Average Episode Reward Averaged per 100 Episodes / 10
Episode Number / 10^3
0
5
10
15
20
25
30
35
40
0
2
4
6
8
10
MDP-based
plan-flag
no-shaping
plan-based
flag-based
optimal
plan optimal