

# 강화학습의 개요

'파이썬과 케라스로 배우는 강화학습' 도서 참고

KAIG 세미나 (2019-02-28)

김수환

# 강화학습 개요

1. 강화학습 개념
2. 강화학습 문제
3. 정리

# 1. 강화학습 개념

# 1. 강화학습 개념

---

행동심리학과 강화학습

기계학습과 강화학습

# 1. 강화학습 개념

---

행동심리학과 강화학습

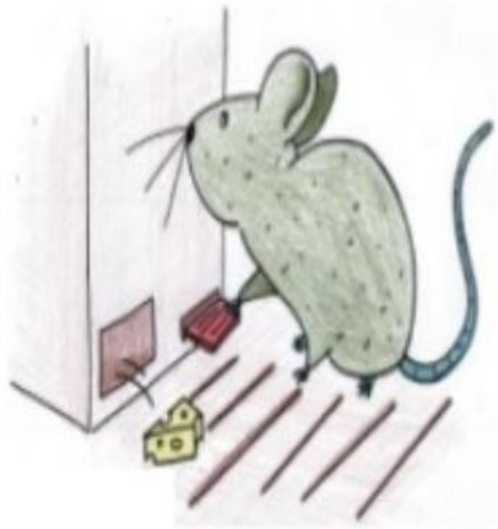
기계학습과 강화학습

강화(Reinforcement)는 동물이  
시행착오(Trial and Error)를 통해 학습하는 방법 중 하나

## 스키너의 쥐 실험

# 행동심리학과 강화학습

• 스키너의 쥐 실험



1. 굶긴 쥐를 상자에 넣는다.
2. 쥐는 돌아다니다가 우연히 상자 안에 있는 지렛대를 누르게 된다.
3. 지렛대를 누르자 먹이가 나온다.
4. 지렛대를 누르는 행동과 먹이와의 상관관계를 모르는 쥐는 다시 돌아다닌다.
5. 그러다가 우연히 쥐가 다시 지렛대를 누르면 쥐는 이제 먹이와 지렛대 사이의 관계를 알게 되고 점점 지렛대를 자주 누르게 된다.
6. 이 과정을 반복하면서 쥐는 지렛대를 누르면 먹이를 먹을 수 있다는 것을 학습한다.



# 행동심리학과 강화학습

---

강화의 핵심은 바로 보상을 얻게 해주는 행동의 빈도 증가

# 행동심리학과 강화학습

---

## • 스키너의 쥐 실험



쥐는 페달을 밟았을 때 왜 먹이가 나오는지 모르지만  
페달을 밟을 때마다 먹이가 나온다는 건 알게 된다.

# 1. 강화학습 개념

---

행동심리학과 강화학습

기계학습과 강화학습

# 기계학습과 강화학습

---

기계학습(Machine Learning)이란?

인공지능의 한 범주로서  
컴퓨터가 스스로 학습하게 하는 알고리즘을 개발하는 연구 분야

종류 : 지도학습, 비지도학습, 강화학습

# 기계학습과 강화학습

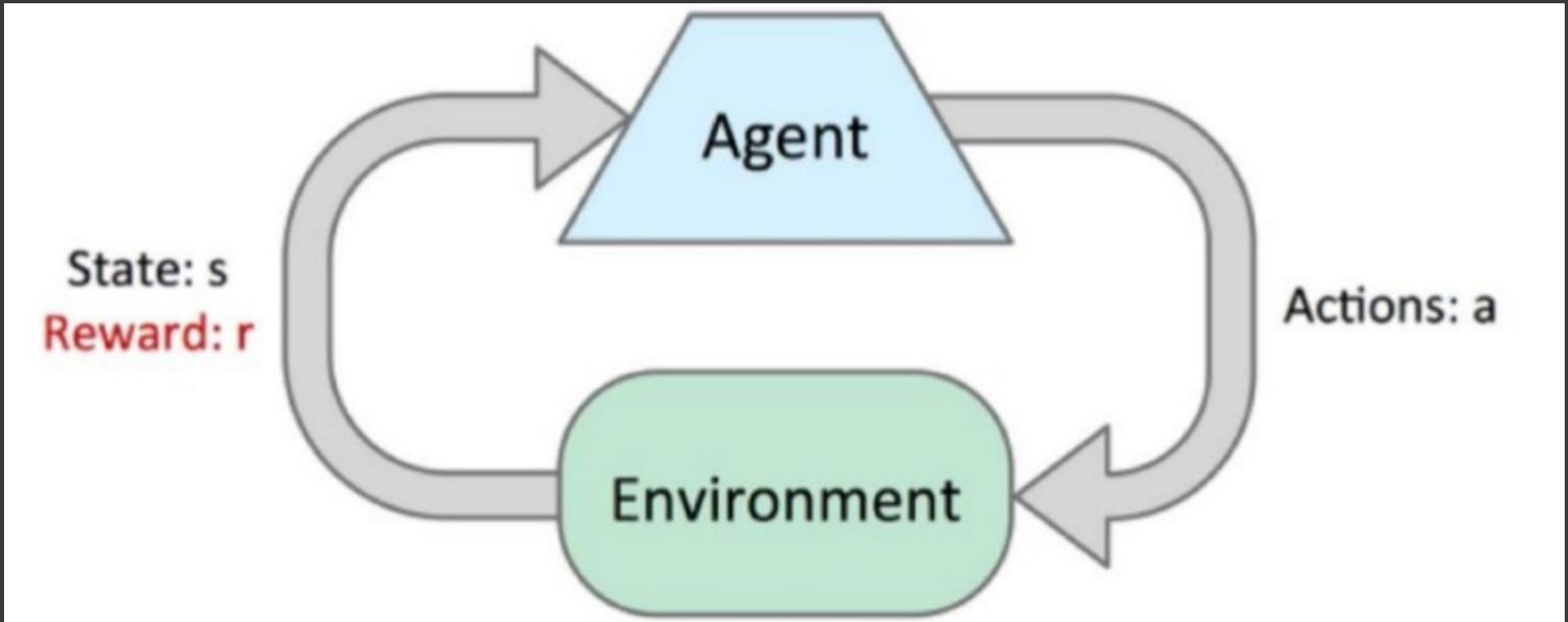
---

## 강화학습(Reinforcement Learning)

- 보상(Reward)을 통해 학습
- 보상은 컴퓨터가 선택한 행동(Action)에 대한 환경의 반응
- 따라서 행동의 결과로 나타나는 보상을 통해 학습
- 보상을 얻게 하는 행동을 점점 많이 하도록 학습

# 기계학습과 강화학습

---



강화학습의 목적 : "최적의 행동양식 또는 정책"을 학습하는 것

강화학습의 장점이란?

환경에 대한 사전지식이 없어도 학습한다는 것



## 2. 강화학습 문제

## 2. 강화학습 문제

---

강화학습은 결정을 순차적으로 내려야 하는 문제에 적용

## 2. 강화학습 문제

---

순차적으로 내려야 하는 문제를 정의할 때 사용하는 방법

MDP(Markov Decision Process)

# MDP(Markov Decision Process) – 마르코프 결정 과정

---

- 상태(State) : 정적인 요소 + 동적인 요소(ex. 속도, 가속도 등)
- 행동(Action) : 어떠한 상태에서 취할 수 있는 행동(ex. 상,하,좌,우)
- 보상(Reward) : 에이전트가 학습할 수 있는 유일한 정보  
(다른 머신러닝 기법과 다른 가장 핵심적인 요소!!)
- 정책(Policy) : 순차적 행동 결정 문제(MDP)에서 구해야 할 답  
모든 상태에 대해 에이전트가 어떤 행동을 해야하는지 정해놓은 것

순차적 행동 결정 문제를 풀었다? -> 'Optimal Policy'(최적정책)를 찾았다!