$\pi \longrightarrow$

s

$\pi(a_1|s)$ $\pi(a_k|s)$

q ...

trans. prob.
$P(s'|s,a)$

s'

$\longrightarrow V(s) = \sum_a q(s,a) \pi^*(a|s)$

$\longrightarrow q(s,a) = \sum_{s',v} \left[\gamma V(s) + r \right] P(s',r|s,a)$

s

$\pi^* \longrightarrow$

$P \longrightarrow$

$\pi^* \longrightarrow$

$\text{argmax}_a \; q_*(s,a)$