



Chapter

6

강화학습(Reinforcement Learning)

1. 강화학습 사례
2. 강화학습
3. 강화학습 실습

강의에 앞서서..

- ❖ 본 문서는 아래의 자료들을 활용하여 만들어 졌음을 알립니다
- ❖ 모두를 위한 딥러닝 강좌
 - <https://hunkim.github.io/ml/>
- ❖ Hands on Machine Learning
 - <https://github.com/ugaemi/handson-ml-1>

강화학습 사례: AlphaGo



Google DeepMind, AlphaGo

- 강화학습의 가장 유명한 사례
- 프로 기사를 맞바둑으로 최초로 이긴 프로그램
- 한국기원 프로 9단
- ELO Rating 세계 1위
- 이세돌과의 대국 이후, 개량된 버전 출시
 - 커제, 박정환 등 프로 기사와의 인터넷 대국에서 60연승
 - 커제 9단과 2017년 5월 23일 3번기, 전승
- 2017년 12월 14일 알파고 개발 완전 종료 선언

강화학습 사례: OSARO



OSARO

- 산업용 로봇을 위한 기계학습 소프트웨어 개발사, 창업 후 실리콘벨리에서 330만 달러 자금 조달
- '18. 6월 일본 FOOMA JAPAN에서 식품 조립 작업이 가능한 FoodPick 공개
- Innotech(반도체 전자 설계 서비스), DENSO WAVE(산업용 로봇 제조 업체)와 파트너십 체결

강화학습 사례: Preferred Network



Preferred Network

- 일본 도쿄의 AI 스타트업, 도요타와 자율주행 기술 개발
- 자동차 센서의 출력 데이터를 이용한 시뮬레이션을 통해 실제 공간에서보다 100만배 빠른 시뮬레이션이 가능

강화학습 사례: Blade and Soul



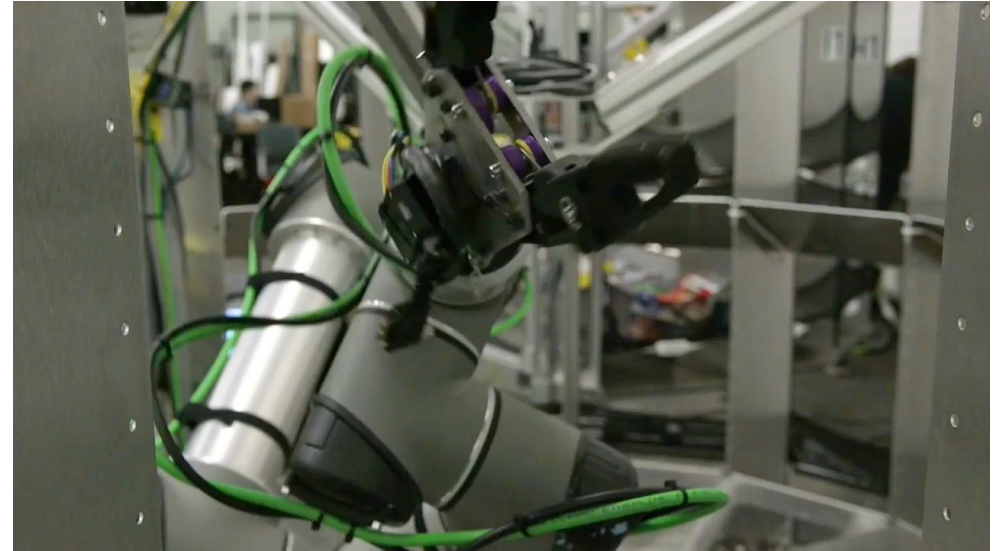
Blade and Soul, NC Soft

- NCSoft의 17회 대한민국 게임대상 수상작
- 무한의 탑이라는 총 100개의 층을 가지는 던전에서, 층마다 강해지는 적을 강화학습을 통해 구현

강화학습 사례: The others

Kindred

- 로봇이 대면하지 않았던 특정 제품을 인식한 후 정리할 수 있는 솔루션을 개발
- 물류 센터 처리 과정 개선에 활용



Microsoft

- MSN.com의 헤드라인 뉴스를 추천하는데 강화학습을 이용
 - 홈페이지 방문자가 제시된 링크를 클릭한 회수를 **보상**으로 활용

강화학습 사례:

- “Playing Atari with Deep Reinforcement Learning”의 Deep Q-Learning



<https://youtu.be/V1eYniJ0Rnk>

강화학습(Reinforcement Learning)

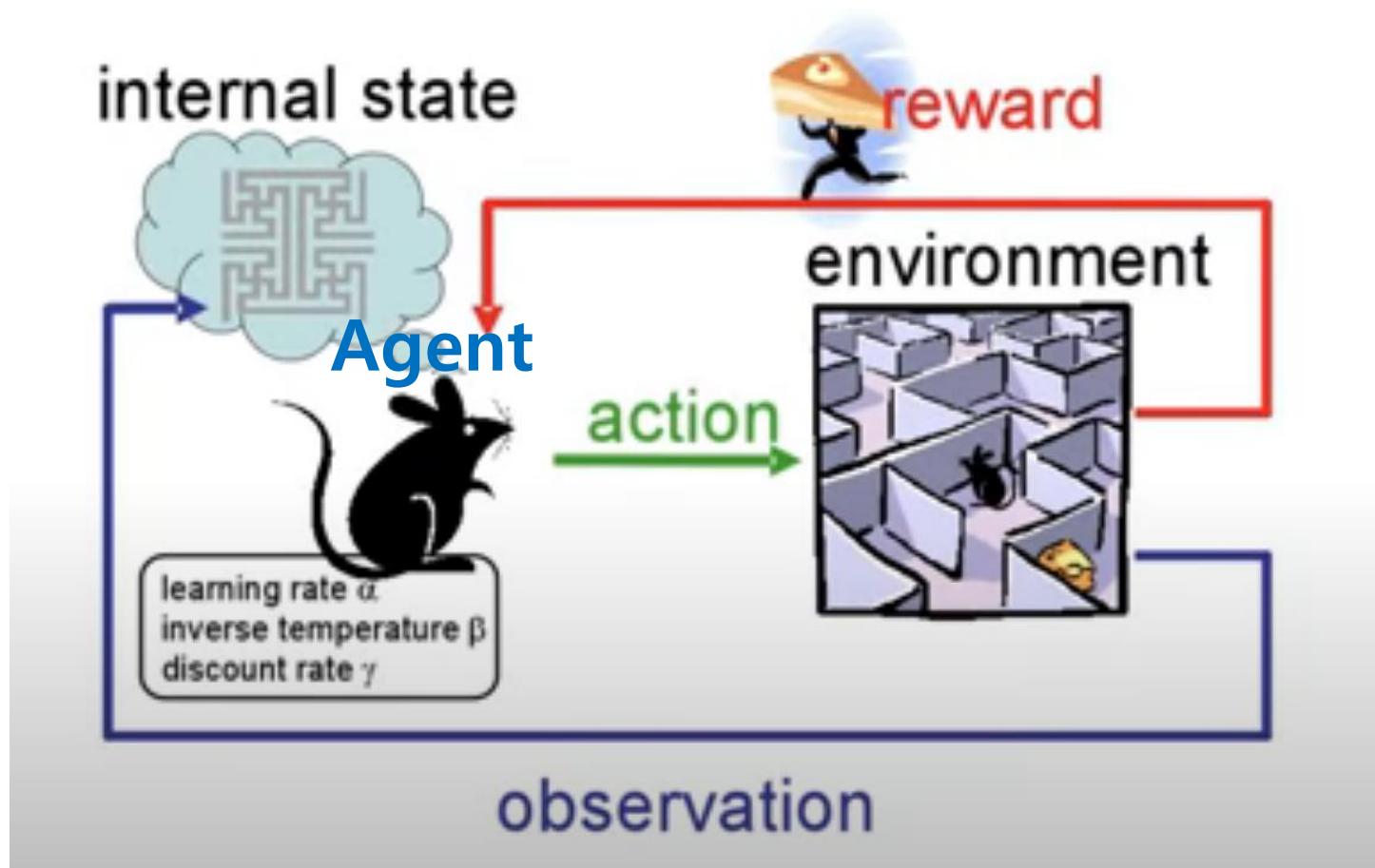
❖ What is Positive Reinforcement Dog Training ?



<https://www.thesprucepets.com/positive-reinforcement-dog-training-1118248>

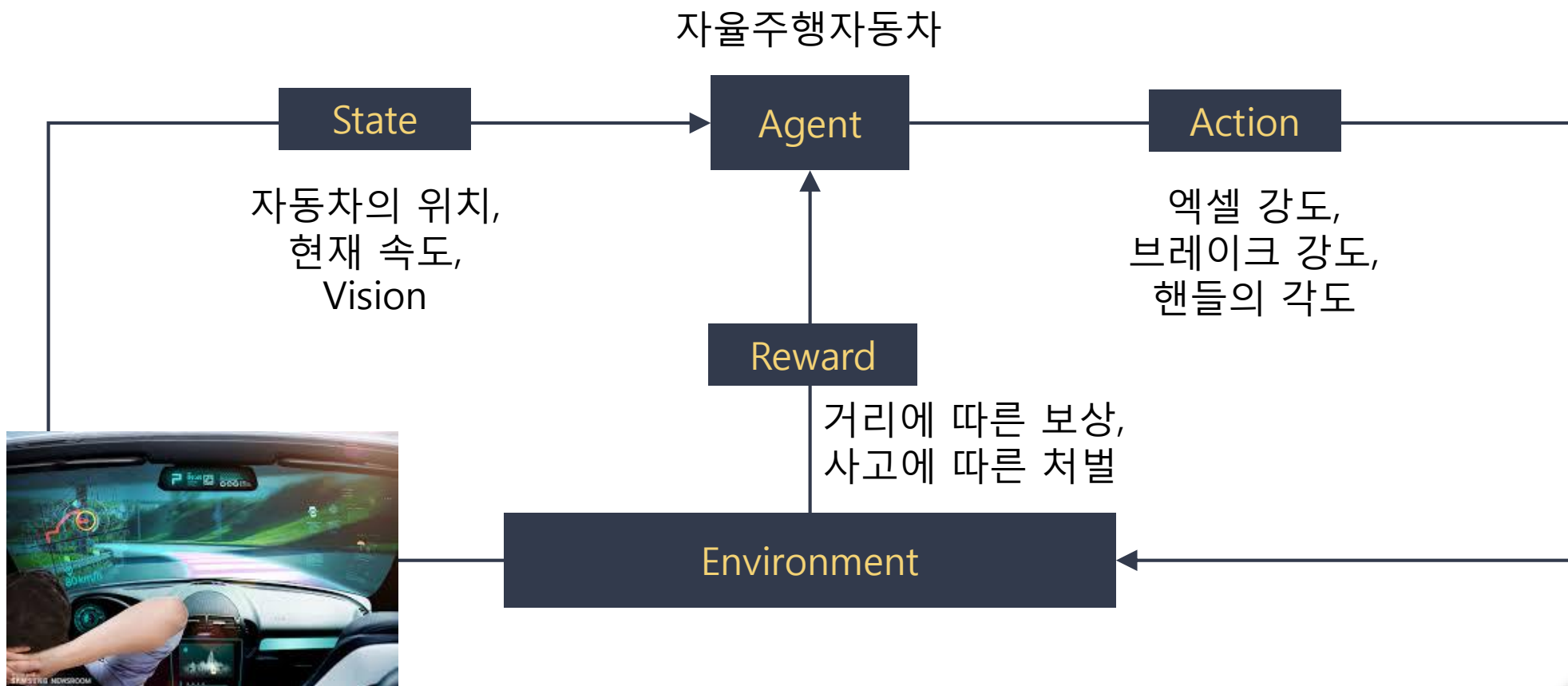
강화학습(Reinforcement Learning)

- Agent 는 주어진 환경(Environment)을 관측(observation)을 하고 상황(State)에 따라 행동(Action)하며 보상(Reward)에 의해 강화 (Reinforcement) 된다.



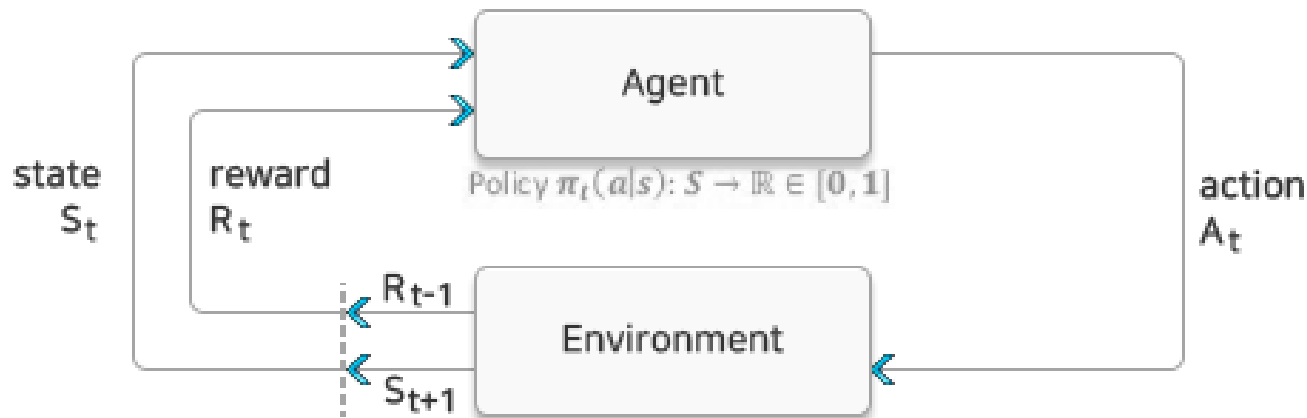
강화학습

자율 주행 자동차



강화학습(Reinforcement Learning)

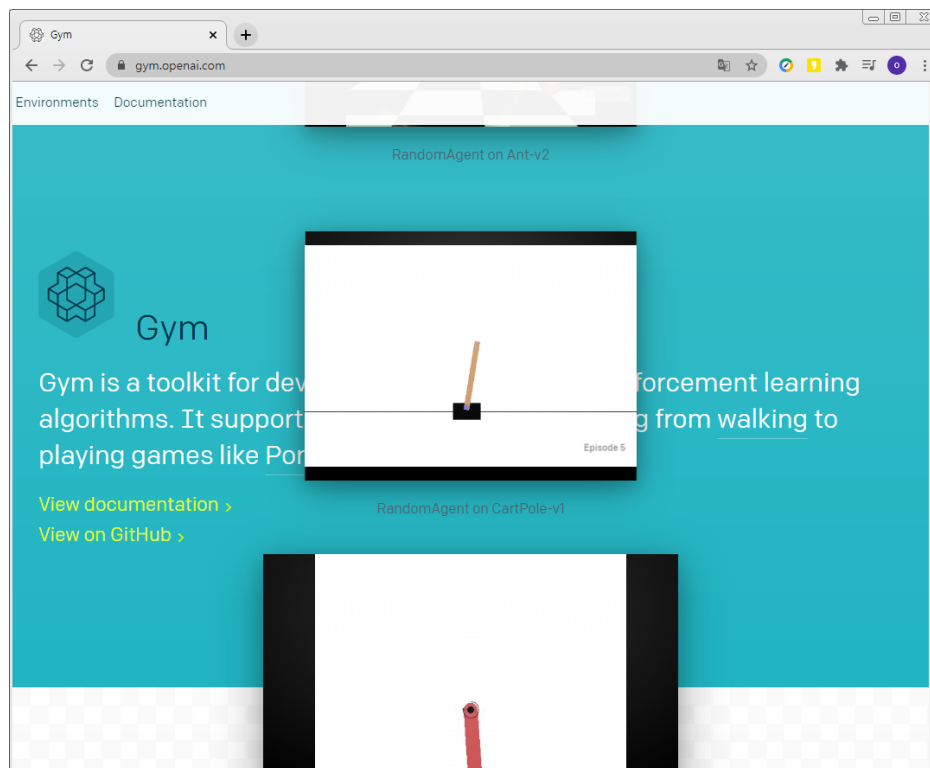
- ❖ 환경 안에서 행동하고 시행착오를 통하여 최적의 행동패턴을 학습
- ❖ Policy(정책)
 - 에이전트가 행동을 결정하기 위한 알고리즘
 - 학습을 통하여 정해진 행동패턴
- ❖ 강화학습의 목적
 - 기대 누적 보상값(expected accumulated reward)이 최대가 되는 정책을 찾는 문제



OpenAI GYM을 활용한 강화학습 실습

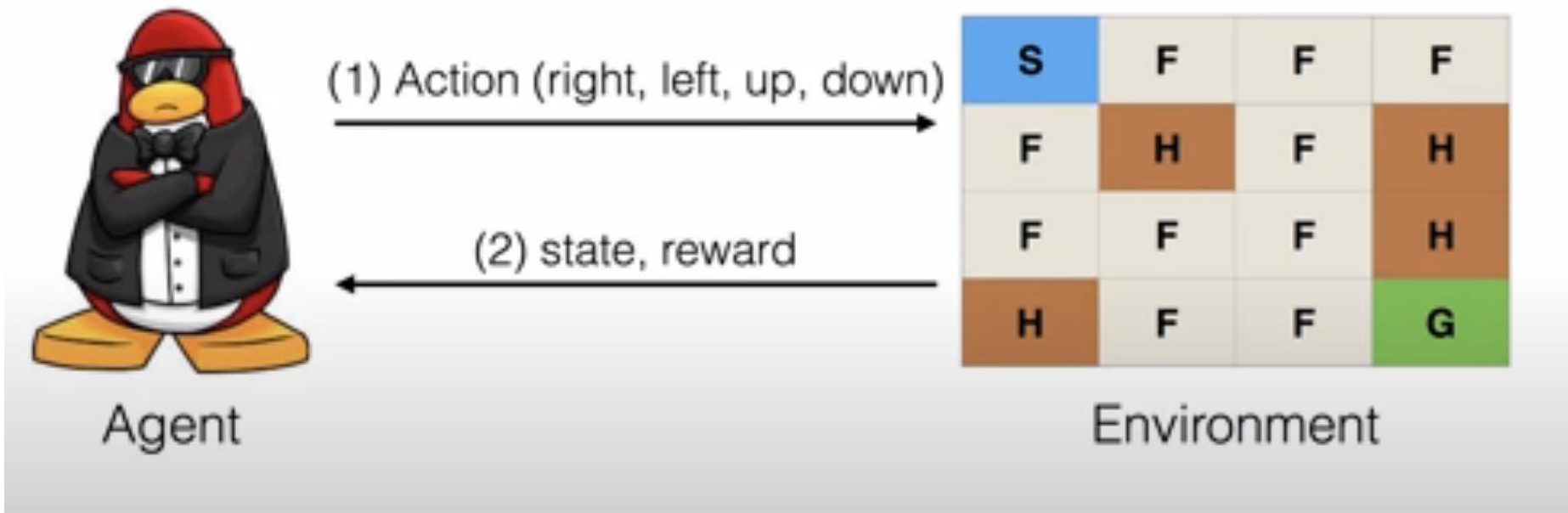
❖ 게임 시뮬레이션 환경 도구

- 에이전트를 학습시킬 시뮬레이션 환경
- <https://gym.openai.com>
- 설치 : `pip install --upgrade gym`



강화학습(Reinforcement Learning)

❖ Frozen Lake World (OpenAI GYM)



강화학습(Reinforcement Learning)

❖ Frozen Lake World (OpenAI GYM)



(1) Action (Right, left, up, down)



(2) state (observation), reward



S	F	F	F
F	H	F	H
F	F	F	H
H	F	F	G

```
import gym
env = gym.make("FrozenLake-v0")
observation = env.reset()
for _ in range(1000):
    env.render()
    action = env.action_space.sample() # your agent here (this takes random actions)
    observation, reward, done, info = env.step(action)
```

강화학습(Reinforcement Learning)

❖ Frozen Lake World (OpenAI GYM)

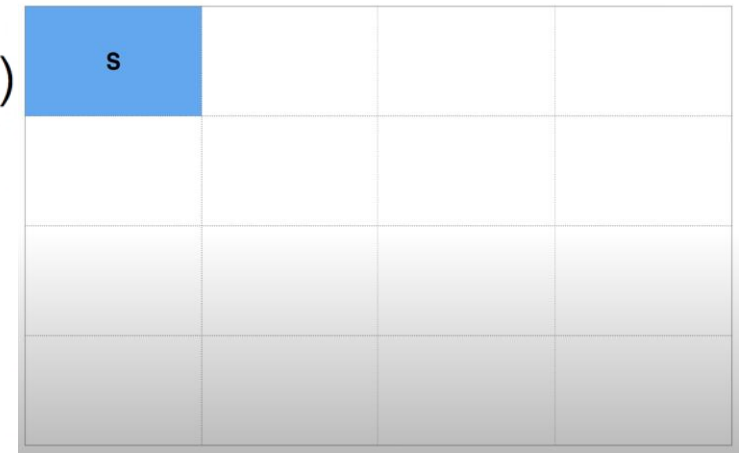


Agent

(1) Action (Right, left, up, down)



(2) state, reward



Environment

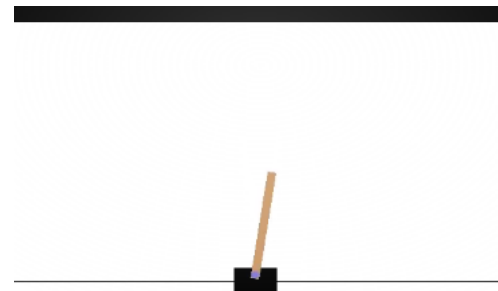
CartPole 시뮬레이션을 이용한 강화 학습

- OpenAI Gym에서 제공하는 CartPole-V1 환경을 사용한 강화학습
- <https://gym.openai.com/envs/CartPole-v1/>

CartPole-v1

카트 위에 막대를 세워놓고, 막대가 쓰러지지 않도록 좌우로 움직이는 게임

- 주요함수
 - `gym.make('CartPole-v1')` : CartPole-v1 게임을 위한 gym 환경 생성
 - `env.reset()` : 환경 초기화
 - `env.step()` : 값을 주어 `obs`, `reward`, `done`, `info` 반환
 - `obs`: 현재 상황 ([카트 위치, 카트 속도, 막대의 각도, 막대의 각속도])
 - `reward`: 보상 (이 게임은 오래 버티는게 목표이므로, 보상은 막대가 떨어지지 않는다면 1을 제공)
 - `done`: 막대가 떨어졌는지를 의미, 막대가 떨어진 경우는 막대의 각도가 12도 이상 기울었거나, 카트가 2.4칸 이상 움직인 경우(화면에서 나간 경우)
 - `env.render()` : 화면에 게임 출력



Episode 5