

R-max agent policy with $m = 36$

