MAS Practice Exam Dec 2023

① Game Theory

Q1

		A (q)	B ($1-q$)
p	A	2, 2	−1, −1
$1-p$	B	−1, −1	1, 1

PNE:
(A, A) , (B, B)

MNE:

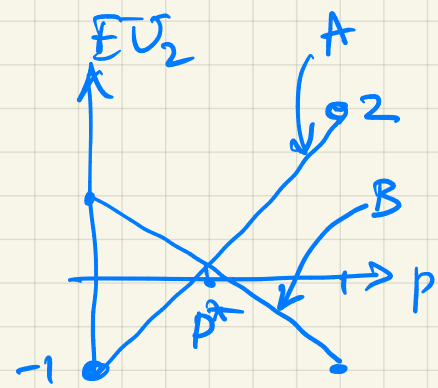
$$EU_2(p, A) = EU_2(p, B)$$

$$2p - (1-p) = -p + (1-p)$$

$$3p - 1 = -2p + 1$$

$$5p = 2 \Rightarrow \boxed{p^* = 2/5} \longrightarrow q^* = 2/5$$

(symmetric game)



Corresponding utility:

$$EU_2(p^*, A) = EU_2(p^*, B) = 2 \cdot \frac{2}{5} - \frac{3}{5} = \frac{1}{5}$$

$$EU_1(A, q^*) = EU_1(B, q^*) = \frac{1}{5}$$

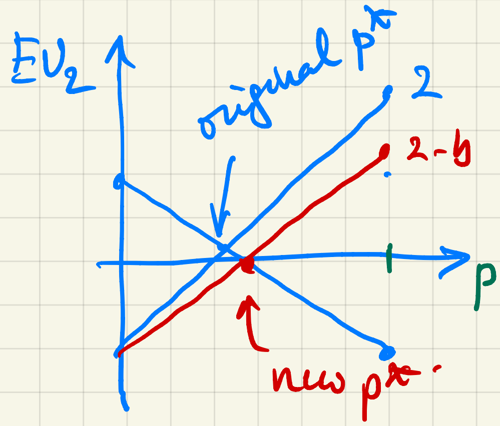
Q2: MNE is pareto dominated by PNE

$$MNE < (B, B) < (A, A)$$

Q3. $u_2(A,A) = 2-b$ $(0<b<1)$

		A q	B $1-q$
A	p	$2, 2-b$	$-1, -1$
B	$1-p$	$-1, -1$	$1, 1$

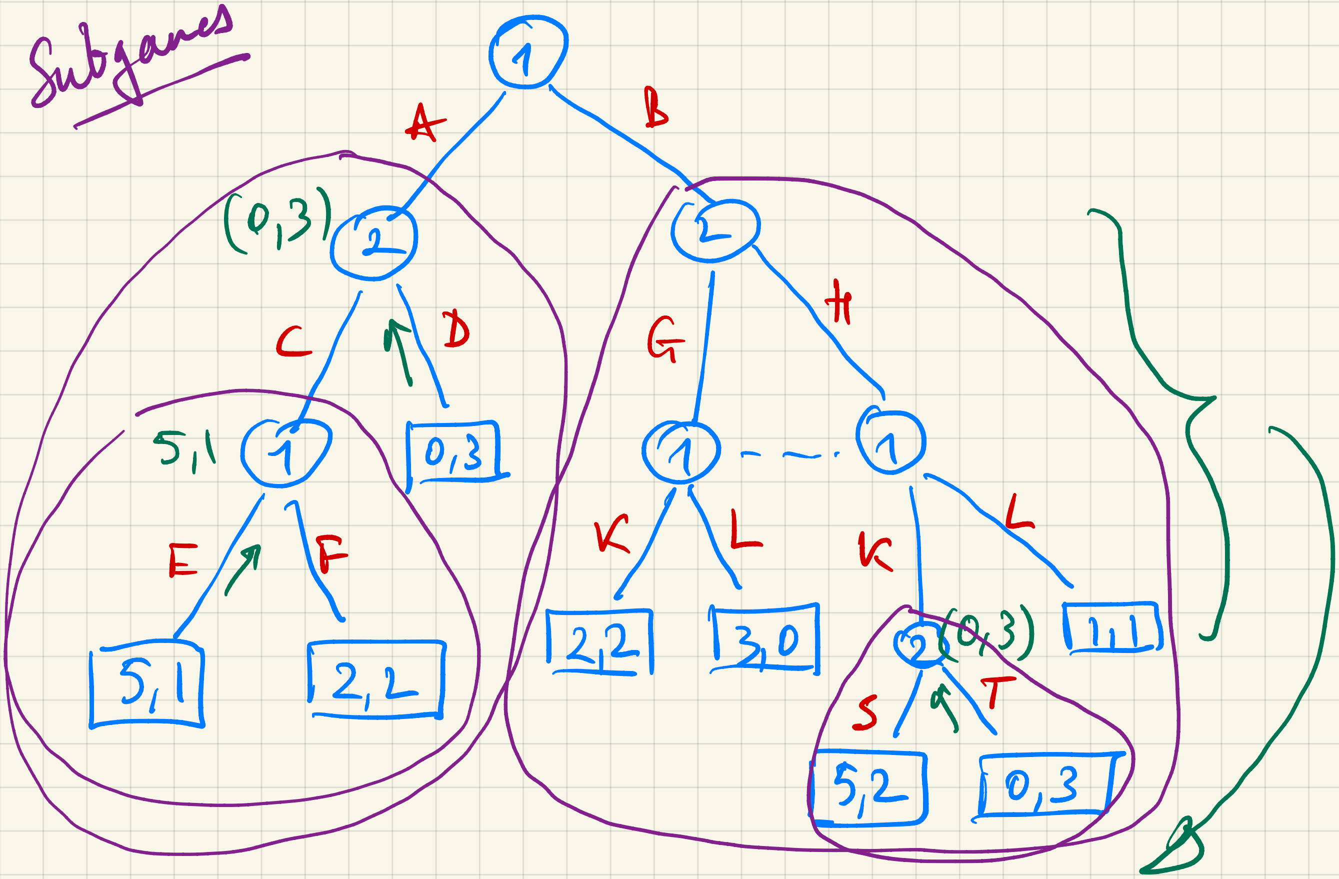
PNE unchanged.



As $b\uparrow \Rightarrow p^*\uparrow$.

NB: q^* is unaffected!

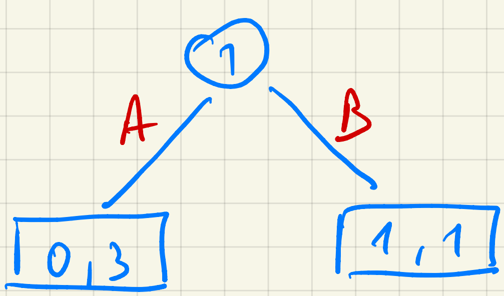
4 Subgames



	G	H
K	2,2	0,3
L	3,0	1,1

→ (L,H) is NE

We can reduce the game tree to:



⇒ 1 will choose B

2 ——— H

1 ——— L

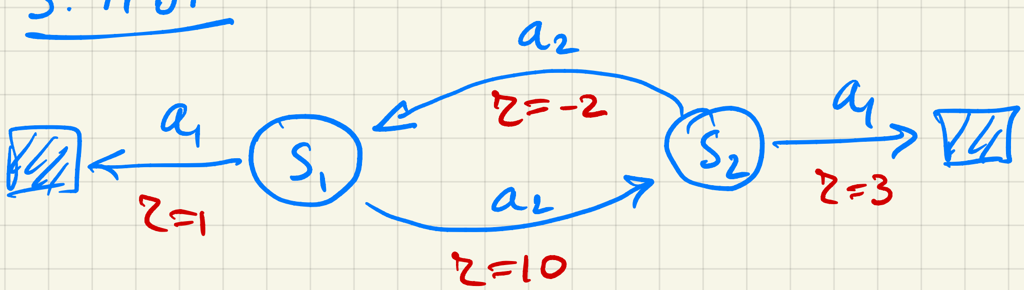
Payoff = (1,1).

Q3: $1 \to \{A,B\} \times \{EF\} \times \{KL\} = \{AEK, AE$

$= \{AEK, AEL, AFK, AFL, BEK, BEL, BFK, BFL\}$

$2 \to \{C,D\} \times \{GH\} \times \{S,T\} = \dots$

3. MDP



Q1: Infinite horizon (could keep on cycling)

Q2: $\gamma = 0.9 \longrightarrow$ optimal policy?

4 possible (deterministic) policies.

π_1: $s_1 \to a_1$, $s_2 \to a_1$

$v(s_1) = 1$, $v(s_2) = 3$

π_2: $s_1 \to a_2$, $s_2 \to a_1$

$v(s_1) = 10 + 3\gamma = 10 + 2.7 = 12.7$

$v(s_2) = 3$

π_3: $s_1 \to a_1$, $s_2 \to a_2$

$v(s_1) = 1$

$v(s_2) = -2 + \gamma = -2 + 0.9 = -1.1$

π_4: $s_1 \to a_2$, $s_2 \to a_2$

$$v(s_1) = 10 + (-2)\gamma + 10\gamma^2 + (-2)\gamma^3 + \dots$$

$$= 10\underbrace{(1+\gamma^2+\gamma^4+\dots)}_{\frac{1}{1-\gamma^2} = \frac{1}{1-0.9^2} = \frac{1}{1-0.81} = \frac{1}{0.19} \simeq 5{,}3} - 2\gamma(1+\gamma^2+\dots) \simeq \underline{\underline{43.2}}$$

$$\overbrace{(10 - 1.8)\cdot 5.3}^{43.2}$$

$$v(S_2) = -2 + \gamma \cdot 10 + (-2)\gamma^2 + 10\gamma^3 + \dots$$

$$= -2(1+\gamma^2+\dots) + 10\gamma(1+\gamma^2+\dots)$$

$$= \underbrace{(10\cdot\gamma - 2)}_{7}\underbrace{(1+\gamma^2+\gamma^4+\dots)}_{\frac{1}{1-\gamma^2} \simeq 5.3} \simeq \underline{\underline{37.1}}$$

__Conclusion__, for $\gamma = 0.9$ optimal policy: $\pi^* = \pi_4$.

4. Reinforcement Learning

(1) Recall: $v(s) = \sum_a q(s,a)\,\pi(a|s)$

$$v(2) = 5 = \underbrace{q(2,R)}_{x}\,\underbrace{\pi(R|2)}_{1/4} + \underbrace{q(2,L)}_{4}\,\underbrace{\pi(L|2)}_{3/4}$$

$$= \frac{1}{4}x + \frac{3}{4}\cdot 4 = 5 \Rightarrow x = 20 - 12 = 8$$

$$\boxed{q(2,R) = 8}$$

Similarly: $q(3,L) = 9$

(2) Q-Learning: $2 \xrightarrow{R} 3$

$$q(2,R) \leftarrow q(2,R) + \alpha\left[r(2,R) + \gamma \underbrace{\max_{a'} q(3,a')}_{9} - q(2,R)\right]$$

$$\leftarrow 8 + 0.9\left[\underbrace{-1 + \frac{2}{3}\cdot 9 - 8}_{-3}\right] = 8 - 2.7 = \underline{\underline{5.3}}$$