A policy *π*:*S*→*A* gives an action for each state. If∣*S*∣=*n*, and ∣*A*∣=*k*, the total number of policies is?

**A**

*nk*

**B**

*kn*

**C**

*kn*

**D**

*k*(*n*2)

@@@@@@@@@@@@@@@@@@

Policy *π*1​ is said to be better than (or equal to) policy *π*2​ if...

**A**

Following policy *π*1​ always return better (or equal) utility values compared to following policy *π*2​.

**B**

At every state, the expected utility received by following policy *π*1​ is better than (or equal to) the expected utility received by following policy *π*2​.

**C**

The sum of rewards received by following policy *π*1​ is guaranteed to be more than (or equal to) the one received by following policy *π*2​.

**D**

If the environment is stochastic, we can't say whether one policy is better than the other or not.

@@@@@@@@@@@@@@

What is the name of the *Vπ* function?

**A**

The action-value function.

**B**

The state-value function.

**C**

The action-value function for policy *π*.

**D**

The state-value function for policy *π*.

A close-up of a logo

Description automatically generated

@@@@@@@@@@@@@@@@@@@@@@@

S is the set of states. What is the complexity of iterative Policy Evaluation?

**A**

*O*(∣*S*∣2)

**B**

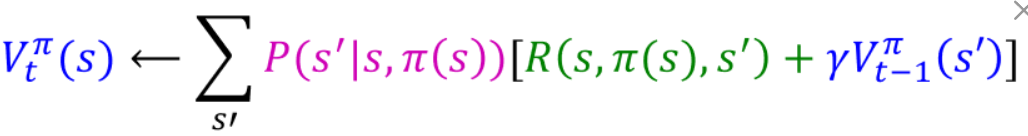
﻿*O*(∣*S*∣2) for each iteration

**C**

﻿*O*(∣*S*∣)

**D**

﻿*O*(∣*S*∣) for each iteration



@@@@@@@@@@@@@@@@@@@@@@@@@@

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.

**B**

Using iterative policy evaluation to update *Vπ* until acceptable convergence.

**C**

Both A and B can be used.

**D**

None can be used.

A close-up of a logo

Description automatically generated

@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@

What is the name of the *Q*⋆ function?

**A**

The optimal action-value function.

**B**

The optimal state-value function.

**C**

The state-value function.

**D**

The action-value function.

A close-up of a math problem

Description automatically generated

@@@@@@@@@@@@@@@@@@@@@@@@@@@@

Can there be more than one optimal policy?

**A**

Yes

**B**

No

A close-up of a math problem

Description automatically generated

@@@@@@@@@@@@@@@@@@@@@

Can there be more than one optimal state-value function?

**A**

Yes

**B**

No

A close-up of a math problem

Description automatically generated

@@@@@@@@@@@@@@@@@@@@@@@@@@

*S* is the set of states. *A* is the set of actions. What is the complexity of Value Iteration?

**A**

*O*(∣*S*∣×∣*A*∣)

**B**

﻿*O*(∣*S*∣×∣*A*∣2)﻿

**C**

﻿*O*(∣*S*∣×∣*A*∣2) for each iteration

**D**

﻿*O*(∣*A*∣×∣*S*∣2) for each iteration

A close-up of a text

Description automatically generated

@@@@@@@@@@@@@@@@@@@@@@@@@@

What are the direct results of Value Iteration?

**A**

The optimal state-value function.

**B**

The expected utility received at each state by following an optimal policy.

**C**

The action that should be taken at each state to act optimally.

**D**

Both A and B are correct.

A close-up of a text

Description automatically generated