

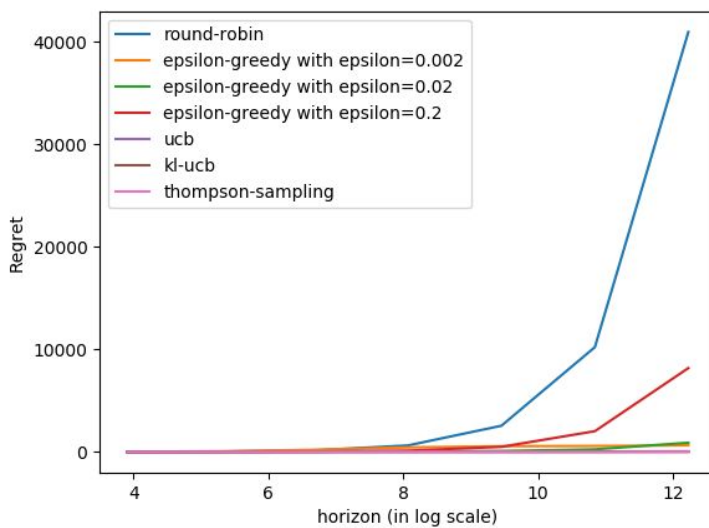
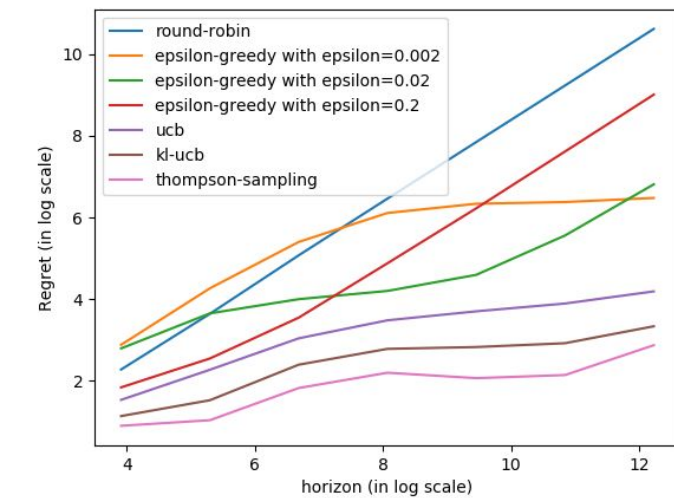
# Assignment 1 - 160050042

## Assumptions:

- Round Robin
  - No Assumption as such.
- E- greedy
  - When multiple values are give max probability in case of exploitation, chosen the first of the instance in order as tie-breaker.
  - Starts Epsilon exploration and exploitation from begining with empirical probability as 0. Rather than doing a fixed constant pull in beginning
- UCB
  - Round Robin for 1 pull each arm
  - For tie-breaker, used first one in the group
- KL-UCB
  - Round robin for 1 run
  - For tie-breaker, used first one in the group
  - For Binary Search, stopped if search window is less than  $10^{-2}$  or the required inequality difference reach  $10^{-4}$  difference
- Thompson
  - No Round Robin in start
  - For tie-breaker, used first one in the group
- **PRINTING** `outputData.txt`
  - The structure of loop is [instance → algo → seed → horizon]
- **Plotting**
  - 2 graphs are plotted for each instance. One where regret is plotted and on another  $\log(\text{regret})$  is plotted.
  - Log used in axes are natural log base e.

## Plots

Instance 1 →

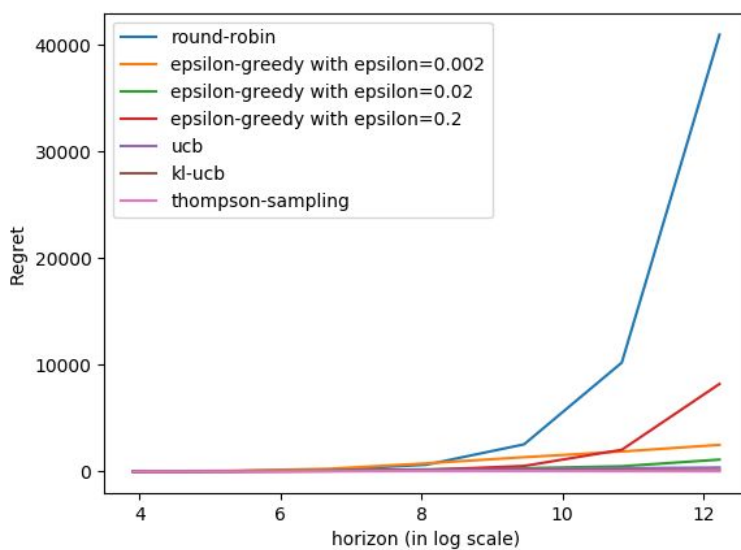
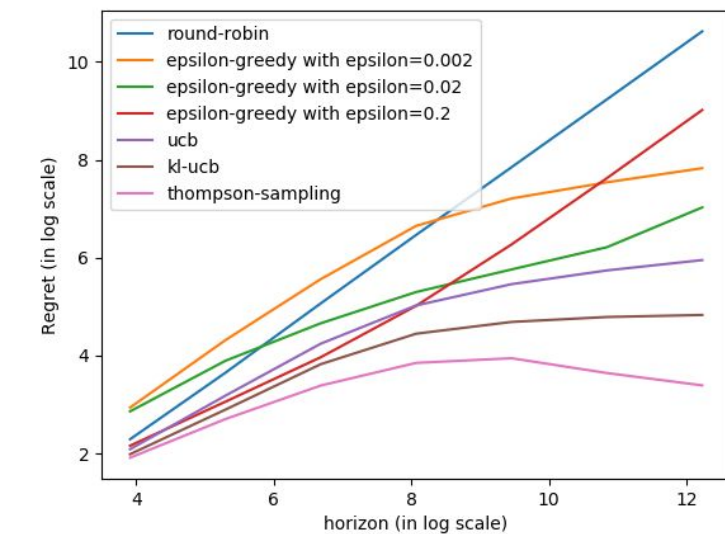


## Observations:

- Thompson sampling is a good performer.
- In Epsilon Greedy, with high epsilon, it performs badly. With low epsilon, it might perform badly in small horizons as it is exploiting high on bad values, but after few horizons, it performs better than the one with high epsilon as having high exploration till end means that not picking the best.
- KL-UCB performs better than UCB

## Plots

### Instance 2 →

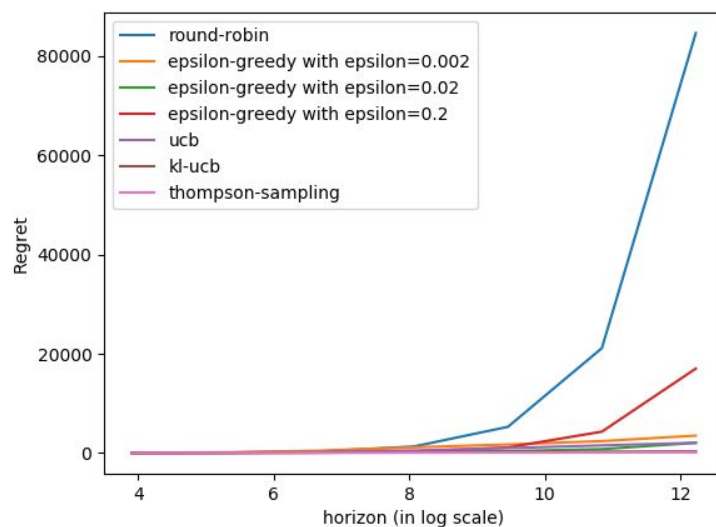
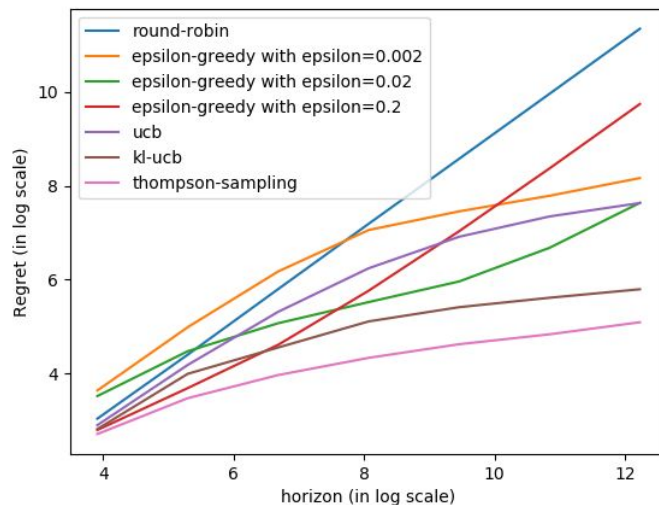


### Observations:

- Thompson sampling is a good performer.
- There is a drop in thompson sampling with increase in horizon which can be just a case where you are performing better than empirically best probability luckily.
- In Epsilon Greedy, with high epsilon, it performs badly. With low epsilon, it might perform badly in small horizons as it is exploiting high on bad values, but after few horizons, it performs better than the one with high epsilon as having high exploration till end means that not picking the best. E-Greedy with 0.02 is a good middle-ground.

## Plots

### Instance 3 →



### Observations:

- Thompson sampling is a good performer.
- In Epsilon Greedy, with high epsilon, it performs badly. With low epsilon, it might perform badly in small horizons as it is exploiting high on bad values, but after few horizons, it performs better than the one with high epsilon as having high exploration till end means that not picking the best. E-Greedy with 0.02 is a good middle-ground.
- KL-UCB performs better than UCB