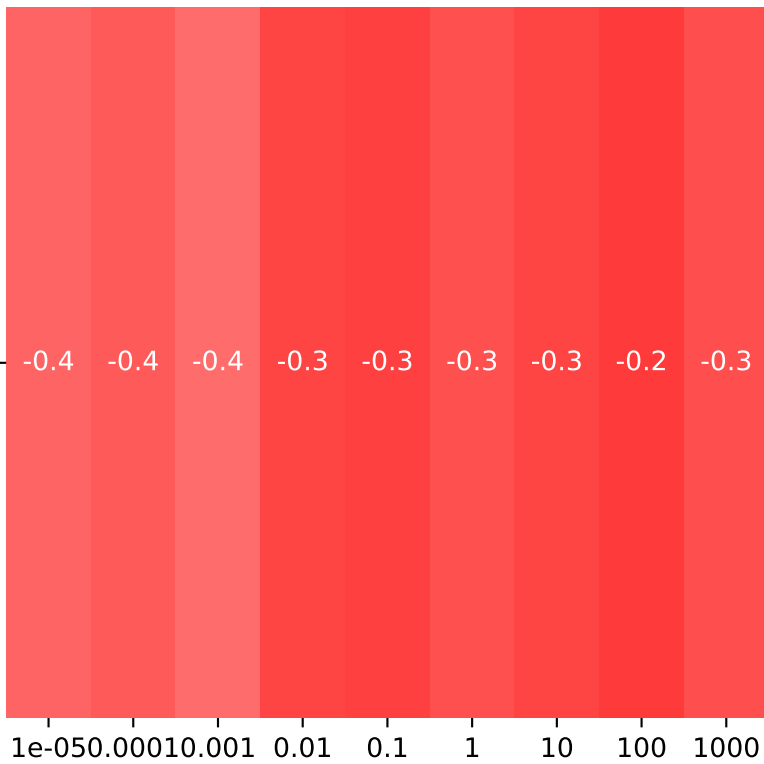


$\zeta$ -R-max agent policy - last 500 steps

gamma

0.95



0

-1

-12

prior\_LP