

시즌 RL - Lecture 04

노트북: 모두를 위한 머신러닝
만든 날짜: 2019-01-17 오후 4:18
작성자: rr
태그: #모두를 위한, .Lecture

수정한 날짜: 2019-01-17 오후 4:34

Lecture 04

= Dummy Q-learning 문제점

- Exploit

현재 있는 값을 이용한다

- Exploration

모험을 한다, 안 가본 곳을 간다

= E-greedy

$\epsilon = 0.1$

ϵ 보다 작으면 랜덤하게, ϵ 보다 크면 알고 있는 길로 간다

= decaying E-greedy

$\epsilon = 0.1 / (i + 1)$

학습 초반에는 랜덤하게 많이, 학습 후반에는 알고 있는 곳에 많이 가게 한다

= add random noise

$a = \operatorname{argmax}(Q(s, a) + \text{random_values} / (i + 1))$

완전한 랜덤이 아니라 1번 말고 2, 3번을 선택할 때 도움이 되는 방식

= Learning Q (s, a) with discounted reward

현재 받는 상의 값어치를 높이고 그 다음에 받을 상은 줄여서 받음

$r + 0.9 * (\max Q(s', a'))$

$$\hat{Q}(s, a) \leftarrow r + \gamma \max_{a'} \hat{Q}(s', a')$$