Bayes correlated equilibria and no-regret dynamics

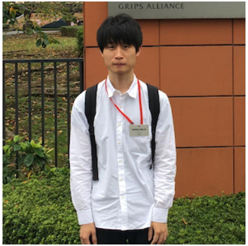
Kaito Fujii (National Institute of Informatics)

16 Feb. 2024 @ UTokyo

Self-introduction

Kaito Fujii (National Institute of Informatics, Tokyo)

- I obtained my PhD from Prof. Iwata (UTokyo) (combinatorial optimization group)
- I joined NII in 2020 and started game theory



Research Topics

Optimization, operations research, machine learning

Goals of algorithmic game theory

Goal 1 **Computing equilibria efficiently**

- Is it possible to compute equilibria of a given game in reasonable time?
- If it is difficult, is it possible to find an evidence for difficulty?

Goal 2 **Guaranteeing quality of equilibria** (price of anarchy)

- In the worst equilibria, how much does social welfare deteriorate?

This study aims to achieve these two goals for Bayesian games

※ There are various other goals (e.g., computing auctions, cooperative games)

Table of Contents

Two-player zero-sum games

A game is zero-sum $\overset{\Delta}{\Leftrightarrow}$ the total payoff is always zero

		🧑 ✊	🧑 ✌️	🧑 🖐️
👮‍♀️	✊	0, 0	1, -1	-1, 1
	✌️	-1, 1	0, 0	1, -1
	🖐️	1, -1	-1, 1	0, 0

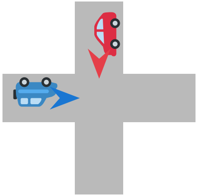
Example:

rock paper scissors

Nash equilibria:
both players
choose every action
with prob. $1/3$ (unique)

Two-player non-zero-sum games

🚙 and 🚗 at an intersection decide whether to go or to stop



		🚗	
		Go	**Stop**
🚙	**Go**	0, 0	4, 3
	Stop	3, 4	1, 1

Nash equilibria:

1. (Go, Stop)
2. (Stop, Go)
3. Both choose Go and Stop with prob. $1/2$

Computing Nash equilibria

Problem Compute any Nash equilibrium given a payoff table

Is there an algorithm that runs in time polynomial in #actions?

- **Two-player zero-sum games**: Yes

  Linear-programming-based algorithm [von Neumann 1928, Khachiyan'79]

- **Two-player non-zero-sum games**: No (probably)

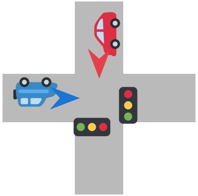
  This problem is PPAD-complete [Chen–Deng–Teng'09]

Computer scientists "believe" that solving it in poly-time is impossible

Question Is there any equilibrium concept easy to compute?

Correlated equilibria

Players’ actions can be correlated via a traffic signal



	Go	Stop
Go	0, 0	4, 3
Stop	3, 4	1, 1

Correlated equilibria:

infinitely many including Nash eq.

e.g.) (Go, Stop) with prob. $1/2$

(Stop, Go) with prob. $1/2$

Correlated equilibria

$N = \{1, 2, \ldots, n\}$ players $\quad N = \{$🚙, 🚗$\}$

A_i finite set of actions for player $i \in N$ $\quad A_i = \{\text{Go}, \text{Stop}\}$

$A = A_1 \times A_2 \times \cdots \times A_n$ set of action profiles

$v_i \colon A \to [0, 1]$ utility function for player $i \in N$ $\quad u_{🚙}(\text{Go}, \text{Stop}) = 4$

Definition

A distribution over action profiles $\pi \in \Delta(A)$ is a correlated equilibrium

$\overset{\triangle}{\Leftrightarrow}$ For any player $i \in N$ and deviation $\phi \colon A_i \to A_i$,

$$\mathop{\mathbb{E}}_{a \sim \pi}[v_i(\phi(a_i), a_{-i})] \le \mathop{\mathbb{E}}_{a \sim \pi}[v_i(a)].$$

※ If π is a product distribution, this definition coincides with Nash equilibria

Correlated equilibria

Definition

A distribution over action profiles $\pi \in \Delta(A)$ is a correlated equilibrium

$\overset{\triangle}{\Leftrightarrow}$ For any player $i \in N$ and deviation $\phi \colon A_i \to A_i$,

$$\mathop{\mathbb{E}}_{a \sim \pi}[v_i(\phi(a_i), a_{-i})] \leq \mathop{\mathbb{E}}_{a \sim \pi}[v_i(a)].$$

	Go	Stop
Go	0, 0	4, 3
Stop	3, 4	1, 1

We can define a CE $\pi \in \Delta(A)$ as follows:

$\pi(\text{Go}, \text{Stop}) = 1/2$, $\pi(\text{Stop}, \text{Go}) = 1/2$

Each player cannot increase the payoff by any ϕ

e.g., $\phi(\text{Go}) = \text{Stop}$, $\phi(\text{Stop}) = \text{Stop}$ decreases it

LP formulation of correlated equilibria

The set of CEs is expressed by linear constraints with $|A|$ variables

$$\mathrm{CE} = \left\{ \pi \in [0,1]^A \;\middle|\; \begin{array}{l} \displaystyle\sum_{\substack{a\in A:\\ a_i=a_i'}} \pi(a)[v_i(a) - v_i(a_i'', a_{-i})] \leq 0 \ (\forall i \in N, \forall a_i', a_i'' \in A_i) \\ \displaystyle\sum_{a\in A} \pi(a) = 1 \end{array} \right\}$$

If the number of players is a constant, the size of this LP is polynomial

$\rightarrow$ The problem of finding (also optimizing) a CE is tractable [Khachiyan'79]

Question How about cases where the number of players is large?

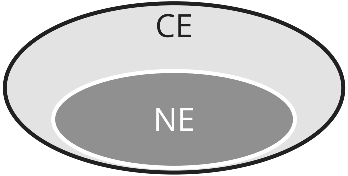
Computing correlated equilibria

Theorem [Foster–Vohra'97, Hart–Mas-Collel'00, Blum–Mansour'07]

There exists a poly-time algo. for computing a CE of n-player games

- Since v_i requires space exponential in n, we assume oracle access to v_i
- An ϵ-approximate CE is obtained in time polynomial in n, $\max_{i \in N} |A_i|$, and $1/\epsilon$

cf. Computing Nash equilibria is PPAD-complete even for two-player games



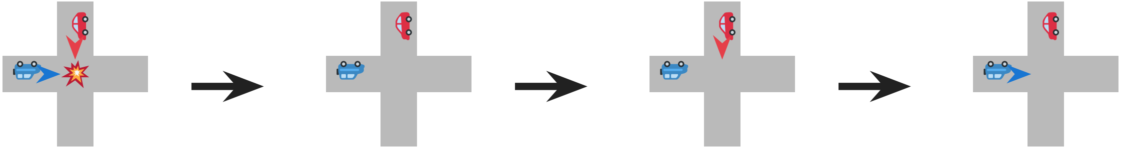
The problem of computing **any** CE is

easier than computing **any** NE

No-regret dynamics

Algorithm Simulate **no-regret dynamics** converging to a CE

Players learn their strategy in repeated play of the same game



for $t = 1, 2, \ldots, T$ **do**

Each player $i \in N$ decides a (mixed) strategy $\pi_i^t \in \Delta(A_i)$

All players' strategies $(\pi_i^t)_{i \in N}$ are revealed to each other

Each player i obtains reward $\mathbb{E}[v_i(a^t)]$, where $a_i^t \sim \pi_i^t$ independently $(\forall i)$

Swap regret [Blum–Mansour'07]

$$\mathrm{SwapRegret}_i^T \triangleq \max_{\phi\colon A_i \to A_i} \sum_{t=1}^{T} \underbrace{\mathbb{E}\left[v_i^t(\phi(a_i^t), a_{-i}^t)\right]}_{\text{reward in round } t \text{ if the actions are replaced according to } \phi} - \sum_{t=1}^{T} \underbrace{\mathbb{E}\left[v_i^t(a^t)\right]}_{\text{reward in round } t}$$

Theorem [Blum–Mansour'07]

If swap regret of every player grows sublinearly in T,

the empirical distribution converges to a correlated equilibrium

↑ The uniform mixture of action profiles of T rounds

※ Another variant called *internal regret* does not work for Bayes correlated equilibria

Table of Contents

Price of anarchy (PoA)

the social welfare achieved by the worst equilibrium

$$\mathrm{PoA} \triangleq \frac{\inf_{\pi:\ \text{equilibrium}} \mathbb{E}_{a\sim\pi}[v_{\mathrm{SW}}(a)]}{\max_{a\in A} v_{\mathrm{SW}}(a)}$$

the optimal social welfare

$v_{\mathrm{SW}}: A \to \mathbb{R}_{\geq 0}$ social welfare

usually $v_{\mathrm{SW}}(a) \triangleq \sum_{i\in N} v_i(a)$

※ **PoA depends on the equilibrium concept** (PoA for NE, etc.)

	Coop.	Defect
Cooperate	10, 10	0, 15
Defect	15, 0	1, 1

In the prisoners' dilemma game, the PoA can be close to 0

the worst equilibrium: 2 at (Defect, Defect)

the optimal: 20 at (Cooperate, Cooperate)

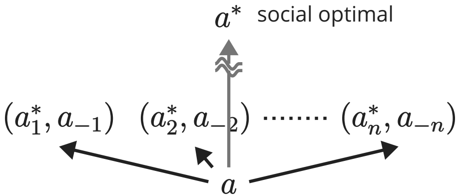
Smoothness framework (1/2) [Roughgarden'15]

Question For what class of games is the PoA lower-bounded?

Definition [Roughgarden'15]

An n-player game is (λ, μ)-smooth

$$\overset{\triangle}{\Leftrightarrow} \forall a, a^* \in A\colon \underbrace{\sum_{i=1}^{n} v_i(a_i^*, a_{-i})}_{\text{Player } i \text{ switches from } a_i \text{ to } a_i^*} \geq \lambda \underbrace{v_{\mathrm{SW}}(a^*)}_{\text{social welfare achieved by } a^*} - \mu \underbrace{v_{\mathrm{SW}}(a)}_{\text{social welfare achieved by } a}$$



The deviations significantly increase social welfare towards the optimal

Smoothness framework (2/2) [Roughgarden'15]

Smooth games are a broad class of games with bounded PoA

Theorem [Roughgarden'15]

PoA for correlated equilibria is at least $\dfrac{\lambda}{1+\mu}$ in (λ, μ)-smooth games

※ Roughgarden further proved this bound for *coarse correlated equilibria*

Examples of smooth games

Congestion games, various auctions, competitive facility location, effort market games, competitive information spread, ...

Table of Contents

👮‍♀️ and 🧑‍🦰 independently decide where to go

👮‍♀️ prefers sea 🌊, while 🧑‍🦰 prefers mountain 🌋

		🧑‍🦰	
		🌊	🌋
👮‍♀️	🌊	4, 3	1, 1
	🌋	0, 0	3, 4

Bayesian games [Harsanyi'67]

Players' types are generated from a common prior distribution

Each of 👮‍♀️ and 🧑‍🦰 prefers 🌊 and 🌋 with prob. $1/2$ for each

(Each player knows the prior distribution only, not the others' types)

w.p. $1/4$		🧑‍🦰 type: 🌊	
		🌊	🌋
👮‍♀️ type: 🌊	🌊	4, 4	1, 0
	🌋	0, 1	3, 3

w.p. $1/4$		🧑‍🦰 type: 🌋	
		🌊	🌋
👮‍♀️ type: 🌊	🌊	4, 3	1, 1
	🌋	0, 0	3, 4

Notations for Bayesian games

$N = \{1, 2, \ldots, n\}$ players — $N = \{$👮, 🧒$\}$

A_i finite set of actions for player $i \in N$ — $A_1 = A_2 = \{$🌊, 🌋$\}$

Θ_i finite set of types for player $i \in N$ — $\Theta_1 = \Theta_2 = \{$type:🌊, type:🌋$\}$

$A = \prod_{i \in N} A_i$ action profiles, $\Theta = \prod_{i \in N} \Theta_i$ type profiles

$\rho \in \Delta(\Theta)$ prior distribution over type profiles — $\rho($type:🌊, type:🌊$) = 1/4$

$v_i : \Theta \times A \to [0, 1]$ utility function for player $i \in N$ — $v_1($type:🌊, type:🌊; 🌊, 🌋$) = 1$

Computational studies on Bayesian games

- **Equilibrium computation**:

  Computing Bayes Nash equilibria (BNE) is PPAD-complete

  Existing algorithms can compute weak equilibria (Bayes coarse CE)

  [Hartline–Syrgkanis–Tardos'15]

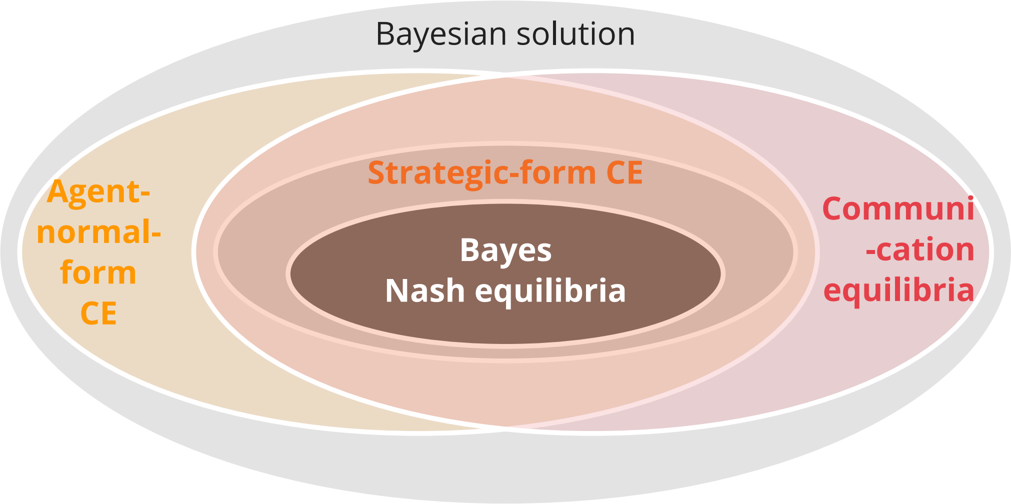
- **Price of anarchy**

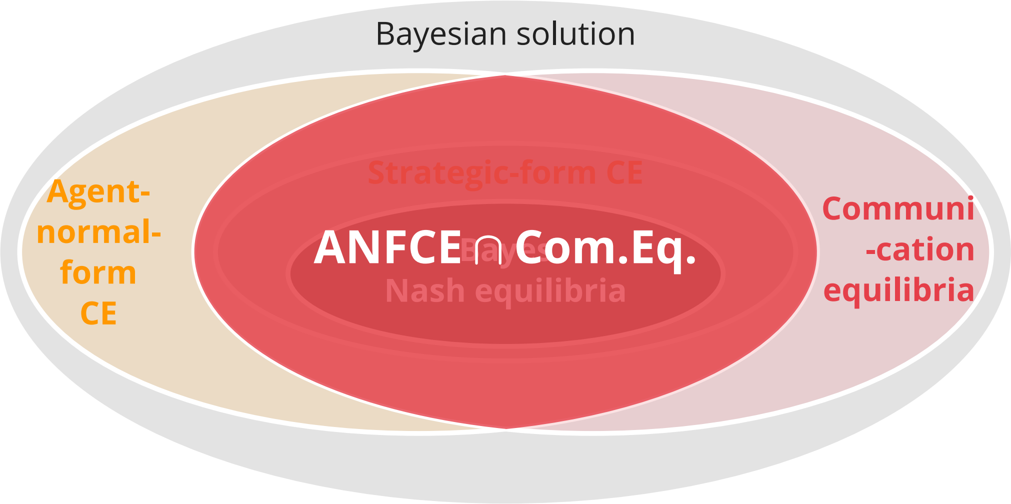
  Smoothness framework provides PoA bounds only for BNE

  [Roughgarden'15b, Syrgkanis–Tardos'13]

Q **Is there any equilibrium concept that has both merits?**

Various Bayes correlated equilibria [Forges'93]



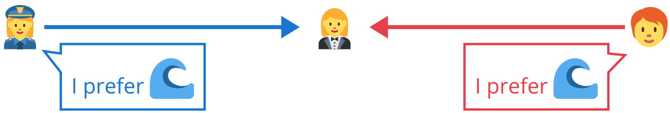


Bayesian solution
Agent-normal-form CE
Strategic-form CE
ANFCE ∩ Com.Eq.
Bayes Nash equilibria
Communi-cation equilibria

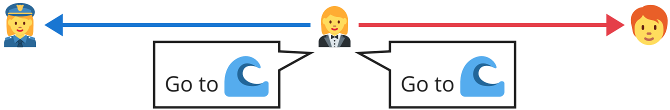
Communication equilibria [Myerson'82, Forges'86]

Equilibria realized by a credible mediator 🤵‍♀️

1. Each player tells the mediator 🤵‍♀️ their types



2. The mediator 🤵‍♀️ sends a recommendation to each



Same type → Recommend their preferred place

Different types → Recommend 🌋 or 🌊 each with prob. 1/2

w.p. 1/4 — 🧑 type: 🌊; 👮‍♀️ type: 🌊

	🌊	🌋
🌊	4, 4	1, 0
🌋	0, 1	3, 3

w.p. 1/4 — 🧑 type: 🌋; 👮‍♀️ type: 🌊

	🌊	🌋
🌊	4, 3	1, 1
🌋	0, 0	3, 4

Communication equilibria [Myerson'82, Forges'86]

Definition

A distribution $\pi \in \Delta(A)^{\Theta}$ is a communication equilibrium

$\overset{\Delta}{\Leftrightarrow}$ For any player $i \in N$, $\psi \colon \Theta_i \to \Theta_i$, and $\phi \colon \Theta_i \times A_i \to A_i$,

$$\mathop{\mathbb{E}}_{\theta \sim \rho}\left[\mathop{\mathbb{E}}_{a \sim \pi(\psi(\theta_i), \theta_{-i})}[v_i(\theta; \phi(\theta_i, a_i), a_{-i})]\right] \leq \mathop{\mathbb{E}}_{\theta \sim \rho}\left[\mathop{\mathbb{E}}_{a \sim \pi(\theta)}[v_i(\theta; a)]\right].$$

Two incentive constraints

1. No incentive to **tell an untrue type** (represented by ψ)
2. No incentive to **disobey the recommendation** (represented by ϕ)

Agent-normal-form correlated equilibria

ANFCE is defined as CE of the agent normal form

Agent normal form of Bayesian games

The same player with different types are regarded as different players

Only (hypothetical) players with realized types play the game

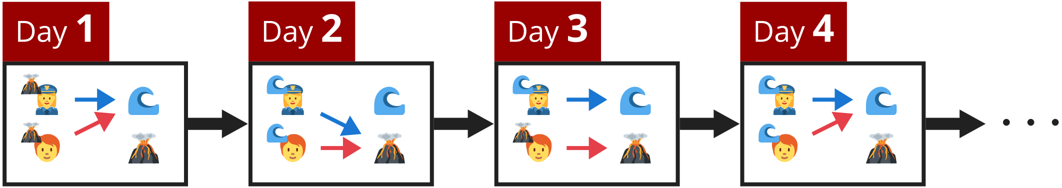
In our example, randomly selected two out of (👮, 🌊), (👮, 🌋), (👦, 🌊), (👦, 🌋) play the game

Difference from communication equilibria:

- No incentive constraint for truthful type telling
- The distribution must satisfy some technical condition called **strategy representability**

Our contribution 1: dynamics

We propose no-regret dynamics converging to ANFCE ∩ Com.Eq.



In repeated play, players aim to minimize **untruthful swap regret**

defined later

Theorem (informal)

Dynamics with $o(T)$ untruthful swap regret converge to ANFCE ∩ Com.Eq. and can be simulated by the proposed algorithm in polynomial time

Our contribution 2: PoA bounds

PoA bounds for ANFCE ∩ Com.Eq. via smoothness arguments

Previous results PoA bounds for **BNE** via smoothness

↓ extend [Roughgarden'15b, Syrgkanis–Tardos'13]

Our results PoA bounds for **ANFCE ∩ Com.Eq.** via smoothness

※ PoA decreases as equilibria get broader (the worst equilibrium considered)

Theorem (informal)

PoA for ANFCE ∩ Com.Eq. is at least $\lambda/(1+\mu)$ if a game for each fixed $\theta \in \Theta$ is (λ, μ)-smooth

Applications:

$v_{\mathrm{SW}} = \sum_i v_i$ case,

various auctions, ...

Table of Contents

No-regret dynamics in Bayesian games

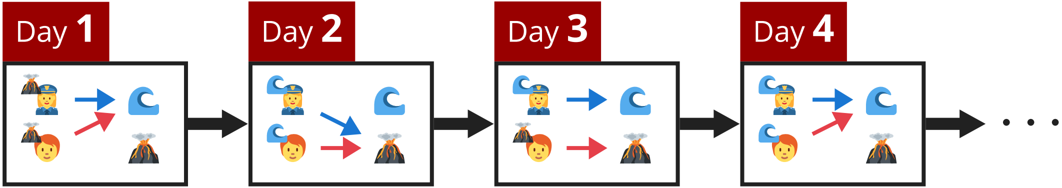
For $t = 1, 2, \ldots, T$**:**

Each player $i \in N$ decides a (mixed) strategy $\pi_i^t \in \Delta(A_i)^{\Theta_i}$

All players' strategies $(\pi_i^t)_{i \in N}$ are revealed to each other

Each player i obtains reward $\mathbb{E}[v_i(\theta; a^t)]$,

where $\theta \sim \rho$ and then $a_i^t \sim \pi_i^t(\theta_i)$ independently for each i



※ We consider the expected value w.r.t. θ and a in each round

Untruthful swap regret

Untruthful swap regret for player $i \in N$

$$R^T_{\mathrm{US},i} = \max_{\substack{\psi:\,\Theta_i \to \Theta_i \\ \phi:\,\Theta_i \times A_i \to A_i}} \sum_{t=1}^{T} \mathbb{E}_{\theta_i \sim \rho_i} \left[\mathbb{E}_{a_i \sim \pi_i^t(\psi(\theta_i))} \left[u_i^t(\theta_i, \phi(\theta_i, a_i)) \right] \right] - \sum_{t=1}^{T} \mathbb{E}_{\theta_i \sim \rho_i} \left[\mathbb{E}_{a_i \sim \pi_i^t(\theta_i)} \left[u_i^t(\theta_i, a_i) \right] \right],$$

where $u_i^t(\theta_i, a_i) \triangleq \mathbb{E}_{\theta_{-i} \sim \rho_{-i}|\theta_i} \left[\mathbb{E}_{a_{-i} \sim \pi_{-i}^t(\theta_{-i})} [v_i(\theta; a)] \right]$ is the reward vector at round t

(ρ_i the marginal distribution, $\rho_{-i}|\theta_i$ the conditional distribution)

Two incentive constraints for communication equilibria

1. No incentive to **tell an untrue type** (represented by ψ)
2. No incentive to **disobey the recommendation** (represented by ϕ)

Untruthful swap regret minimization

Suppose each player minimizes USR against adversarial players

Upper bound Φ-regret minimization framework + decomposition

Theorem

The proposed algo. achieves $R_{\mathrm{US},i} = O\left(\sqrt{T \max\{|A_i| \log |A_i|, \log |\Theta_i|\}}\right)$

Lower bound Analyze a hard instance with optimal stopping theory

Theorem

Any algorithm satisfies $R_{\mathrm{US},i} = \Omega\left(\sqrt{T \max\{|A_i| \log |A_i|, \log |\Theta_i|\}}\right)$

External regret minimization algo.

$u^t \in [0,1]^A$ reward vector in round $t \in [T]$

$\pi^t \in \Delta(A)$ mixed strategy in round $t \in [T]$ ※ Subscript i is omitted from now on

$$\mathrm{ExternalRegret}^T \triangleq \max_{a^* \in A} \sum_{t=1}^{T} u^t(a^*) - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{a^t \sim \pi^t} \left[u^t(a^t) \right]$$

Multiplicative Weights Update method: Initialize $\pi^1(a) = 1/|A|$ ($\forall a \in A$),

For each $t \in [T]$: Update $\pi^{t+1}(a) \propto \pi^t(a) \exp(\eta u^t(a))$ ($\forall a \in A$)

Theorem [Cesa-Bianchi–Lugosi'07]

If $\eta = \sqrt{\frac{\log |A|}{T}}$, MWU achieves $\mathrm{ExternalRegret}^T = O\left(\sqrt{T \log |A|}\right)$

Swap regret minimization algo. [Blum–Mansour'07]

$$\mathrm{SwapRegret}^T \triangleq \max_{\phi:\, A_i \to A_i} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{a^t \sim \pi^t}\left[u^t(\phi(a^t))\right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{a^t \sim \pi^t}\left[u^t(a^t)\right]$$

$$\mathrm{SwapRegret}^T \triangleq \max_{Q \in \mathcal{Q}} \sum_{t=1}^{T} \langle Q\pi^t, u^t \rangle - \sum_{t=1}^{T} \langle \pi^t, u^t \rangle,$$

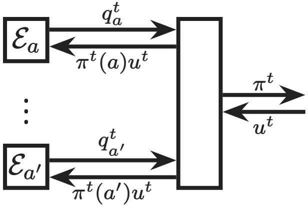
where $\mathcal{Q} = \left\{ Q \in [0,1]^{A \times A} \;\middle|\; \mathbf{1}Q = \mathbf{1} \right\}$

$$\mathrm{SwapRegret}^T \triangleq \max_{Q \in \mathcal{Q}} \sum_{t=1}^{T} \langle Q, \pi^t \otimes u^t \rangle - \sum_{t=1}^{T} \langle Q^t, \pi^t \otimes u^t \rangle \text{ if } Q^t \pi^t = \pi^t \text{ for all } t \in [T]$$

Swap regret minimization algo. [Blum–Mansour'07]

$$\text{SwapRegret}^T \triangleq \max_{Q \in \mathcal{Q}} \sum_{t=1}^{T} \langle Q, \pi^t \otimes u^t \rangle - \sum_{t=1}^{T} \langle Q^t, \pi^t \otimes u^t \rangle \text{ if } Q^t \pi^t = \pi^t \text{ for all } t \in [T]$$

1: Initialize subroutines $(\mathcal{E}_a)_{a \in A}$ for external regret minimization with actions A
2: **for** $t = 1, 2, \ldots, T$ **do**
3: Let $q_a^t \in \Delta(A)$ be the output of subroutine $\mathcal{E}_a$ for each $a \in A$
4: Let Q^t be an $|A| \times |A|$ matrix with each column q_a^t
5: Find $\pi^t \in \Delta(A)$ such that $\pi^t = Q^t \pi^t$
6: Observe u^t and feed $\pi^t(a)u^t$ to subroutine $\mathcal{E}_a$



Untruthful swap regret minimization algo.

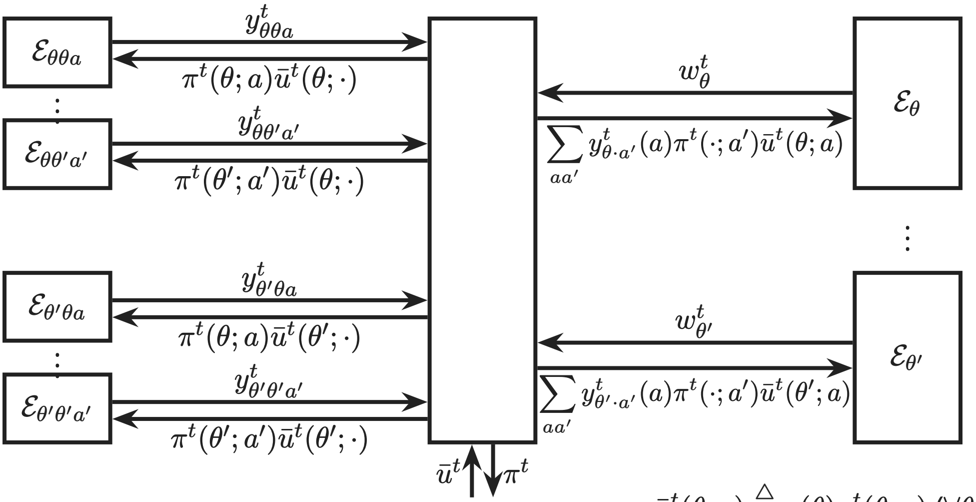
$$R^T_{\mathrm{US},i} = \max_{\substack{\psi:\,\Theta\to\Theta \\ \phi:\,\Theta\times A\to A}} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{a\sim\pi^t(\psi(\theta))} [u^t(\theta, \phi(\theta, a))] \right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{a\sim\pi^t(\theta)} [u^t(\theta, a)] \right]$$

$$\mathrm{SwapRegret}^T \triangleq \max_{Q\in\mathcal{Q}} \sum_{t=1}^{T} \langle Q\pi^t, u^t\rangle - \sum_{t=1}^{T} \langle \pi^t, u^t\rangle, \text{ where}$$

$$\mathcal{Q} = \left\{ Q \in [0,1]^{(\Theta\times A)\times(\Theta\times A)} \;\middle|\; \begin{array}{l} \text{there exists some } W \in [0,1]^{\Theta\times\Theta} \text{ such that} \\ \sum_{\theta'\in\Theta} W(\theta,\theta') = 1 \ (\forall\theta\in\Theta) \text{ and} \\ \sum_{a\in A} Q((\theta,a),(\theta',a')) = W(\theta,\theta') \ (\forall\theta,\theta'\in\Theta, a'\in A) \end{array} \right\}$$

※ π^t and u^t are flattened to be a $|\Theta| \times |A|$ dimensional vector

Untruthful swap regret minimization algo.



$$\bar{u}^t(\theta, a) \triangleq \rho(\theta)u^t(\theta, a) \ (\forall \theta, a)$$

Full description of the algorithm

The set of types Θ_i and the set of actions A_i are specified in advance. The reward vector $u_i^t \in [0,1]^{\Theta_i \times A_i}$ is given at the end of each round $t \in [T]$.

Initialize subroutines as follows:

- let $\mathcal{E}_{\theta_i}$ be a multiplicative weights algorithm with decision space Θ_i for each $\theta_i \in \Theta_i$
- let $\mathcal{E}_{\theta_i, \theta_i', a_i'}$ be AdaHedge with decision space A_i for each $\theta_i, \theta_i' \in \Theta_i$ and $a_i' \in A_i$

for each round $t = 1, \ldots, T$ **do**

Let $w_{\theta_i}^t \in \Delta(\Theta_i)$ be the output of $\mathcal{E}_{\theta_i}$ in round t for each $\theta_i \in \Theta_i$

Let $y_{\theta_i, \theta_i', a_i'}^t \in \Delta(A_i)$ be the output of $\mathcal{E}_{\theta_i, \theta_i', a_i'}$ in round t for each $\theta_i, \theta_i' \in \Theta_i$ and $a_i' \in A_i$

Define $Q^t \in [0,1]^{(\Theta_i \times A_i) \times (\Theta_i \times A_i)}$ by $Q^t((\theta_i, a_i), (\theta_i', a_i')) = w_{\theta_i}^t(\theta_i') y_{\theta_i, \theta_i', a_i'}^t(a_i)$ for each $\theta_i, \theta_i' \in \Theta_i$ and $a_i, a_i' \in A_i$

Compute an eigenvector $x^t \in \mathbb{R}^{\Theta_i \times A_i}$ of Q^t such that $Q^t x^t = x^t$ and $(x^t)^\top \mathbf{1} = |\Theta_i|$

Decide the output $\pi_i^t \in \Delta(A_i)^{\Theta_i}$ by $\pi_i^t(\theta_i; a_i) = x^t(\theta_i, a_i)$ for each $\theta_i \in \Theta_i$ and $a_i \in A_i$

Observe reward vector $u_i^t \in [0,1]^{\Theta_i \times A_i}$ and feed reward vectors to subroutines as follows:

- feed $\displaystyle\sum_{a_i, a_i' \in A_i} y_{\theta_i, \theta_i', a_i'}^t(a_i) \pi_i^t(\theta_i'; a_i') \rho_i(\theta_i) u_i^t(\theta_i, a_i)$ as the reward for decision $\theta_i' \in \Theta_i$ to subroutine $\mathcal{E}_{\theta_i}$ for each $\theta_i \in \Theta_i$
- feed $\pi_i^t(\theta_i'; a_i') \rho_i(\theta_i) u_i^t(\theta_i, a_i)$ as the reward for decision $a_i \in A_i$ to subroutine $\mathcal{E}_{\theta_i, \theta_i', a_i'}$ for each $\theta_i, \theta_i' \in \Theta_i$ and $a_i' \in A_i$

Summary

ANFCE ∩ Com.Eq. in Bayesian games satisfies the following goals

Goal 1 **Efficient computation**

- No-regret dynamics converging to ANFCE ∩ Com.Eq.
- Algorithm for simulating the dynamics with the optimal convergence rate

Goal 2 **PoA bounds**

- Extension of the smoothness framework from BNE to ANFCE ∩ Com.Eq.

Reference

- Avrim Blum and Yishay Mansour. 2007. From External to Internal Regret. *Journal of Machine Learning Research* 8, 1307–1324.
- Xi Chen, Xiaotie Deng, and Shang-Hua Teng. 2009. Settling the complexity of computing two- player Nash equilibria. *Journal of the ACM* 56, 3, 14:1–14:57.
- Cesa-Bianch and Lugosi. 2007. *Prediction, Learning, and Games*, Cambridge University Press.
- Françoise Forges. 1986. An approach to communication equilibria. *Econometrica*, 1375–1385.
- Françoise Forges. 1993. Five legitimate definitions of correlated equilibrium in games with incomplete information. *Theory and Decision* 35, 277–310.
- Dean P Foster and Rakesh V Vohra. 1997. Calibrated learning and correlated equilibrium. *Games and Economic Behavior* 21(1-2), 40–55.
- John C. Harsanyi. 1967. Games with Incomplete Information Played by “Bayesian” Players, I–III. *Management Science* 14(3):159–182, 14(5):320–334, 14(7):486–502.
- Sergiu Hart and Andreu Mas-Colell. 2000. A simple adaptive procedure leading to correlated equilibrium. *Econometrica* 68(5), 1127–1150.
- Jason D. Hartline, Vasilis Syrgkanis, and Éva Tardos. 2015. No-Regret Learning in Bayesian Games. In *NIPS 2015*, 3061–3069.
- Tim Roughgarden. 2015a. Intrinsic Robustness of the Price of Anarchy. *Journal of the ACM* 62(5), 32:1–32:42.
- Tim Roughgarden. 2015b. The Price of Anarchy in Games of Incomplete Information. *ACM Transactions on Economics and Computation* 3(1), 6:1–6:20.
- Vasilis Syrgkanis and Éva Tardos. 2013. Composable and efficient mechanisms. In *STOC 2013*. 211–220.

Definition

We call a type-wise distribution $\pi \in \Delta(A)^\Theta$ *strategy-representable* if there exists $\sigma \in \Delta(S)$ such that $\pi(\theta; a) = \Pr_{s \sim \sigma}(s(\theta) = a)$ for each $\theta \in \Theta$ and $a \in A$.

Definition

For any $\epsilon \geq 0$, a distribution $\sigma \in \Delta(S)$ is an ϵ-approximate agent-normal-form correlated equilibrium if for any $i \in N$ and $\phi \colon \Theta_i \times A_i \to A_i$, it holds that

$$\mathop{\mathbb{E}}_{\theta \sim \rho}\left[\mathop{\mathbb{E}}_{s \sim \sigma}[v_i(\theta; s(\theta))]\right] \geq \mathop{\mathbb{E}}_{\theta \sim \rho}\left[\mathop{\mathbb{E}}_{s \sim \sigma}[v_i(\theta; \phi(\theta_i, s_i(\theta_i)), s_{-i}(\theta_{-i}))]\right] - \epsilon. \quad (1)$$