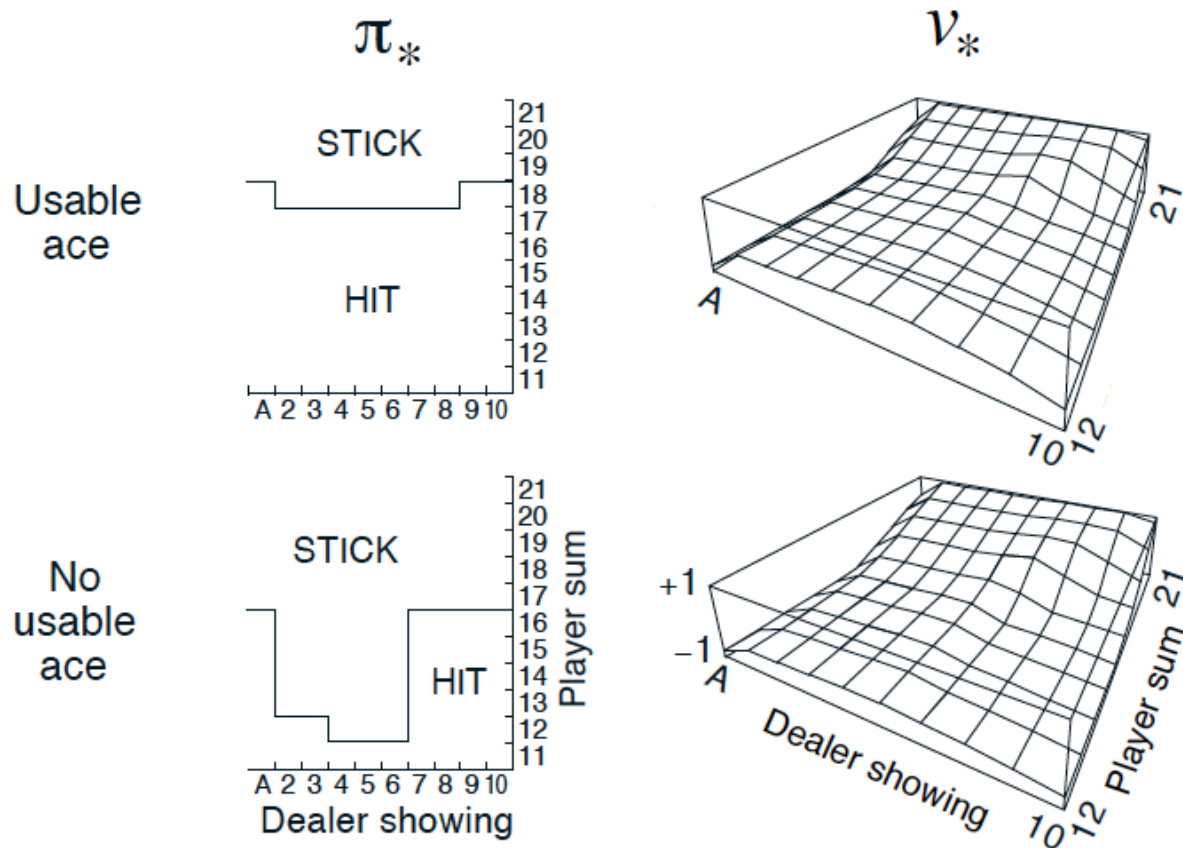# Summary

Optimal policy and state-value function in blackjack (Sutton and Barto, 2017)

MC Prediction: State Values

- Algorithms that solve the **prediction problem** determine the value function v_π (or q_π) corresponding to a policy π.
- Methods that evaluate a policy π from interaction with the environment fall under one of two categories:
  - **On-policy** methods have the agent interact with the environment by following the same policy π that it seeks to evaluate (or improve).

  - **Off-policy** methods have the agent interact with the environment by following a policy b (where $b \neq \pi$) that is different from the policy that it