Game Redesign in No-regret Game Playing

Yuzhe Ma, Young Wu, Xiaojin (Jerry) Zhu

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Motivation

- Many real-world problems are intrinsically multi-agent games
  - Rock-Paper-Scissors
  - Gambling
  - Decision making in economic or societal fields.

- Players are selfish: Nash Equilibrium might lead to suboptimal global objective.
- Shape the behavior (selected actions) of the players.

Mechanism Design

- Designer is the rule maker
  - Designer may not have full control over the game
- Assume agents are rational players
  - In case of multiple NE, which NE is adopted by rational players

Game Redesign

- The original loss function is $\ell^o(a) = \left(\ell_1^o(a), \dots, \ell_M^o(a)\right), \ell_i^o(a) \in [L, U], \forall i$

- Players apply no-regret learning algorithms (e.g., EXP3.P) to play the game T rounds

- In round $t = 1, \dots, T$:
  - Players take actions $a^t = (a_1^t, \dots, a_M^t)$
  - Original loss is $\ell^o(a^t)$
  - Designer changes the loss to $\ell(a^t)$
  - Player i observes loss $\ell_i(a^t)$ instead of $\ell_i^o(a^t)$
  - Designer incurs redesign cost $C(\ell^o, \ell, a^t)$ (e.g., $||\ell^o(a^t) - \ell(a^t)||_1$)

Game Redesign Goal

- Force all players to take a target action profile $a^\dagger$ as often as possible

$$\sum_{t=1}^{T} 1\{a^t = a^\dagger\}$$

- Small cumulative redesign cost

$$\sum_{t=1}^{T} C(\ell^o, \ell, a^t)$$

Interior Design

Assumption: $\ell_i^o\left(a^\dagger\right) \in [L+\rho, U-\rho]$ for some $\rho \in (0, \frac{U-L}{2})$

Redesign strategy:

$$\forall i, a, \ell_i(a) = \begin{cases} \ell_i^o\left(a^\dagger\right) - \left(1 - \frac{d(a)}{M}\right)\rho & if\ a_i = a_i^\dagger, \\ \ell_i^o\left(a^\dagger\right) + \frac{d(a)}{M}\rho & if\ a_i \neq a_i^\dagger \end{cases}$$

where $d(a) = \sum_{j=1}^{M} 1\{a_j = a_j^\dagger\}$

Key Ideas Behind Our Redesign

$$\forall i, a, \ell_i(a) = \begin{cases} \ell_i^o\left(a^\dagger\right) - \left(1 - \frac{d(a)}{M}\right)\rho & if\ a_i = a_i^\dagger, \\ \ell_i^o\left(a^\dagger\right) + \frac{d(a)}{M}\rho & if\ a_i \neq a_i^\dagger \end{cases}$$

(1). For player i, $\ell\left(a_i^\dagger, a_{-i}\right) = \ell(a_i, a_{-i}) - \left(1 - \frac{1}{M}\right)\rho\ (induced\ regret)$

(2). $\ell^o\left(a^\dagger\right) = \ell\left(a^\dagger\right)\ (no\ design\ cost\ when\ target\ is\ selected)$

The designer can force all players to follow a target action profile in almost every but $O(T^\alpha)$ $(\alpha < 1)$ rounds while incurring $O(T^\alpha)$ redesign cost.

Boundary Design

Assumption: $\exists i, \ell_i^o\left(a^\dagger\right) \in \{L, U\}$

The designer can force all players to follow a target action profile in almost every but $O\left(T^{\frac{1+\alpha}{2}}\right)$ $(\alpha < 1)$ rounds while incurring $O\left(T^{\frac{1+\alpha}{2}}\right)$ redesign cost.

The Tragedy of Commons

- 2 farmers, each can farm 0 to 15 sheep
- The price of a sheep is $\sqrt{30-(a_1+a_2)}$
- Payoff of farmer 1 is $a_1 \times \sqrt{30-(a_1+a_2)}$ (similar for farmer 2)

$$Nash\ Equilibrium:\ a^* = (12, 12)$$

- Social welfare: $(a_1+a_2) \times \sqrt{30-(a_1+a_2)}$ maximized at $a_1 + a_2 = 20$
- Social equality: $a_1 = a_2 = 10$
- Designer goal: $a^\dagger = (10, 10)$
- Redesign forces $a^\dagger$ in 98% of rounds when $T = 10^7$.
- The average design cost in each round is 0.5 (loss range is $[-15\sqrt{15}, 0]$)

Thanks!

Contact: ma234@wisc.edu