

R-max agent policy with  $m = 0.5$

