General-Sum Games and Correlated Equilibrium

Chen-Yu Wei

General-Sum Games

Normal-Form Game

$$[M], (\mathcal{A}_j)_{j\in[M]}, (u_j)_{j\in[M]}$$

- Player $1, 2, \dots, M$
- Player j choose action $a_j \in \mathcal{A}_j$
- Player j's payoff/utility is $u_j(a_1, \dots, a_M) = u_j(a_j, a_{-j})$

$$a_{-j} \coloneqq (a_1, a_2, \dots, a_{j-1}, a_{j+1}, \dots, a_M)$$

Two-player zero-sum game is a special case when $M = 2$ and $u_1 = -u_2$

Example: Prisoner's Dilemma

	Cooperate	Defect
Cooperate	-1,-1	-10,0
Defect	0,-10	-6,-6

Prisoner's Dilemma

	High	Low
High	5,5	0,6
Low	6,0	3,3

Pricing game

General form:

	C	D
C	x,x	z,w
D	w,z	y,y

$w > x > y > z$

Example: Chicken Game

	Chicken	Dare
Chicken	6,6	2,7
Dare	7,2	0,0

Chicken game

	Compromise	Stand firm
Compromise	0,0	-5,5
Stand firm	5,-5	-10,-10

Military confrontation / negotiation

General form:

	C	D
C	x,x	z,w
D	w,z	y,y

w > x > z > y

Example: Stag Hunt

	Stag (C)	Hare (D)
Stag (C)	6,6	2,7
Hare (D)	7,2	4,4

Stag Hunt

	Wait	Withdraw
Wait	5,5	0,4
Withdraw	4,0	2,2

Bank runs

General form:

	C	D
C	x,x	z,w
D	w,z	y,y

$x > w > y > z$

Equilibrium

- An equilibrium is a **steady state** where no player wants to unilaterally deviate
  - Different strategy spaces and allowed deviations give different equilibrium concepts

- Traditional game theory studies equilibria. Studying equilibria is meaningful because it allows a player to
  - Predict other players' moves
  - Reason about their own move

  It also allows the game designer (e.g. government) to
  - Design the rules of the game so that the equilibrium outcome is desirable **(mechanism design)**

Nash Equilibrium (NE)

$(p_1, p_2, \cdots, p_M) \in \Delta_{A_1} \times \Delta_{A_2} \times \cdots \Delta_{A_N}$ is called a NE if

$$u_j(p_j, p_{-j}) \geq u_j(p_j', p_{-j}) \qquad \forall\, p_j' \in \Delta_{A_j}$$

$$u_j(p_1, p_2, \cdots, p_M) = \sum_{a_1, \cdots, a_M} p_1(a_1)\, p_2(a_2) \cdots p_M(a_M)\, u_j(a_1, \cdots, a_M)$$

	A	B
A	1, 1	0, 0
B	0, 0	2, 2

(A, A)

(B, B)

Incomplete-information game and Bayesian NE

In a incomplete-information game:

The environment first draws $\theta = (\theta_1, \theta_2, \cdots, \theta_m) \sim p$ (known to all players)

↑ observation

Player i observes θ_i

Player i chooses p_i

Bayesian NE

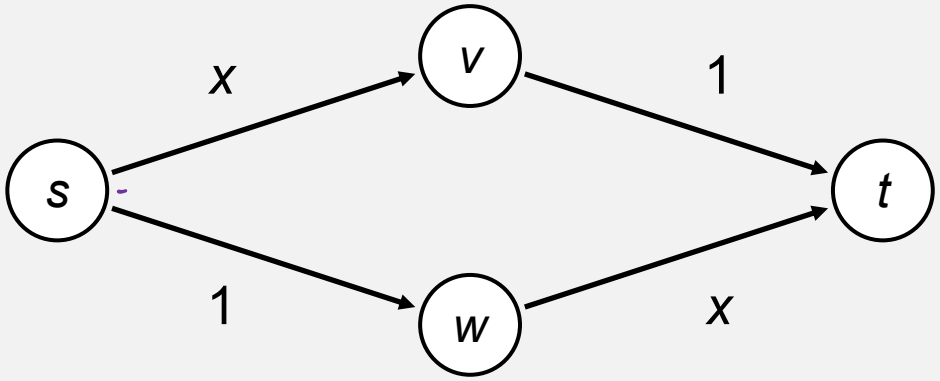
$$\mathop{\mathbb{E}}_{\theta \sim p}\left[u_j\left(p_j(\cdot|\theta_j), p_{-j}(\cdot|\theta_{-j}), \theta\right)\right] \geq \mathop{\mathbb{E}}_{\theta \sim p}\left[u_j\left(p_j'(\cdot|\theta_j), p_{-j}(\cdot|\theta_{-j}), \theta\right)\right]$$

$\forall p_j'$

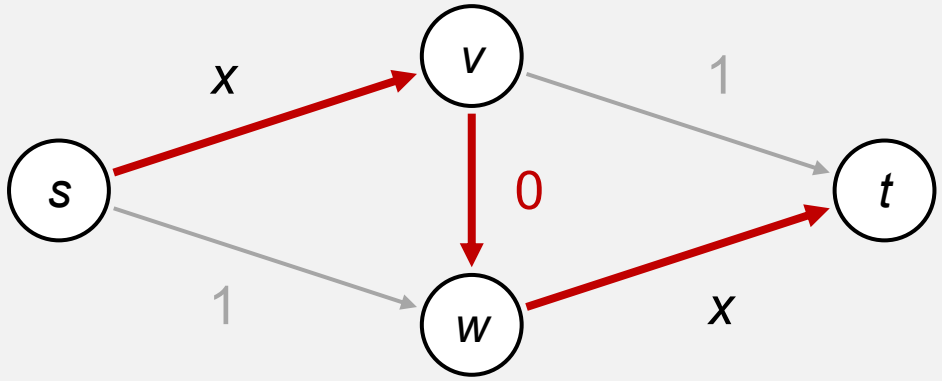
NE Example: Braess's paradox

1 unit of traffic goes from s to t, and every driver takes a fastest route.
Edge labels are travel times, where *x* = fraction of drivers using that edge.

Before



After: add a road v → w with travel time 0



NE Example: Second-Price Auction

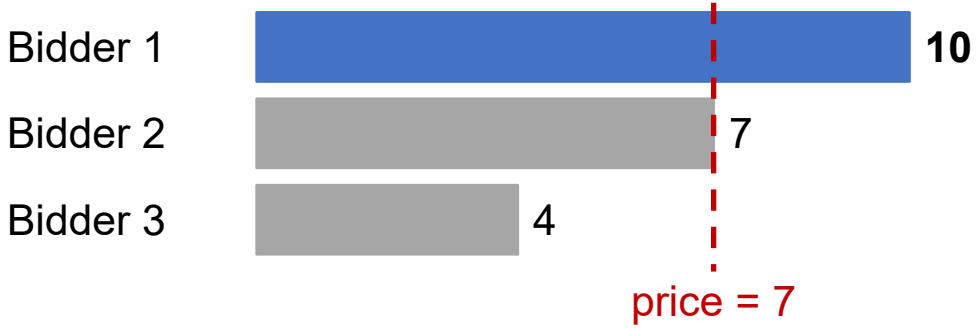
$b_i = v_i$ (NE)

Rules

- Each bidder i privately knows their value v_i
- Bidders submit sealed bids b_i simultaneously
- The highest bid wins
- The winner pays the second-highest bid

$$u_i(b_i, b_{-i}) = \begin{cases} v_i - \max_{j \neq i} b_j & \text{if } b_i > \max_{j \neq i} b_j \\ 0 & \text{otherwise} \end{cases}$$

Example (everyone bids truthfully)



Why bid truthfully?

Fix bidder i (value v, bid b) and let $B = \max_{j \neq i} b_j$ be the highest competing bid.

	Truthful b = v	**Overbid b > v**	**Underbid b < v**
B < v	Win: utility v − B	Same as truthful	May lose: utility 0
B > v	Lose: utility 0	May win: utility v − B < 0	Same as truthful

In every case, b = v is at least as good, so truthful bidding is a dominant strategy

Does NE Really Reflect the Players' Behaviors

- Human do not always act according to NE strategy. For example, in the second price auction, it is found that people tends to overbid
- Algorithmically, independent utility maximization algorithms do not always lead to NE
- Computing NE in a general-sum game is PPAD-hard
  - Existing algorithms requires $\left(M \sum_j A_j\right)^{O\left(\sum_j A_j\right)}$ time find an NE

$P \subseteq PPAD \subseteq NP$

Kagel, Harstad & Levin (1987), "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study

Gradient Descent/Ascent in Two-Player Zero-Sum Games

Example. Gradient descent /ascent in two-player zero-sum games leads to *divergence* from the NE.

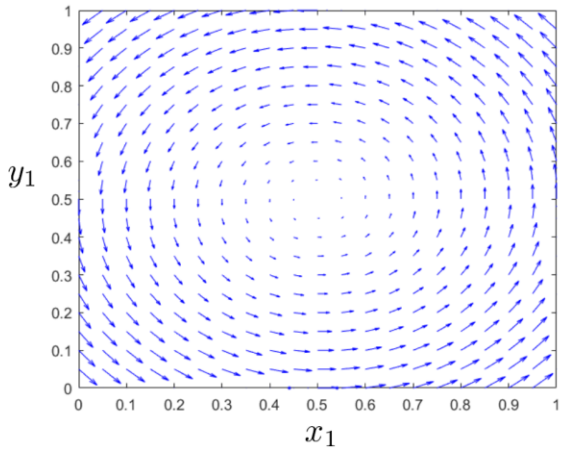
$$x_{t+1} = x_t + \eta \nabla_x f(x_t, y_t)$$

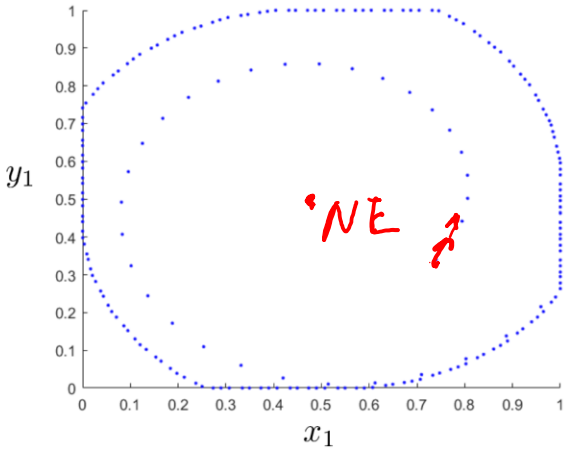
$$y_{t+1} = y_t - \eta \nabla_y f(x_t, y_t)$$

$\bar{x} = \sum_{t=1}^{T} x_t \qquad \bar{y} = \sum_{t=1}^{T} y_t$

$(\bar{x}, \bar{y}) \to NE.$

(x_t, y_t)





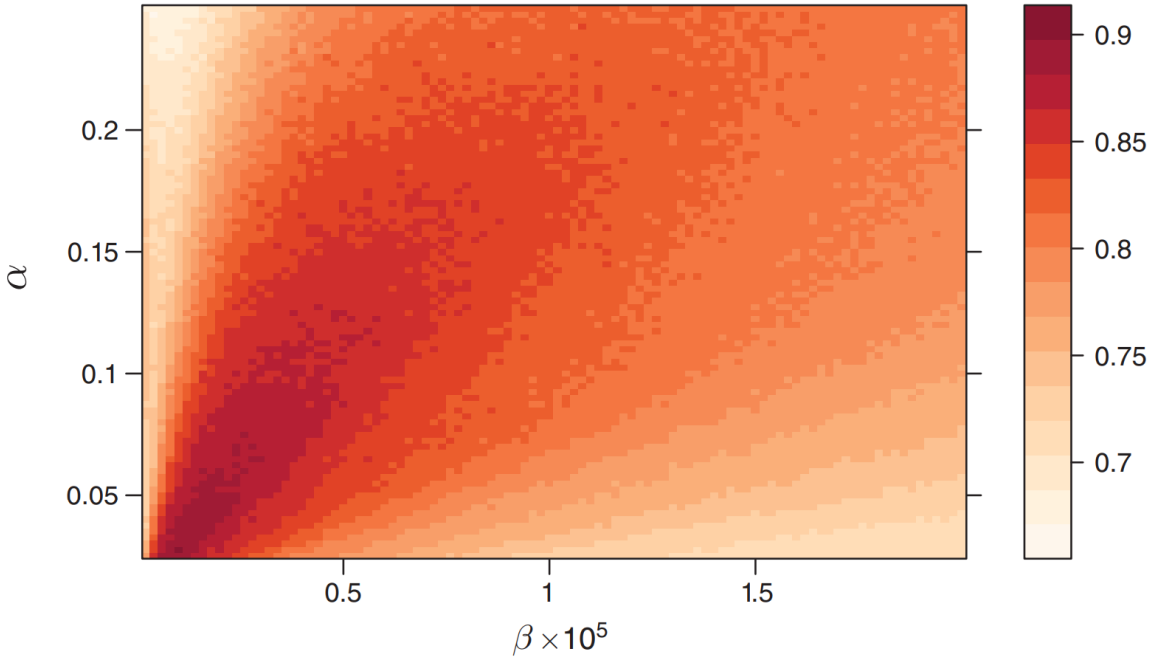
Independent Q-Learning in Pricing Game

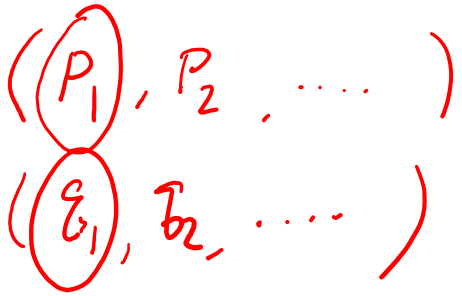
$$\Delta = \frac{\bar{R} - R^{\mathrm{NE}}}{R^{\mathrm{Monopoly}} - R^{\mathrm{NE}}}$$

R^{Monopoly}: profit in full cooperation

R^{NE}: profit in full competition
(i.e., equilibrium profit in one-shot game)

collusion





Calvano et al. Artificial Intelligence, Algorithmic Pricing, and Collusion. 2020.

Learning in Games

General questions:

$(a_{t,1}, a_{t,2}, \cdots, a_{t,m})$

- What does the empirical distribution of joint play converge to, or approach, when players interact using specific learning algorithms?
- What can a player using a specific algorithm guarantee about their average utility?

For example,

- If every player uses Q-learning
- If every player uses no-(swap)-regret algorithm (symmetric)
- If one player uses no-(swap)-regret algorithm and the other player knows it and optimizes over it

Correlated Equilibrium

and its connection to no-swap-regret dynamics

Correlated Equilibrium

Setting is as before

Additionally, there is "mediator" that will give suggested actions to every player.

Correlated Equilibrium. $\sigma \in \Delta(A_1 \times A_2 \times \cdots \times A_M)$ is a CE iff

$$\mathop{\mathbb{E}}_{a_{-j}} \left[u_j(a_j, a_{-j}) \mid a_j \right] \geq \mathop{\mathbb{E}}_{a_{-j}} \left[u_j(a_j', a_{-j}) \mid a_j \right] - \varepsilon$$

$$\| \qquad\qquad \geq$$

$$\forall a_j \in \text{supp}(\sigma) \quad \mathop{\mathbb{E}}_{a_{-j} \sim \sigma(\cdot \mid a_j)} \left[u_j(a_j, a_{-j}) \right] \qquad \mathop{\mathbb{E}}_{a_{-j} \sim \sigma(\cdot \mid a_j)} \left[u_j(a_j', a_{-j}) \right] - \varepsilon$$

↑ expected payoff if i'm recommended a_j

↑ expect payoff if we don't follow the recommendation a_j

Correlated Equilibrium Example: Traffic Light

Two drivers meet at an intersection (a Chicken game). A traffic light picks one of the three cells marked 1/3 at random, and shows each driver only their own light.

	Stop	**Go**
Stop	6, 6 prob. 1/3	2, 7 prob. 1/3
Go	7, 2 prob. 1/3	0, 0 prob. 0

If we are recommended "Stop":

$\Rightarrow \Pr(\text{the other player is recommended "Stop"}) = \frac{1/3}{1/3+1/3} = 1/2$

$\Pr(\text{_____ "Go"}) = 1/2$

$\Rightarrow \mathbb{E}[u_1(\text{Stop}, a_2) \mid \text{stop}] = \frac{6+2}{2} = 4$ (BR)

$\mathbb{E}[u_1(\text{go}, a_2) \mid \text{stop}] = \frac{7+0}{2} = 3.5$

If _____ "Go":

$\mathbb{E}[u_1(\text{Go}, a_2) \mid \text{Go}] = 7$ (BR)

$\mathbb{E}[u_1(\text{Stop}, a_2) \mid \text{Go}] = 6$

No-Swap-Regret Learning converges to Correlated Equilibria

Let M players repeated play a normal-form game with each other using no-swap-regret algorithms.

Let the action profile be $(a_{t,1}, \dots, a_{t,M})$ in round t.

Then their ~~average joint distribution~~ converges to a CE.

$$\sigma(a_{-j} \mid a_j) = \frac{\sum_t \mathbb{1}(a_{t,j} = a_j, a_{t,-j} = a_{-j})}{\sum_t \mathbb{1}(a_{t,j} = a_j)}$$

$$\sigma(a_1, a_2, \cdots, a_M) = \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}(a_{t,1} = a_1, \cdots, a_{t,M} = a_M)$$

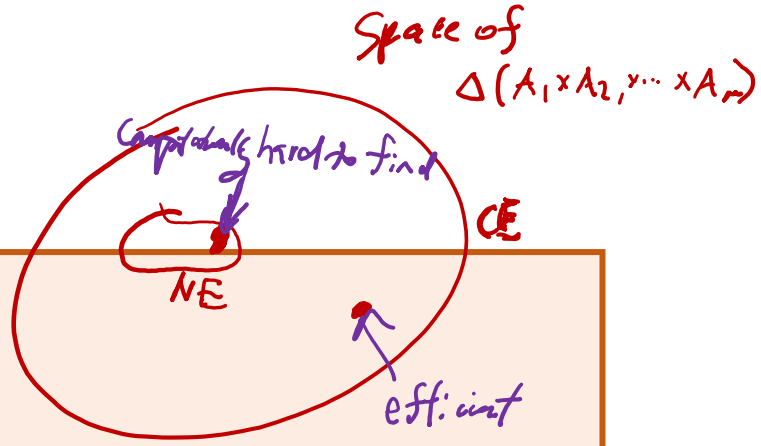
P.S. by swap regret, $\forall j, \forall a_j'$

$$\sum_{t=1}^{T} \mathbb{1}(a_{t,j} = a_j) \left(u_j(a_j', a_{t,-j}) - u_j(a_j, a_{t,-j}) \right) \le o(T)$$

$$\Rightarrow \sum_{t=1}^{T} \sum_{a_{-j}} \mathbb{1}(a_{t,j} = a_j, a_{t,-j} = a_{-j}) \left(u_j(a_j', a_{-j}) - u_j(a_j, a_{-j}) \right) \le o(T)$$

$$\Rightarrow \sum_{t=1}^{T} \mathbb{1}(a_{t,j} = a_j) \mathop{\mathbb{E}}_{a_{-j} \sim \sigma(\cdot \mid a_j)} \left(u_j(a_j', a_{-j}) - u_j(a_j, a_{-j}) \right) \le o(T)$$

Nash Equilibria ⊆ Correlated Equilibria



$\sigma \in \Delta(\mathcal{A}_1 \times \cdots \times \mathcal{A}_M)$ is a Correlated Equilibrium iff

$\forall j \quad \forall\, a_j, a_j' \in \mathcal{A}_j$ with $a_j \in \text{support}(\sigma)$

$$\mathbb{E}_{a_{-j} \sim \sigma(\cdot|a_j)}\left[u_j(a_j, a_{-j})\right] \geq \mathbb{E}_{a_{-j} \sim \sigma(\cdot|a_j)}\left[u_j(a_j', a_{-j})\right]$$

$\sigma(a_1, a_2, \cdots, a_M) = P_1(a_1)\, P_2(a_2) \cdots P_M(a_M)$

$(p_1, p_2, \cdots, p_M) \in \Delta(\mathcal{A}_1) \times \cdots \times \Delta(\mathcal{A}_M)$ is a Nash Equilibrium iff

$\forall j \quad \forall\, p_j' \in \Delta(\mathcal{A}_j) \qquad u_j(p_j, p_{-j}) \geq u_j(p_j', p_{-j})$

$u_j(p_j, p_{-j}) = \mathbb{E}_{a_j \sim p_j} \mathbb{E}_{a_{-j} \sim p_{-j}} \left(u_j(a_j, a_{-j})\right)$

Coarse Correlated Equilibrium

and its connection to no-regret dynamics

Coarse Correlated Equilibrium (CCE)

CE: $$\mathbb{E}_{a_{-j}\sim\sigma(\cdot|a_j)}\left[u_j(a_j, a_{-j})\right] \geq \mathbb{E}_{a_{-j}\sim\sigma(\cdot|a_j)}\left[u_j(a'_j, a_{-j})\right]$$

CCE: $$\mathbb{E}_{a\sim\sigma}\left[u_j(a_j, a_{-j})\right] \geq \mathbb{E}_{a_{-j}\sim\sigma}\left[u_j(a'_j, a_{-j})\right]$$

Check : For every player. follow the recommendation is no worse than choosing a fixed action

Coarse Correlated Equilibrium Example

	Rock	Paper	Scissor
Rock	0,0 prob. 0	-1, 1 prob. 1/6	1, -1 prob. 1/6
Paper	1, -1 prob. 1/6	0, 0 prob. 0	-1, 1 prob. 1/6
Scissor	-1, 1 prob. 1/6	1, -1 prob. 1/6	0, 0 prob. 0

No-Regret Learning converges to CCE