ROB311 Quiz 3

Hanhee Lee

March 30, 2025

Contents

One-Shot Multi-Agent Decision Problems

1 Multi-Agent Problems

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 $\mathrm{tr} : \mathcal{T} \to \mathcal{S}$	$\mathrm{tr}(s, a) =$ state transition from s under a
Reward to each agent, i defined by r_i: $\mathcal{Q} \times \mathcal{S} \to \mathbb{R}_+$	$r_i(s, a, \mathrm{tr}(s, a)) =$ 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 Action Equilibria

1.1.1 Finding Action Equilibria

1.2 Strategy Equilibria

1.2.1 Finding Strategy Equilibria

1.2.2 Existence of Stategy Equilibria

1.2.3 Convergence of Stategy Equilibria

1.3 Examples

1.3.1 Finding Action Equilibria

1.3.2 Optimal Action Profiles

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:

$$x = \begin{cases} 1 & \text{if } y > \frac{1}{6} \text{ i.e. } c > 0 \\ [0,1] & \text{if } y = \frac{1}{6} \text{ i.e. } c = 0 \\ 0 & \text{if } y < \frac{1}{6} \text{ i.e. } c < 0 \end{cases}$$

   - P2:

$$y = \begin{cases} 1 & \text{if } x > \frac{3}{6} \text{ i.e. } c > 0 \\ [0,1] & \text{if } x = \frac{3}{6} \text{ i.e. } c = 0 \\ 0 & \text{if } x < \frac{3}{6} \text{ i.e. } c < 0 \end{cases}$$

   (d)