

SAC agent's characteristics

$E[Q(s, \pi(s))] \approx 275$, where s is a state right before the first hurdle ($x \approx 0.5$)

