

ζ -R-max agent policy with $\alpha = 0.3$

