SAC agent's caracteristics

$E[Q(s, \pi(s)] \approx 245$, where s is a state right before the first hurdle (x $\approx$ 0.5)

$Q(s, a)$

$\pi(s)$