

02

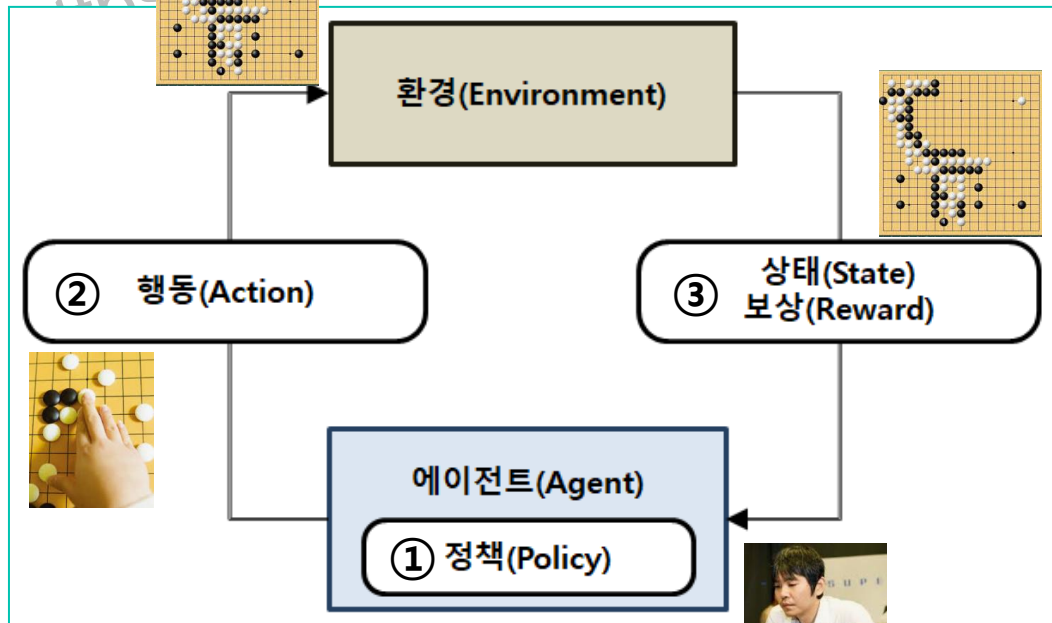
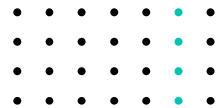
강화학습 기본개념

1. 기본개념과 확률과정



강화학습 개념

강화학습 구성요소



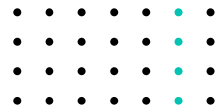
강화학습이란?

적절히 설계된 보상 체계를
활용해 에이전트가 긍정적인
행동을 할 수 있도록 에이전트
행동을 제어하는 정책을
찾아내는 최적화 기법



강화학습 개념

강화학습 구성요소



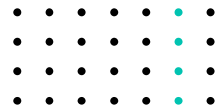
강화학습이란?

강화 학습은 아이가 발걸음을 배우는 과정과 같다. 칭찬과 꾸중을 통해 스스로 올바른 판단을 할 수 있는 정책을 만들어 간다.



강화학습 개념

Open AI Gym



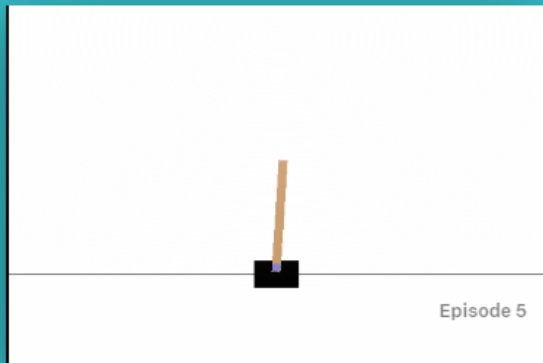
Open AI Gym

- 강화학습을 테스트할 수 있는 다양한 게임을 제공하는 사이트
- Cartpole 게임을 활용해서 강화학습을 설명

<https://gym.openai.com/envs/CartPole-v1/>

CartPole-v1

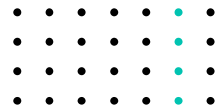
A pole is attached by an un-actuated joint to a cart, which moves along a frictionless track. The system is controlled by applying a force of +1 or -1 to the cart. The pendulum starts upright, and the goal is to prevent it from falling over. A reward of +1 is provided for every timestep that the pole remains upright. The episode ends when the pole is more than 15 degrees from vertical, or the cart moves more than 2.4 units from the center.



RandomAgent on CartPole-v1

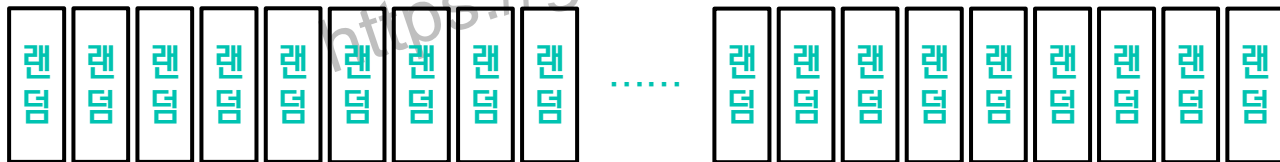


확률과 확률과정 확률



확률

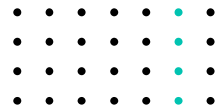
- 어떤 사건이 실제로 일어날 것인지 혹은 일어났는지에 대한 **지식** 혹은 **믿음**을 표현하는 방법
- 같은 원인에서 특정한 결과가 나타나는 **비율**
- 짧은 시간에서는 랜덤, 긴 시간에서는 확률(규칙)**



확률



확률과 확률과정 조건부확률



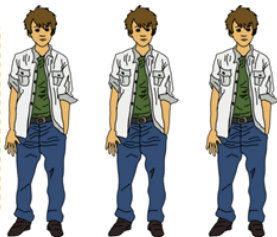
조건부확률

확률의 특수한 경우로 어떤 **특정한 조건** 아래에서 발생하는 확률

$\frac{2(\text{노트북})}{5(\text{남학생})}$



$\frac{5(\text{노트북})}{10(\text{전체학생})}$



조건부 확률
 $P = (B|A)$
(노트북|남학생)

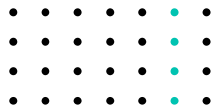


전체 학생 중 한 명을 선택했을 때 노트북을 가지고 있을 확률

남학생 중 한 명을 선택했을 때 노트북을 가지고 있을 확률



확률과 확률과정 확률과정



확률과정

- 확률(Stochastic)의 개념과 과정(Process)의 개념이 합쳐진 용어
- 시간의 흐름에 따라 확률적(무작위적)으로 움직이는 상태

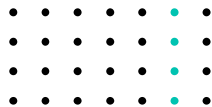
$$\{ X_t \}$$

- X : 랜덤 변수
- t : 시간
- $\{ \}$: 집합

시간의 흐름에 따라 발생하는 랜덤 변수의 집합



확률과 확률과정 확률과정



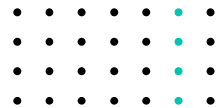
브라운 운동

- 1827년 스코틀랜드 식물학자 로버트 브라운(Robert Brown)이 발견한 현상
- 물위에 꽃가루 입자가 불규칙적으로 운동하는 현상을 이론적으로 설명한 것
- 아인슈타인은 브라운 운동의 원인이 액체 분자에 의한 충돌 때문이라는 것을 밝혀냈으며 이러한 현상을 수학적식 정리
- 통계학, 열역학, 경제학 분야에 활용

https://en.wikipedia.org/wiki/Brownian_motion



정리



강화학습 개념

Open AI Gym

확률

조건부 확률

확률과정

브라운운동

