R-max optimal policy - last 500 steps -0.1 -0.1 -0.7 -1.4 -0.1 -0.1 -0.7 30 -1.4 -0.8 -0.1

5 10 15 20 25 30 35 40 m uncertain states