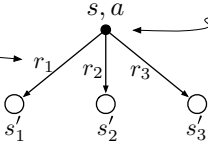


期望奖赏



$q_{\pi}(s, a)$

$v_{\pi}(s')$