Average Episode Reward Averaged per 100 Episodes / 10^2
5.5
5
4.5
4
3.5
3
0
2
4
6
8
10
Episode Number / 10^4
MDP-based
plan-flag
no-shaping
plan-based
flag-based
optimal
plan optimal