吴徒遥

517370910257

Homework 3 Written

June 10rd, 2020 at 11:59pm

(a)
By iterated dominance equilibrium, we eliminate less optimal strategies, leaving the best strategy for one player. That means that with this policy, the player would not like to switch to other strategy if he plays optimally no matter the choices of others, which is the definition of a Nash equilibrium

1 Nash Equilibrium and Iterated Dominance Equilibrium

(a) Show that every iterated dominance equilibrium s^* is a Nash equilibrium.

(b) Show by a counter example that not every Nash equilibrium can be generated by iterated dominance.

(b)

	C	D
A	-39, -39	39, -17
B	-17, 39	0, 0

2 Game in Matrix

(a) (a) Consider the game with the following bimatrix:

(b, A) $2x<1$ or $2x<2 \Rightarrow x<1$
(a, B) $x<1$ or $x<0 \Rightarrow x<1$
(c, C) $x^2<3$ or $x^2<2 \Rightarrow x^2<3$
or $x>4$

P2

P1	A	B	C
a	$\cancel{1, 1}$	$\cancel{3, x}$	$\cancel{2, 0}$
b	$\cancel{2x, 3}$	$\cancel{2, 2}$	$\cancel{3, 1}$
c	$\cancel{2, 1}$	$\cancel{1, x}$	x^2, 4

(b) $x^2 \geq 3$, $x^2 \geq 2$, $4 \geq x$, $4 \geq 1$

(b) $x \in (-\infty, -\sqrt{3}] \cup [\sqrt{3}, 4]$

(a) $x \in (-\sqrt{3}, 1)$

(a) Find x so that the game has no pure Nash equilibrium.

(b) Find x so that the game has (c, C) as pure Nash equilibrium.

3 Nash Equilibrium

Consider the game in which two players choose nonnegative integers no greater than 1000. Player 1 must choose an even integer, while player 2 must choose an odd integer. When they announce their number, the player who chose the lower number wins the number she announced in dollars. Find the Nash equilibrium.

Player 1 chooses 0, player 2 chooses 1.

For any other situations, there are no Nash equilibriums. When player 1 chooses $2n$ ($n \in Z^+$), player 2 can get better utilities by choosing $2n-1$. When player 2 chooses $2m+1$ ($m \in Z^+$), player 1 can get better with $2m$.

4 MDPs: Dice Bonanza

A casino is considering adding a new game to their collection, but need to analyze it before releasing it on their floor.

They have hired you to execute the analysis. On each round of the game, the player has the option of rolling a fair 6-sided die. That is, the die lands on values 1 through 6 with equal probability. Each roll costs 1 dollar, and the player **must** roll the very first round. Each time the player rolls the die, the player has two possible actions:

i. *Stop*: Stop playing by collecting the dollar value that the die lands on;

ii. *Roll*: Roll again, paying another 1 dollar.

Having taken VE 492, you decide to model this problem using an infinite horizon Markov Decision Process (MDP). The player initially starts in state *Start*, where the player only has one possible action: *Roll*. State s_i denotes the state where the die lands on i. Once a player decides to *Stop*, the game is over, transitioning the player to the *End* state.

(a) In solving this problem, you consider using policy iteration. Your initial policy π is in the table below. Evaluate the policy at each state, with $\gamma = 1$.

State	s_1	s_2	s_3	s_4	s_5	s_6
$\pi(s)$	Roll	Roll	Stop	Stop	Stop	Stop
$V^\pi(s)$	3	3	3	4	5	6

(b) Old policy π and has filled in parts of the updated policy π' for you. If both *Roll* and *Stop* are viable new actions for a state, write down both *Roll/Stop*. In this part as well, we have $\gamma = 1$.

State	s_1	s_2	s_3	s_4	s_5	s_6
$\pi(s)$	Roll	Roll	Stop	Stop	Stop	Stop
$\pi'(s)$	Roll	Roll	Roll/Stop	Stop	Stop	Stop

(c) Is $\pi(s)$ from part (a) optimal? Explain why or why not.

Yes. It is optimal. After doing $\pi(s)$, for the two new policies $\pi(s')$, one is the same as $\pi(s)$, another has the same value as $\pi(s)$. Thus, the policy iteration converges to the optimal policy. So $\pi(s)$ from part(a) is optimal.

(d) Suppose that we were now working with some $\gamma \in [0, 1)$ and wanted to run **value iteration**. Select the one statement that would hold true at convergence, or write the correct answer next to Other if none of the options are correct.

I

A. $V^*(s_i) = \max\Big\{ -1 + \frac{i}{6},\ \sum_j \gamma V^*(s_j) \Big\}$

B. $V^*(s_i) = \max\Big\{ i,\ \frac{1}{6}\big[-1 + \sum_j \gamma V^*(s_j) \big] \Big\}$

C. $V^*(s_i) = \max\Big\{ -\frac{1}{6} + i, \sum_j \gamma V^*(s_j) \Big\}$

D. $V^*(s_i) = \max\Big\{ i,\ -\frac{1}{6} + \sum_j \gamma V^*(s_j) \Big\}$

E. $V^*(s_i) = \frac{1}{6} \sum_j \max\{ i,\ -1 + \gamma V^*(s_j) \}$

F. $V^*(s_i) = \frac{1}{6} \sum_j \max\{ -1 + i,\ \sum_k V^*(s_j) \}$

G. $V^*(s_i) = \sum_j \max\Big\{ -1 + i,\ \frac{1}{6} \gamma V^*(s_j) \Big\}$

H. $V^*(s_i) = \sum_j \max\Big\{ \frac{i}{6},\ -1 + \gamma V^*(s_j) \Big\}$

I. $V^*(s_i) = \max\Big\{ i,\ -1 + \frac{1}{6} \gamma \sum_j V^*(s_j) \Big\}$

J. $V^*(s_i) = \sum_j \max\Big\{ i,\ -\frac{1}{6} + \gamma V^*(s_j) \Big\}$

K. $V^*(s_i) = \sum_j \max\Big\{ -\frac{i}{6},\ -1 + \gamma V^*(s_j) \Big\}$