D3_Jeong,wonryeol

Jeong, wonryeol

2021-01-19

Contents

Exercise 1

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

- $S = \{1, 2, 3, 4....N\}$
- $A = \{a_{m1}, a_{m1} + 1, a_{m1} + 2, a_{m1} + 3, ...a_{m2}\}$
- $a_m =$ "increment by m
- $S_{next} = S_{now} + a_{now}$

Exercise 2

Two players are to play a game. The two players take turns to call out integers. The rules are as follows. Describe A's winning strategy.

- A must call out an integer between 4 and 8, inclusive.
- B must call out a number by adding A's last number and an integer between 5 and 9, inclusive.
- A must call out a number by adding B's last number and an integer between 2 and 6, inclusive.
- Keep playing until the number larger than or equal to 100 is called by the winner of this game.

To get optimal starategy, we should analysis number

B can add at least 5 and at most 9 and A can add at least 2, at most 6

it means if one step is gone, at least 7 is increased and at most 15 is increased

For winning A, A should find out one step increasing. that is 11

Exercise 3

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

Answer is $|A|^{|S|}$

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.

A and B are to play a game. They take turn to call out integers. \ 1. The serving player must call out an integer between 1 or 2. \ 2. The opponent player 1) takes the other player's number and 2) increments it by 1 or 2, then 3) call out the number. \ 3. Keep playing back and forth until someone calling out the number 31. The person calling out 31 is winner. }% }

- State = $1 \leq S_t \leq 31$
- a_1 is increased by 1
- a_2 is increased by 2

Policy

---

if S_t % 3 == 1: a_2 elif S_t % 3 == 2: a_1 else: you loose

Reward

- $R(29, a_1) = 1$
- $R(28, a_1) = 1$
- otherwise 0

Transition

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

Exercise 5

From the first example

- Assume that your opponent increments by 1 with prob 0.5 and by 2 with prob 0.5
- Assume that the winning number is 0 instead of 31
- your opponent played first and she called out 1
- your current a policy π_0 is that
  - if the current state s 5 then increment by 2
  - if the current state s > 5 then increment by 1

Evaluate $V^{\pi_0}(1)$

$S_1 = 1$

while $S_t \leq 10$

P = random Prob

if P > 0.5

$S_t = S_t + 1$

elif P 0.5

$S_t = S_t + 2$

If $S_t == 10$ # when you lose

return 0

If $S_t \leq 5$

$S_{t+1} = S_t + 2$

If $S_t > 5$

$S_{t+1} = S_t + 1$

If $S_t == 10$ # when you win

return 1