

ζ -R-max optimal policy with alpha = 2.7

