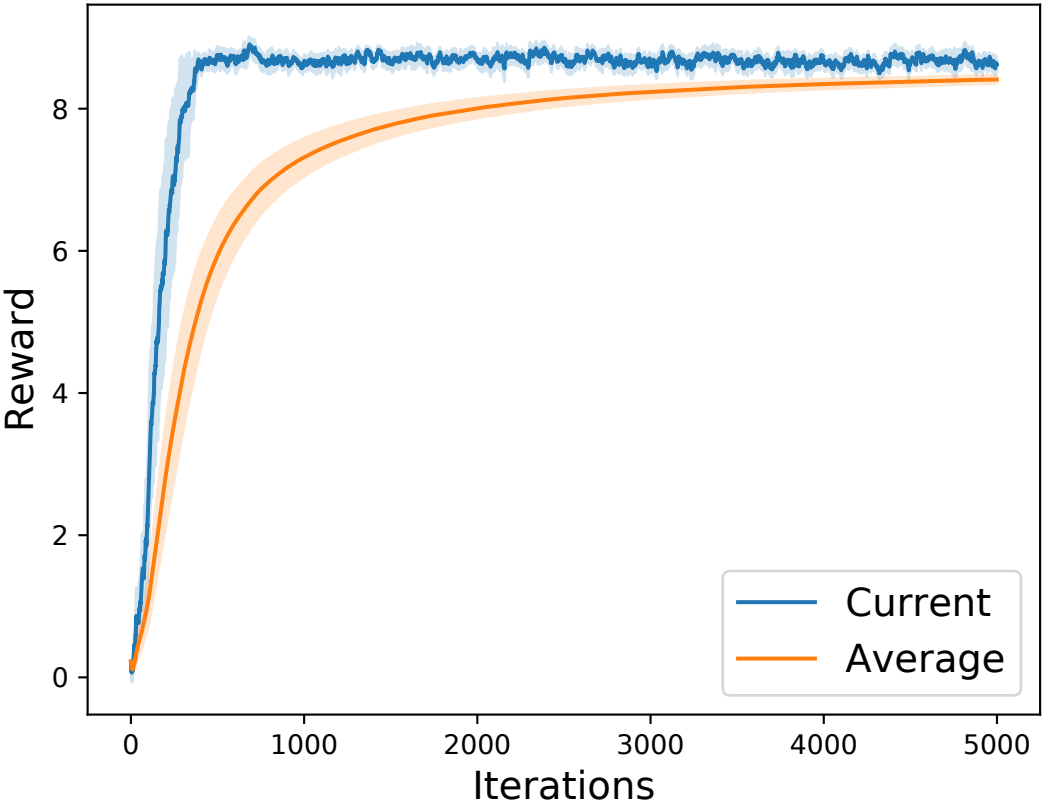
Reward
Iterations
0
2
4
6
8
0
1000
2000
3000
4000
5000
Current
Average