D3_Exercises

Kwon do yun

2021-01-18

차 례

Exercise 1

How would you genalize this game with arbitraty value of m_1 (minimum increment),m_2 (maximum increment), and N (the winning number)?

if $m_1 = 1$ $m_2 = 2$ $N = 31$ You have to be in 28, 25, $\cdots$ states to win the game.

if $m_1 = 2$ $m_2 = 5$ $N = 50$ You have to be in 43, 36, $\cdots$ states to win the game.

If you generalize the expression,

$$\pi^*(S) = N - k(m_1 + m_2) - S,\ where\ \ N - k(m1 + m2) - s \in [m1, m2]$$

Exercise 3

There is only finite number of $deterministic\ stationary$ policy. How many is it?

There is a fixed state for each action.

So, $|A|^{|S|}$

Exercise 4

Formulate the first example in this lecture note using the terminology including state,action, reward, policy, transition. Describe the optimal policy using the terminology as well.

State

$S = \{1, 2, ..., 31\}$

Action

$A = \{a_1, a_2\}$

Reward

$R(30, a_1) = R(29, a_2) = 1$ all ohter $R(s, a)$ = 0

Transition

$P^a_{ss'} = P(S_{t+1} = S'|S_t = s, A_t = a) = 1$

$s' = s + 1,\ if(a = a1)$

$s' = s + 2,\ if(a = a2)$

otherwise 0.

Optimal Policy

$\pi^* = argmax_\pi V_t(s)^\pi$

$S(3n-1) = a_2 $

$S(3n) = a_1$

```
"D3_Exercises"
```