

R-max agent policy with  $m = 3.5$

