ROB311 Quiz 3

Hanhee Lee

April 13, 2025

Contents

One-Shot Multi-Agent Decision Problems

1 Multi-Agent Problems - Game Theory

Summary: In a **Multi-Agent problem**, we assume that:

- Set of states for environment is $\mathcal{S}$
- P agents within environment.
- For each state $s \in \mathcal{S}$:
  - possible actions for agent i is $\mathcal{A}_i(s)$
  - set of action profiles is $\mathcal{A}(s) = \prod_{i=1}^{P} \mathcal{A}_i(s)$
- possible state-action pairs are $\mathcal{T} = \{(s, a) \text{ s.t. } s \in \mathcal{S}, a \in \mathcal{A}(s)\}$
- environment in some origin state, s_0
- environment destroyed after N transitions
- agent j wants to find policy $\pi_j(a_j \mid s)$ so that $\mathbb{E}[r_j(p)]$ is maximized
- agents act independently given the environment's state: $\pi(a \mid s) = \prod_{j \in [P]} \pi_j(a_j \mid s)$

Name	**Function:**
State transition given state-action pair defined by $\text{tr} : \mathcal{T} \to \mathcal{S}$	$\text{tr}(s, a) =$ state transition from s under a
Reward to each agent, i defined by r_i: $\mathcal{Q} \times \mathcal{S} \to \mathbb{R}_+$	$r_i(s, a, \text{tr}(s, a)) =$ rwd to agent i for $(s, a, tr(s, a))$
State evolution of environment after N transitions • Given sequence of actions: $p.a = \langle a^{(1)}, \dots, a^{(n)} \rangle$ • $s_N = \tau(s_{n-1}, a^{(n)})$	$p = \langle (s_0, a^{(1)}, s_1), \dots, (s_{N-1}, a^{(N)}, s_N) \rangle$
reward to agent i	$r_i(p) = \sum_{n=1}^{N} r_i(s_{n-1}, a^{(n)}, s_n)$
expected-reward (value) of playing a from s for agent j • $\mathcal{A}(s) = \emptyset$ if $s \in \mathcal{G}$	$q_j(s, a) = r_j(s, a, \tau(s, a)) + \sum_{a' \in \mathcal{A}(\tau(s,a))} \pi(a' \mid \tau(s, a)) q_j(\tau(s, a), a')$

1.1 Policy Equilibria

Notes:

- **No Regret:** π is no-regret if π_j maximizes q_j when π_{-j} is fixed.
- If all agents play perfectly, then we expect

$$\pi(a \mid s) = \begin{cases} 1 & \text{if } a = a^*(s) \\ 0 & \text{otherwise} \end{cases}$$

  - $a_j^*(s) = \arg \max_{a_j \in \mathcal{A}_j(s)} q_j(s, a_j, a_{-j}^*)$ is the best action for agent j given the other agents' policies.

Warning: No regret if it got the highest reward given the other players' action.

1.2 Single Action Games

Summary: In a **Single Action Game**, we assume that:
- $N = 1$ (one-shot game)
- Initial state is $s_0 \in \mathcal{S}$
- Agent j wants to find policy, $\pi_i(a_i \mid s_0)$ so $\mathbb{E}[r_i(p)]$ is maximized

1.3 Actions (Deterministic)

Summary: Allow each agent to choose action deterministically.

Name	**Function:**
Action j for agent i • One-hot vector of M_i components, $\mathbf{e}_{i,j}$	$[0 \cdots 0\ 1\ 0 \cdots 0]^T$
Agent i's set of possible actions • Agent i's chosen action with $a_i \in \mathcal{A}_i$	$\mathcal{A}_i = \left\{ a_i \in \{0,1\}^{M_i} \mid \sum_{j \in [M_i]} a_{i,j} = 1 \right\}$
Action profile is a tuple of actions • **Notational Convenience:** $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_P)$ so that $a = (a_i, a_{-i})$.	$a = (a_1, \ldots, a_P)$
Optimal action profile • An action profile w.r.t. which any other action profile leaves at least one player worse off.	a^+ s.t. $\forall a \exists i$ s.t. $r_i(a) < r_i(a^+)$
Set of optimal action profiles	$\text{aOpt} = \{a^+ \mid \forall a \exists i : r_i(a) < r_i(a^+)\}$
Best-action mapping, ba_i: $\mathcal{A}_{-i} \to \mathcal{A}_i$	$\text{ba}_i(a_{-i}) = \arg\max_{a_i \in \mathcal{A}_i} r_i(a_i, a_{-i})$ $= \{a_i \in \mathcal{A}_i \mid r_i(a_i, a_{-i}) = \max_{a_i' \in \mathcal{A}_i} r_i(a_i', a_{-i})\}$
Agent i will **not regret** playing a_i^* when others play a_{-i}^* if	$r_i(a_i^*, a_{-i}^*) \geq r_i(a_i, a_{-i}^*)\ \forall a_i \in \mathcal{A}_i$ or $a_i^* \in \text{ba}_i(a_{-i}^*)$
Action equilibria is any action, a^* in which no agent regrets	$a_i^* \in \text{ba}_i(a_{-i}^*)\ \forall i \in [P]$
Existence of action equilibria	May not always exist, i.e., it may be that $\text{aEq} = \emptyset$

1.4 Strategies (Probabilistic)

Summary: Allow each agent to choose action based on a distribution/strategy.

Name	**Function:**
Stategy for agent i • Vector of M_i components, that are non-negative and sum to 1	$[0.05 \cdots 0.2\ 0.7\ 0 \cdots 0.05]^T$
Agent i's set of possible strategies • Agent i's chosen strategy with $x_i \in \Delta_i$	$\Delta_i = \Delta^{M_i} = \left\{ x_i \in [0,1]^{M_i}, \sum_{j \in [M_i]} x_{i,j} = 1 \right\}$
Expected reward	$\bar{r}_i(x_1, \ldots, x_P) = \mathbb{E}[r_i(a)] = \sum_{a_i \in \mathcal{A}_i} \pi(a) r_i(a)$
Stategy profile is a tuple of strategies • **Notational Convenience:** $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_P)$ so that $x = (x_i, x_{-i})$.	$x = (x_1, \ldots, x_P)$
No-regret strategies • Agent i will not regret using x_i^* when others use x_{-i}^*.	$\bar{r}_i(x_i^*, x_{-i}^*) \geq \bar{r}_i(x_i, x_{-i}^*)\ \forall x_i \in \Delta_i$ $x_i^* \in \text{bs}_i(x_{-i}^*)$
Best strategy mapping bs_i: $\Delta_{-i} \to \Delta_i$	$\text{bs}_i(x_{-i}) = \arg\max_{x_i \in \Delta_i} \bar{r}_i(x_i, x_{-i})$ $= \{x_i \in \Delta_i \mid \bar{r}_i(x_i, x_{-i}) = \max_{x_i' \in \Delta_i} \bar{r}_i(x_i', x_{-i})\}$
Strategy equilibria is any strategy, x^* in which no agent regrets	$x_i^* \in \text{bs}_i(x_{-i}^*)\ \forall i \in [P]$
Joint best-strategy mapping $\text{bs} : \Delta \to \Delta$	$\text{bs}(x) = (\text{bs}_1(x_{-1}), \ldots, \text{bs}_P(x_{-P}))$
Existence of strategy equilibria • Fixed pt. always exists.	Any strategy equilibrium, x^* is a fixed pt. of $x^* = \text{bs}(x^*)$

1.4.1 Simplifying Games

Notes: May be able to reduce M_i by eliminating useless actions/strategies:

- **Equivalent Stategies:** $x_i^{(1)} \equiv x_i^{(2)}$

$$\bar{r}_i(x_i^{(1)}, x_{-i}) = \bar{r}_i(x_i^{(2)}, x_{-i})\ \forall x_{-i}$$

- **Dominated Strategies:** x_i

$$\exists x_i' \text{ s.t. } \bar{r}_i(x_i', x_{-i}) \leq \bar{r}_i(x_i, x_{-i})\ \forall x_{-i}$$

Can remove dominated and equilvalent strategies w/o changing the game.

1.5 Examples

1.5.1 Finding Action Equilibria

Process: To find action equilibria:

1. For each i, compute $\text{ba}_i(a_{-i})$ for all a_{-i}
2. Define bap_i so that $\text{bap}_i = \{(a_i', a_{-i}),\ \forall a_i' \in \text{ba}_i(a_{-i}),\ \forall a_{-i} \in \mathcal{A}_{-i}\}$
3. Action equilibria are then $\text{aEq} = \bigcap_{i\in[P]} \text{bap}_i$.

Process:

1. Fix strategy (i.e. prob. 1) for other player, then find best move for current player by getting max reward.
2. See if there's intersection between best responses.

Warning:

- Action equilibrium is a pure equilibrium (i.e. prob. 1 or 0)
  - aEq doesn't mean that the actual outcome is fight,fight.
  - aEq doesn't mean it is socially optimal (i.e. equilibria may not be optimal).
- Mixed equilibrium is a probabilistic equilibrium (i.e. prob. p)
  - Every action equilibrium is a mixed equilibrium, but not vice versa.

Example:

1. **Given:** Consider a 2-player single-action game, in which each player has 2 actions. Let a_1 and a_2 be their chosen actions and $r_1(a_1, a_2), r_2(a_1, a_2)$ be the resulting rewards.

a_1	a_2	$r_1(a_1, a_2)$	$r_2(a_1, a_2)$
1	1	2	2
1	2	4	1
2	1	1	4
2	2	3	3

2. **Solution:**
   - Convert:

	$a_2 = 1$	$a_2 = 2$
$a_1 = 1$	$(2, 2)$	$(4, 1)$
$a_1 = 2$	$(1, 4)$	$(3, 3)$

   - **Fix a_2, choose a_1:**

$$\begin{aligned} &\text{If } a_2 = 1 \Rightarrow a_1 = 1 \Rightarrow (1, 1) \\ &\text{If } a_2 = 2 \Rightarrow a_1 = 1 \Rightarrow (1, 2) \\ &\Rightarrow \text{bap}_1 = \{(1, 1), (1, 2)\} \end{aligned}$$

   - **Fix a_1, choose a_2:**

$$\begin{aligned} &\text{If } a_1 = 1 \Rightarrow a_2 = 1 \Rightarrow (1, 1) \\ &\text{If } a_1 = 2 \Rightarrow a_2 = 1 \Rightarrow (2, 1) \\ &\Rightarrow \text{bap}_2 = \{(1, 1), (2, 2)\} \end{aligned}$$

   Intersection:

$$\text{aEq} = \text{bap}_1 \cap \text{bap}_2 = \{(1, 1)\}$$

3. Therefore, the action equilibria is $(1, 1)$.

Example:

1. **Given:** Suppose lion and cavemen both want meat. Each must decide whether to fight for the food or share it.

		Fight	Share
	Fight	$(1,1)$	$(4,0)$
	Share	$(0,4)$	$(2,2)$

Figure 1

2. **Problem:** Find the action equilibria
3. **Solution:**
   (a) **Best Action Profiles:**
   - Cavemen: $\text{bap}_{\text{cavemen}} = \{(\text{Fight, Fight}), (\text{Fight, Share})\}$. Cavemen fights no matter what.
     - If lion fights, then cavemen fights to get maximum reward in this scenario of +1.
     - If lion shares, then caveman fights to get maximum reward in this scenario of +4.
   - Lion: $\text{bap}_{\text{lion}} = \{(\text{Fight, Fight}), (\text{Share, Fight})\}$. Lion fights no matter what.
     - If caveman fights, then lion fights to get maximum reward in this scenario of +1.
     - If caveman shares, then lion fights to get maximum reward in this scenario of +4.

   (b) **Best Action Equilibria:** Intersection of the best action profiles.
   - $\text{aEq} = \text{bap}_{\text{cavemen}} \cap \text{bap}_{\text{lion}} = \{(\text{Fight, Fight})\}$

Example:

1. **Given:** Suppose lion and cavemen both want meat. Each must decide whether to fight for the food or share it.

		Fight	Share
	Fight	$(3,0)$	$(0,3)$
	Share	$(0,1)$	$(1,0)$

Figure 2

2. **Problem:** Find the action equilibria
3. **Solution:**
   (a) **Best Action Profiles:**
   - Cavemen: $\text{bap}_{\text{cavemen}} = \{(\text{Fight, Fight}), (\text{Share, Share})\}$. Cavemen fights no matter what.
     - If lion fights, then cavemen fights to get maximum reward in this scenario of +3.
     - If lion shares, then caveman shares to get maximum reward in this scenario of +1.
   - Lion: $\text{bap}_{\text{lion}} = \{(\text{Fight, Share}), (\text{Share, Fight})\}$. Lion fights no matter what.
     - If caveman fights, then lion shares to get maximum reward in this scenario of +3.
     - If caveman shares, then lion fights to get maximum reward in this scenario of +1.

   (b) **Best Action Equilibria:** Intersection of the best action profiles.
   - $\text{aEq} = \text{bap}_{\text{cavemen}} \cap \text{bap}_{\text{lion}} = \emptyset$

1.5.2 Optimal Action Profiles

Process:
1. Switching action profiles will leave at least one player worse off.
   - That is, no unilateral deviation by any player should result in both players being better off.

Example:
1. **Given:**

a_1	a_2	$r_1(a_1, a_2)$	$r_2(a_1, a_2)$
1	1	2	2
1	2	4	1
2	1	1	4
2	2	3	3

2. **Solution:**
   - **At action profile** $(1, 1)$:
   $$\text{Switch to } (2, 2): \text{ both players are better off} \Rightarrow \text{not optimal}$$
   - **At action profile** $(1, 2)$:
   $$\text{Switch to } (1, 1), (2, 1), (2, 2): \text{ Player 1 will be worse off for all switches} \Rightarrow \text{optimal}$$
   - **At action profile** $(2, 1)$:
   $$\text{Switch to } (1, 1), (1, 2), (2, 2): \text{ Player 2 will be worse off for all switches} \Rightarrow \text{optimal}$$
   - **At action profile** $(2, 2)$:
   $$\begin{aligned}&\text{Switch to } (1, 1): \text{Both Player 1 and Player 2 will be worse off}\\&\text{Switch to } (2, 1): \text{Player 1 will be worse off}\\&\text{Switch to } (1, 2): \text{Player 2 will be worse off} \Rightarrow \text{optimal}\end{aligned}$$
   - Therefore, $(1, 2)$, $(2, 1)$, and $(2, 2)$ are optimal action profiles.

Example:

1. **Given:** Suppose lion and cavemen both want meat. Each must decide whether to fight for the food or share it.

		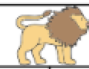	
		Fight	Share
	Fight	$(1, 1)$	$(4, 0)$
	Share	$(0, 4)$	$(2, 2)$

Figure 3

2. **Problem:** Find the optimal action profiles:
3. **Solution:**
   (a) aOpt $= \{(\text{Share, Share}), (\text{Fight, Share}), (\text{Share, Fight})\}$
   - **(Fight, Fight) → (1, 1):**
     Both players could be better off by choosing (Share, Share), which gives (2, 2).
     Since there exists another outcome where no one is worse off and at least one player is better off, this is not an optimal action profile.
     ⇒ **Not optimal**.
   - **(Fight, Share) → (4, 0):**
     The caveman gets the highest possible payoff (4), but the lion gets 0.
     Any other outcome that gives the lion more than 0 will reduce the caveman's payoff.
     So, every other option makes at least one player worse off.
     ⇒ **Optimal**.
   - **(Share, Fight) → (0, 4):**
     The lion gets the highest possible payoff (4), but the caveman gets 0.
     Any other outcome that gives the caveman more than 0 will reduce the lion's payoff.
     So, every other option makes at least one player worse off.
     ⇒ **Optimal**.
   - **(Share, Share) → (2, 2):**
     Both players receive a fair and equal payoff. No other outcome makes both players better off at the same time.
     So, changing this would hurt at least one of them.
     ⇒ **Optimal**.

1.5.3 Finding/Convergence Strategy Equilibria

Process: Finding

1. For each i, compute $\text{bs}_i(x_{-i})$ for all x_{-i}
2. Define bsp_i so that $\text{bsp}_i = \{(x'_i, x_{-i}),\ \forall x'_i \in \text{bs}_i(x_{-i}),\ \forall x_{-i} \in \Delta_{-i}\}$
3. Strategy equilibria are then $\text{sEq} = \bigcap_{i \in [P]} \text{bsp}_i$.

- Requires each agent, j, to know $\bar{r}_1, \ldots, \bar{r}_P$

Warning:

- If on the line, then don't move for the player, then it's optimal, so don't move it.

Example:
1. 4-1, pg. 29-35

Example:
1. 4-1, pg. 46-47

Example:
1. **Given/Problem:** Find all equilibria of the following one-shot game or state that none exist.

	B1 (y)	**B2 (1-y)**
A1 (x)	(5, 3)	(1, 0)
A2 (1-x)	(0, 1)	(2, 4)

   - (#,#) is the payoff to P1 and P2 respectively for a given action profile.
2. **Solution:**
   (a) **Define Probabilities:**
   - Let y be the probability that B1 plays action B1 so $1-y$ is the probability that B1 plays action B2.
   - Let x be the probability that A1 plays action A1 so $1-x$ is the probability that A1 plays action A2.

   (b) **Expected Rewards:**
   - P1:

$$\begin{aligned} E[x] = 5xy + 1x(1-y) + 0(1-x)y + 2(1-x)(1-y) &= 5xy + x - xy + 2 - 2x - 2y + 2xy \\ &= 5xy - xy + 2xy + x - 2x - 2y + 2 \\ &= 6xy - x - 2y + 2 \quad \text{simplify} \\ &= \underbrace{(6y-1)}_{c} x + 2 - 2y \quad \text{linear in } x \end{aligned}$$

   - P2:

$$\begin{aligned} E[y] = 3xy + 0x(1-y) + 1(1-x)y + 4(1-x)(1-y) &= 3xy + 0 + y - xy + 4 - 4x - 4y + 4xy \\ &= 3xy - xy + 4xy + y - 4x - 4y + 4 \\ &= 6xy - 4x - 3y + 4 \quad \text{simplify} \\ &= \underbrace{(6x-3)}_{c} y + 4 - 4x \quad \text{linear in } y \end{aligned}$$

   - **Note:** $E[x]$ is linear in x and $E[y]$ is linear in y.

   (c) **Constrained Argmax Expected Rewards w.r.t $x \in [0,1]$ (since P1):** If it was cost, then minimize. Also don't care about constant term in y since we are derivating w.r.t x.
   - P1:

$$\text{bs}_\text{A}(x) = \begin{cases} 1 & \text{if } y > \dfrac{1}{6} \text{ i.e. } c > 0 \text{ since positive want maximum positive} \\ [0,1] & \text{if } y = \dfrac{1}{6} \text{ i.e. } c = 0 \text{ doesn't matter since } 0 \\ 0 & \text{if } y < \dfrac{1}{6} \text{ i.e. } c < 0 \text{ since negative want maximum negative} \end{cases}$$

   - P2:

$$\text{bs}_\text{B}(y) = \begin{cases} 1 & \text{if } x > \dfrac{3}{6} \text{ i.e. } c > 0 \text{ since positive want maximum positive} \\ [0,1] & \text{if } x = \dfrac{3}{6} \text{ i.e. } c = 0 \text{ doesn't matter since } 0 \\ 0 & \text{if } x < \dfrac{3}{6} \text{ i.e. } c < 0 \text{ since negative want maximum negative} \end{cases}$$

(d) **Finding all equilibrium:** Lines on the graph represents where your reward is maximized.

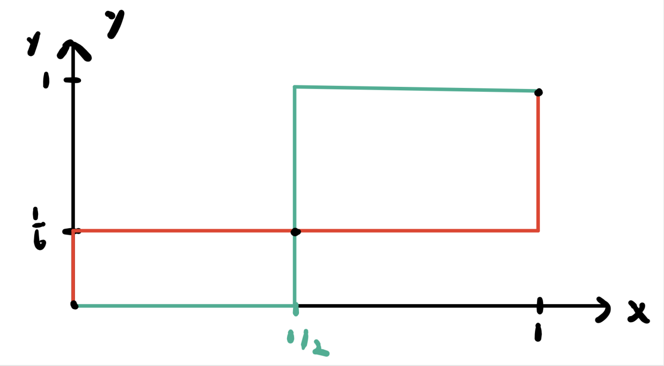

Figure 4

- **Case 1:** $x = 0$ and $y = 0$
  - $P(\text{P1 chooses A1}) = 0$
  - $P(\text{P1 chooses A2}) = 1$
  - $P(\text{P2 chooses B1}) = 0$
  - $P(\text{P2 chooses B2}) = 1$
- **Case 2:** $x = 1/2$ and $y = 1/6$
  - $P(\text{P1 chooses A1}) = 1/2$
  - $P(\text{P1 chooses A2}) = 1/2$
  - $P(\text{P2 chooses B1}) = 1/6$
  - $P(\text{P2 chooses B2}) = 5/6$
- **Case 3:** $x = 1$ and $y = 1$
  - $P(\text{P1 chooses A1}) = 1$
  - $P(\text{P1 chooses A2}) = 0$
  - $P(\text{P2 chooses B1}) = 1$
  - $P(\text{P2 chooses B2}) = 0$

(e) **Unstable Equilibrium:** P1 moves left and right b/c x is associated with x-axis. P2 moves up and down b/c y is associated with y-axis.

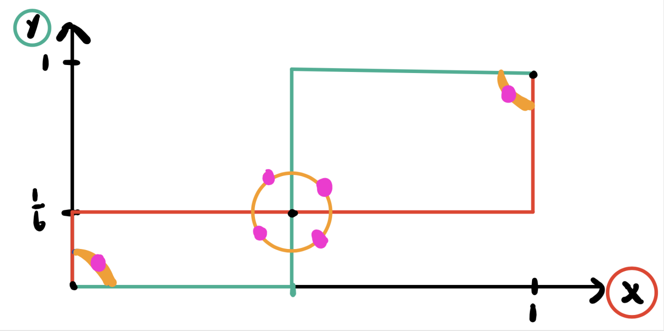

Figure 5

- Stability means that in a radius disc around the equilibrium, if you move a little bit, you will still be in the equilibrium (have to check all relevant quadrants)
  - If one quadrant is unstable, then don't need to check the other quadrants as the equilibrium point is unstable.
  - Simulatenous (both players move at the same time) and sequential (one player moves first and the other player moves second)
- **Case 1:** $x = 0$ and $y = 0$ is stable
  - Q1: Always converges to $(0, 0)$ since P1 moves left to red and P2 moves down to turquoise.
- **Case 2:** $x = 1/2$ and $y = 1/6$ is unstable
  - Q1 (Top Left): P1 moves right to red and P2 moves up to turquoise $\implies$ $(1, 1)$

- Q2 (Top Right): P1 moves right to red and P2 moves up to turquoise $\implies (1, 1)$
- Q3 (Bottom Left): P1 moves left to red and P2 moves down to turquoise $\implies (0, 0)$
- Q4 (Bottom Right): P1 moves left to red and P2 moves down to turquoise $\implies (0, 0)$

- **Case 3:** $x = 1$ and $y = 1$ is stable
  - Q1: Always converges to $(1, 1)$ since P1 moves left to red and P2 moves down to turquoise.

Example:

1. **Given:** A 2-player game where each player has two actions. Let a_1 and a_2 denote the actions of Player 1 and Player 2, respectively. Let $r_1(a_1, a_2)$ and $r_2(a_1, a_2)$ denote the payoffs to Player 1 and Player 2.

a_1	a_2	$r_1(a_1, a_2)$	$r_2(a_1, a_2)$
1	1	2	0
1	2	0	1
2	1	0	1
2	2	4	0

   Define strategy variables:
   - Let Player 1 play $a_1 = 1$ with probability x, and $a_1 = 2$ with probability $1 - x$.
   - Let Player 2 play $a_2 = 1$ with probability y, and $a_2 = 2$ with probability $1 - y$.
2. **Problem:**
   Find the mixed strategy Nash equilibrium (x^*, y^*) for this game.
3. **Solution:**
   (a) **Convert:**

	$a_2 = 1$	$a_2 = 2$
$a_1 = 1$	$(2, 0)$	$(0, 1)$
$a_1 = 2$	$(0, 1)$	$(4, 0)$

   (b) **Expected payoff for Player 1:**

$$\begin{aligned}\mathbb{E}[r_1] &= 2xy + 4(1-x)(1-y) \\ &= 2xy + 4 - 4x - 4y + 4xy \\ &= 6xy - 4x - 4y + 4 \\ &= (6y - 4)x + 4 - 4y\end{aligned}$$

   (c) **Expected payoff for Player 2:**

$$\begin{aligned}\mathbb{E}[r_2] &= x(1-y) + (1-x)y \\ &= x - xy + y - xy \\ &= -2xy + x + y \\ &= (1 - 2x)y + x\end{aligned}$$

   (d) **Best response conditions:**

$$x = \begin{cases} 1 & \text{if } y > \frac{2}{3} \\ [0,1] & \text{if } y = \frac{2}{3} \\ 0 & \text{if } y < \frac{2}{3} \end{cases} \qquad y = \begin{cases} 0 & \text{if } x > \frac{1}{2} \\ [0,1] & \text{if } x = \frac{1}{2} \\ 1 & \text{if } x < \frac{1}{2} \end{cases}$$

   (e) **Equilibrium:**

$$\text{At } x = \frac{1}{2}, \quad y = \frac{2}{3} \Rightarrow \text{Mixed Strategy Nash Equilibrium is } \left(\frac{1}{2}, \frac{2}{3}\right)$$

Example:

1. **Given:** Consider a 2-player single-action game, where each player has 2 actions. Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$ denote the mixed strategies of Player 1 and Player 2, respectively.

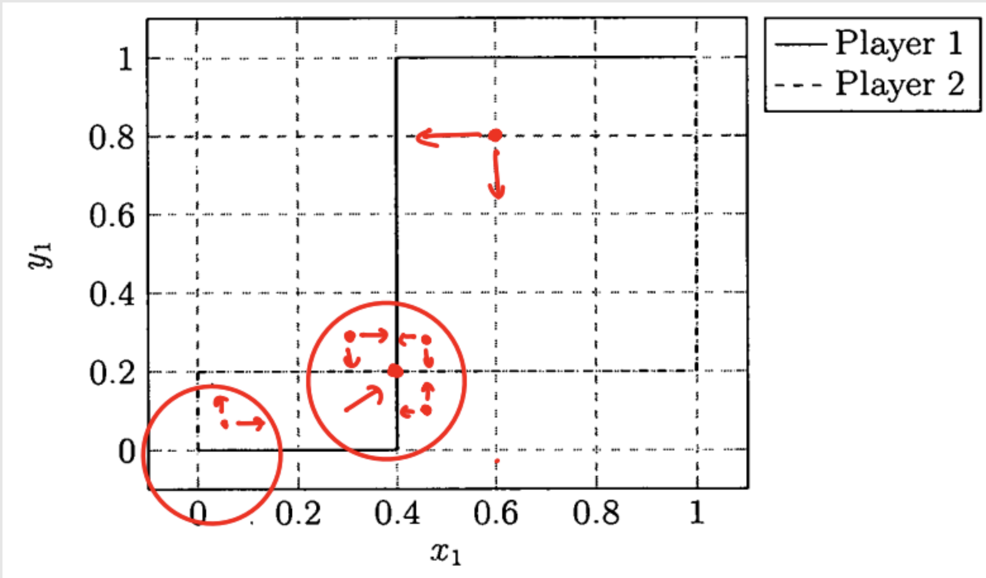

Figure 6: Best response functions $\mathrm{bsp}_1(y_1)$ and $\mathrm{bsp}_2(x_1)$

2. **Problem:** For each of the following strategy profiles expressed as $((x_1, x_2), (y_1, y_2))$, determine whether it is:
   - an equilibrium of the game or not, and
   - if it is an equilibrium, whether it is stable or unstable under best-response dynamics.
3. **Solution:**
   i. $((0, 1), (0, 1))$
   Explanation: As shown in the plot, arrows diverge from this point rather than converge. Therefore:

   $$\boxed{\text{Unstable Equilibrium}}$$

   ii. $((0.6, 0.4), (0.8, 0.2))$
   Explanation: This point does not lie on either player's best response curve. Therefore:

   $$\boxed{\text{Non-equilibrium}}$$

   iii. $((0.4, 0.6), (0.2, 0.8))$
   Explanation: As shown in the plot, arrows converge towards this point. Therefore:

   $$\boxed{\text{Stable Equilibrium}}$$

1.5.4 Simplifying Games

Process:

1. Fix action of other player, find best action for current player that maximizes its own reward.
2. Repeat by fixing all actions of other player.
3. Cross out actions that don't get used.

Example:

1. **Given:** A 2-player action game where each player has 3 actions. Let a_1 and a_2 represent the actions chosen by Player 1 and Player 2, respectively. The payoffs for each action profile are as follows:

a_1	a_2	$r_1(a_1, a_2)$	$r_2(a_1, a_2)$
1	1	3	0
1	2	2	1
1	3	1	2
2	1	4	2
2	2	3	3
2	3	5	5
3	1	6	1
3	2	4	4
3	3	3	3

2. **Problem:** Determine which action(s) are strictly dominated for Player 1. Eliminate them and write the reduced payoff matrix.
3. **Solution:**
   (a) **Step 1:** Fix the opponent's actions and determine Player 1's best response for each:
   - Fix $a_2 = 1$: $\max\{3, 4, 6\} \Rightarrow a_1 = 3$
   - Fix $a_2 = 2$: $\max\{2, 3, 4\} \Rightarrow a_1 = 3$
   - Fix $a_2 = 3$: $\max\{1, 5, 3\} \Rightarrow a_1 = 2$

   (b) **Step 2:** Observe that $a_1 = 1$ is never a best response for any a_2.

$$\Rightarrow a_1 = 1 \text{ is strictly dominated. Eliminate it.}$$

Reduced Payoff Matrix:

$a_1 \backslash a_2$	1	2	3
2	$(4, 2)$	$(3, 3)$	$(5, 5)$
3	$(6, 1)$	$(4, 4)$	$(3, 3)$

Example:

1. **Given:**

	Fight	Share
Fight	$(1,1)$	$(4,0)$
Share	$(0,4)$	$(2,2)$
Leave	$(-4,8)$	$(-2,8)$

Figure 7

2. **Problem:** Simplify the game
3. **Solution:**
   (a) 'Leave' action for Cavemen is dominated so it will never choose this action b/c it can get more reward by choosing Fight or Share.
   (b) As aresult, we can remove it from the game and have a 2 action, 2 player game, rather than a 2-3 action, 2 player game.

1.5.5 Iterative Game

Example:

1. **Given:** The best response curves $\text{bsp}_1(y_1)$ and $\text{bsp}_2(x_1)$ for Player 1 and Player 2 are given below:

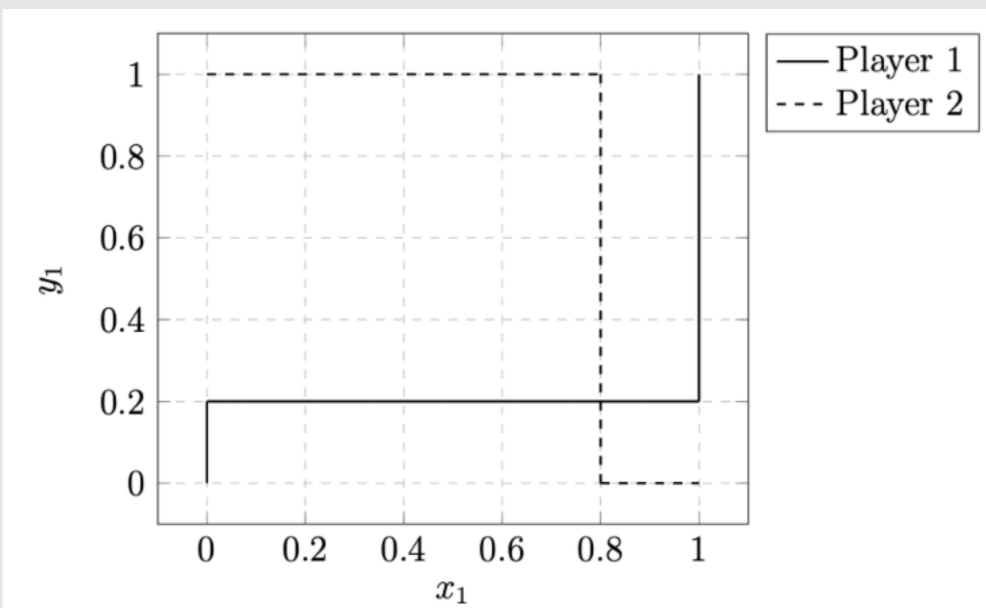

Figure 8

Let $x_i^{(k)}$ and $y_i^{(k)}$ represent the probabilities with which Player 1 and Player 2 choose Action $i \in \{1, 2\}$ after k iterations of best-response play.

2. **Problem:** Determine the specified probability values of Player 1 or Player 2's mixed strategies after a certain number of iterations of best-response dynamics, based on the initial probabilities and the given best-response plot.
3. **Solution:**

i. Find $x_2^{(3)}$ when $x_1^{(0)} = 0.2$ and $y_1^{(0)} = 0.4$.

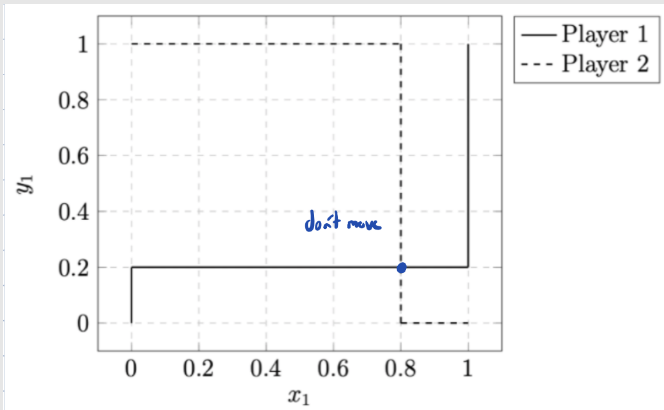

Figure 9

$$x_2^{(3)} = 1/1$$

ii. Find $x_2^{(19)}$ when $x_1^{(0)} = 0.8$ and $y_1^{(0)} = 0.2$.

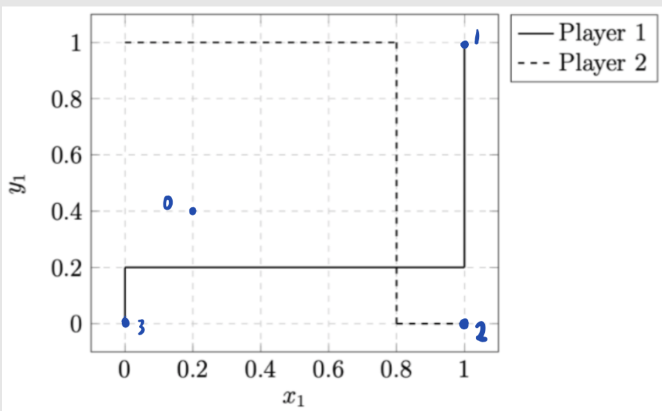

Figure 10

$$x_2^{(19)} = 1/5$$

iii. Find $y_2^{(221)}$ when $x_2^{(0)} = 1.0$ and $y_1^{(0)} = 0.2$.

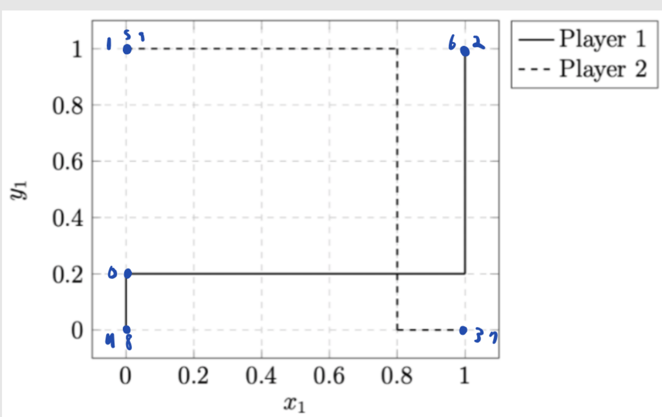

Figure 11

- Cycle is 4 $\implies$ 221 mod 4 = 1, so ends up in top left w/ $x_1^{(221)} = 0$, $y_1^{(221)} = 1 \implies y_2^{(221)} = 0$.

$$y_2^{(221)} = 0/1$$