

ζ -R-max optimal policy with gamma = 0.95

