Average Episode Reward Averaged per 100 Episodes / 10^2
6
4
2
0
-2
-4
-6
-8
-10
-12
0
2
4
6
8
10
Episode Number / 10^4
MDP-based
plan-flag
no-shaping
plan-based
flag-based
optimal