

2025 march 31

Monday, March 31, 2025 12:57

$$L_{\text{actor}} = - \mathbb{E} \left[\begin{array}{c|c} \text{sum of} & \\ \text{all} & \\ \text{future} & \\ \text{rewards} & \end{array} \middle| s \right]$$

$$z = \sum_a \pi_a(s) q(s, a)$$

$\pi_a(s)$	s		$q(s, a)$	s	
	0	1		0	1
p	0.8	0.9	a^p	90	30
a^w	0.2	0.1	w	40	50

$s=0!$

$$u(0) = \sum_a \pi_a(0) q(0, a)$$

$$= 0.8 \cdot 90 + 0.2 \cdot 40$$