BEB-AO β=0.3 m_B=+∞
MBIE-EB-AO β=0.1 m_B=5
GO-AO
Rmax-AO
Rmax
Average cumulative reward / 10^3
Number of episodes / 10^2
0
-10
-20
-30
-40
-50
-60
0
2
4
6
8
10
12
14