

R-max agent policy with  $m = 2.5$

