taken with probability $\pi(a|s)$

$v_\pi(s)$

$q_\pi(s, a)$

s

a_1 a_2 a_3