ζ-EB optimal policy - last 500 steps -1 -1.1 -0.9 -1 -0.7 -1.3 -1 -0.8 -0.6 -0.6 -0.9 -0.9 -1.1 -1.1 1.5 -0.6 -0.8 -0.7 -0.7 -0.6 -0.8 -0.8 2.5 -0.6 -0.7 3.0 -0.7 -0.7 3.5 -0.6 1.0 2.0 0.1 0.5 1.5

alpha

beta