COMP4418

Knowledge Representation and Reasoning

Assignment 3

Pengcheng Liu

z5128626

1.
a)

Uncovered set:

{a, b, c}

Top Cycle:

{a, b, c, d, e, f}

Copeland Winner:

{c}

Banks winners:

{b, c}

Condorcet winners:

{}
There's no Condorcet winner because no alternative dominates all others pair-wisely.

b)
Pure Nash equilibrium:

	D	E
A		8,5
B	6,6	

Mixed Nash equilibrium:

Player1 has chance x to take action A, then 1-x chance to take action B:

$2x+6(1-x) = 8x+4(1-x)$

So x=1/4
Player2 has chance y to take action D then 1-y chance to take action E:

$4y + 5(1 - y) = 6y + 4(1 - y)$

So y=1/3

The mixed Nash equilibrium is Player1:(1/4,3/4) Player2:(1/3,2/3).

2.
a)

1. Blackjack POMDP

The player does not know the coming card and the bank can not take action, so it is pomdp.

2. Candy Crush MDP

The player can observe the full board and need to take actions, so it is mdp.

3. Chess
None/Other
There are two players involved, so it is not Markov process.

4. Minesweeper
POMDP
Player need to take action and the board is not fully observable.

5. Snakes and Ladders
Markov process
Player can not take actions. All actions are purely from the dice with known probability.

vi. Texas Hold'em Poker
None/Other
There are multiple players involved, so it is not Markov process.

b)

If discount factor is very high, The optimal policy for s3 is either Leave or Stay because the reward will remain the same for both action. But for s2 and s3 the optimal policy is to choose Stay because if they leave, they will have a chance to get to s3 which the reward in that state is negative.

If discount factor is very low, The optimal policy for s3 is the same because either Leave or Stay have the same reward. The optimal policy for s1 will be Stay since it brings reward 1 which is larger than Leave, for s2 will be Leave since the reward is 5 which is larger than Stay.

d)

For each step we have:

Initialization:

$V(s_{i+1}) = \max\{u(S) + \delta \sum P(s, S, s^{'})V_i(s^{'})\}$ $V_0(s_1) = 0$; $V_0(s_2) = 0$; $V_0(s_3) = 0$

In first step:

$V_1(s_1) = \max\{1; 0\} = 1$ $V_1(s_2) = \max\{0; 5\} = 5$ $V_1(s_3) = \max\{-2; -2\} = -2$

In second step:

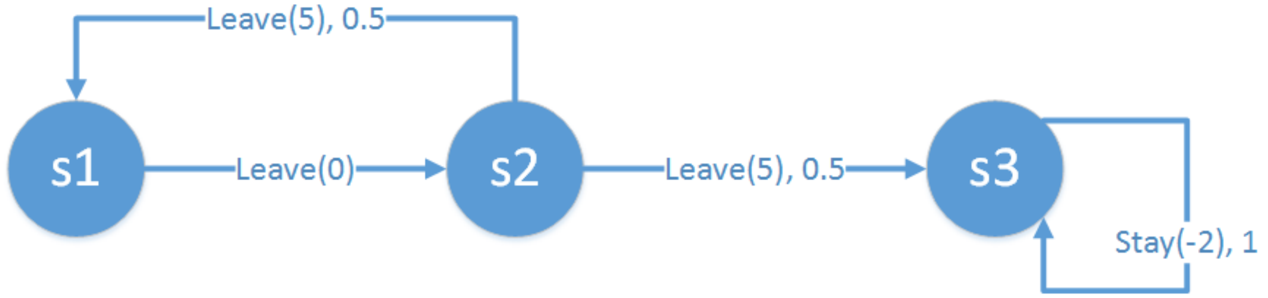
$V_2(s_1)$=max{1+0.6∗1; 0+0.6∗5}=3 $V_2(s_2)$=max{0+0.6∗5; 5+0.6∗(0.5∗1+0.5∗(−2))} =4.7 $V_2(s_3) = \max\{-2 + 0.6 * (-2); -2 + 0.6 * (-2)\} = -3.2$

In third step:

$V_3(s_1)$=max{1+0.6∗3; 0+0.6∗4.7}=2.82 $V_3(s_2)$=max{0+0.6∗4.7; 5+0.6∗(0.5∗3+0.5∗(−3.2))}=4.94 $V_3(s_3) = \max\{-2 + 0.6 * (-3.2); -2 + 0.6 * (-3.2)\} = -3.92$

	$V_0(s)$	$V_0(s, S)$	$V_0(s, L)$	$V_1(s)$	$V_1(s, S)$	$V_1(s, L)$	$V_2(s)$	$V_2(s, S)$	$V_2(s, L)$	$V_3(s)$
s1	0	1	0	1	1.6	3	3	2.8	2.82	2.82
s2	0	0	5	5	3	4.7	4.7	2.82	4.94	4.94
s3	0	-2	-2	-2	-3.2	-3.2	-3.2	-3.92	-3.92	-3.92

e)



According to the process above, we have the following formulas.

$V(s_{i+1}) = u(S) + \delta \sum P(s, S, s^{'}) V_i(s^{'})$

	$V_0(s)$	$V_1(s)$	$V_2(s)$	$V_3(s)$
s1	0	1	3	2.4
s2	0	5	4	4.9
s3	0	-2	-3.2	-3.72

f)

$V_{i+1}(s_1) = \delta V_i(s_2)$

$V_{i+1}(s_2) = 5 + \delta(0.5 * V_i(s_1) + 0.5 * V_i(s_3))$ $V_{i+1}(s_3) = -2 + \delta V_i(s_3)$

From the equation above, we can solve s3 first:

$V_{n\to\infty}(s_3)$=1/1−δ

Then the others:

$V_{n \to \infty}(s_2) = 5 - 6\,\delta/(1-\delta)(1-0.5\delta^2)$

$V_{n\to\infty}(s_1) = \delta(5-6\delta)/(1-\delta)(1-0.5\delta^2)$

Now we can tell from the equation above that if δ is very high, both $V_{n\to\infty}(s_2)$ and $V_{n\to\infty}(s_1)$ will become negative infinity.So we can not choose Leave for s1 and s2. Meanwhile, if δ is very low, $V_{n\to\infty}(s_2)$ will become 5 and $V_{n\to\infty}(s_1)$ will become 0,so for s2 we should choose Leave and for s1 we should choose Stay.The conclusion is the same with the answer in the previous questions.