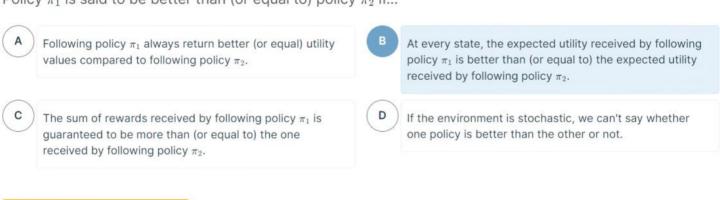
A policy $\pi:S\to A$ gives an action for each state. If |S|=n, and |A|=k, the total number of policies is?



2 of 10

Policy π_1 is said to be better than (or equal to) policy π_2 if...



UBMIT ANSWER

$V^{\pi}(s) = \sum_{s'} P(s' s,\pi(s))[R(s,\pi(s),s') + \gamma V^{\pi}(s')]$ $Q \text{ Zoom}$	What is the name of the V^π function?
$\begin{tabular}{ c c c c c c c c c c c c c c c c c c c$	B The state-value function. The state-value function for policy π .
4 of 10	
$V_t^{\pi}(s) \leftarrow \sum_{s'} P(s' s,\pi(s))[R(s,\pi(s),s') + \gamma V_{t-1}^{\pi}(s')]$ $Q \text{ Zoom}$	S is the set of states. What is the complexity of iterative Policy Evaluation?
$egin{picture}(A) & O(S ^2) & \\ \hline$	$O(S ^2)$ for each iteration
	O(S) for each iteration

$V^{\pi}(s) =$	$=\sum_{s'} p$	(s' s,π((s))[R(s, π(s),	$s') + \gamma I$	$V^{\pi}(s')$
		С	\ Zoor	n		

How can we determine the expected utility received at each state by following π ?

- A Solving a system of linear equations, where each equation is a state-value function for policy π when starting at a state.
- ${\bf B}$ Using iterative policy evaluation to update V^π until acceptable convergence.

Both A and B can be used.

D None can be used.

SUBMIT ANSWER

6 of 10

$$Q^*(s,a) = \sum_{s'} P(s'|s,a) [R(s,a,s') + \gamma V^*(s')]$$

What is the name of the Q^* function?

Q Zoom

A The optimal action-value function.

B The optimal state-value function.

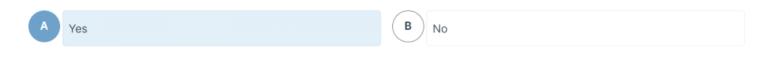
C The state-value function.

D The action-value function.

UBMIT ANSWER

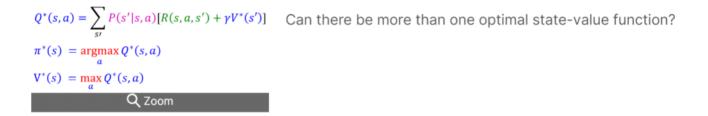
$$Q^*(s,a) = \sum_{s'} P(s'|s,a)[R(s,a,s') + \gamma V^*(s')]$$
 Can there be more than one optimal policy?
$$\pi^*(s) = \operatorname*{argmax}_a Q^*(s,a)$$

$$V^*(s) = \operatorname*{max}_a Q^*(s,a)$$
 Q Zoom



SUBMIT ANSWER

8 of 10





SUBMIT ANSWER

$V^*(s) \leftarrow \max_{a} \sum_{s'} P(s' s,a) [R(s,a,s') + \gamma V^*(s')]$ $Q \text{ Zoom}$	S is the set of state of Value Iteration?	es. A is the set of actions. What is the comple	exity
	В	$O(S imes A ^2)$	
$oldsymbol{C}$ $O(S imes A ^2)$ for each iteration	D	$O(A imes S ^2)$ for each iteration	
SUBMIT ANSWER 10 of 10			
$V^*(s) \leftarrow \max_{a} \sum_{s'} P(s' s,a) [R(s,a,s') + \gamma V^*(s')]$ $Q \text{ Zoom}$	What are the direct r	results of Value Iteration?	
A The optimal state-value function.	В	The expected utility received at each state by following an optimal policy.	g
C The action that should be taken at each optimally.	state to act	Both A and B are correct.	