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