

R-max optimal policy with  $m = 12$

