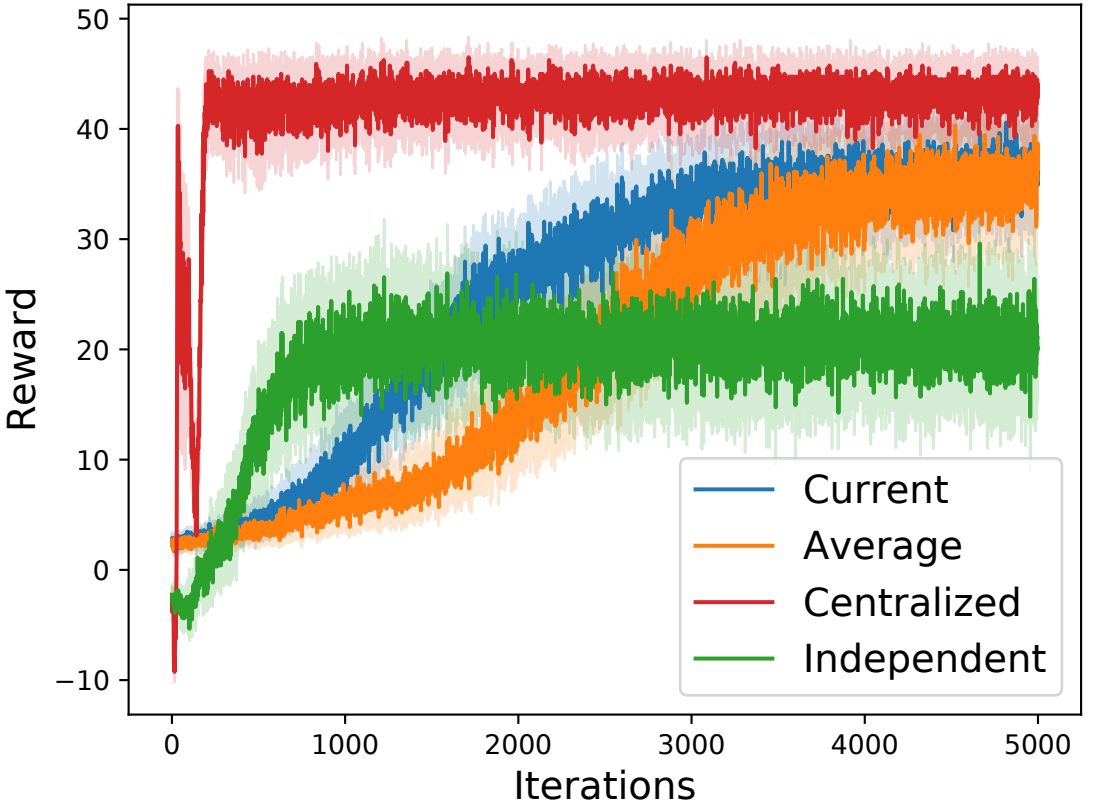
Reward
Iterations
50
40
30
20
10
0
−10
0
1000
2000
3000
4000
5000
Current
Average
Centralized
Independent