

ζ -R-max optimal policy with alpha = 0.3

