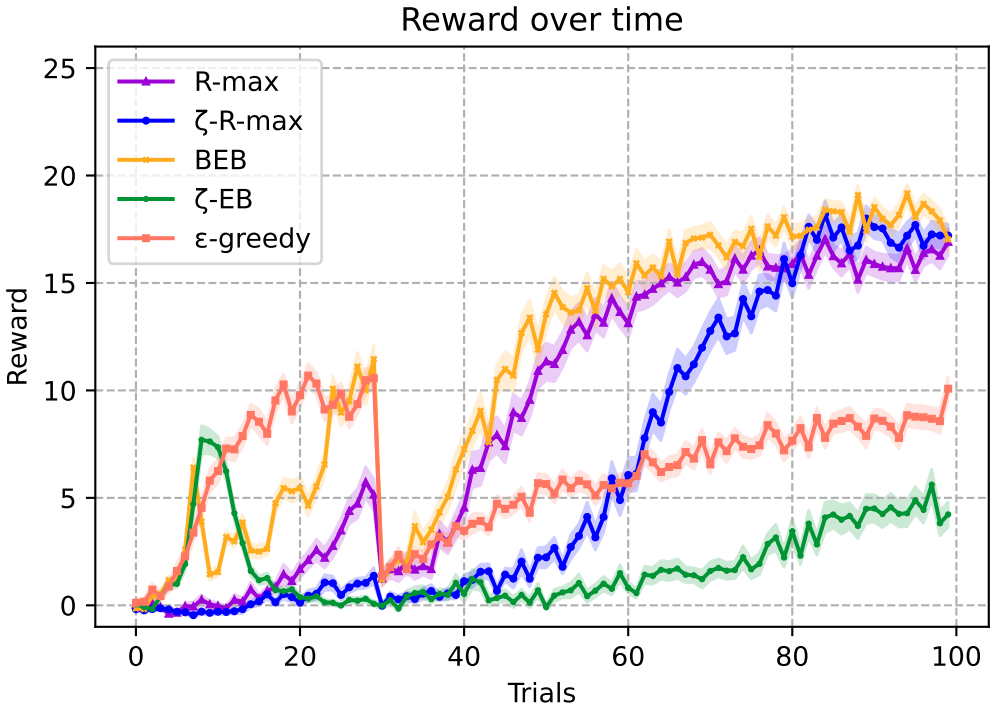
Reward over time
R-max
ζ-R-max
BEB
ζ-EB
ε-greedy
Reward
Trials
0
5
10
15
20
25
0
20
40
60
80
100