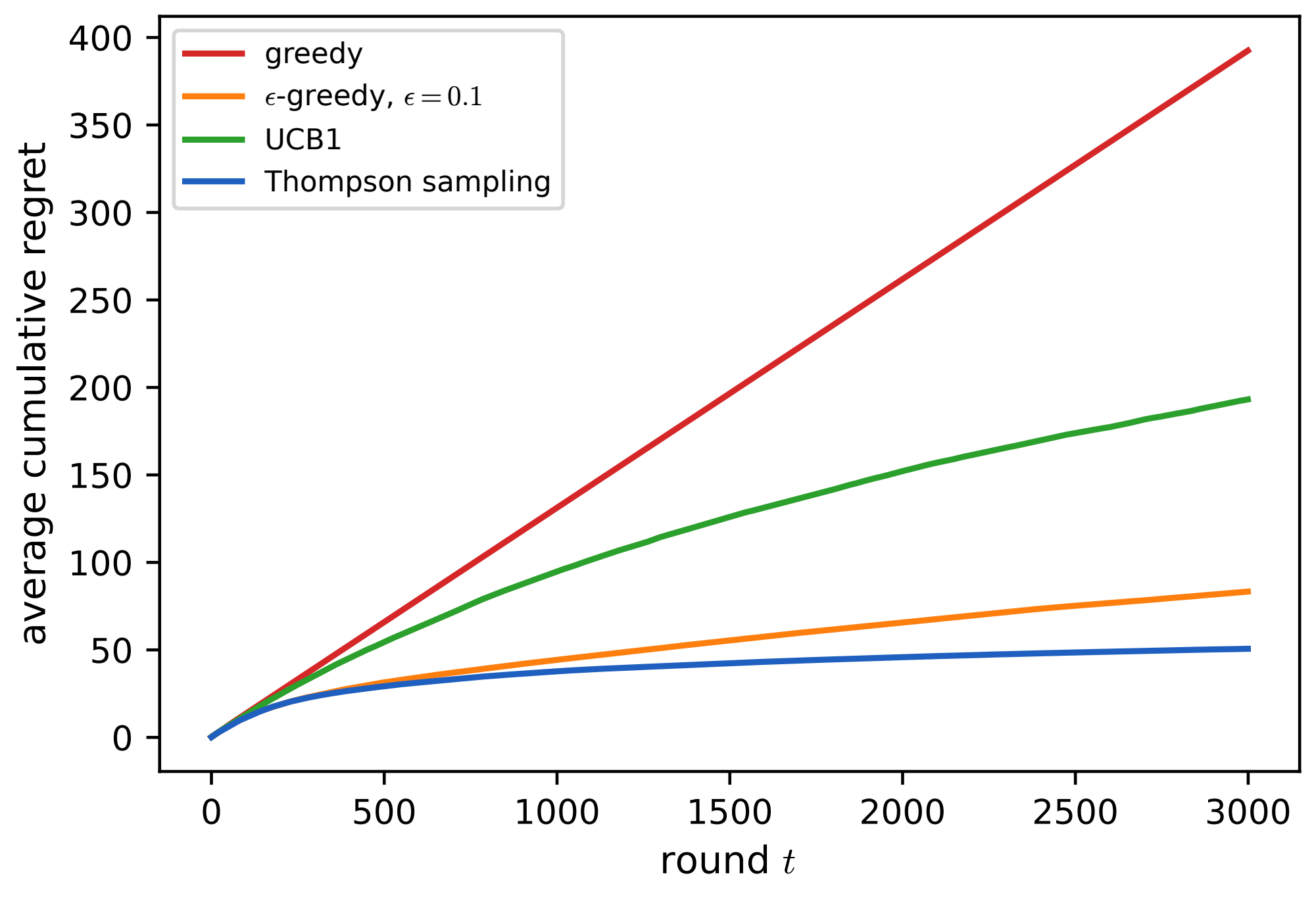

greedy
ϵ-greedy, $\epsilon = 0.1$
UCB1
Thompson sampling
average cumulative regret
round t
0
50
100
150
200
250
300
350
400
0
500
1000
1500
2000
2500
3000