ζ-R-max agent policy - last 500 steps -0.6 -0.1 -0.6 -1.2 -1.3 -2.7 -1.2 -1.1 -0.8 -0.7 -1.5 -0.7 -0.1 -0.6 -0.7

-0.8 -1.5 -0.7 -0.7 -0.6 -0.1 -0.8

-0.8 -0.6 1.5 2.5 3.0 3.5 0.1 0.5 1.0 2.0

alpha