

# Deep Q Learning

- ‘Playing Atari with Deep Reinforcement Learning’, NIPS 2013
- ‘Human-level control through deep reinforcement learning’, Nature 2015
  - Author: David Silver, et al.



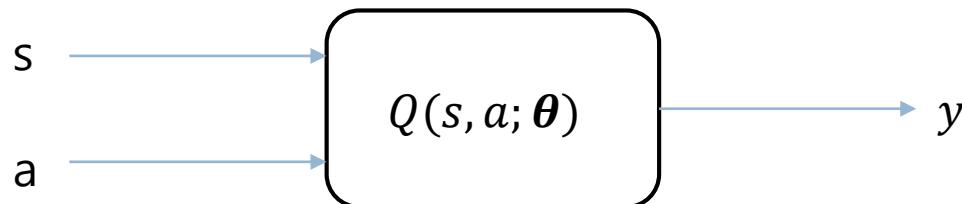
# Deep Q Learning

## Definition: Deep Q Learning

- 기존의 Q Learning을 딥러닝을 확장한 방법
- State가 입력되면 행동의 가치를 예측

$$Q_{\pi}(s, a) \approx Q(s, a; \theta)$$

- 상태가치  $V$ 는 사용하지 않음 (행동가치  $Q$ 만 사용)
  - 행동 비교 불가: 상태  $S$ 에서 여러 행동이 있을 때, 어떤 행동이 더 좋은지 알 수 없음
  - 정책 개선 불가: 주어진 정책이 얼마나 좋은지만 측정할 수 있고, 새로운 정책으로 업데이트 어려움
  - ->  $Q$ 를 사용하는 이유는 샘플링으로 부터 즉시 정책개선에 대한 학습이 가능



# Deep Q Learning

- 학습 방법

알고리즘	업데이트 방식	핵심 오차 개념
TD(0)	$V(s) = V(s) + \alpha[R + \gamma V(s') - V(s)]$	TD error
SARSA	$Q(s, a) = Q(s, a) + \alpha[R + \gamma Q(s', a') - Q(s, a)]$	TD error
Q-learning	$Q(s, a) = Q(s, a) + \alpha \left[ R + \gamma \max_{a'} Q(s', a') - Q(s, a) \right]$	TD error
DQN	$L(\theta) = [(R + \gamma \max_{a'} Q(s', a'; \theta^-)) - Q(s, a; \theta)]^2$	Loss (target   prediction)