周维凯 518021911039

1. (a). If s^* is not a Nash equilibrium, then we can find some i such that $u_i(s^*) < u_i(s, s^*_{-i})$. We can see that (s', s^*_{-i}) is not dominated by s^*. Therefore, s^* is not a dominance equilibrium, which is a contradiction.

Homework 3 Written

(b).

	A	B
X	3,4	2,3
Y	2,2	3,5

There is no dominance but (Y,B) and (X,A) are Nash equilibria.

June 16th, 2021 at 23:59

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.

2 Game in Matrix

Consider the game with the following bimatrix:

	A	B	C
a	1, 1	3, x	2, 0
b	$2x$, 3	2, 2	3, 1
c	2, 1	1, x	x^2, 4

(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.

2. (a). $\begin{cases} 2 > 2x \\ x < 1 \\ x^2 < 3 \text{ or } x > 4 \end{cases}$

$\Rightarrow \boxed{-\sqrt{3} < x < 1}$

(b). $\begin{cases} x^2 \geq 3 \\ x \leq 4 \end{cases} \Rightarrow \boxed{x \leq -\sqrt{3} \text{ or } \sqrt{3} \leq x \leq 4}$

3 Nash Equilibrium

Consider the zero-sum game in which two players choose nonnegative integers no greater than 1000. Player 1 must choose an odd integer, while player 2 must choose an even 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 2

Player 1

	0	2	4	...	1000
1	0, 0	1, -1	1, -1		1, -1
3	0, 0	-2, 2	3, -3		3, -3
5	0, 0	-2, 2	-4, 4		5, -5
⋮					
999	0, 0	-2, 2	-4, 4		999, -999

The Nash equilibrium is $\boxed{(1, 0)}$.

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. $\pi'(s)$ and $\pi(s)$ are almost the same. The difference is S_3. But their utilities are the same, which means the policy iteration has converged. Therefore, $\pi(s)$ from part (a) is optimal.

B.

(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.

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

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

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

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

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

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