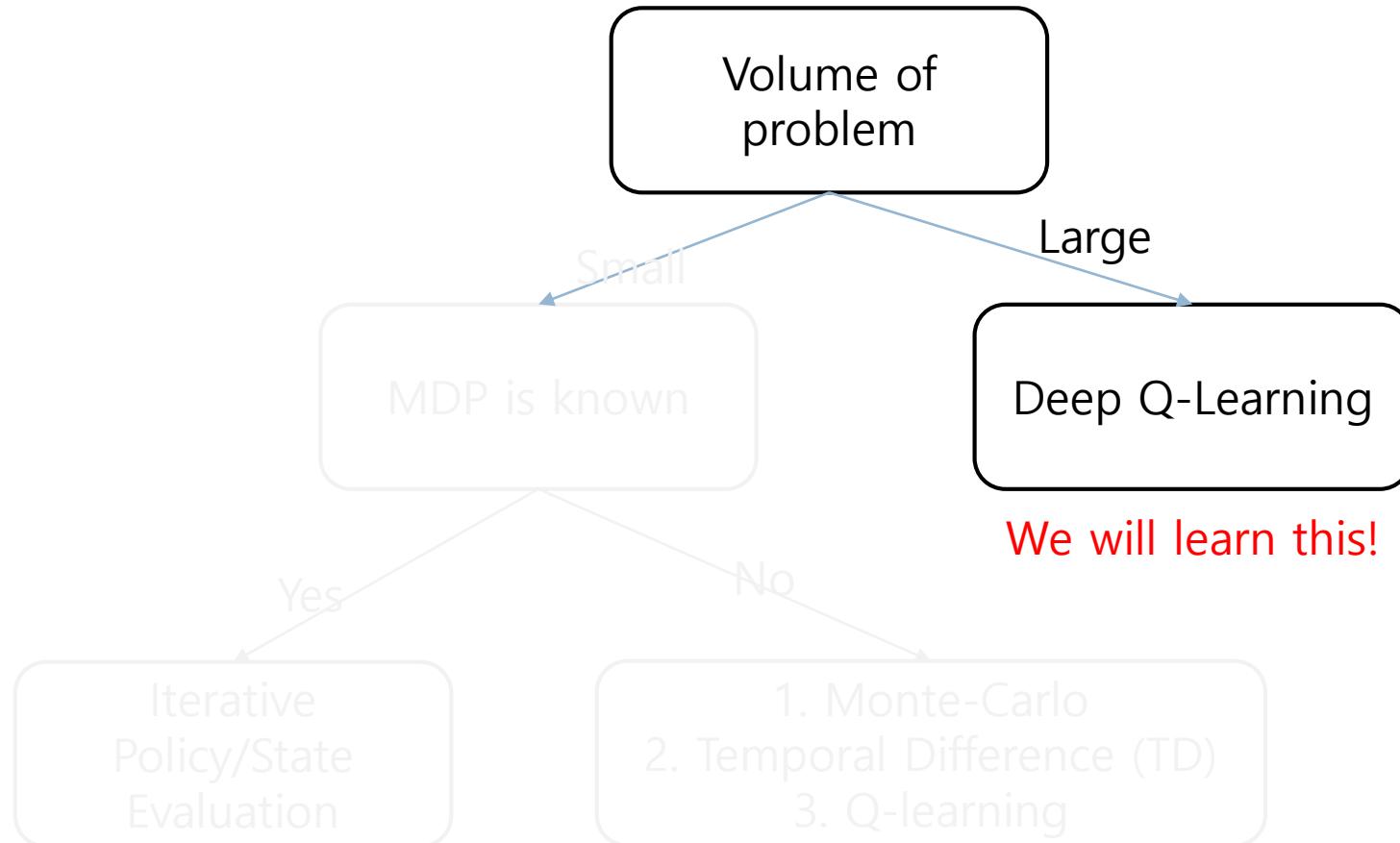
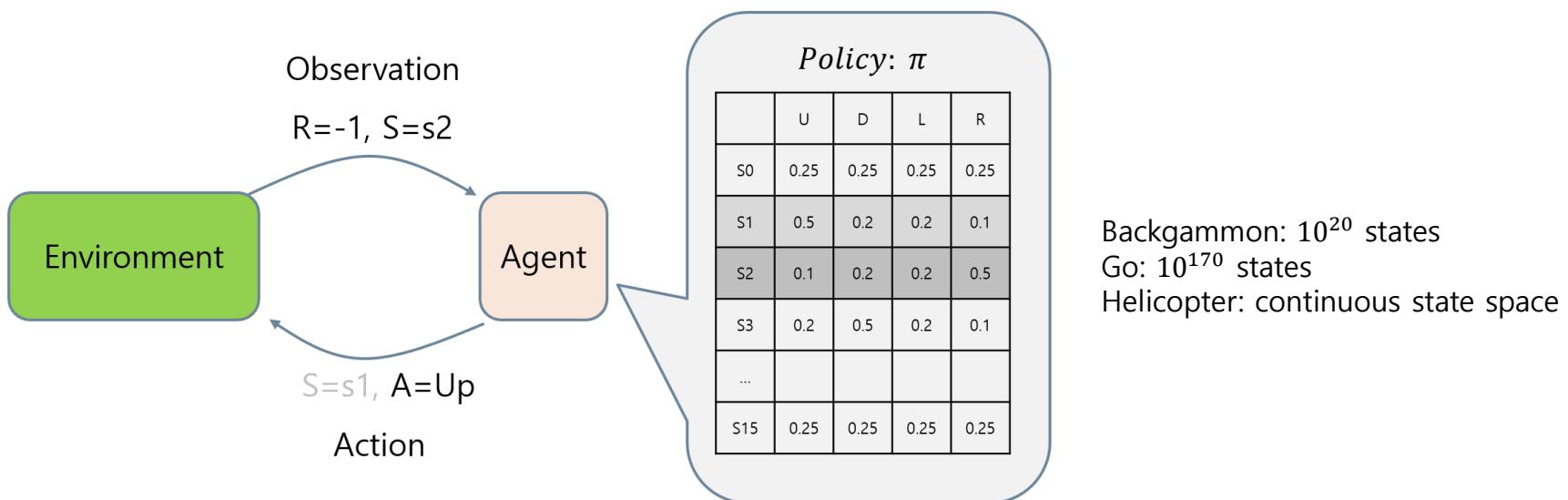


How to solve MDP problem



Limitation of Tabular Method

- Previous models use tabular method for planning and learning
 - They require a memory to store the policies for each state
 - But our real problem is more huge and complex
 - 큰 테이블을 사용하는 것은 비효율적
 - 실제 문제에서는 학습되지 않은 칸이 발생할 수 있음



Approximate Methods

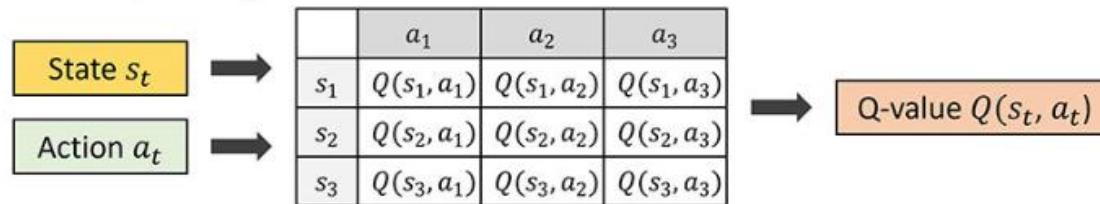
- **Solution for large MDPs:**

- Estimate value function with *function approximation*

$$\begin{array}{ccc} \text{Array} & \xleftarrow{\hspace{1cm}} & v_{\pi}(s) \approx \hat{v}(s; \theta) \\ & & \text{or } q_{\pi}(s, a) \approx \hat{q}(s, a; \theta) \\ & \xrightarrow{\hspace{1cm}} & \text{Neural Net} \end{array}$$

- Generalize from seen states to unseen states
- Update parameter θ using MC or TD learning

Classic Q-learning



Deep Q-learning

