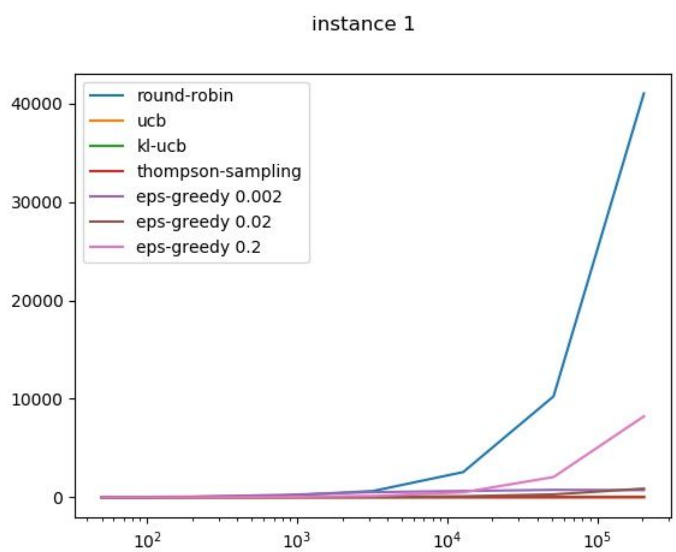
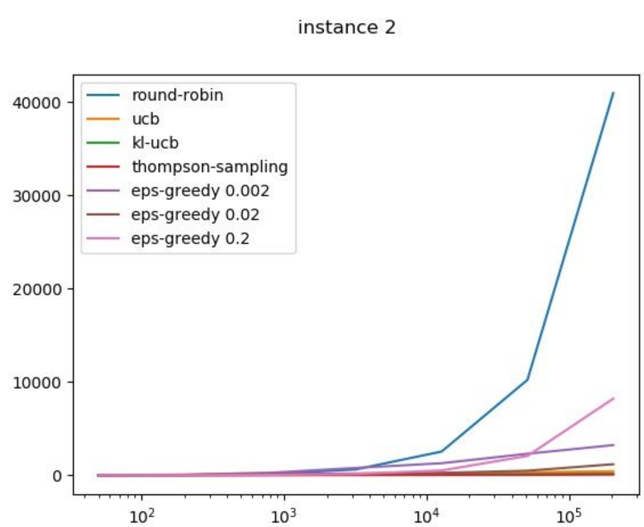
CS 747

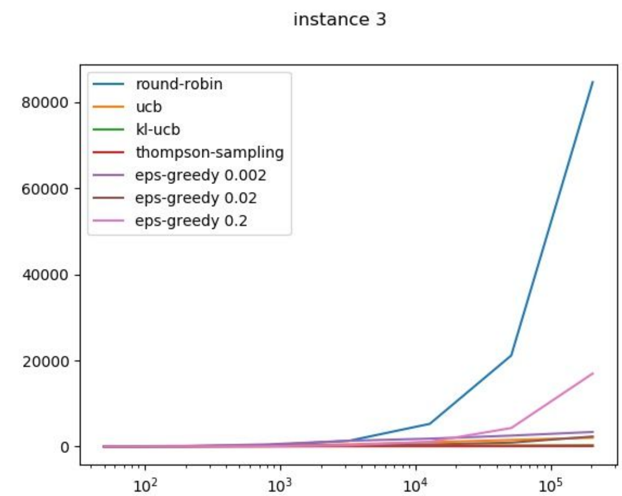
ASSIGNMENT-1

160050099

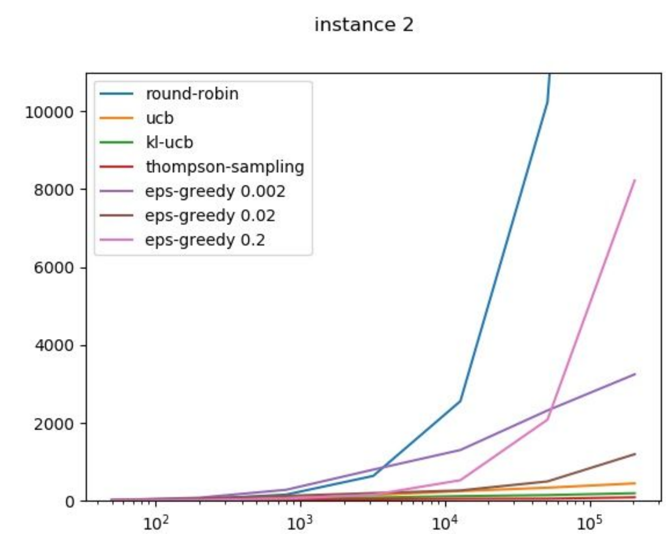
PLOTS:

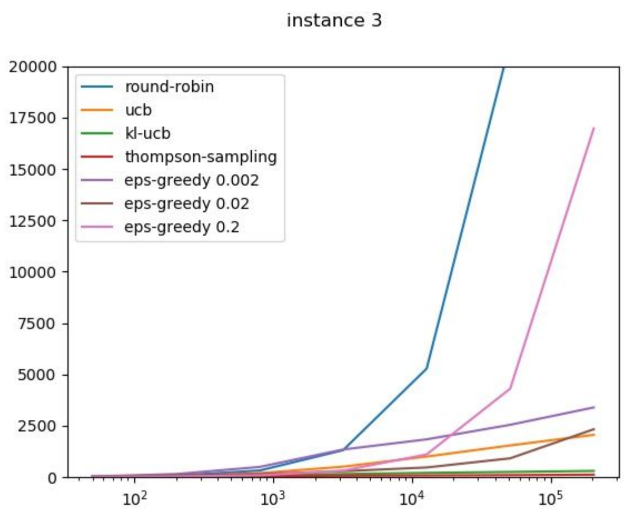






trimmed (yrange) plots of instance -2,3





Observations:

Thompson sampling gave best performance in all cases, round-robin the worst

Average Performance order: (for large number of pulls)

> Thompson-sampling > KL-UCB > UCB > eps-greedy 0.02 > eps-greedy 0.002 > epsilon-greedy 0.2 > round-robin

From the data, 0.02 appears to be optimal epsilon for epsilon greedy among the three values of epsilons experimented with.

In the initial stages, epsilon-greedy 0.002 performed the worst, but as the number of pulls increase, it did better than epsilon-greedy 0.2 and round robin.

Among the epsilon-greedy algorithms, initially eps=0.2 does well because of better exploration, in the limit it doesn't do well because at that stage exploitation is better. eps=0.02 does the best at higher horizons due to optimal tradeoff between exploring and exploiting. eps=0.002 doesn't do well initially because of too little exploration, but it takes some time to do enough exploration.

We observe negative values of regret sometimes because, empirical means might be greater than actual means at that time instant because probabilistic nature of the algorithms, although in the limit empirical values converge to actual values

The avg regret usually increases as the number of pulls increase although rarely there might be slight dips due to aforementioned negative regrets

Implementation details:

Used bisection method to solve the inequality in KL-UCB algorithm to solve for ucb upto 7 significant digits

Used numpy to generate most of the distributions