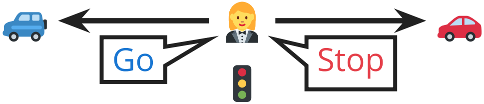
Bayes correlated equilibria and no-regret dynamics

Kaito Fujii (Kyoto University)

20 June 2026 @ SUFE ITCS

Overview

Correlated equilibrium

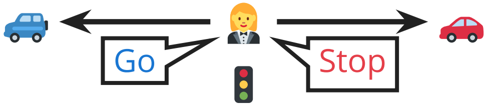


Go

Stop

Overview

Correlated equilibrium ← **Swap regret minimization**

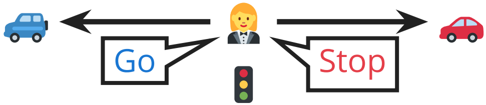


[Blum–Mansour’07]

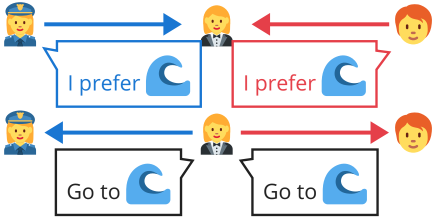
Overview

Correlated equilibrium ← Swap regret minimization

[Blum–Mansour’07]



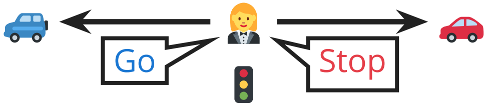
Communication equilibrium



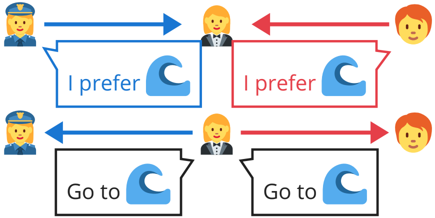
Overview

Correlated equilibrium ← Swap regret minimization

[Blum–Mansour’07]



Communication equilibrium ← Untruthful swap regret minimization



$O(\sqrt{T \max\{|A_i| \log |A_i|, \log |\Theta_i|\}})$ UB

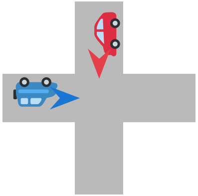
$\Omega(\sqrt{T \log |\Theta_i|})$ LB

Price-of-anarchy bounds

Table of Contents

Intersection game

🚙 and 🚗 independently decide whether to go or stop



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

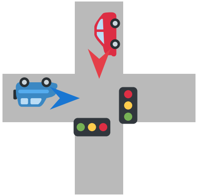
Nash equilibria

A state where no one can improve their expected payoff by deviating

- (Go, Stop)
- (Stop, Go)
- Players independently choose Go and Stop with probability $1/2$

Correlated equilibria

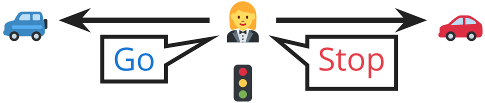
Players' actions can be arbitrarily correlated via a traffic signal



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

Correlated equilibria

Mediator 🤵‍♀️ recommends actions



cf. Players independently decide in NE

Infinitely many (including Nash eq.)
E.g.) (Go, Stop) and (Stop, Go) w.p. $1/2$

Correlated equilibria

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

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

$A = A_1 \times A_2 \times \cdots \times A_n$ set of action profiles $(\text{Go}, \text{Stop}) \in A$

$v_i \colon A \to \mathbb{R}$ utility function for player $i \in N$ $v_{🚙}(\text{Go}, \text{Stop}) = 4$

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 $\phi\colon A_i \to A_i$,

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

※ 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	3 4
Stop	4 3	1 1

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

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

Each player cannot increase the payoff by any ϕ

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

Table of Contents

Battle of the sexes (complete information)

👮‍♀️ and 🧑 choose their destinations independently

👮‍♀️ prefers sea 🌊, while 🧑 prefers mountain 🌋

		🧑	
		🌊	🌋
👮‍♀️	🌊	3 4	1 1
	🌋	0 0	4 3

Same place: 3 points

Preferred place: 1 point

Bayesian games (incomplete info. + common prior) [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)

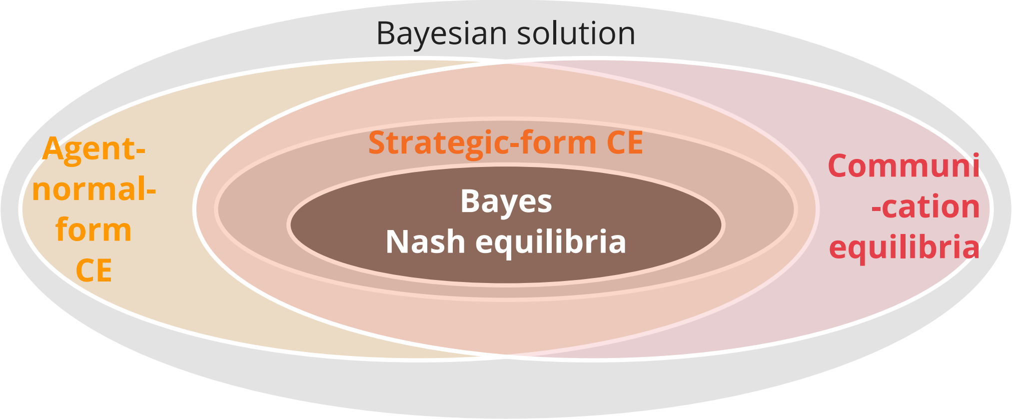
w.p. $1/4$		🧑 type: 🌊	
		🌊	🌋
👮‍♀️ type: 🌊	🌊	4 4	0 1
	🌋	1 0	3 3

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

Various Bayes correlated equilibria [Forges'93]

Bayes correlated equilibria (= correlated eq. in Bayesian games)

have many variants with various communication protocols



Bayesian solution

Agent-normal-form CE

Strategic-form CE

Bayes Nash equilibria

Communi-cation equilibria

Strategic-form correlated equilibria

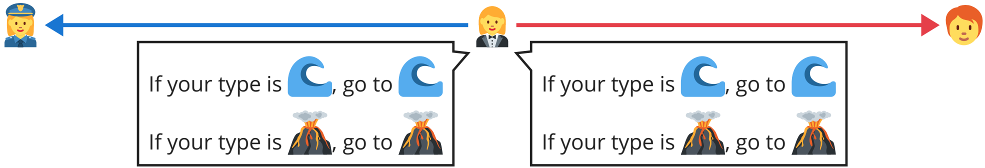
Strategic form of Bayesian games

A **strategy** $s_i : \Theta_i \to A_i$ is interpreted as an action

The set of all actions in this interpretation is $S_i := A_i^{\Theta_i}$

🤵‍♀️ **privately recommends an action for each type separately**

$\leftarrow$ **No incentive to disobey the recommendation**



If your type is 🌊, go to 🌊

If your type is 🌋, go to 🌋

If your type is 🌊, go to 🌊

If your type is 🌋, go to 🌋

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 SFCE

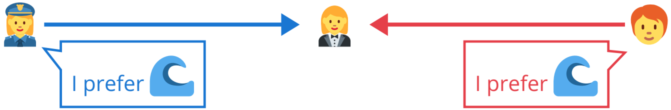
Each player cannot observe the recommendation to unrealized types

※ No realistic scenario involving a mediator 🤵‍♀️

Bayesian solutions [Forges'93]

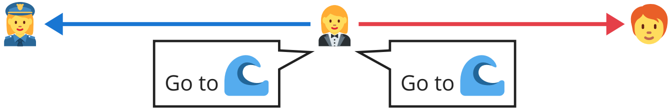
Mediator 🤵‍♀️ knows the true types in advance

1. Each player privately tells their **true** types to the mediator 🤵‍♀️



2. The mediator 🤵‍♀️ privately sends a recommendation to each player

← **No incentive to disobey the recommendation**



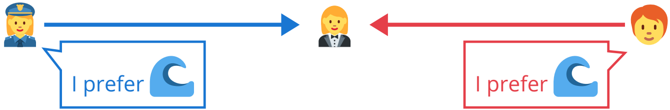
Communication equilibria [Myerson'82, Forges'86]

Equilibria realized by a credible third-party mediator 🤵‍♀️

Communication equilibria [Myerson’82, Forges’86]

Equilibria realized by a credible third-party mediator 🤵‍♀️

① Each player privately tells their types to the mediator 🤵‍♀️



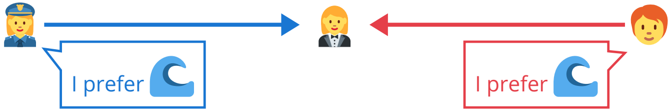
I prefer 🌊

I prefer 🌊

Communication equilibria [Myerson’82, Forges’86]

Equilibria realized by a credible third-party mediator 🤵‍♀️

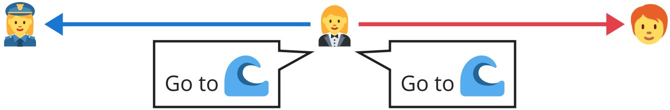
① Each player privately tells their types to the mediator 🤵‍♀️



I prefer 🌊

I prefer 🌊

② The mediator 🤵‍♀️ privately sends a recommendation to each player



Go to 🌊

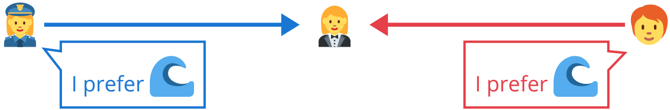
Go to 🌊

Communication equilibria [Myerson’82, Forges’86]

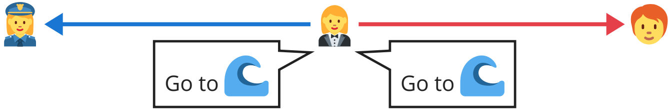
Equilibria realized by a credible third-party mediator 🤵‍♀️

① Each player privately tells their types to the mediator 🤵‍♀️

← **No incentive to tell an untrue type**



② The mediator 🤵‍♀️ privately sends a recommendation to each player

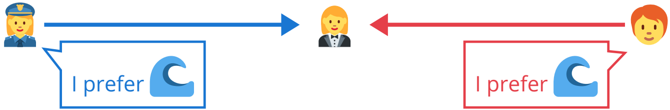


Communication equilibria [Myerson’82, Forges’86]

Equilibria realized by a credible third-party mediator 🤵‍♀️

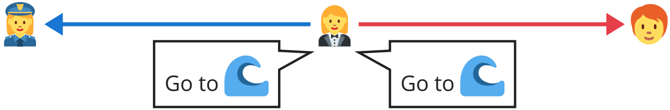
① Each player privately tells their types to the mediator 🤵‍♀️

← **No incentive to tell an untrue type**



② The mediator 🤵‍♀️ privately sends a recommendation to each player

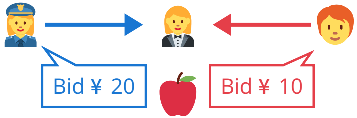
← **No incentive to disobey the recommendation**



Communication equilibrium combines ...

Mechanism design

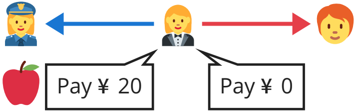
① Each player tells their types
← **No incentive to lie**



Bid ¥ 20

Bid ¥ 10

② 🤵‍♀️ decides the outcome
← This decision is binding



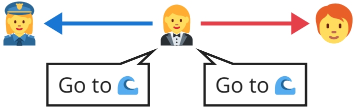
Pay ¥ 20

Pay ¥ 0

Correlated equilibria

① No type (complete info.)

② 🤵‍♀️ recommends actions
← **No incentive to deviate**



Go to 🌊

Go to 🌊

Communication equilibria [Myerson’82, Forges’86]

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

$N = \{$👮‍♀️, 🧑$\}$

Communication equilibria [Myerson'82, Forges'86]

$N = \{1, 2, \ldots, n\}$ players $N = \{$👮‍♀️, 🧑‍🦰$\}$

A_i action set, Θ_i type set for $i \in N$ $A_1 = A_2 = \{$🌊, 🌋$\}$, $\Theta_1 = \Theta_2 = \{$type:🌊, type:🌋$\}$

Communication equilibria [Myerson’82, Forges’86]

$N = \{1, 2, \ldots, n\}$ players $N = \{$👮‍♀️, 🧑‍🦰$\}$

A_i action set, Θ_i type set for $i \in N$ $A_1 = A_2 = \{$🌊, 🌋$\}$, $\Theta_1 = \Theta_2 = \{$type:🌊, type:🌋$\}$

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

Communication equilibria [Myerson’82, Forges’86]

$N = \{1, 2, \ldots, n\}$ players $N = \{$👮‍♀️, 🧑‍🦰$\}$

A_i action set, Θ_i type set for $i \in N$ $A_1 = A_2 = \{$🌊, 🌋$\}$, $\Theta_1 = \Theta_2 = \{$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:🌊$) = 1/4$

Communication equilibria [Myerson’82, Forges’86]

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

$N = \{$👮, 🧑$\}$

A_i action set, Θ_i type set for $i \in N$

$A_1 = A_2 = \{$🌊, 🌋$\}$, $\Theta_1 = \Theta_2 = \{$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:🌊$) = 1/4$

$v_i : \Theta \times A \to \mathbb{R}$ utility function for player $i \in N$

$v_1($type:🌊, type:🌊; 🌊, 🌋$) = 1$

Communication equilibria [Myerson'82, Forges'86]

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

A_i action set, Θ_i type set for $i \in N$ $A_1 = A_2 = \{$🌊, 🌋$\}$, $\Theta_1 = \Theta_2 = \{$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:🌊$) = 1/4$

$v_i \colon \Theta \times A \to \mathbb{R}$ utility function for player $i \in N$ $v_1($type:🌊, type:🌊; 🌊, 🌋$) = 1$

Definition

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

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

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

Communication equilibria [Myerson’82, Forges’86]

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

$N = \{$👮, 🧑$\}$

A_i action set, Θ_i type set for $i \in N$

$A_1 = A_2 = \{$🌊, 🌋$\}$, $\Theta_1 = \Theta_2 = \{$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:🌊$) = 1/4$

$v_i \colon \Theta \times A \to \mathbb{R}$ utility function for player $i \in N$

$v_1($type:🌊, type:🌊; 🌊, 🌋$) = 1$

Definition

A distribution $\pi \in \Delta(A)^{\Theta}$ i

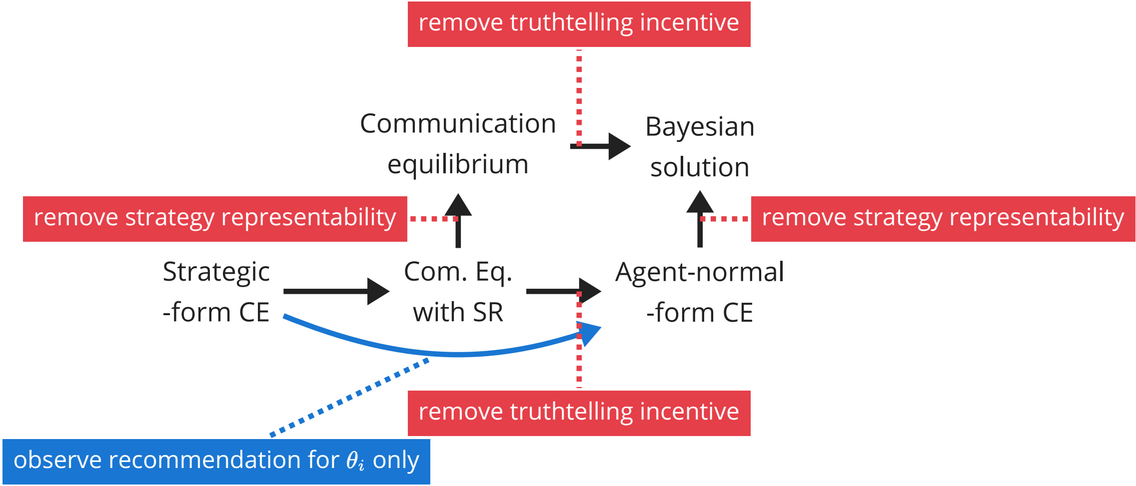
Misreporting $\psi(\theta_i)$ instead of true type θ_i

Choosing action $\phi(\theta_i, a_i)$ instead of recommended a_i

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

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

Relations among BCE concepts



remove truthtelling incentive

Communication equilibrium

Bayesian solution

remove strategy representability

remove strategy representability

Strategic -form CE

Com. Eq. with SR

Agent-normal -form CE

remove truthtelling incentive

observe recommendation for θ_i only

Table of Contents

ϵ-Approximate correlated equilibria

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

A_i finite set of actions for player $i \in N$

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

$v_i \colon A \to [0, 1]$ utility function for player $i \in N$

Definition

$\pi \in \Delta(A)$ is an ϵ-approximate correlated equilibrium

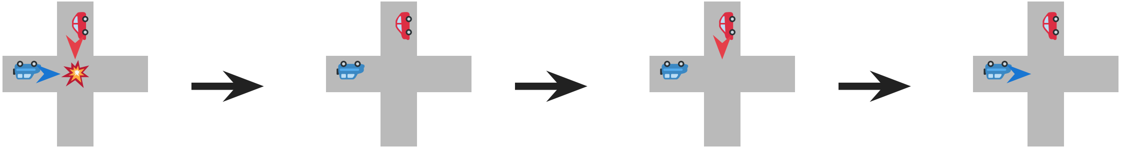
$\overset{\triangle}{\Leftrightarrow}$ For any player $i \in N$ and deviation $\phi \colon A_i \to A_i$,

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

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 $\pi_i^t \in \Delta(A_i)$
    Each player $i$ observes the reward vector $u_i^t(\cdot) \triangleq \mathbb{E}_{a_j^t \sim \pi_j^t (\forall j)} \left[ v_i(\cdot, a_{-i}^t) \right]$
    Each player $i$ obtains the expected reward $\mathbb{E}_{a_i^t \sim \pi_i^t}[u_i^t(a_i^t)]$
```

Swap regret [Blum–Mansour'07]

If 🚙 chooses **Stop** instead of **Go** and **Go** instead of **Stop**...

t	1	2	3	4	5	6
🚙 (reality)	**Stop**	**Stop**	**Go**	**Stop**	**Stop**	**Go**

🚗	**Stop**	**Stop**	**Go**	**Go**	**Stop**	**Go**
Reward	1	1	0	3	1	0

Swap regret is **the total regret under the optimal replacement**

Swap regret [Blum–Mansour'07]

If 🚙 chooses **Stop** instead of **Go** and **Go** instead of **Stop**...

t	1	2	3	4	5	6
🚙 (reality)	**Stop**	**Stop**	**Go**	**Stop**	**Stop**	**Go**
🚙 (hypothetical)	**Go**	**Go**	**Stop**	**Go**	**Go**	**Stop**
🚗	**Stop**	**Stop**	**Go**	**Go**	**Stop**	**Go**
Reward	$1 \to \mathbf{4}$	$1 \to \mathbf{4}$	$0 \to \mathbf{3}$	$3 \to \mathbf{0}$	$1 \to \mathbf{4}$	$0 \to \mathbf{3}$

Swap regret is **the total regret under the optimal replacement**

Swap regret [Blum–Mansour'07]

$$R^T_{\mathrm{swap},i} \triangleq \max_{\phi:\, A_i \to A_i} \sum_{t=1}^{T} \underbrace{\mathbb{E}_{a_i^t \sim \pi_i^t}\left[u_i^t(\phi(a_i^t))\right]}_{\text{reward in round } t \text{ if the actions are replaced according to } \phi} - \sum_{t=1}^{T} \underbrace{\mathbb{E}_{a_i^t \sim \pi_i^t}\left[u_i^t(a_i^t)\right]}_{\text{reward in round } t}$$

cf. (external) regret $R_i^T \triangleq \max_{a_i^* \in A_i} \sum_{t=1}^{T} u_i^t(a_i^*) - \sum_{t=1}^{T} \mathbb{E}_{a_i^t \sim \pi_i^t}\left[u_i^t(a_i^t)\right]$

Theorem [Blum–Mansour'07]

The empirical distribution $\frac{1}{T} \bigotimes_{i \in N} \pi_i^t$ is a $\left(\max_{i \in N} R^T_{\mathrm{swap},i}/T\right)$-approximate CE

Swap regret minimization [Blum–Mansour'07]

Step 1 **Express $\phi\colon A_i \to A_i$ using a stochastic matrix**

$$R^T_{\text{swap},i} \triangleq \max_{\phi\colon A_i \to A_i} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{a_i^t \sim \pi_i^t} \left[u_i^t(\phi(a_i^t)) \right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{a_i^t \sim \pi_i^t} \left[u_i^t(a_i^t) \right]$$

Using left stochastic matrices $\mathcal{Q} = \left\{ Q \in [0,1]^{A_i \times A_i} \,\middle|\, \mathbf{1}Q = \mathbf{1} \right\}$

$$R^T_{\text{swap},i} = \max_{Q \in \mathcal{Q}} \sum_{t=1}^{T} \langle Q\pi_i^t, u_i^t \rangle - \sum_{t=1}^{T} \langle \pi_i^t, u_i^t \rangle$$

Swap regret minimization [Blum–Mansour'07]

Step 2 **Reduction using a stationary distribution of Q**

$$R^T_{\mathrm{swap},i} = \max_{Q \in \mathcal{Q}} \sum_{t=1}^{T} \langle Q\pi_i^t, u_i^t \rangle - \sum_{t=1}^{T} \langle \pi_i^t, u_i^t \rangle$$

Reduce selection of π_i^t to selection of Q^t

Decide π_i^t from Q^t such that $Q^t\pi_i^t = \pi_i^t$ for each $t \in [T]$

$$\begin{aligned} R^T_{\mathrm{swap},i} &= \max_{Q \in \mathcal{Q}} \sum_{t=1}^{T} \langle Q\pi_i^t, u_i^t \rangle - \sum_{t=1}^{T} \langle Q^t\pi_i^t, u_i^t \rangle \\ &= \max_{Q \in \mathcal{Q}} \sum_{t=1}^{T} \langle Q, \pi_i^t \otimes u_i^t \rangle - \sum_{t=1}^{T} \langle Q^t, \pi_i^t \otimes u_i^t \rangle \end{aligned}$$

Swap regret minimization [Blum–Mansour’07]

Step 3 Decompose into $|A_i|$ external regret minimization

$$R^T_{\mathrm{swap},i} = \max_{Q \in \mathcal{Q}} \sum_{t=1}^{T} \langle Q, \pi_i^t \otimes u_i^t \rangle - \sum_{t=1}^{T} \langle Q^t, \pi_i^t \otimes u_i^t \rangle$$

Decompose Q^t into each column $q^t_{a_i}$

$$R^T_{\mathrm{swap},i} = \sum_{a_i \in A_i} \left[\max_{q^*_{a_i} \in \Delta(A_i)} \sum_{t=1}^{T} \langle q^*_{a_i}, \pi_i^t(a_i) u_i^t \rangle - \sum_{t=1}^{T} \langle q^t_{a_i}, \pi_i^t(a_i) u_i^t \rangle \right]$$

$R^T_{\mathrm{swap},i} = O\left(\sqrt{T|A_i| \log |A_i|}\right)$ from external regret min. bounds

Table of Contents

Learning dynamics in Bayesian games

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

Learning dynamics in Bayesian games

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

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

Learning 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

Learning 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

Learning 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

→ **Online learning with reward vector $u_i^t \in [0, 1]^{\Theta_i \times A_i}$** defined by

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

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

Untruthful swap regret

The regret definition corresponding to communication equilibria

Untruthful swap regret for player $i \in N$

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

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

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

Untruthful swap regret minimization [Fujii'25]

Step 1 Express $\psi\colon \Theta_i \to \Theta_i$ and $\phi\colon \Theta_i \times A_i \to A_i$ as a single matrix

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

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

$\bar{u}^t(\theta,a) \triangleq \rho(\theta)u^t(\theta,a)$ $(\forall\theta\in\Theta, a\in A)$ (i is omitted for simplicity)

$$R^T_{\mathrm{US},i} = \max_{Q\in\mathcal{Q}} \sum_{t=1}^T \langle Qx^t, \bar{u}^t\rangle - \sum_{t=1}^T \langle x^t, \bar{u}^t\rangle$$

Untruthful swap regret minimization [Fujii'25]

Step 2 **Reduction using an eigenvector of Q**

$$R^T_{\mathrm{US},i} = \max_{Q \in \mathcal{Q}} \sum_{t=1}^T \langle Q x^t, \bar{u}^t \rangle - \sum_{t=1}^T \langle x^t, \bar{u}^t \rangle$$

Reduce selection of Q^t to selection of π_i^t

Decide x^t from Q^t such that $Q^t x^t = x^t$ for $t \in [T]$

$$R^T_{\mathrm{US},i} = \max_{Q \in \mathcal{Q}} \sum_{t=1}^T \langle Q x^t, \bar{u}^t \rangle - \sum_{t=1}^T \langle Q^t x^t, \bar{u}^t \rangle$$
$$= \max_{Q \in \mathcal{Q}} \sum_{t=1}^T \langle Q, x^t \otimes \bar{u}^t \rangle - \sum_{t=1}^T \langle Q^t, x^t \otimes \bar{u}^t \rangle$$

Untruthful swap regret minimization [Fujii'25]

Step 3 **Decompose into $|\Theta_i|^2|A_i| + |\Theta_i|$ regret minimization**

$$R^T_{\mathrm{US},i} = \max_{Q\in\mathcal{Q}} \sum_{t=1}^{T} \langle Q, x^t \otimes \bar{u}^t \rangle - \sum_{t=1}^{T} \langle Q^t, x^t \otimes \bar{u}^t \rangle$$

Decompose into regret $R^T_{\theta_i}$ for each $\theta_i \in \Theta_i$ and

regret $R^T_{\theta_i,\theta'_i,a'_i}$ for each $\theta_i, \theta'_i \in \Theta_i$ and $a'_i \in A_i$

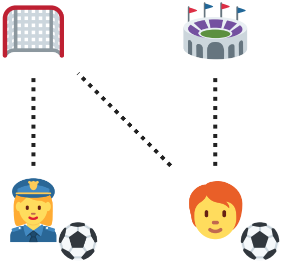
$$R^T_{\mathrm{US},i} \leq \sum_{\theta_i\in\Theta_i} R^T_{\theta_i} + \sum_{\theta_i\in\Theta_i} \max_{\theta'_i\in\Theta_i} \sum_{a'_i\in A_i} R^T_{\theta_i,\theta'_i,a'_i}$$

⇝ Upper bound $R^T_{\mathrm{US},i} = O\left(\sqrt{T \max\{|A_i| \log |A_i|, \log |\Theta_i|\}}\right)$

Table of Contents

Example of valid utility games

Players simultaneously choose a resource to share

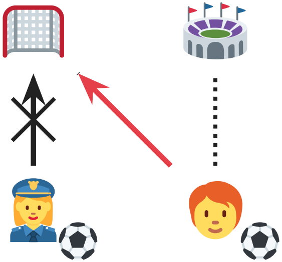


Resources chosen by multiple players are partitioned in a prespecified way

Example: 🧑 is prioritized over 👮‍♀️

Example of valid utility games

Players simultaneously choose a resource to share



Resources chosen by multiple players are partitioned in a prespecified way

Example: 🧑 is prioritized over 👮‍♀️

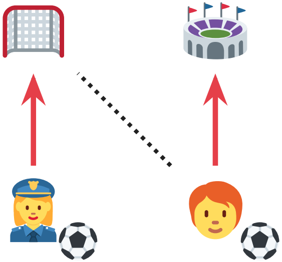
No player can benefit from deviations

◁⇕

Worst **Nash equilibrium** $= 1$

Example of valid utility games

Players simultaneously choose a resource to share



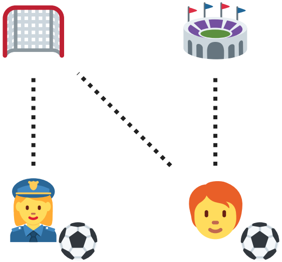
Resources chosen by multiple players are partitioned in a prespecified way

Example: 🧑 is prioritized over 👮‍♀️

Optimal social welfare $= 2$

Example of valid utility games

Players simultaneously choose a resource to share



Resources chosen by multiple players are partitioned in a prespecified way

Example: 🧑‍🦰 is prioritized over 👮‍♀️

No player can benefit from deviations

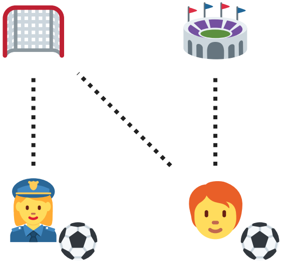
$$\triangleleft\Updownarrow$$

$$\text{PoA} := \frac{\text{Worst } \textbf{Nash equilibrium}}{\textbf{Optimal social welfare}} = \frac{1}{2}$$

(price of anarchy)

Example of valid utility games

Players simultaneously choose a resource to share



Resources chosen by multiple players are partitioned in a prespecified way

Example: 🧑 is prioritized over 👮‍♀️

No player can benefit from deviations

$$\triangleleft\Updownarrow$$

$$\text{PoA} := \frac{\text{Worst } \textbf{Nash equilibrium}}{\textbf{Optimal social welfare}} = \frac{1}{2}$$

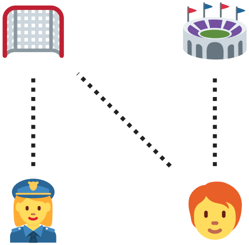
(price of anarchy)

Theorem [Vetta'02]

PoA ≥ 0.5 in any valid utility game

Example of valid utility games

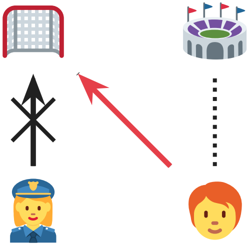
Q **How good or bad social welfare can be achieved by mediators**



A mediator 🤵‍♀️ sends recommendations
(🤵‍♀️ realizes correlated equilibrium)

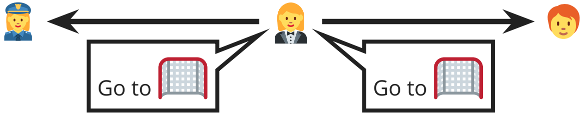
Example of valid utility games

Q **How good or bad social welfare can be achieved by mediators**



A mediator 🤵‍♀️ sends recommendations

(🤵‍♀️ realizes correlated equilibrium)

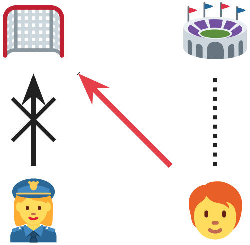


Go to 🥅

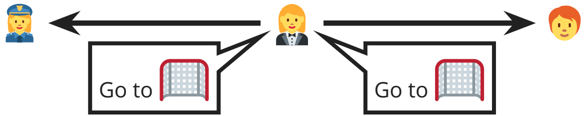
Go to 🥅

Example of valid utility games

Q **How good or bad social welfare can be achieved by mediators**



A mediator 🤵‍♀️ sends recommendations
(🤵‍♀️ realizes correlated equilibrium)

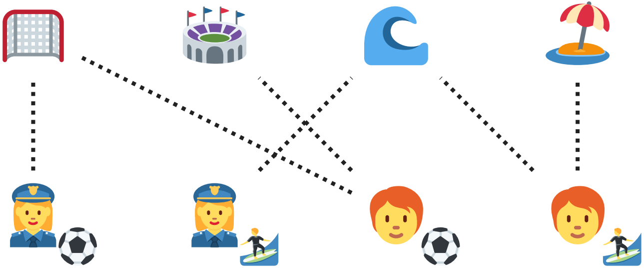


Theorem **[Roughgarden'15a]**

PoA ≥ 0.5 in any valid utility game for **correlated equilibria**

Example of Bayesian valid utility games

The set of actions for each player changes depending on their type

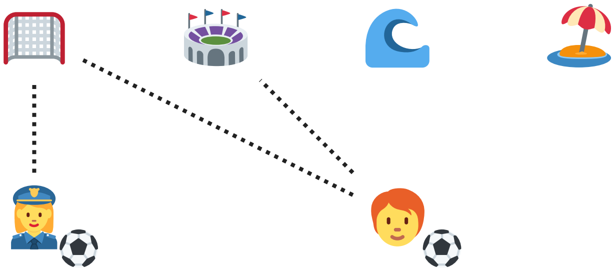


with prob. 0.5 with prob. 0.5 with prob. 0.5 with prob. 0.5

Q **How do mediators 🤵‍♀️ work in Bayesian games?**

Example of Bayesian valid utility games

The set of actions for each player changes depending on their type



with probability $1/4$

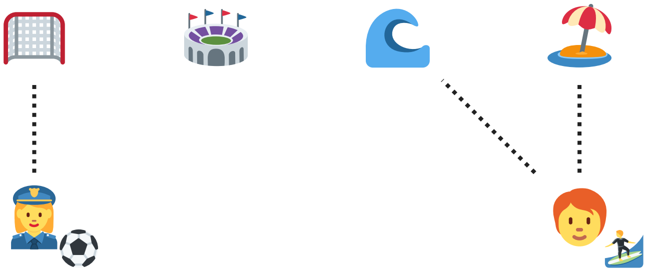
👮‍♀️ does not know 🧑‍🦰's type

🧑‍🦰 does not know 👮‍♀️'s type

Q **How do mediators 🤵‍♀️ work in Bayesian games?**

Example of Bayesian valid utility games

The set of actions for each player changes depending on their type



with probability $1/4$

👮‍♀️ does not know 🧑‍🦰's type

🧑‍🦰 does not know 👮‍♀️'s type

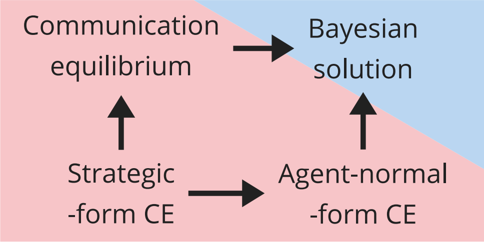
Q How do mediators 🤵‍♀️ work in Bayesian games?

Our results

For various equilibrium concepts, we provide PoA and PoS bounds

PoA bounds for independent priors

PoA $\in [0.316, 0.441]$



Communication equilibrium

Bayesian solution

Strategic -form CE

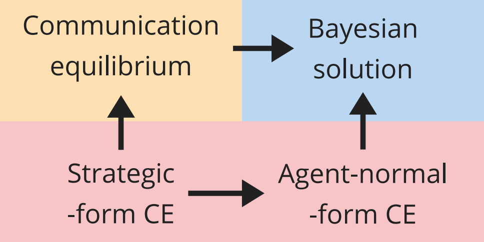
Agent-normal -form CE

PoA $= 0.5$

PoS bounds for independent priors

PoS $\in [1 - 1/e, 0.8]$

PoS $= 1$



Communication equilibrium

Bayesian solution

Strategic -form CE

Agent-normal -form CE

PoS $= 1 - 1/e$

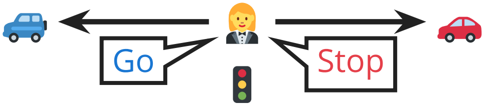
under the basic utility assumption

We also obtained PoA and PoS bounds for the correlated prior case

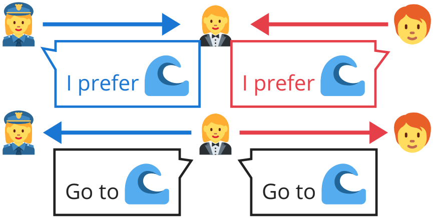
Overview

Correlated equilibrium ← **Swap regret minimization**

[Blum–Mansour’07]



Communication equilibrium ← **Untruthful swap regret minimization**



$O(\sqrt{T \max\{|A_i| \log |A_i|, \log |\Theta_i|\}})$ UB

$\Omega(\sqrt{T \log |\Theta_i|})$ LB

Price-of-anarchy bounds

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.
- Illustrations: ”Twemoji” by Twitter, Inc and other contributors is licensed under CC BY 4.0

Table of Contents

SFCE & Strategy swap regret

Definition

A distribution $\sigma \in \Delta(S_1 \times \cdots \times S_n)$ is an SFCE

$\overset{\triangle}{\Leftrightarrow}$ For any player $i \in N$, $\phi_{\mathsf{SF}} : S_i \to S_i$,

$$\mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{s\sim\sigma}[v_i(\theta; s_1(\theta_1), \ldots, s_n(\theta_n))]\right] \geq \mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{s\sim\sigma}[v_i(\theta; \phi_{\mathsf{SF}}(s_i)(\theta_i), s_{-i}(\theta_{-i}))]\right].$$

$$R^T_{\mathrm{SS},i} \triangleq \max_{\phi_{\mathsf{SF}} : S_i \to S_i} \sum_{t=1}^{T} \underbrace{\mathbb{E}\left[v_i(\phi_{\mathsf{SF}}(s_i^t)(\theta_i^t), a_{-i}^t)\right]}_{\text{reward in round } t \text{ if the strategies are replaced according to } \phi_{\mathsf{SF}}} - \sum_{t=1}^{T} \underbrace{\mathbb{E}\left[v_i(s_i^t(\theta_i^t), a_{-i}^t)\right]}_{\text{reward in round } t}$$

※ Each player chooses $\sigma_i^t \in \Delta(S_i)$ and generates $s_i^t \sim \sigma_i^t$

ANFCE & Type-wise swap regret

Definition

A distribution $\sigma \in \Delta(S_1 \times \cdots \times S_n)$ is an ANFCE

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

$$\mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{s\sim\sigma}[v_i(\theta; s_1(\theta_1), \ldots, s_n(\theta_n))]\right] \geq \mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{s\sim\sigma}[v_i(\theta; \phi(\theta_i, s_i(\theta_i)), s_{-i}(\theta_{-i}))]\right].$$

$$R^T_{\mathrm{TS},i} \triangleq \max_{\phi\colon \Theta_i \times A_i \to A_i} \sum_{t=1}^{T} \underbrace{\mathbb{E}\left[v_i^t(\phi(\theta_i, s_i^t(\theta_i)), a_{-i}^t)\right]}_{\text{reward in round } t \text{ if the actions are replaced according to } \phi} - \sum_{t=1}^{T} \underbrace{\mathbb{E}\left[v_i^t(s_i^t(\theta_i), a_{-i}^t)\right]}_{\text{reward in round } t}$$

Bayesian solutions

Definition

A distribution $\pi \in \Delta(A)^{\Theta}$ is a Bayesian solution

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

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

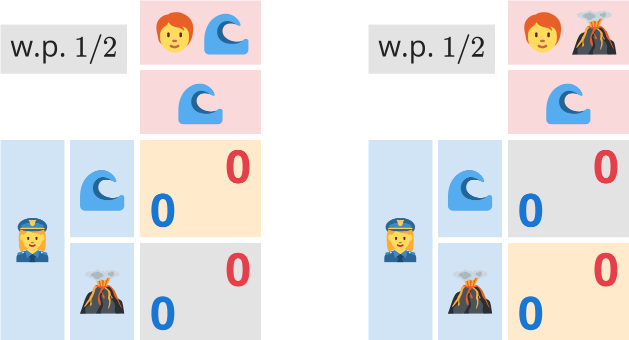
Difference from ANFCE

$\pi \in \Delta(A)^{\Theta}$ can express broader distributions than $\sigma \in \Delta(S)$,

which we call **strategy representability**

Example of non-SR distribution

This distribution cannot be realized by any strategy-profile distribution



w.p. 1/2		🧑 🌊
		🌊
👮‍♀️	🌊	0 0
	🌋	0 0

w.p. 1/2		🧑 🌋
		🌊
👮‍♀️	🌊	0 0
	🌋	0 0

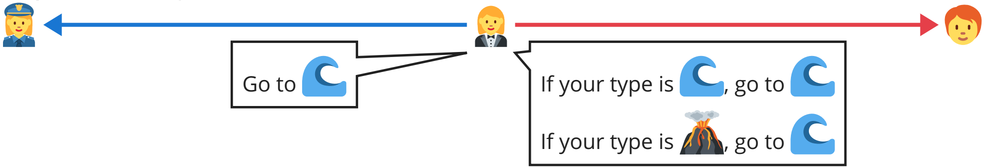
If 👮‍♀️ ’s type is 🌊,
🤵‍♀️ recommends 🌊

If 👮‍♀️ ’s type is 🌋,
🤵‍♀️ recommends 🌋

Example of ANFCE but not SFCE

We consider the following strategy-profile distribution

with probability 0.5,

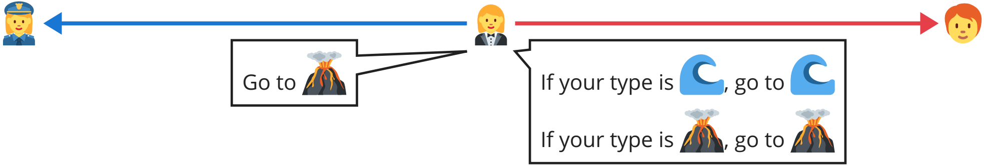


Go to 🌊

If your type is 🌊, go to 🌊

If your type is 🌋, go to 🌊

with probability 0.5,



Go to 🌋

If your type is 🌊, go to 🌊

If your type is 🌋, go to 🌋

ANFCE but not SFCE

In this distribution, 👮‍♀️ and 🧑‍🦰 (type: 🌋) are always recommended the same place

This distribution is not an SFCE because 🧑‍🦰 (type: 🌊) is always recommended 🌊 but can deviate to the same action as 👮‍♀️ by observing the recommendation to 🧑‍🦰 (type: 🌋)

On the other hand, this is an ANFCE because 🧑‍🦰 (type: 🌊) can observe only the recommendation to himself (always 🌊)

Table of Contents

Notations for Bayesian games

$N = \{1, 2, \ldots, n\}$ players $N = \{\text{👮‍♀️}, \text{👦}\}$

Θ_i finite set of types for player $i \in N$ $\Theta_{\text{👮‍♀️}} = \Theta_{\text{👦}} = \{\text{⚽}, \text{🏄}\}$

$A_i^{\theta_i}$ finite set of actions for player $i \in N$ with type $\theta_i \in \Theta_i$ $A_{\text{👦}}^{\text{⚽}} = \{\text{🥅}, \text{🏟️}\}$

$\Theta = \prod_{i \in N} \Theta_i$ type profiles

$\rho \in \Delta(\Theta)$ prior distribution over type profiles $\rho(\text{⚽}, \text{⚽}) = 1/4$

$v_i : A \to \mathbb{R}_{\geq 0}$ utility function for each player $i \in N$,

where $A = \prod_{i \in N} \left(\bigcup_{\theta_i \in \Theta_i} A_i^{\theta_i} \right)$ is the set of all possible action profiles

Equivalence of two formulations

Original formulation

A_i finite set of actions for player $i \in N$

$v_i : \Theta \times A \to \mathbb{R}$ utility function for player $i \in N$

(θ_i, a_i) as an action ↓ ↑ $A_i := \bigcup_{\theta_i} A_i^{\theta_i}$ and ignore actions for $\forall \theta_i' \neq \theta_i$

Type-dependent-action formulation

$A_i^{\theta_i}$ finite set of actions for player $i \in N$ with type $\theta_i \in \Theta_i$

$v_i : A \to \mathbb{R}_{\geq 0}$ utility function for each player $i \in N$,

where $A = \prod_{i \in N} \left(\bigcup_{\theta_i \in \Theta_i} A_i^{\theta_i} \right)$ is the set of all possible action profiles

Submodular social welfare functions

Let $E = \bigcup_{i \in N} \bigcup_{\theta_i \in \Theta_i} A_i^{\theta_i}$ be the set of all possible actions

Assumption **[Vetta'02]**

The social welfare function $f\colon 2^E \to \mathbb{R}$ is assumed to be

- **non-negative**: $f(X) \geq 0$ for any $X \subseteq E$
- **monotone**: $f(X \cup \{v\}) \geq f(X)$ for any $X \subseteq E$ and $v \in E$
- **submodular**: $f(X \cup \{v\}) - f(X) \geq f(Y \cup \{v\}) - f(Y)$

for any $X \subseteq Y \subseteq E$ and $v \in E \setminus Y$

Submodular social welfare functions

The marginal contribution to social welfare of each action decreases as other actions are added

$$\underbrace{f(\{\text{🥅👮‍♀️}\}) - f(\{\})}_{\substack{\text{The increase in social welfare} \\ \text{when no one attended yet}}} \geq \underbrace{f(\{\text{🥅👮‍♀️}, \text{🥅🧒}\}) - f(\{\text{🥅🧒}\})}_{\substack{\text{The increase in social welfare} \\ \text{when other players already attended}}}$$

Intuitively, this assumption is **substitutability** among players’ actions

※ Note that we assume this property even among the same player’s actions

Bayesian valid utility games

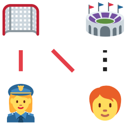
$v_i \colon A \to \mathbb{R}_{\geq 0}$ utility function for each player $i \in N$,

where $A = \prod_{i \in N} \left(\bigcup_{\theta_i \in \Theta_i} A_i^{\theta_i} \right)$ is the set of all possible action profiles

Assumption [Vetta'02]

- $\sum_{i \in N} v_i(a) \leq f(\{a_1, \ldots, a_n\})$ for any $a \in A$ (total utility condition)
- $v_i(a) \geq f(\{a_1, \ldots, a_n\}) - f(\{a_j \mid j \in N \setminus \{i\}\})$ for any $i \in N$ and $a \in A$

(marginal contribution condition)



The sum of utility values is at most $f($🥅$)$

The contribution of 👮‍♀️ is at least $f($🥅$) - f($🥅$) = 0$

Example: 🧒 gets all, 👮‍♀️ gets all, two players share equally, or both get 0