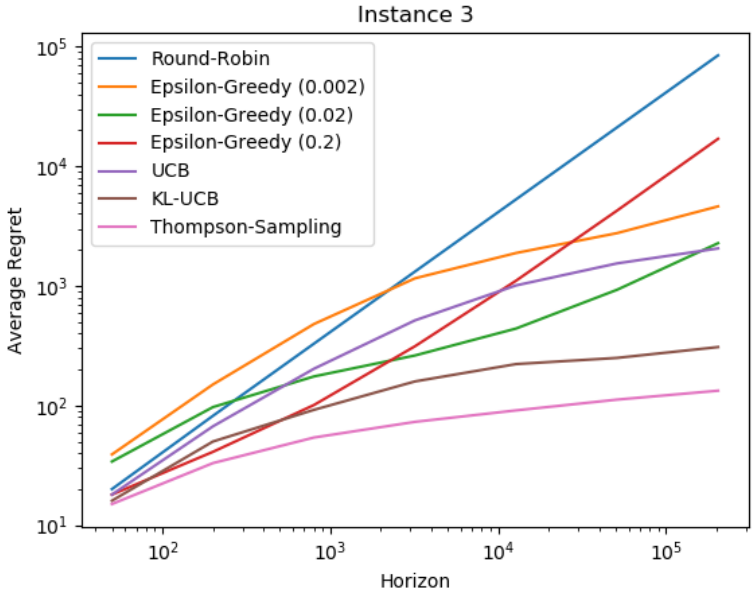
1 Plots



Plot for Instance 1



Plot for Instance 2



Plot for Instance 3

2 Assumptions in the Implementation

- In the implementation of KL-UCB, we need to find largest q which satisfies the inequality. This search in the continuous interval $[p_a, \mathbf{1})$ is achieved using binary search. In this search, it is assumed that once the difference between the upper and lower limit is less that **0.05**, the search finishes. Hence this has a maximum error of 0.05.

3 Observations

- Thompson Sampling is the best of the implemented algorithms. This is followed by KL-UCB. Round Robin is has the highest regret.
- In Thompson Sampling, for one horizon the regret decreased with an increase in horizon. This is possible due to the fact that Thompson Sampling is the best of the implemented algorithms.
- In the limit, the regret with smaller epsilon is smaller, that is, the regret "almost" saturates, as visible in the plots.
- Round Robin is a linear line in log vs log plot. This is expected because we are doing infinite exploration.
- KL-UCB is the most compute intensive due to the fact that we have to search for an optimal value in a range, thus proving as a bottleneck in the implementation.