

Average reward using actor-critic RL

$\gamma = 0.9$, $\alpha_{\theta} = 0.001$, $\alpha_v = 0.0001$

