

ζ -R-max optimal policy with alpha = 1.2

