COMP4418, 2017 – Assignment 3

Z5045582

Yunhe Zhang

1. Social Choice and Game Theory

(a)

the domination set of each alternative
a {b, d, f, g}
b {c, d, e, g}
c {a, d, e, f, g}
d {e, f, g}
e {a, f, g}
f {b, g}
g {}

- **the uncovered set:**

  Alternative in two steps:

  a {b, c, d, e, f, g} ✓
  b {a, c, d, e, f, g} ✓
  c {a, b, d, e, f, g} ✓
  d {a, b, e, f, g} ×
  e {a, b, d, f, g} ×
  f {b, c, d, e, g} ×
  g {} ×

  The uncovered set is **{a, b, c}**

- **the top cycle:**

  The top cycle is **{a, b, c, d, e, f}**

- **the set of Copeland winners:**

  the set of Copeland winners is **{c}**

- **the set of Banks winners:**

maximal acyclic subgraph length is 5
b ⇨ c ⇨ d ⇨ e ⇨ g
c ⇨ a ⇨ d ⇨ f ⇨ g
the set of Banks winners is **{b, c}**

- **the set of Condorcet winners:**

no alternative is satisfied, so the set of Condorcet winners is **{}**

(b)
Pure Nash equilibria
(A, E)
(B, D)

Mixed Nash equilibria
Player 1 plays A with p, plays B with 1 – p
$4p + 6(1 - p) = 5p + 4(1 - p)$
$4p + 6 - 6p = 5p + 4 - 4p$
$3p = 2$
$p = 2 / 3$

Player 2 plays D with q, plays E with 1 – q
$2q + 8(1 - q) = 6q + 4(1 - q)$
$2q + 8 - 8q = 6q + 4 - 4q$
$8q = 4$
$q = 1 / 2$

Mixed Nash equilibria
(A, B) = (2 / 3, 1 / 3)
(D, E) = (1 / 2, 1 / 2)

2. Decision Making

(a)

1. Blackjack: (D) POMDP
2. Candy Crush: (B) MDP
3. Chess: (E) None/Other
4. Minesweeper: (D) POMDP
5. Snakes and Ladders: (A) MP
6. Texas Hold ’em Poker: (E) None/Other

(b)
π(S1) = Stay
π(S2) = Stay
π(S3) = Stay or Leave
If the discount factor is very high, the value of state is hard to converge, S3'two actions are same, and will result in the negative infinity, however for S1 and S2, chose 'Stay' the S1 will be infinity and S2 will be 0, chose 'Leave' S2 might be negative infinity because of S3 and S1 might be 0 or negative infinity because of S2.

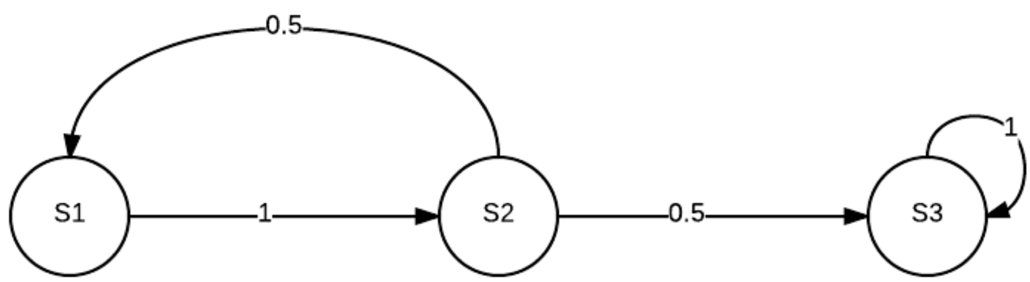
(c)
π(S1) = Stay
π(S2) = Leave
π(S3) = Stay or Leave

Since the discount factor is very low, the future part could be ignored, thus just chose the action whose associated reward is higher.

(d)

	V0(s)	V0(s, S)	V0(s, L)	V1(s)	V1(s, S)	V1(s, L)	V2(s)	V2(s, S)	V2(s, L)	V3(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	7.82	4.94	7.82
s3	0	-2	-2	-2	-3.2	-3.2	-3.2	-3.92	-3.92	-3.92

(e)



(f)

Since agent uses π this could see as Markov process problem

when i = 0
$v_0(s1) = v_0(s2) = v_0(s3) = 0$

when i > 0
$v_i(s1)$
$= u(s1, L) + \delta P(s1, L, s2)v_i(s2)$
$= \delta v_i(s2)$

$v_i(s2)$
$= u(s2, L) + \delta(P(s2, L, s1)*v_{i-1}(s1)+P(s2, L, s3)*v_{i-1}(s3))$
$= 5 + \delta(0.5*v_i(s1) + 0.5*v_i(s3))$

$v_i(s3)$
$= u(s3, S) + \delta P(s3, S, s3)v_i(s3)$
$= -2 + \delta v_i(s3)$

According to
(1) $v_i(s1) = \delta v_i(s2)$
(2) $v_i(s2) = 5 + \delta(0.5*v_i(s1) + 0.5*v_i(s3))$
(3) $v_i(s3) = -2 + \delta v_i(s3)$

Get:
(1) $v_i(s1) = \frac{\delta(10-12\delta)}{(1-\delta)*(2-\delta^2)}$
(2) $v_i(s2) = \frac{10-12\delta}{(1-\delta)*(2-\delta^2)}$
(3) $v_i(s3) = \frac{-2}{1-\delta}$

When δ is very high like $\delta = 0.999$

Vi of all 3 states will be close to negative infinity

This could partly support my intuition of question 2b since vi of all 3 states will be negative infinity as I mentioned in my explanation in 2b.

When δ is very low like $\delta = 0.001$

Vi of all 3 states will be close to values of their reward of action in π ($v_i(s1)$ close to 0, $v_i(s2)$ close to 5, $v_i(s3)$ close to -2)

It supports my intuition of question 2c since vi is the action associated reward, so the higher associated reward action will be the optional policy.