Reinforcement Learning

(feat. SARSA, Q-Learning)

HY-KIERA

What is Reinforcement Learning?

What is SARSA / Q-Learning?

Let’s code them with PyTorch!

What is Reinforcement Learning?

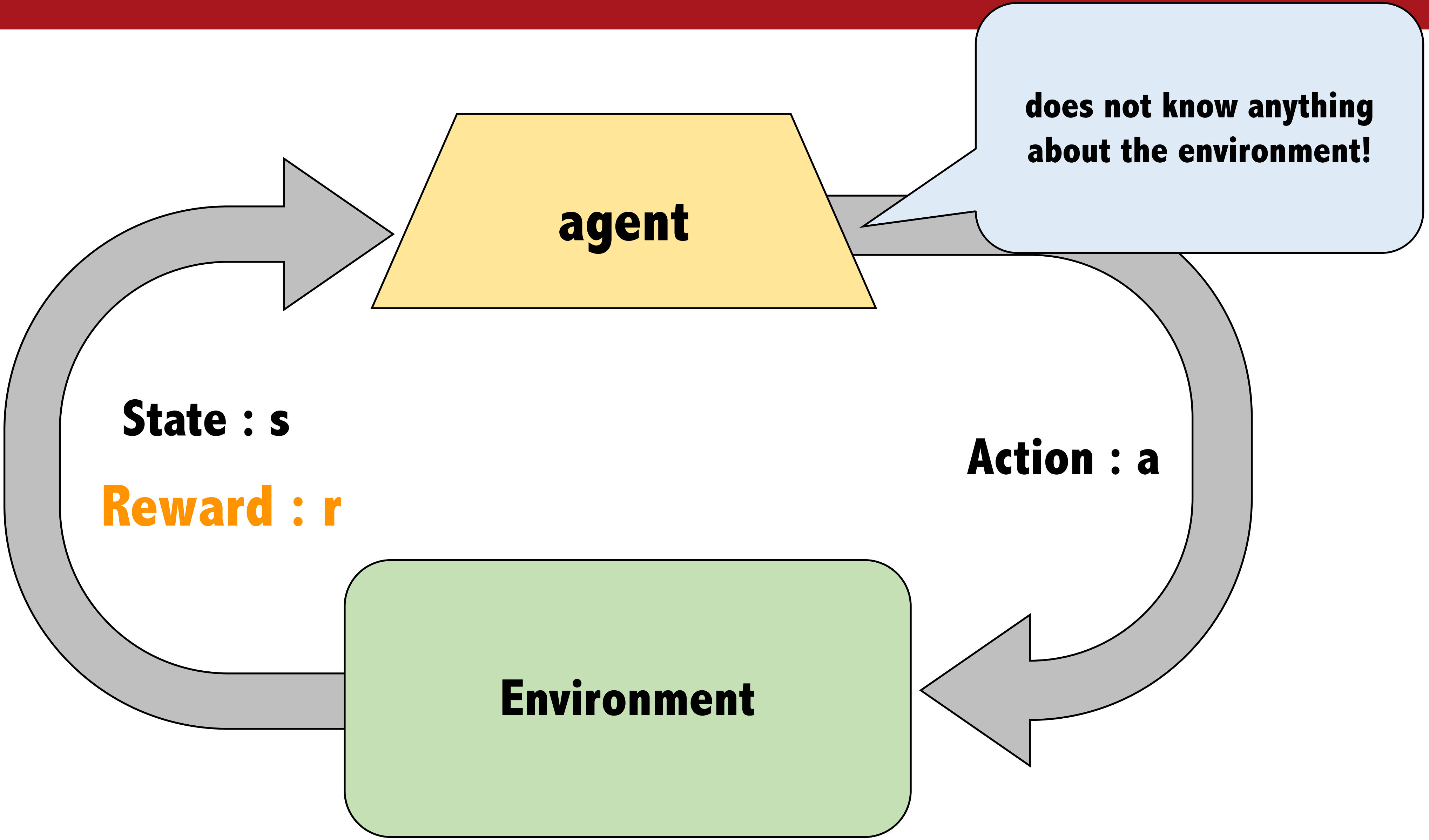
What is Reinforcement Learning?

This is Google's DeepMind AI teaching itself how to walk

https://www.youtube.com/watch?v=gn4nRCC9TwQ

What is Reinforcement Learning?

What is Reinforcement Learning?



What is Reinforcement Learning?

Action Value Function

Q(State,Action)

(Q-Table)

Explore + Exploit

ε-greedy

What is SARSA / Q-Learning?

What is SARSA / Q-Learning?

Monte Carlo + DP

Sarsa

an On-Policy algorithm for TD(Temporal-Difference) Learning

TD will learn using actual rewards
and future estimated values for the next step

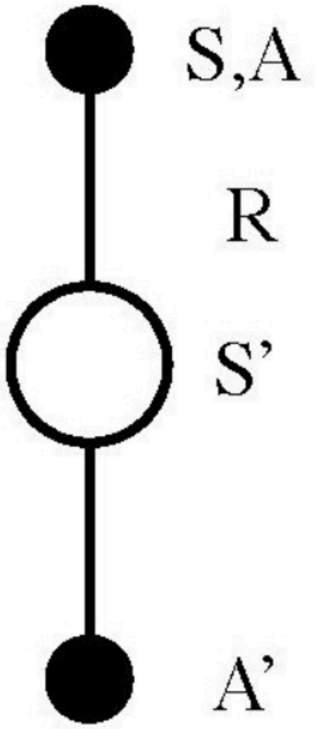
What is SARSA / Q-Learning?

Sarsa

an On-Policy algorithm for TD(Temporal-Difference) Learning

Learning can be done only if
learning policy and action policy are the same.

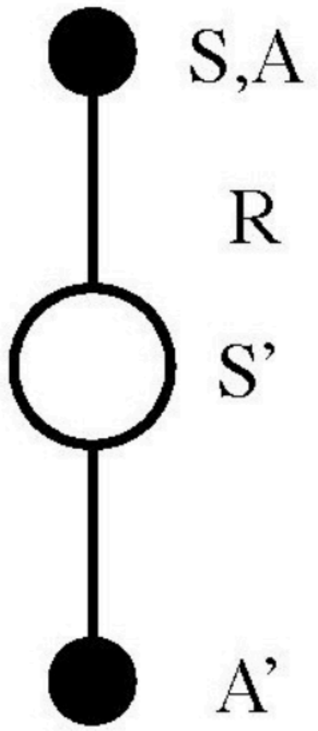
What is SARSA / Q-Learning?



$$Q(S_t,A_t) <- Q(S_t,A_t) + \underbrace{\alpha}_{\text{learning rate}}(R_{t+1} + \overset{\text{discount rate}}{\gamma}\underbrace{Q(S_{t+1},A_{t+1})}_{\text{Expected future reward}} - Q(S_t,A_t))$$

$$[S_t, A_t,\ R_{t+1},\ S_{t+1}, A_{t+1}]$$

What is SARSA / Q-Learning?

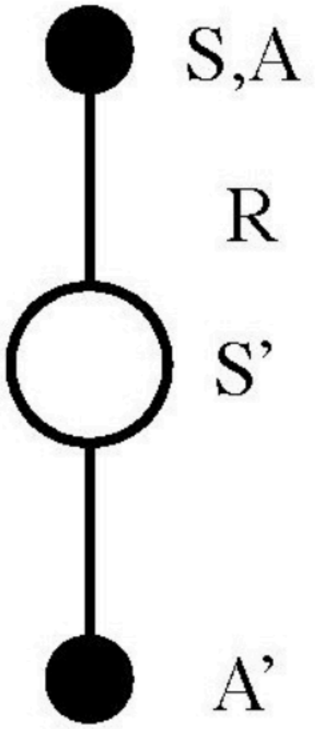


TD error

$$Q(S_t, A_t) \text{ <- } Q(S_t, A_t) + \alpha(\underbrace{R_{t+1} + \gamma Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)}_{\text{TD error}})$$

$$[S_t, A_t, R_{t+1}, S_{t+1}, A_{t+1}]$$

What is SARSA / Q-Learning?



→The agent starts in S, performs A,
and gets R, and goes to S’

→Now the agent chooses another action A’ from S’

→Then updates the value of A performed in S.

What is SARSA / Q-Learning?

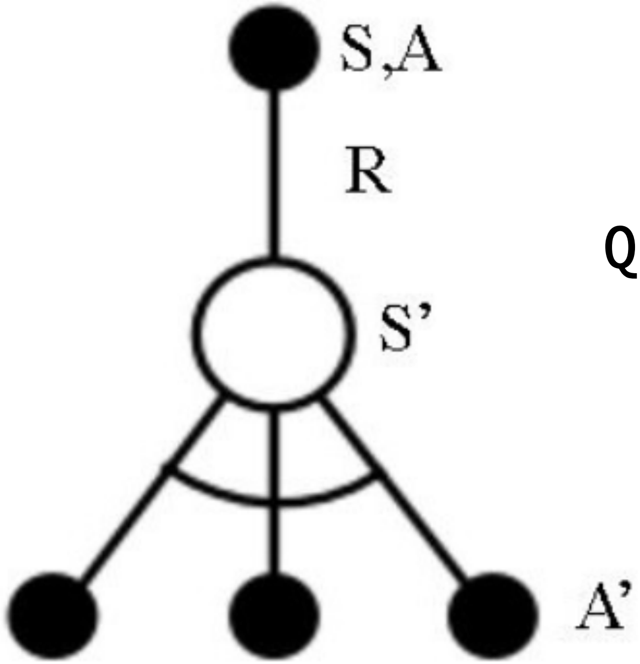
Q-Learning

an **Off-Policy** algorithm for TD(Temporal Difference) learning

Learning can be done if

learning policy and action policy aren’t the same.

What is SARSA / Q-Learning?

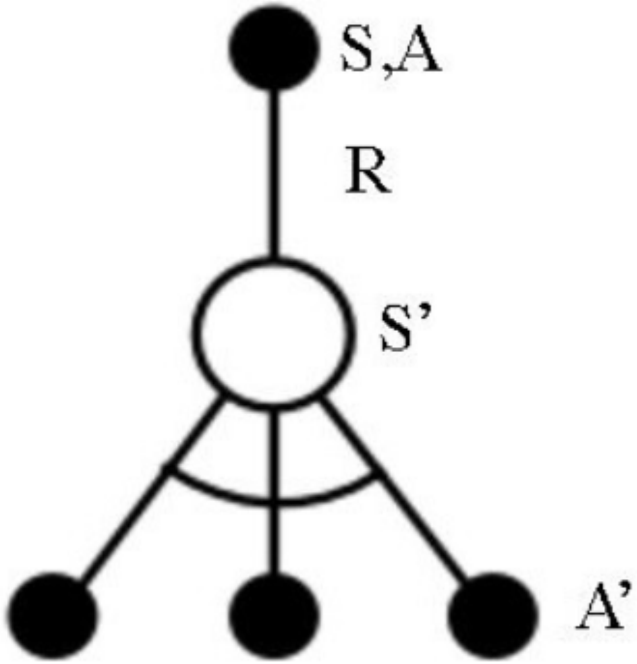


Maximum expected future reward

$$Q(S_t, A_t) \text{ <- } Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t))$$

$$[S_t, A_t, R_{t+1}, S_{t+1}]$$

What is SARSA / Q-Learning?

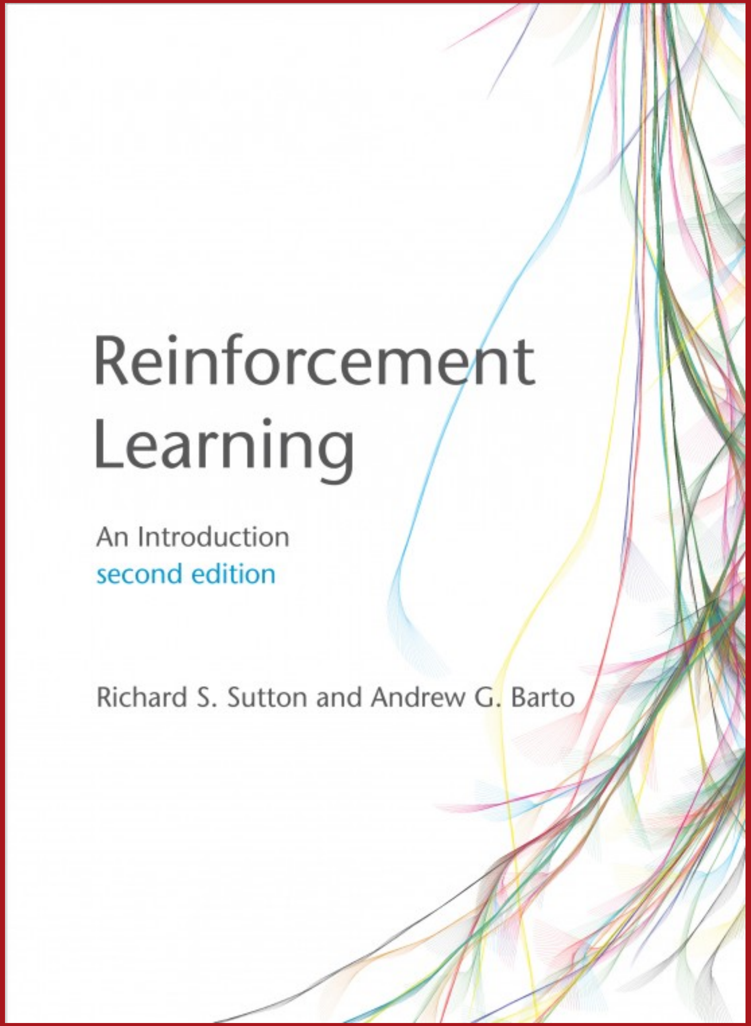


→The agent starts in S, performs A, and gets R, and goes to S'

→Now the agent chooses maximum action A' from S'

→Then updates the value of A performed in S.

Reference



Reinforcement Learning

An Introduction
second edition

Richard S. Sutton and Andrew G. Barto

PyTorch를 활용한
강화학습/
심층강화학습
실전 입문

파이토치로 익히는
기초 강화학습 및
심층강화학습
알고리즘의 원리와 구현

오가와 유타로 지음
/
심효섭 옮김

DS 데이터 사이언스 시리즈_025

Q러닝, Sarsa,
DQN, DDQN,
A3C, A2C

위키북스

Let’s code them with PyTorch!

https://github.com/hy-kiera/my-ai-study/tree/master/pytorch_drl

QnA