

R-max agent policy with  $\gamma = 0.95$

