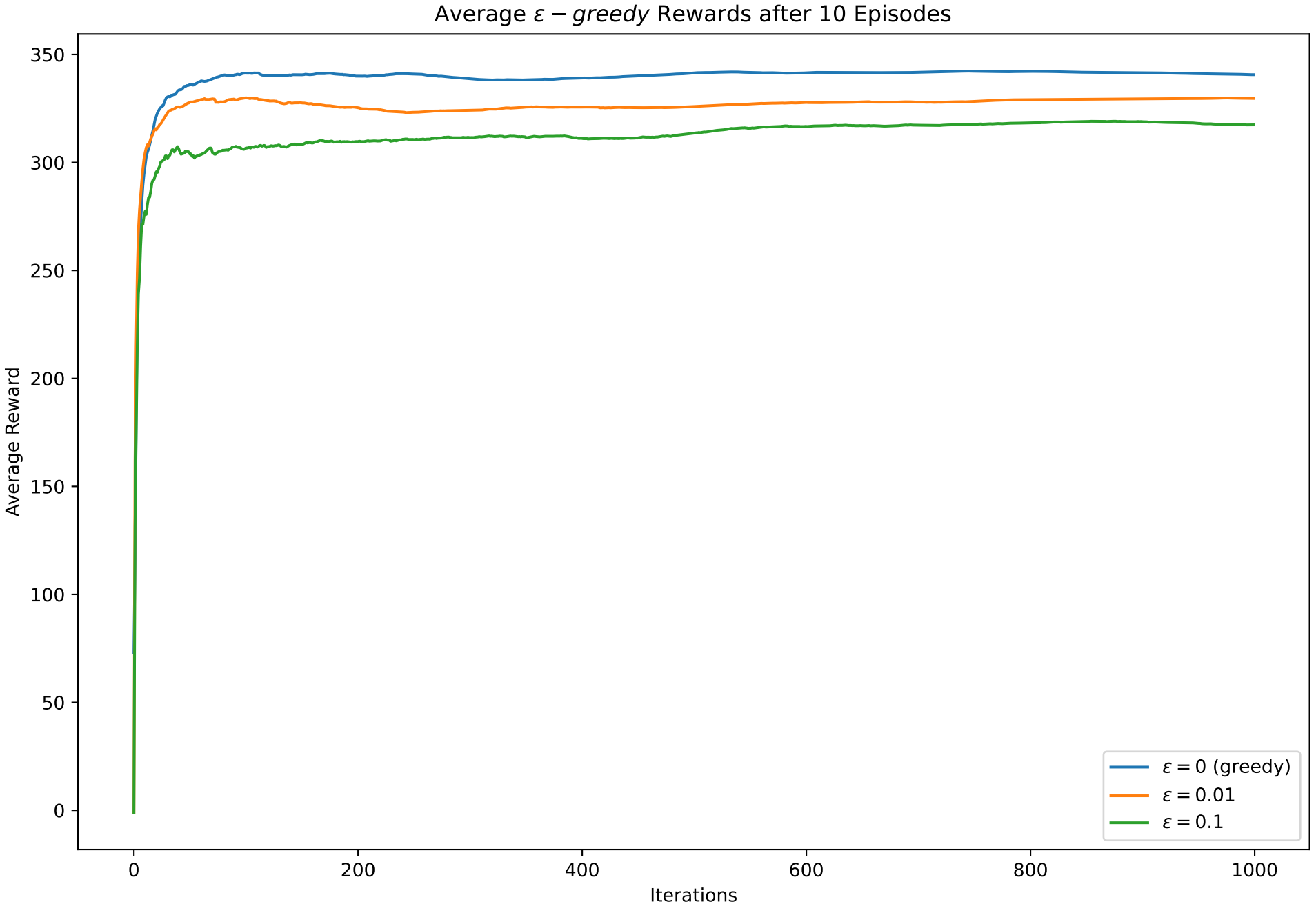
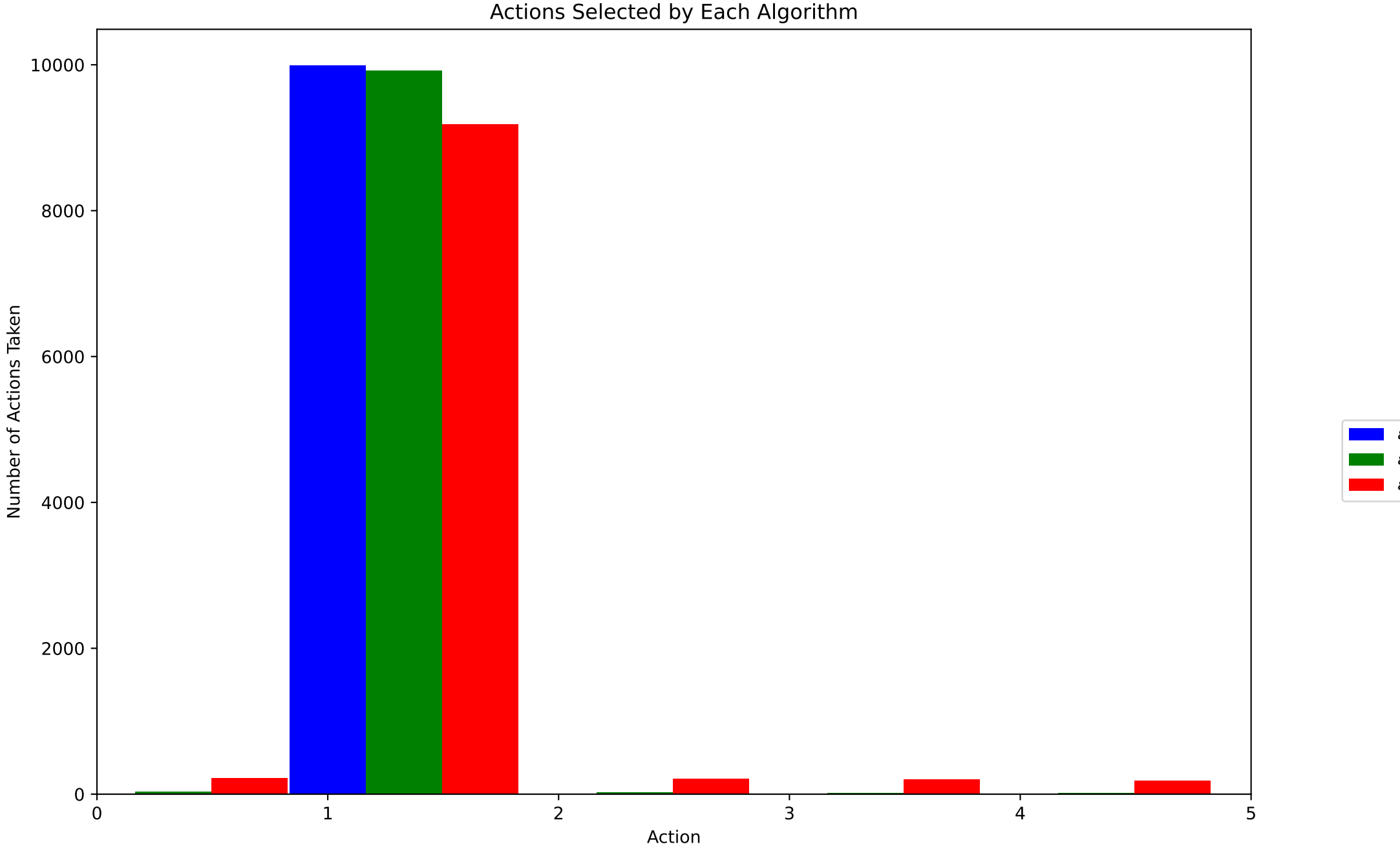

Average $\varepsilon - greedy$ Rewards after 10 Episodes
Average Reward
Iterations
0
50
100
150
200
250
300
350
0
200
400
600
800
1000
$\varepsilon = 0$ (greedy)
$\varepsilon = 0.01$
$\varepsilon = 0.1$


Actions Selected by Each Algorithm
Number of Actions Taken
Action
0
2000
4000
6000
8000
10000
0
1
2
3
4
5