ベイズ相関均衡とリグレット最小化ダイナミクス

藤井海斗　(国立情報学研究所)

RAMP 2025 @ 九州大学西新プラザ

本発表の概要

1 **ベイジアンゲームにおけるさまざまな相関均衡の定義**

信頼できる第三者（仲介者）を介した情報交換によって相関均衡が実現

ベイジアンゲームでは相関均衡の自然な（非等価な）定義が複数知られている

2 **学習ダイナミクスによる相関均衡の計算**

ゲームの繰り返しの中でプレイヤーたちがリグレット最小化すると均衡に収束

このダイナミクスをシミュレートすることで相関均衡を効率的に計算

3 **整合的効用（劣モジュラ最大化）ゲームにおける社会厚生の保証**

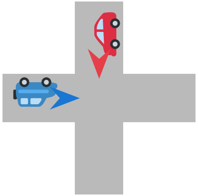
相関均衡において社会厚生（社会的な望ましさ）が近似的に最適

目次

目次

交差点のゲーム

交差点に進入する 2 台の車🚙と🚗がそれぞれ進むか停まるか決める



		🚗 進	🚗 停
🚙	進	0, 0	4, 3
🚙	停	3, 4	1, 1

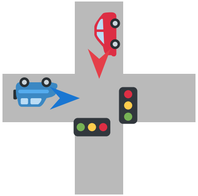
ナッシュ均衡

誰も逸脱しても期待利得が
改善しない安定した状態

- (進, 停)
- (停, 進)
- それぞれ独立に確率 1/2 ずつ

相関均衡 [Aumann’78]

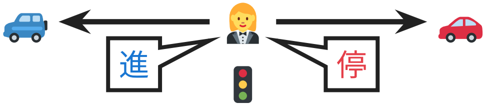
仲介者（交通信号など）によってプレイヤーたちの行動が任意に相関



		進	停
	進	0, 0	4, 3
	停	3, 4	1, 1

相関均衡

仲介者 🤵‍♀️ が各プレイヤーに行動推薦



cf. ナッシュ均衡では各自独立に行動選択

ナッシュ均衡を含む無限個存在

例）(進, 停) と (停, 進) を確率 1/2 ずつ

相関均衡 [Aumann'78]

$N = \{1, 2, \ldots, n\}$ プレイヤーの集合　　$N = \{$🚙, 🚗$\}$

A_i プレイヤー $i \in N$ の行動の集合（有限）　　$A_i = \{$進, 停$\}$

$A = A_1 \times A_2 \times \cdots \times A_n$ 行動の組（action profile）全体の集合

$v_i \colon A \to \mathbb{R}$ プレイヤー $i \in N$ の効用関数　　$v_{\text{🚙}}(\text{進}, \text{停}) = 4$

定義

行動の組の分布 $\pi \in \Delta(A)$ が相関均衡　　$a_{-i} = (a_j)_{j \in N \setminus \{i\}}$ i 以外の行動の組

$\overset{\triangle}{\Leftrightarrow}$ 任意のプレイヤー $i \in N$ と逸脱 $\phi \colon A_i \to A_i$ について

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

（逸脱しても期待利得が改善しない）

※ この定義において π を直積分布に制限すれば、ナッシュ均衡の定義に一致

相関均衡

定義

行動の組の分布 $\pi \in \Delta(A)$ が相関均衡

$\overset{\triangle}{\Leftrightarrow}$ 任意のプレイヤー $i \in N$ と逸脱 $\phi\colon A_i \to A_i$ について

$$\mathop{\mathbb{E}}_{a\sim\pi}\left[v_i(\phi(a_i), a_{-i})\right] \leq \mathop{\mathbb{E}}_{a\sim\pi}\left[v_i(a)\right] \qquad \text{（逸脱しても期待利得が改善しない）}$$

	進	停
進	0, 0	4, 3
停	3, 4	1, 1

前頁の相関均衡の例は以下のような分布 $\pi \in \Delta(A)$

$\pi(\text{進}, \text{停}) = 1/2, \quad \pi(\text{停}, \text{進}) = 1/2$

どのような ϕ を使っても期待利得は非増加

例）$\phi(\text{進}) = \text{停}, \phi(\text{停}) = \text{停}$ を使うと 3.5 から 2 へ減少

相関均衡の LP 定式化

相関均衡全体の集合は $|A| = \prod_{i \in N} |A_i|$ 変数の線形不等式系で表現される

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

プレイヤー数が**定数**の場合、この線形不等式系は多項式サイズ

$\rightarrow$ （最適な）相関均衡を計算する問題は効率的に解ける

cf. ナッシュ均衡全体の集合は非凸（独立性の制約が扱いづらい）

相関均衡のメリット

メリット 1　**仲介者**を用いた自然な解釈（前述）

メリット 2　**学習ダイナミクス**によって効率的に計算可能（後述）

- ゲームの繰り返しでプレイヤーたちがリグレット最小化すると自然に到達
- n 人ゲームの相関均衡を n の多項式時間で計算可能
（cf. ナッシュ均衡の計算は 2 人でも PPAD 困難 [Chen–Deng–Teng'09]）

メリット 3　さまざまなゲームにおける**社会厚生**の保証（後述）

- 都市交通、オークション、資源配分などの応用上重要なゲームで
相関均衡の社会厚生（社会的な望ましさ）は近似的に最適 [Roughgarden'15]

目次

行き先を決めるゲーム（完備情報版）

休日に一緒に出かけたい 👮‍♀️ と 🧑 がそれぞれ行き先を決める

👮‍♀️ は海 🌊 に行きたいが、🧑 は山 🌋 に行きたい

		🧑	
		海 🌊	山 🌋
👮‍♀️	海 🌊	4, 3	1, 1
	山 🌋	0, 0	3, 4

行きたい場所 → 1 点

同じ場所 → それぞれ 3 点ずつ

ベイジアンゲーム （不完備情報ゲーム + 共有事前分布） [Harsanyi'67]

各プレイヤーのタイプ（好み）が共有された事前分布に従って決まるゲーム

👮‍♀️ と 🧑 はそれぞれ確率 1/2 で海 🌊 派、確率 1/2 で山 🌋 派

（互いに相手のタイプは知らないが分布は知っている）

確率 1/4		🧑 🌊派	
		海 🌊	山 🌋
👮‍♀️ 🌊派	海 🌊	4 4	0 1
	山 🌋	1 0	3 3

確率 1/4		🧑 🌋派	
		海 🌊	山 🌋
👮‍♀️ 🌊派	海 🌊	3 4	1 1
	山 🌋	0 0	4 3

ベイジアンゲーム

$N = \{1, 2, \ldots, n\}$ プレイヤーの集合 $N = \{$👮, 👦$\}$

A_i プレイヤー $i \in N$ の行動の集合（有限） $A_1 = A_2 = \{$🌊, 🌋$\}$

Θ_i プレイヤー $i \in N$ のタイプの集合（有限） $\Theta_1 = \Theta_2 = \{$🌊派, 🌋派$\}$

$A = \prod_{i \in N} A_i$ 行動の組全体の集合, $\Theta = \prod_{i \in N} \Theta_i$ タイプの組全体の集合

$\rho \in \Delta(\Theta)$ タイプの組の事前分布（共有知識） $\rho($🌊派, 🌊派$) = 1/4$

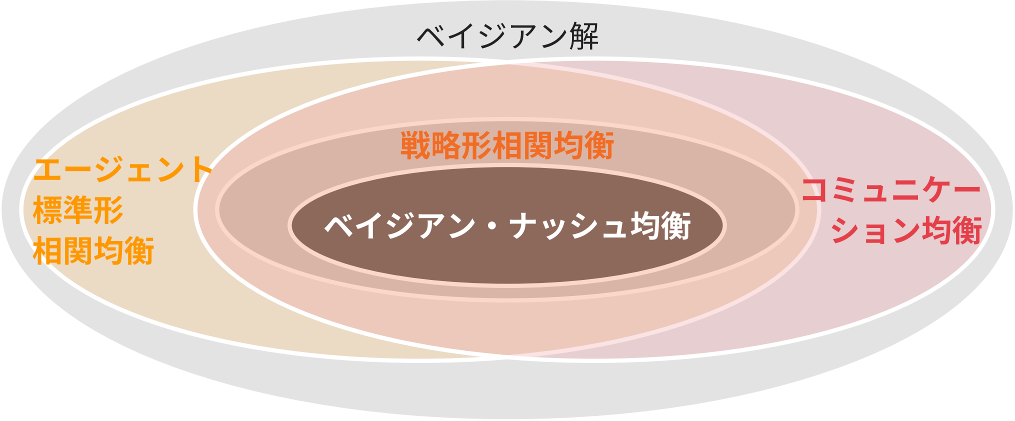
$v_i \colon \Theta \times A \to \mathbb{R}$ プレイヤー $i \in N$ の効用関数 $v_1($🌊派, 🌊派; 🌊, 🌋$) = 1$

1. タイプの組 $\theta \sim \rho$ が生成され、各 $i \in N$ が自分のタイプ θ_i を知る
2. 各 $i \in N$ は他人のタイプ θ_{-i} を知らないまま行動 $a_i \in A_i$ を選択

さまざまなベイズ相関均衡 [Forges'93]

ベイズ相関均衡（ベイジアンゲームにおける相関均衡）には

複数の自然な定義があり、それぞれ異なる情報交換の仕組みをもつ



戦略形相関均衡

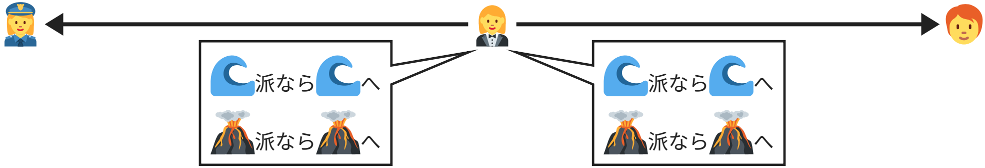
ベイジアンゲームの戦略形

タイプから行動への写像 $s_i\colon \Theta_i \to A_i$（**戦略**と呼ぶ）を一つの行動と解釈

各プレイヤーは戦略全体 $S_i \coloneqq A_i^{\Theta_i}$ から戦略を一つ選択

仲介者🤵‍♀️は各プレイヤーに private channel でタイプごとの行動を推薦

$\leftarrow$ **推薦から逸脱するインセンティブがない**



エージェント標準形相関均衡

ベイジアンゲームのエージェント標準形

各プレイヤーをタイプごとに異なる仮想的なプレイヤーとみなす

実現したタイプに対応する仮想的プレイヤーだけがゲームに参加する

例）仮想的プレイヤー４人 (👮‍♀️, 🌊), (👮‍♀️, 🌋), (🧑‍🦰, 🌊), (🧑‍🦰, 🌋) のうち２人がゲームに参加

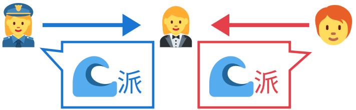
戦略形相関均衡との違い：

各プレイヤーは自分のタイプ以外への推薦は観測しない

※ 仲介者🤵‍♀️によって実現する自然なシナリオは知られていない

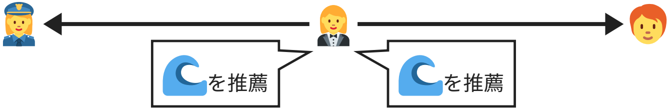
ベイジアン解 [Forges’93]

1. 仲介者 は各プレイヤーの**真の**タイプを観測



2. 仲介者 は各プレイヤーに private channel で行動を推薦

← **推薦から逸脱するインセンティブがない**



エージェント標準形相関均衡との違い：

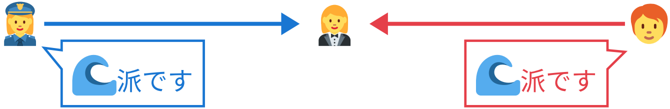
仲介者 はタイプの組に基づいて推薦を決められる（戦略表現性）

コミュニケーション均衡 [Myerson’82, Forges’86]

仲介者とプレイヤーたちの双方向の情報交換で実現

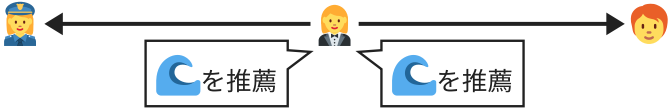
1. 各プレイヤーは仲介者に private channel でタイプを報告

← **嘘のタイプを報告するインセンティブがない**



2. 仲介者は private channel で各プレイヤーに行動を推薦

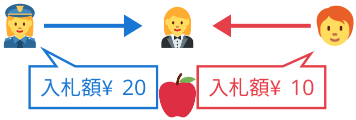
← **推薦から逸脱するインセンティブがない**



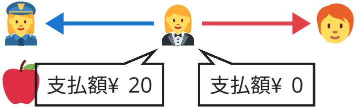
コミュニケーション均衡は二つの概念の融合

メカニズムデザイン

1. 各プレイヤーがタイプを報告
← **嘘をついても得しない**



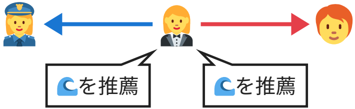
2. が結果を決定
← この決定は逸脱不可



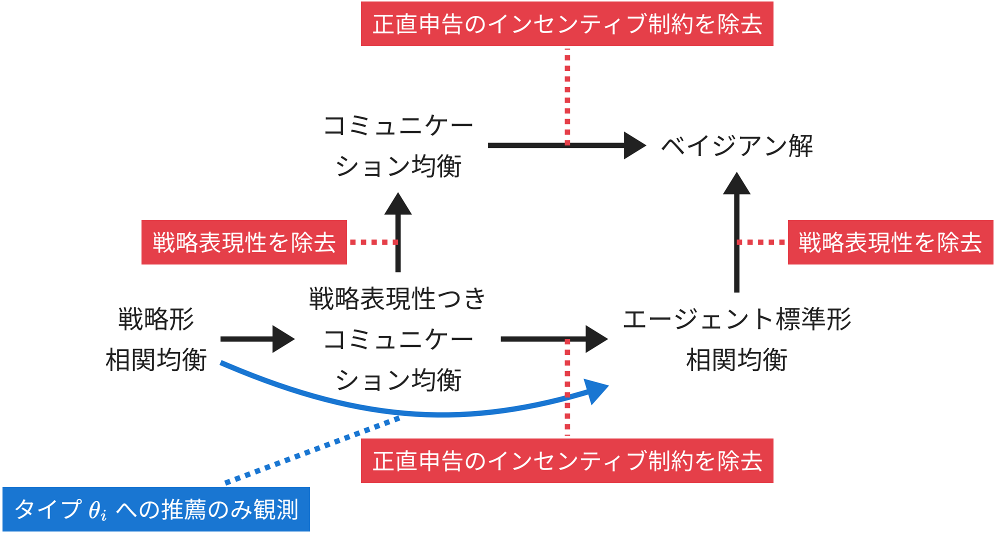
相関均衡

1. タイプなし（完備情報）

2. が行動を推薦
← **逸脱しても得しない**



ベイジアンゲームの相関均衡概念の関係



目次

目次

ϵ 近似相関均衡

$N = \{1, 2, \ldots, n\}$ プレイヤーの集合 $\qquad N = \{$🚙, 🚗$\}$

A_i プレイヤー $i \in N$ の行動の集合（有限） $\qquad A_i = \{$ 進, 停 $\}$

$A = A_1 \times A_2 \times \cdots \times A_n$ 行動の組（action profile）全体の集合

$v_i \colon A \to [0, 1]$ プレイヤー $i \in N$ の効用関数 $\qquad v_{🚙}($進, 停$) = 4$

定義

行動の組の分布 $\pi \in \Delta(A)$ が ϵ 近似相関均衡

$\overset{\triangle}{\Leftrightarrow}$ どのプレイヤー $i \in N$ と逸脱 $\phi \colon A_i \to A_i$ についても

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

（逸脱しても期待利得が高々 ϵ しか改善しない）

相関均衡の計算

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

n 人ゲームで ϵ 近似相関均衡を計算する効率的なアルゴリズム存在

※ 利得関数 v_i の値オラクルを仮定（v_i の表現には n の指数サイズ必要）

※ ϵ 近似相関均衡が $n, \max_{i \in N} |A_i|, 1/\epsilon$ の多項式時間で確率的に求まる

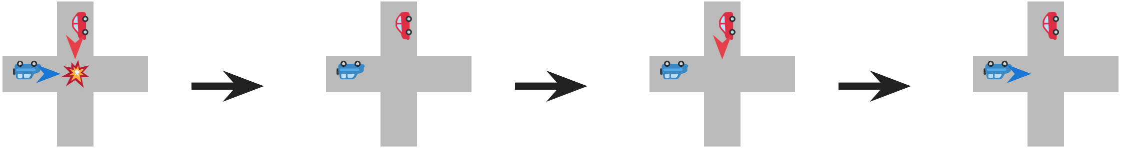
cf. ナッシュ均衡の計算は n に関する指数回オラクル呼び出しが必要 [Babichenko'16]

"どれでもいいから一つ相関均衡を計算" は

"どれでもいいから一つナッシュ均衡を計算" より簡単

相関均衡へと収束するリグレット最小化ダイナミクスをシミュレート

同じゲームの繰り返しのなかでプレイヤーたちが戦略を更新（学習基準は次頁）



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

各プレイヤー $i \in N$ が混合（確率的）戦略 $\pi_i^t \in \Delta(A_i)$ を決める

i の報酬ベクトル $u_i^t(\cdot) \triangleq \mathbb{E}[v_i(\cdot, a_{-i}^t)]$ が決まる（$j \in N$ ごと独立に $a_j^t \sim \pi_j^t$）

各プレイヤー i は期待報酬 $\mathbb{E}_{a_i^t \sim \pi_i^t}[u_i^t(a_i^t)]$ を受けとる

スワップリグレット [Blum–Mansour’07]

もし**進**の代わりに**停**を、**停**の代わりに**進**を選んでいたら……

t	1	2	3	4	5	6
🚙（現実）	停	停	進	停	停	進

🚗	停	停	進	進	停	進
報酬	1	1	0	3	1	0

スワップリグレットは**最適な置換による後悔の量**（この例では **12**）

スワップリグレット [Blum–Mansour'07]

もし**進**の代わりに**停**を、**停**の代わりに**進**を選んでいたら……

t	1	2	3	4	5	6
🚙（現実）	**停**	**停**	**進**	**停**	**停**	**進**
🚙（理想）	**進**	**進**	**停**	**進**	**進**	**停**
🚗	**停**	**停**	**進**	**進**	**停**	**進**
報酬	1 → **4**	1 → **4**	0 → **3**	3 → **0**	1 → **4**	0 → **3**

スワップリグレットは**最適な置換による後悔の量**（この例では **12**）

スワップリグレット [Blum–Mansour'07]

$$R^T_{\mathrm{swap},i} \triangleq \max_{\phi\colon A_i \to A_i} \underbrace{\sum_{t=1}^T \mathbb{E}_{a_i^t \sim \pi_i^t}\left[u_i^t(\phi(a_i^t))\right]}_{\phi\text{ で行動を置換したときの } t \text{ ラウンド目の期待報酬}} - \underbrace{\sum_{t=1}^T \mathbb{E}_{a_i^t \sim \pi_i^t}\left[u_i^t(a_i^t)\right]}_{t\text{ ラウンド目の期待報酬}}$$

cf. 普通のリグレット
(external regret)

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

定理 [Blum–Mansour'07]

T ラウンドの平均（経験分布）$\frac{1}{T} \bigotimes_{i \in N} \pi_i^t$ は $\left(\max_{i \in N} R^T_{\mathrm{swap},i}/T\right)$ 近似相関均衡

スワップリグレット最小化 [Blum–Mansour'07]

ステップ 1 置換 $\phi\colon A_i \to A_i$ を確率行列を使って簡潔に表現

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

左確率行列全体 $\mathcal{Q} = \left\{ Q \in [0,1]^{A_i \times A_i} \,\middle|\, \mathbf{1}Q = \mathbf{1} \right\}$ を使って表現

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

スワップリグレット最小化 [Blum–Mansour'07]

ステップ 2 左確率行列の定常分布（固有ベクトル）を利用して帰着

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

π_i^t の選択問題を Q^t の選択問題に帰着

各 $t \in [T]$ で $Q^t \pi_i^t = \pi_i^t$ が成立するように Q^t から π_i^t を決める

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

スワップリグレット最小化 [Blum–Mansour'07]

ステップ 3 $|A_i|$ 個のリグレット（external regret）最小化へと分解

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

Q^t の選択問題を列ごとに確率ベクトル $q^t_{a_i}$ の選択問題に帰着

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

リグレット最小化アルゴリズムを使えば $R^T_{\mathrm{swap},i} = O\left(\sqrt{T|A_i| \log |A_i|}\right)$

目次

ϵ 近似コミュニケーション均衡

$N = \{1, 2, \ldots, n\}$ プレイヤーの集合　　$N = \{$👮, 🧑$\}$

A_i 行動の集合, Θ_i タイプの集合　　$A_1 = A_2 = \{$🌊, 🌋$\}$, $\Theta_1 = \Theta_2 = \{$🌊派, 🌋派$\}$

$A = \prod_{i \in N} A_i$ 行動の組全体の集合, $\Theta = \prod_{i \in N} \Theta_i$ タイプの組全体の集合

$\rho \in \Delta(\Theta)$ タイプの組の事前分布（共有知識）　　$\rho($🌊派, 🌊派$) = 1/4$

$v_i \colon \Theta \times A \to [0, 1]$ プレイヤー $i \in N$ の効用関数　　$v_1($🌊派, 🌊派; 🌊, 🌋$) = 1$

定義

タイプの組ごとの分布 $\pi \in \Delta(A)^\Theta$ は ϵ 近似コミュニケーション均衡

$\overset{\triangle}{\Leftrightarrow}$ 任意の $i \in N$, $\psi \colon \Theta_i \to \Theta_i$, $\phi \colon \Theta_i \times A_i \to A_i$ に対して

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

ϵ 近似コミュニケーション均衡

$N = \{1, 2, \dots, n\}$ プレイヤーの集合 $\qquad N = \{$👮, 🧑$\}$

A_i 行動の集合, Θ_i タイプの集合 $\qquad A_1 = A_2 = \{$🌊, 🌋$\}$, $\Theta_1 = \Theta_2 = \{$🌊派, 🌋派$\}$

$A = \prod_{i \in N} A_i$ 行動の組全体の集合, $\Theta = \prod_{i \in N} \Theta_i$ タイプの組全体の集合

$\rho \in \Delta(\Theta)$ タイプの組の事前分布（共有知識） $\qquad \rho($🌊派, 🌊派$) = 1/4$

$v_i \colon \Theta \times A \to [0, 1]$ プレイヤー $i \in N$ の効用関数 $\qquad v_1($🌊派, 🌊派; 🌊, 🌋$) = 1$

定義

タイプの組ごとの … コ …

真のタイプ θ_i の代わりに $\psi(\theta_i)$ を報告

推薦された a_i の代わりに $\phi(\theta_i, a_i)$ を選択

$\stackrel{\triangle}{\Leftrightarrow}$ 任意の $i \in N$, $\psi \colon \Theta_i \to \Theta_i$, $\phi \colon \Theta_i \times A_i \to A_i$ に対して

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

コミュニケーション均衡の計算

定理 [Fujii'25a]

n 人ベイジアンゲームにおいて ϵ 近似コミュニケーション均衡を計算する効率的なアルゴリズムが存在する

※ 利得関数 v_i の関数値オラクルと ρ からのサンプリングオラクルを仮定

※ n, 各プレイヤーの行動数 $\max_{i \in N} |A_i|$, タイプ数 $\max_{i \in N} |\Theta_i|$, $1/\epsilon$ に関する多項式時間

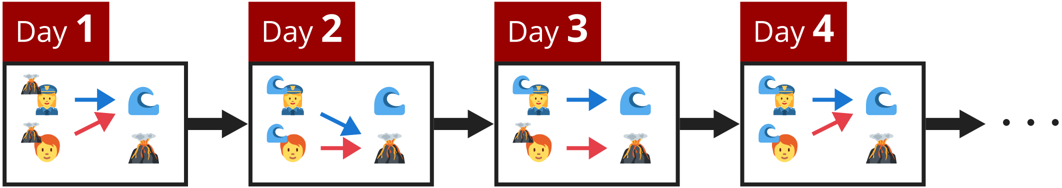
- **嘘つきスワップリグレット**を定義
- 効率的なアルゴリズムによる嘘つきスワップリグレットの $O\left(\sqrt{T \max\{|A_i| \log |A_i|, \log |\Theta_i|\}}\right)$ 上界と $\Omega\left(\sqrt{T \log |\Theta_i|}\right)$ 下界を証明

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

各プレイヤー $i \in N$ がタイプごとの混合戦略 $\pi_i^t \in \Delta(A_i)^{\Theta_i}$ を決める

i の報酬ベクトル $u_i^t(\theta_i, a_i) \triangleq \mathbb{E}[v_i(\theta, a^t)]$ が決まる（$\theta \sim \rho$, $a_{-i}^t \sim \pi_{-i}^t(\theta)$）

各プレイヤー i は期待報酬 $\underset{\theta_i \sim \rho_i}{\mathbb{E}} \left[\underset{a_i^t \sim \pi_i^t(\theta_i)}{\mathbb{E}} [u_i^t(\theta_i, a_i^t)] \right]$ を受けとる



嘘つきスワップリグレット [Fujii'25a]

タイプ虚偽申告 ψ と行動推薦からの逸脱 ϕ を考慮

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

定理 [Fujii'25a]

経験分布 $\frac{1}{T} \bigotimes_{i \in N} \pi_i^t$ は $\left(\max_{i \in N} R^T_{\mathrm{US},i} / T \right)$ 近似コミュニケーション均衡

ステップ 1 $\psi\colon \Theta_i \to \Theta_i$ と $\phi\colon \Theta_i \times A_i \to A_i$ を一つの行列を使って表現

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

ψ と ϕ の組合せ全体を行列の集合 $\mathcal{Q}$ を使って表現

各 $Q \in \mathcal{Q}$ は右確率行列の各要素に左確率行列を入れたブロック行列

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

嘘つきスワップリグレット最小化 [Fujii'25a]

ステップ 2　行列 Q の固有ベクトルを利用して帰着（Φ-regret）

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

π_i^t の選択問題を Q^t の選択問題に帰着

各 $t \in [T]$ で $Q^t x^t = x^t$ が成立するように Q^t から x^t を決める

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

嘘つきスワップリグレット最小化 [Fujii'25a]

ステップ 3 $|\Theta_i|^2|A_i| + |\Theta_i|$ **個のリグレット最小化へと分解**

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

各 $\theta_i \in \Theta_i$ ごとのリグレット $R^T_{\theta_i}$ と

各 $\theta_i, \theta'_i \in \Theta_i$, $a'_i \in A_i$ ごとのリグレット $R^T_{\theta_i,\theta'_i,a'_i}$ に分解

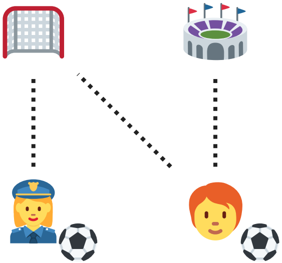
$$R^T_{\mathrm{US},i} \le \sum_{\theta_i\in\Theta_i} R^T_{\theta_i} + \sum_{\theta_i\in\Theta_i} \max_{\theta'_i\in\Theta_i} \sum_{a'_i\in A_i} R^T_{\theta_i,\theta'_i,a'_i}$$

上界 $R^T_{\mathrm{US},i} = O\left(\sqrt{T \max\{|A_i| \log |A_i|, \log |\Theta_i|\}}\right)$ が得られる

目次

整合的効用ゲーム（valid utility game）の例

プレイヤーたちがそれぞれサッカーする場所を選ぶ

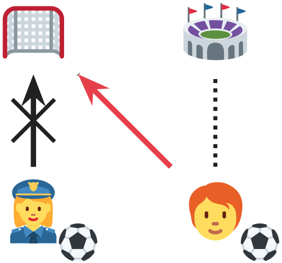


複数のプレイヤーが同じ場所を選んだ場合は
あらかじめ決めたルールに従って分割

例）🧑が👮‍♀️に優先される

整合的効用ゲーム（valid utility game）の例

プレイヤーたちがそれぞれサッカーする場所を選ぶ



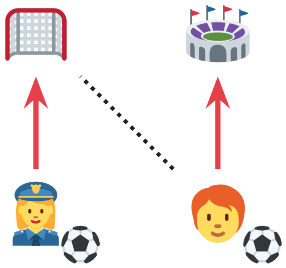
複数のプレイヤーが同じ場所を選んだ場合は
あらかじめ決めたルールに従って分割

例）🧑が👮‍♀️に優先される

最悪の**相関均衡** $= 1$

整合的効用ゲーム（valid utility game）の例

プレイヤーたちがそれぞれサッカーする場所を選ぶ



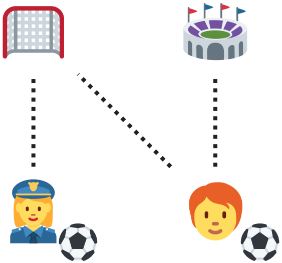
複数のプレイヤーが同じ場所を選んだ場合は
あらかじめ決めたルールに従って分割

例）🧑が👮‍♀️に優先される

最適な社会厚生 $= 2$

整合的効用ゲーム（valid utility game）の例

プレイヤーたちがそれぞれサッカーする場所を選ぶ



複数のプレイヤーが同じ場所を選んだ場合はあらかじめ決めたルールに従って分割

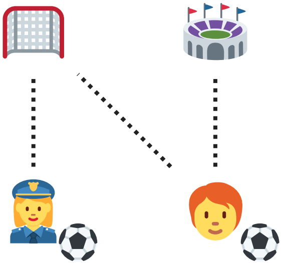
例）🧑が👮‍♀️に優先される

$$\mathrm{PoA} := \frac{\text{最悪の}\textbf{相関均衡}}{\textbf{最適な社会厚生}} = \frac{1}{2}$$

(price of anarchy)

整合的効用ゲーム（valid utility game）の例

プレイヤーたちがそれぞれサッカーする場所を選ぶ



複数のプレイヤーが同じ場所を選んだ場合は
あらかじめ決めたルールに従って分割

例）🧑が👮‍♀️に優先される

$$\mathrm{PoA} := \frac{\text{最悪の}\textbf{相関均衡}}{\textbf{最適な社会厚生}} = \frac{1}{2}$$

(price of anarchy)

定理 [Vetta 2002, Roughgarden’15]

整合的効用ゲームにおける相関均衡の $\mathrm{PoA} \geq 0.5$

整合的効用ゲーム

社会厚生（$\approx$ 全員の利得の和）を集合関数 $f\colon 2^E \to \mathbb{R}$ で表す

（台集合は全プレイヤーの行動全体 $E = \bigcup_{i \in N} A_i$）

整合的効用ゲーム $\overset{\triangle}{\Longleftrightarrow}$ 社会厚生関数が単調劣モジュラ

直感的には、劣モジュラ性はプレイヤーの行動のあいだの**負の外部性**に対応

整合的効用ゲーム

社会厚生（≈ 全員の利得の和）を集合関数 $f\colon 2^E \to \mathbb{R}$ で表す

（台集合は全プレイヤーの行動全体 $E = \bigcup_{i \in N} A_i$）

整合的効用ゲーム $\overset{\triangle}{\Leftrightarrow}$ **社会厚生関数が単調劣モジュラ**

$$\underbrace{f(\{\text{🥅👮‍♀️}\}) - f(\{\})}_{\text{まだ誰も参加していないときの社会厚生の増分}}$$

直感的には、劣モジュラ性はプレイヤーの行動のあいだの**負の外部性**に対応

整合的効用ゲーム

社会厚生（≈ 全員の利得の和）を集合関数 $f\colon 2^E \to \mathbb{R}$ で表す

（台集合は全プレイヤーの行動全体 $E = \bigcup_{i \in N} A_i$）

整合的効用ゲーム $\overset{\triangle}{\Longleftrightarrow}$ 社会厚生関数が単調劣モジュラ

$\underbrace{f(\{\text{🥅}_{\text{👮‍♀️}}\}) - f(\{\})}_{}$

まだ誰も参加していないときの
社会厚生の増分

$\underbrace{f(\{\text{🥅}_{\text{👮‍♀️}}, \text{🥅}_{\text{🧑}}\}) - f(\{\text{🥅}_{\text{🧑}}\})}_{}$

他の人が既に参加しているときの
社会厚生の増分

直感的には、劣モジュラ性はプレイヤーの行動のあいだの**負の外部性**に対応

整合的効用ゲーム

社会厚生（≈ 全員の利得の和）を集合関数 $f\colon 2^E \to \mathbb{R}$ で表す

（台集合は全プレイヤーの行動全体 $E = \bigcup_{i \in N} A_i$）

整合的効用ゲーム $\overset{\triangle}{\Leftrightarrow}$ 社会厚生関数が単調劣モジュラ

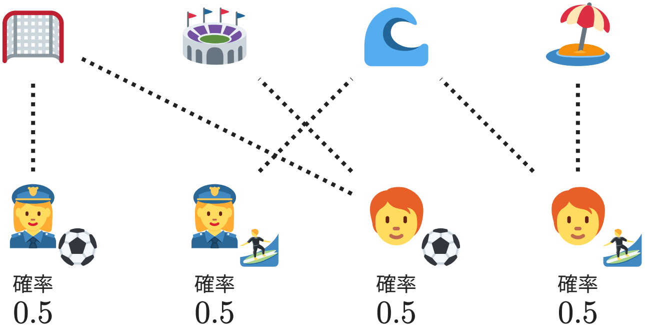
$f(\{$🥅👮‍♀️$\}) - f(\{\})$ $\geq$ $f(\{$🥅👮‍♀️, 🥅🧒$\}) - f(\{$🥅🧒$\})$

まだ誰も参加していないときの社会厚生の増分 ≥ 他の人が既に参加しているときの社会厚生の増分

直感的には、劣モジュラ性はプレイヤーの行動のあいだの**負の外部性**に対応

ベイジアン整合的効用ゲームの例

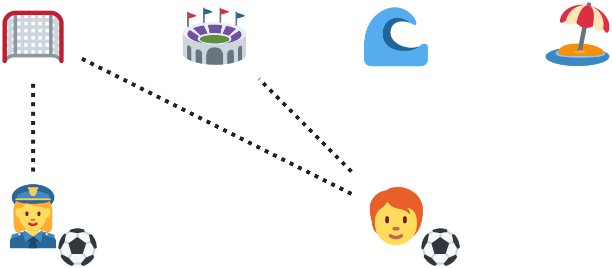
タイプごとに別のプレイヤーとみなして社会厚生関数を定義



Q ベイズ相関均衡の各概念のあいだで PoA の保証に差はあるか？

ベイジアン整合的効用ゲームの例

タイプごとに別のプレイヤーとみなして社会厚生関数を定義



確率 $1/4$

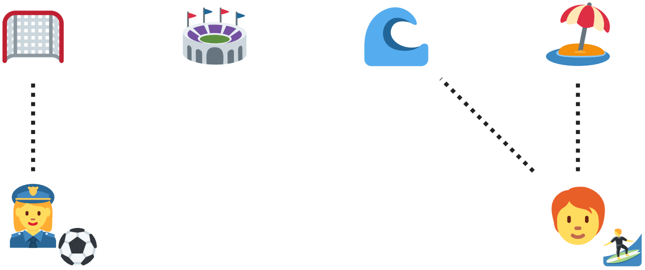
👮‍♀️は🧑のタイプを知らない

🧑は👮‍♀️のタイプを知らない

Q　ベイズ相関均衡の各概念のあいだで PoA の保証に差はあるか？

ベイジアン整合的効用ゲームの例

タイプごとに別のプレイヤーとみなして社会厚生関数を定義



確率 $1/4$

👮‍♀️は🧑のタイプを知らない

🧑は👮‍♀️のタイプを知らない

Q　ベイズ相関均衡の各概念のあいだで PoA の保証に差はあるか？

ベイジアンゲームにおける PoA と PoS の定義

均衡の集合 $\Pi \subseteq \Delta(A)^\Theta$ に対して PoA と Price of Stability (PoS) は

$$\mathrm{PoA}_\Pi \triangleq \frac{\underbrace{\inf_{\pi\in\Pi} \mathop{\mathbb{E}}_{\theta\sim\rho} \left[\mathop{\mathbb{E}}_{a\sim\pi(\theta)} [f(a)] \right]}_{\text{最悪の均衡による社会厚生の値}}}{\underbrace{\mathop{\mathbb{E}}_{\theta\sim\rho} \left[\max_{a^*\in A^\theta} f(a^*) \right]}_{\text{社会厚生の最大値}}}$$

記法の乱用により

$f(a) \triangleq f(\{a_1, \ldots, a_n\})$ と表記

完備情報との違い 最適な行動 a_i^* が他プレイヤーのタイプ θ_{-i} に依存

ベイジアンゲームにおける PoA と PoS の定義

均衡の集合 $\Pi \subseteq \Delta(A)^\Theta$ に対して PoA と Price of Stability (PoS) は

$$\mathrm{PoA}_\Pi \triangleq \frac{\overbrace{\inf_{\pi\in\Pi} \mathbb{E}_{\theta\sim\rho}\left[\overbrace{\mathbb{E}_{a\sim\pi(\theta)}[f(a)]}^{\text{社会厚生の値}}\right]}^{\text{最悪の均衡による}}}{\mathbb{E}_{\theta\sim\rho}\left[\underbrace{\max_{a^*\in A^\theta} f(a^*)}_{\text{社会厚生の最大値}}\right]}$$

$$\mathrm{PoS}_\Pi \triangleq \frac{\overbrace{\sup_{\pi\in\Pi} \mathbb{E}_{\theta\sim\rho}\left[\overbrace{\mathbb{E}_{a\sim\pi(\theta)}[f(a)]}^{\text{社会厚生の値}}\right]}^{\text{最良の均衡による}}}{\mathbb{E}_{\theta\sim\rho}\left[\underbrace{\max_{a^*\in A^\theta} f(a^*)}_{\text{社会厚生の最大値}}\right]}$$

完備情報との違い 最適な行動 a_i^* が他プレイヤーのタイプ θ_{-i} に依存

戦略表現性ギャップ (SR gap) [Fujii'25b]

最悪の均衡による

社会厚生の値

$$\mathrm{PoA}_{\Pi} \triangleq \frac{\inf_{\pi \in \Pi} \mathbb{E}_{\theta \sim \rho}\left[\mathbb{E}_{a \sim \pi(\theta)}[f(a)]\right]}{\mathbb{E}_{\theta \sim \rho}\left[\max_{a^* \in A^\theta} f(a^*)\right]}$$

社会厚生の最大値

戦略表現性ギャップ (SR gap) [Fujii'25b]

最悪の均衡による

社会厚生の値

$$\mathrm{PoA}_\Pi \triangleq \frac{\inf_{\pi\in\Pi}\mathbb{E}_{\theta\sim\rho}\left[\mathbb{E}_{a\sim\pi(\theta)}[f(a)]\right]}{\mathbb{E}_{\theta\sim\rho}\left[\max_{a^*\in A^\theta} f(a^*)\right]}$$

社会厚生の最大値

最適な戦略の組による

社会厚生の値

$$= \underbrace{\frac{\inf_{\pi\in\Pi}\mathbb{E}_{\theta\sim\rho}\left[\mathbb{E}_{a\sim\pi(\theta)}[f(a)]\right]}{\max_{s^*\in S}\mathbb{E}_{\theta\sim\rho}[f(s^*(\theta))]}}_{\text{完備情報の場合に帰着}} \cdot \underbrace{\frac{\max_{s^*\in S}\mathbb{E}_{\theta\sim\rho}[f(s^*(\theta))]}{\mathbb{E}_{\theta\sim\rho}\left[\max_{a\in A^\theta} f(a)\right]}}_{\textbf{SR gap}}$$

戦略表現性ギャップ (SR gap) [Fujii'25b]

$S_i = \prod_{\theta_i \in \Theta_i} A_i^{\theta_i}$ 各 $i \in N$ の戦略全体の集合

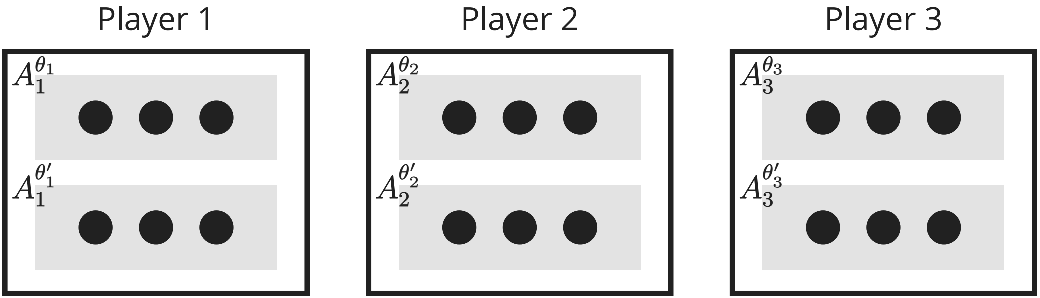
$s_i \in S_i$ はタイプ θ_i ごとに行動 $s_i(\theta_i)$ を決定

$S \triangleq \prod_{i \in N} S_i, \quad s(\theta) \triangleq (s_1(\theta_1), \ldots, s_n(\theta_n))$

$$\mathrm{PoA}_\Pi \triangleq \frac{\overbrace{\inf_{\pi \in \Pi} \mathop{\mathbb{E}}_{\theta \sim \rho} \left[\mathop{\mathbb{E}}_{a \sim \pi(\theta)} [f(a)] \right]}^{\text{最悪の均衡による社会厚生の値}}}{\underbrace{\mathop{\mathbb{E}}_{\theta \sim \rho} \left[\max_{a^* \in A^\theta} f(a^*) \right]}_{\text{社会厚生の最大値}}}$$

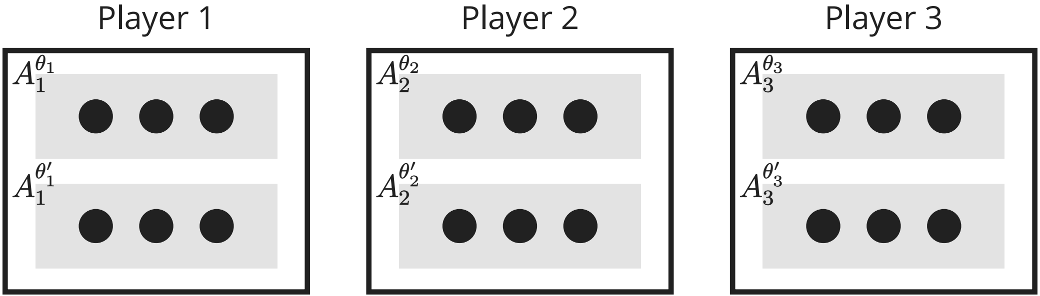
$$= \underbrace{\frac{\inf_{\pi \in \Pi} \mathop{\mathbb{E}}_{\theta \sim \rho} \left[\mathop{\mathbb{E}}_{a \sim \pi(\theta)} [f(a)] \right]}{\max_{s^* \in S} \mathop{\mathbb{E}}_{\theta \sim \rho} [f(s^*(\theta))]}}_{\text{完備情報の場合に帰着}} \cdot \underbrace{\frac{\overbrace{\max_{s^* \in S} \mathop{\mathbb{E}}_{\theta \sim \rho} [f(s^*(\theta))]}^{\text{最適な戦略の組による社会厚生の値}}}{\mathop{\mathbb{E}}_{\theta \sim \rho} \left[\max_{a \in A^\theta} f(a) \right]}}_{\text{SR gap}}$$

戦略表現性ギャップ (SR gap) [Fujii'25b]



プレイヤーごとに一つのブロックが事前分布に従って決定

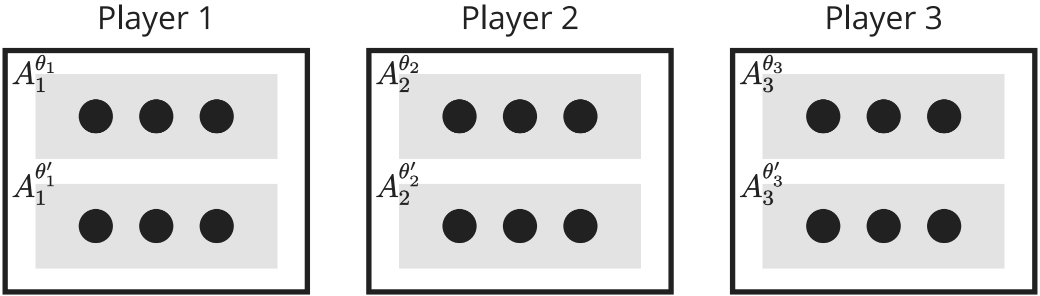
戦略表現性ギャップ (SR gap) [Fujii'25b]



プレイヤーごとに一つのブロックが事前分布に従って決定

$$\mathrm{SRgap}_f \triangleq \frac{\text{各ブロックから要素を一つ選んでからブロックがランダムに選ばれる}}{\text{ブロックがランダムに選ばれてから各ブロックから要素を一つ選ぶ}}$$

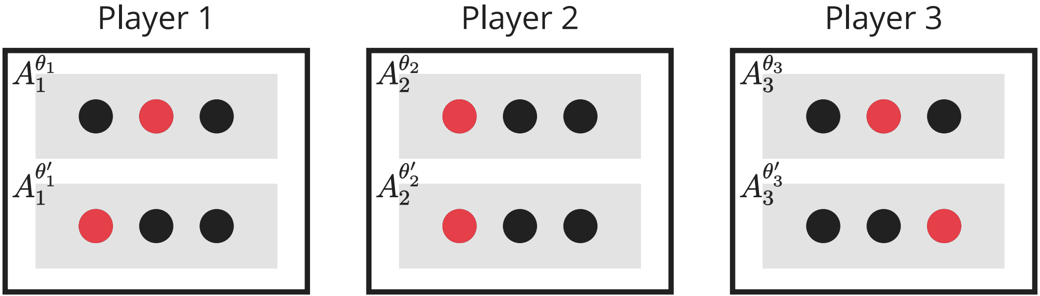
戦略表現性ギャップ (SR gap) [Fujii'25b]



プレイヤーごとに一つのブロックが事前分布に従って決定

$$\mathrm{SRgap}_f \triangleq \frac{\text{各ブロックから要素を一つ選んでからブロックがランダムに選ばれる}}{\text{ブロックがランダムに選ばれてから各ブロックから要素を一つ選ぶ}}$$

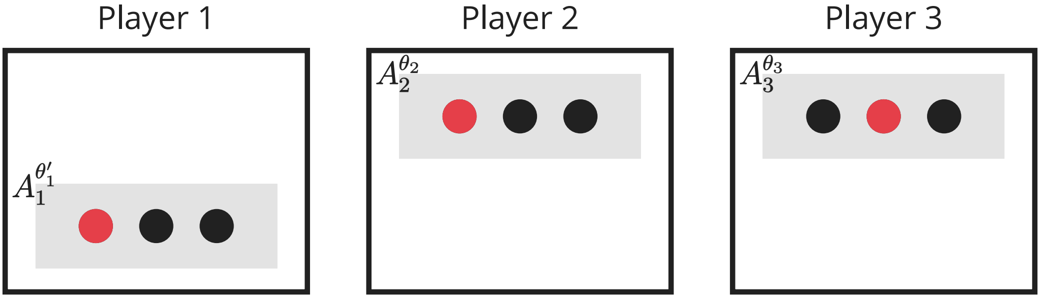
戦略表現性ギャップ (SR gap) [Fujii'25b]



プレイヤーごとに一つのブロックが事前分布に従って決定

$$\mathrm{SRgap}_f \triangleq \frac{\text{各ブロックから要素を一つ選んでからブロックがランダムに選ばれる}}{\text{ブロックがランダムに選ばれてから各ブロックから要素を一つ選ぶ}}$$

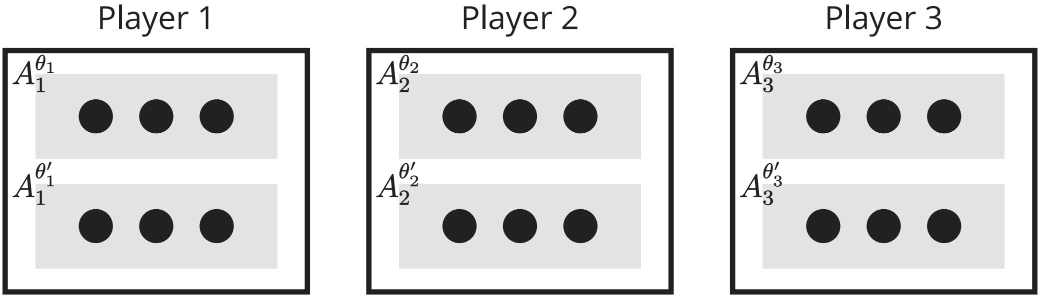
戦略表現性ギャップ (SR gap) [Fujii'25b]



プレイヤーごとに一つのブロックが事前分布に従って決定

$$\mathrm{SRgap}_f \triangleq \frac{\text{各ブロックから要素を一つ選んでからブロックがランダムに選ばれる}}{\text{ブロックがランダムに選ばれてから各ブロックから要素を一つ選ぶ}}$$

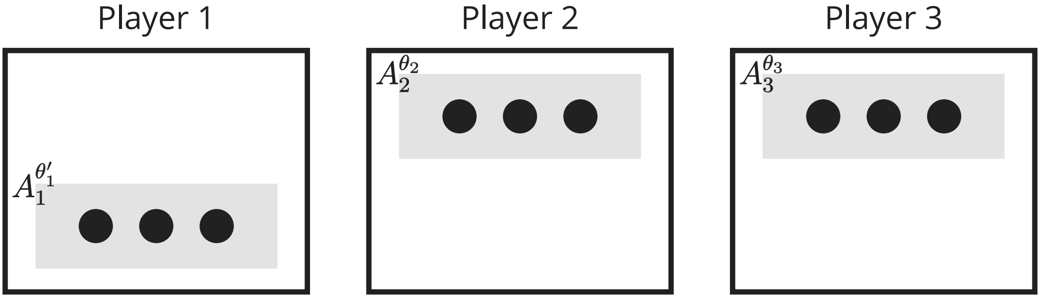
戦略表現性ギャップ (SR gap) [Fujii'25b]



プレイヤーごとに一つのブロックが事前分布に従って決定

$$\mathrm{SRgap}_f \triangleq \frac{\text{各ブロックから要素を一つ選んでからブロックがランダムに選ばれる}}{\text{ブロックがランダムに選ばれてから各ブロックから要素を一つ選ぶ}}$$

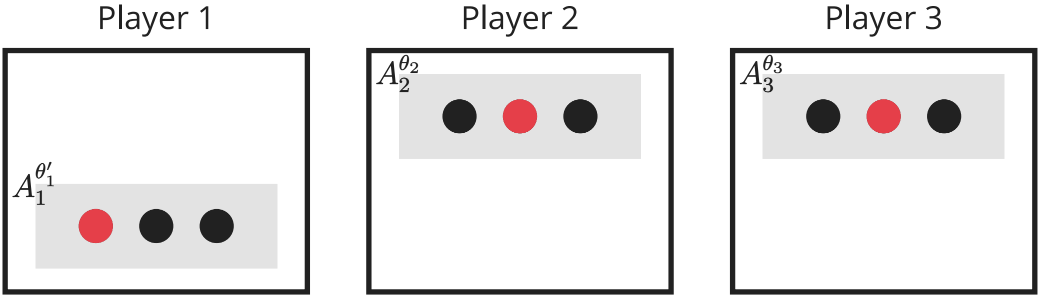
戦略表現性ギャップ (SR gap) [Fujii'25b]



プレイヤーごとに一つのブロックが事前分布に従って決定

$$\mathrm{SRgap}_f \triangleq \frac{\text{各ブロックから要素を一つ選んでからブロックがランダムに選ばれる}}{\text{ブロックがランダムに選ばれてから各ブロックから要素を一つ選ぶ}}$$

戦略表現性ギャップ (SR gap) [Fujii'25b]



プレイヤーごとに一つのブロックが事前分布に従って決定

$$\mathrm{SRgap}_f \triangleq \frac{\text{各ブロックから要素を一つ選んでからブロックがランダムに選ばれる}}{\text{ブロックがランダムに選ばれてから各ブロックから要素を一つ選ぶ}}$$

戦略表現性ギャップの上下界 [Fujii'25b]

Q 非負単調劣モジュラ関数における SR gap の最小値は？

$$\mathrm{SRgap}_f = \frac{\max_{s^* \in S} \mathbb{E}_{\theta \sim \rho}[f(s^*(\theta))]}{\mathbb{E}_{\theta \sim \rho}\left[\max_{a \in A^\theta} f(a)\right]}$$

- 事前分布 ρ がプレイヤーごとに**独立**な場合

$\min_f \mathrm{SRgap}_f = 1 - 1/\mathrm{e}$ ← 相関ギャップ + 弱負回帰 [Qiu–Singla'22]

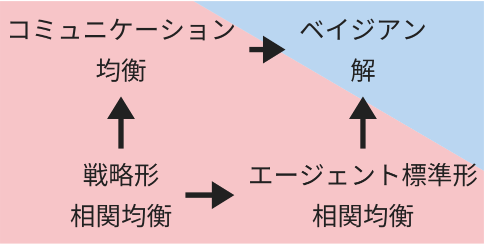
- 事前分布 ρ は**相関してもよい**場合（任意の分布）

$\min_f \mathrm{SRgap}_f = \Theta(1/\sqrt{n})$ ← 多重線形拡張 + 要素を大小に分けて解析

本研究の貢献 [Fujii'25b]

PoA の上下界

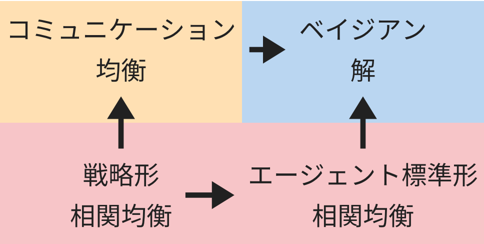
$\mathrm{PoA} \in [0.316, 0.441]$



$\mathrm{PoA} = 0.5$

PoS の上下界

$\mathrm{PoS} \in [1 - 1/e, 0.8]$ $\mathrm{PoS} = 1$



$\mathrm{PoS} = 1 - 1/e$

PoS 結果はより強い条件（basic utility）下

※ 上記の結果はタイプ事前分布 ρ のプレイヤーごとの独立性を仮定

本発表の概要

1 **ベイジアンゲームにおけるさまざまな相関均衡の定義**

信頼できる第三者（仲介者）を介した情報交換によって相関均衡が実現

ベイジアンゲームでは相関均衡の自然な（非等価な）定義が複数知られている

2 **学習ダイナミクスによる相関均衡の計算**

ゲームの繰り返しの中でプレイヤーたちがリグレット最小化すると均衡に収束

このダイナミクスをシミュレートすることで相関均衡を効率的に計算

3 **整合的効用（劣モジュラ最大化）ゲームにおける社会厚生の保証**

相関均衡において社会厚生（社会的な望ましさ）が近似的に最適

参考文献

- Robert J. Aumann. 1974. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1(1), 67–96.
- 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.
- 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.
- Kaito Fujii. 2025a. Bayes correlated equilibria, no-regret dynamics in Bayesian games, and the price of anarchy. Proceedings of Machine Learning Research vol, 291, 1-2.
- Kaito Fujii. 2025b. The Power of Mediators: Price of Anarchy and Stability in Bayesian Games with Submodular Social Welfare. In Proceedings of the 26th ACM Conference on Economics and Computation, pp. 251-251.
- 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.
- Roger B. Myerson. 1982. Optimal coordination mechanisms in generalized principal – agent problems. *Journal of Mathematical Economics*, 10(1), 67–81.
- Tim Roughgarden. 2015. Intrinsic Robustness of the Price of Anarchy. *Journal of the ACM* 62(5), 32:1–32:42.

目次

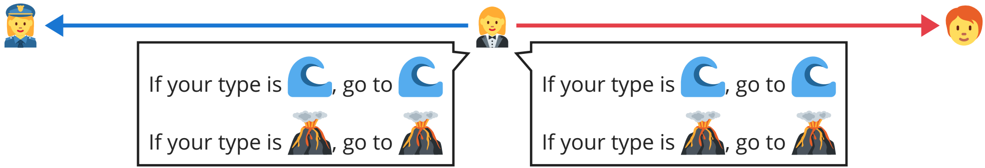
Strategic-form correlated equilibria

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



Strategic-form correlated equilibria

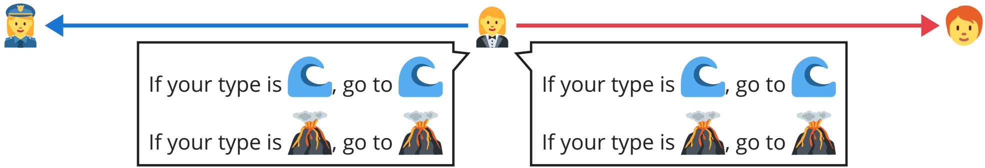
Strategic form of Bayesian games

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**



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 $i \in N$, $\phi_{\mathsf{SF}} : S_i \to S_i$,

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

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

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

SFCE & Strategy swap regret

Definition

A distribution $\sigma \in \Delta(S_1 \times \cdots \times$

Choosing strategy $\phi_{\mathsf{SF}}(s_i)$ instead of recommended s_i

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

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

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

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

目次

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 🤵‍♀️

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 $i \in N$, $\phi\colon \Theta_i \times A_i \to A_i$,

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

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

ANFCE & Type-wise swap regret

Definition

A distribution $\sigma \in \Delta(S_1 \times \cdots \times$

Choosing strategy $\phi(\theta_i, s_i(\theta_i))$ instead of recommended $s_i(\theta_i)$

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

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

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

Example of Bayesian game

🧑‍🦰 (type: 🌊) is trying to choose the same action as 👮‍♀️

🧑‍🦰 (type: 🌋) and 👮‍♀️ do not have any preference

w.p. 1/2

		🧑‍🦰 type: 🌊	
		🌊	🌋
👮‍♀️	🌊	0, 1	0, 0
	🌋	0, 0	0, 1

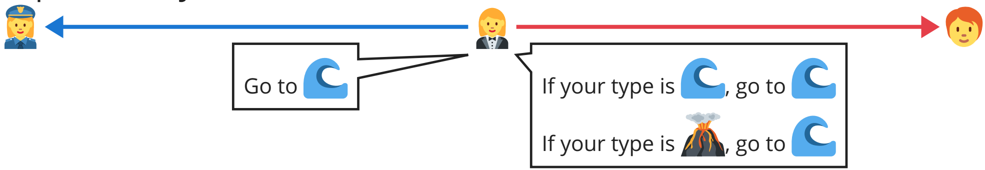
w.p. 1/2

		🧑‍🦰 type: 🌋	
		🌊	🌋
👮‍♀️	🌊	0, 0	0, 0
	🌋	0, 0	0, 0

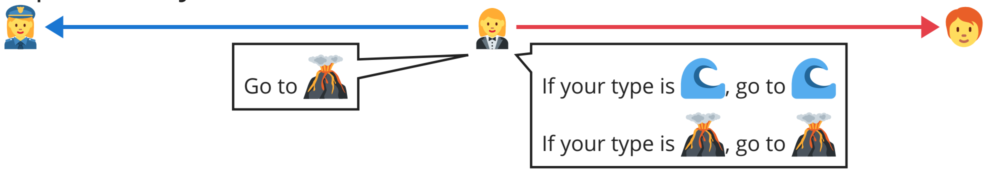
Example of ANFCE but not SFCE

We consider the following strategy-profile distribution

with probability 0.5,



with probability 0.5,



ANFCE but not SFCE

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

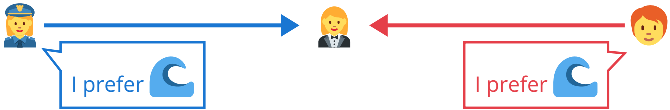
This distribution is not an SFCE because 🧑‍🦰 (type: 🌊) is always recommended 🌊 but can deviate to the same action as 👮‍♀️ by observing the recommendation to 🧑‍🦰 (type: 🌋)

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

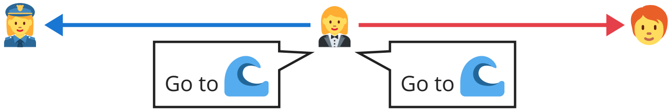
Bayesian solutions [Forges'93]

Mediator knows the true types in advance

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



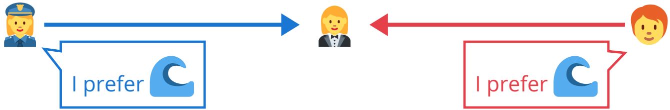
2. The mediator privately sends a recommendation to each player



Bayesian solutions [Forges'93]

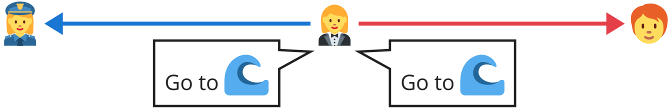
Mediator knows the true types in advance

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**



Bayesian solutions

Definition

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

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

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

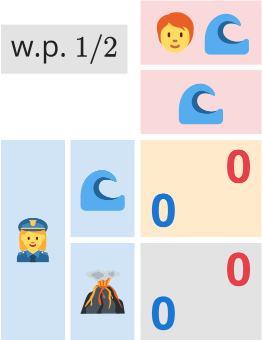
Difference from ANFCE

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

which we call **strategy representability** (e.g., π in the previous page)

Example of non-SR distribution

This distribution cannot be realized by any strategy-profile distribution



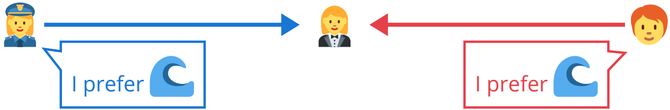
If 👮‍♀️'s type is 🌊, 🤵‍♀️ recommends 🌊

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

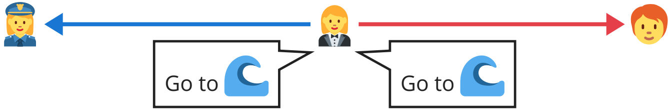
Communication equilibria [Myerson'82, Forges'86]

Equilibria realized by with bidirectional communication

1. Each player privately tells their types to the mediator



2. The mediator privately sends a recommendation to each player

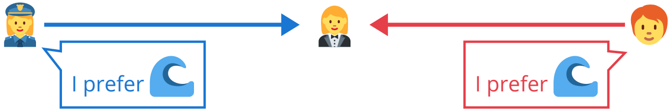


Communication equilibria [Myerson'82, Forges'86]

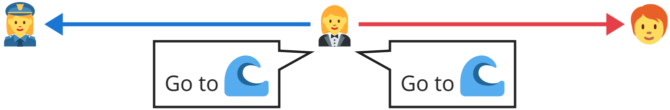
Equilibria realized by with bidirectional communication

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

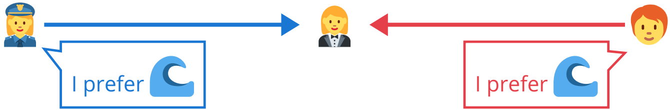


Communication equilibria [Myerson'82, Forges'86]

Equilibria realized by with bidirectional communication

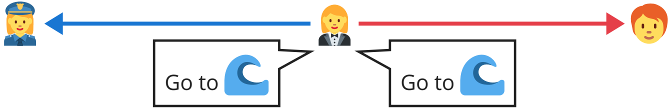
1. Each player privately tells their types to the mediator

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



2. The mediator privately sends a recommendation to each player

   ← **No incentive to disobey the recommendation**



Communication 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 $\phi\colon \Theta_i \times A_i \to A_i$,

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

1. Each player $i \in N$ privately tells θ_i (possibly $\psi(\theta_i)$) to 🤵‍♀️
2. 🤵‍♀️ privately sends recommendations $a \sim \pi(\theta)$ to each player
3. Each player i chooses their action a_i (possibly deviates to $\phi(\theta_i, a_i)$)

Communication equilibria [Myerson'82, Forges'86]

Definition

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

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

A distribution $\pi \in \Delta(A)^\Theta$ i… qu…

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

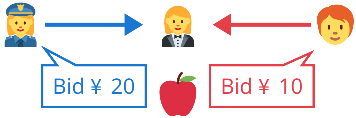
$$\mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{a\sim\pi(\psi(\theta_i),\theta_{-i})}[v_i(\theta;\phi(\theta_i,a_i),a_{-i})]\right] \le \mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{a\sim\pi(\theta)}[v_i(\theta;a)]\right].$$

1. Each player $i \in N$ privately tells θ_i (possibly $\psi(\theta_i)$) to 🤵‍♀️
2. 🤵‍♀️ privately sends recommendations $a \sim \pi(\theta)$ to each player
3. Each player i chooses their action a_i (possibly deviates to $\phi(\theta_i, a_i)$)

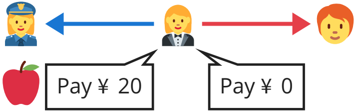
Communication equilibrium combines ...

Mechanism design

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



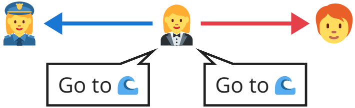
2 👩‍💼 decides the outcome
← This decision is binding



Correlated equilibria

1 No type (complete info.)

2 👩‍💼 recommends actions
← **No incentive to deviate**



目次

No-regret dynamics in Bayesian games

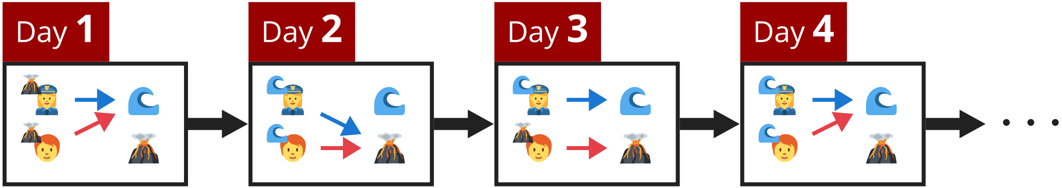
For $t = 1, 2, \dots, T$**:**

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 $i \in N$

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

where $u_i^t(\theta_i, a_i) \triangleq \mathop{\mathbb{E}}_{\theta_{-i} \sim \rho_{-i}|\theta_i} \left[\mathop{\mathbb{E}}_{a_{-i} \sim \pi^t_{-i}(\theta_{-i})} [v_i(\theta; a)] \right]$ 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 $\mathrm{ExternalRegret}^T = O\left(\sqrt{T \log |A|}\right)$

Untruthful swap regret minimization algo.

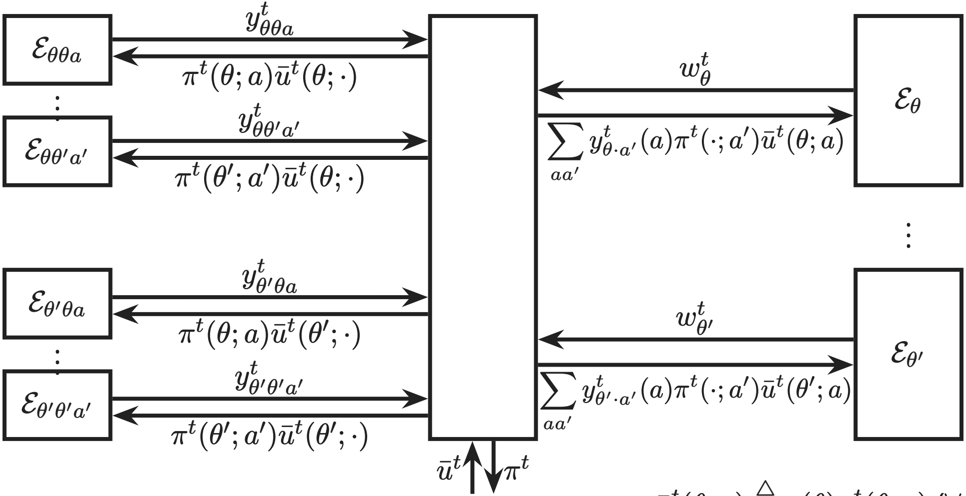
$$R^T_{\mathrm{US},i} = \max_{\substack{\psi:\,\Theta\to\Theta \\ \phi:\,\Theta\times A\to A}} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{a\sim\pi^t(\psi(\theta))}[u^t(\theta,\phi(\theta,a))]\right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\theta\sim\rho}\left[\mathop{\mathbb{E}}_{a\sim\pi^t(\theta)}[u^t(\theta,a)]\right]$$

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

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

※ 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 $a_i' \in A_i$

目次

Price of anarchy (PoA)

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

最悪の均衡によって達成される社会厚生

最良の社会厚生

社会厚生 $v_{\mathrm{SW}} \colon A \to \mathbb{R}_{\geq 0}$ は社会的な望ましさを表す

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

※ **PoA は均衡概念ごとに決まる**（ナッシュ均衡の PoA、相関均衡の PoA など)

	C	D
C	10, 10	0, 15
D	15, 0	1, 1

PoA が抑えられないゲームも存在する

左のゲームでは $\mathrm{PoA} \approx 0$

最悪の均衡での利得和：2 at (D, D)

最良の利得和：20 at (C, C)

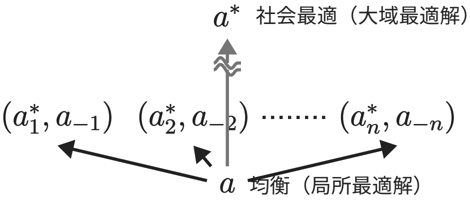
平滑性（smoothness） [Roughgarden’15]

Q どのようなゲームにおいて PoA は抑えられるか？

定義 [Roughgarden’15]

n 人ゲームが (λ, μ) 平滑

$$\overset{\triangle}{\Leftrightarrow} \forall a, a^* \in A: \quad \underbrace{\sum_{i=1}^{n} v_i(a_i^*, a_{-i})}_{i \text{ が } a_i \text{ から } a_i^* \text{ へと逸脱}} \geq \lambda \underbrace{v_{\mathrm{SW}}(a^*)}_{\substack{a^* \text{ によって} \\ \text{達成される社会厚生}}} - \mu \underbrace{v_{\mathrm{SW}}(a)}_{\substack{a \text{ によって} \\ \text{達成される社会厚生}}}$$



逸脱によって社会厚生が
社会最適へと十分近づく

平滑性（smoothness） [Roughgarden’15]

Smooth games are a broad class of games with bounded PoA

Theorem [Roughgarden’15]

(λ, μ) 平滑ゲームにおける相関均衡の PoA は $\dfrac{\lambda}{1+\mu}$ 以上

※ 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, ...

目次

目次

Notations for Bayesian games

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

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

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

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

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

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

where $A = \displaystyle\prod_{i \in N} \left(\bigcup_{\theta_i \in \Theta_i} A_i^{\theta_i} \right)$ 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

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

Submodular social welfare functions

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

$$f(\{\text{🥅👮‍♀️}\}) - f(\{\})$$

The increase in social welfare
when no one attended yet

Submodular social welfare functions

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

$$f(\{🥅_{👮‍♀️}\}) - f(\{\})$$

The increase in social welfare when no one attended yet

$$f(\{🥅_{👮‍♀️}, 🥅_{🧑‍🦰}\}) - f(\{🥅_{🧑‍🦰}\})$$

The increase in social welfare when other players already attended

Submodular social welfare functions

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

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

Submodular social welfare functions

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

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

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

Bayesian valid utility games

$v_i : A \to \mathbb{R}_{\geq 0}$ 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

Assumption [Vetta'02]

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

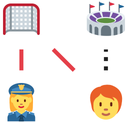
Bayesian valid utility games

$v_i : A \to \mathbb{R}_{\geq 0}$ 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

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

目次

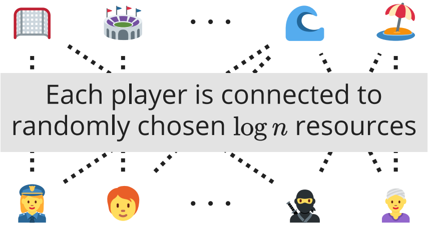
SR gap lower bound (independent case)

Theorem

If ρ is independent, $\mathrm{SRgap} \geq 1 - 1/e$, and this bound is tight

Lower bound based on the correlation gap bound [Vondrák'07]

Upper bound



Optimal social welfare: n

$\because$ There exists a perfect matching w.h.p.

Optimal strategy profile: $\approx (1 - 1/e)n$

$\because$ The expected probability that each resource is chosen can be upper-bounded

SR gap lower bound (correlated case)

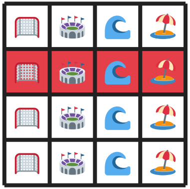
Theorem

$\mathrm{SRgap} = \Omega(1/\sqrt{n})$, and this bound is tight

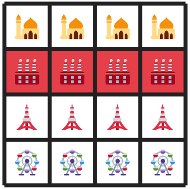
Lower bound (complicated)

Upper bound $\Theta_1 = \cdots = \Theta_n = [n]^k$, where $k = \sqrt{n}$ $\quad j \sim [k]$ and $\ell_1, \ldots, \ell_k \sim [n]$

Types $\{(\ell_1, \ldots, \ell_{j-1}, t, \ell_{j+1}, \ldots, \ell_k) \mid t \in [n]\}$ are randomly assigned to n players



1st action



2nd action

$E = [k] \times [n]$ set of resources

The hth action of type ℓ is to choose $(h, \ell_h) \in E$

Optimal social welfare: n

Optimal strategy profile: $\leq k + n/k = 2\sqrt{n}$

Improved PoA lower bound for com. eq.

Proposition

If ρ is independent, $\mathrm{PoA}_{\mathrm{Com.Eq.}} \geq 0.5$, which improves on the SR-gap approach

Based on the smoothness arguments for Bayes–Nash equilibria

[Roughgarden'15b, Syrgkanis'12]

The key step of their proof

Swapping θ_i and θ'_i in $\theta \sim \rho$ and $\theta' \sim \rho$ using the independence of ρ

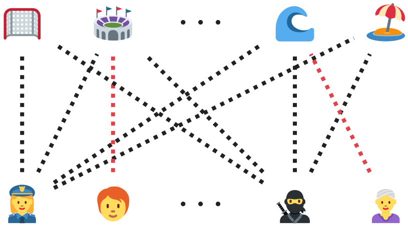
$\leftarrow$ Incentive constraints for misreporting θ'_i instead of θ_i can be used

Remark The same result also holds for agent-normal-form CE

PoA upper bound for Bayesian solutions

Proposition

$\mathrm{PoA_{BS}} \le \dfrac{1 - 1/\sqrt{e}}{3/2 - 1/\sqrt{e}} \approx 0.4403$ for some example with independent ρ



Odd players are connected to all resources

Even players are connected to random one

Odd players are prioritized over even ones

Bad Bayesian solution:

Each $(2k-1)$th player is recommended to choose the $(2k)$th player's action

Optimal: $\approx \underbrace{n/2}_{\text{odd}} + \underbrace{(1-1/\sqrt{e})n}_{\text{even}}$, **Bayesian solution**: $\approx (1-1/\sqrt{e})n$