**ECE-GY 6263 Game Theory** Fall 2019

Lecture 4 — September 27, 2019

*Prof. Quanyan Zhu* *Scribe: Yunfei Ge*

# 1 Overview

In the last lecture, we introduced nonzero-sum finite games, proved the existence of mixed strategy Nash equilibrium, and presented computational methods (best response, fictitious play) for finding it. We also introduced some interpretations of Nash Equilibrium as follows.

**Interpretation of Nash Equilibrium**

- Repeated learning process – (e.g. Fictitious Play)
- Recommendation
- Rationality – (e.g. Prisoner's Lemma)
- ...

In this lecture, we will talk about Braess's paradox[1], correlated equilibrium[2] and its learning method (No-regret learning[3]), and finally the form of extensive games.

# 2 Braess's Paradox ([BO] Chap. 4.7)

In this section, we present a nonzero-sum game for which the Nash equilibrium solution leads to a surprising phenomenon, called the "Braess paradox". This equilibrium is also called Wardrop Equilibrium in traffic networks.

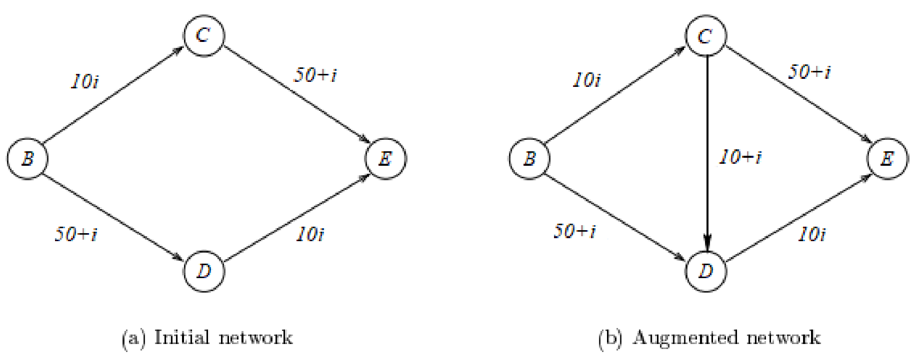


Figure 1: The routes of the Baess's paradox

## 2.1 Initial Network

Consider a network of one-way roads as given in Fig 1.(a). Travelers on this network all want to go from $B$ to $E$. There are two possible routes $\{BCE, BDE\}$. The time needed to get through each segment depends on the unit of travelers $i$.

Suppose there are 6 units of traffic in this network, $x_1$ of them choose route $BCE$ and $x_2$ of them choose route $BDE$. At the equilibrium, no traveler has the intention to leave their routes, which lead to the total time for both routes being equal.

$$\begin{cases} 10x_1 + 50 + x_1 = 50 + x_2 + 10x_2 & (t_{BCE} = t_{BDE}) \\ x_1 + x_2 = 6 & \text{(total traffic)} \end{cases}$$

$$\Longrightarrow \text{Equilibrium: } x_1 = x_2 = 3$$
$$\text{Total traveling time: } T = (10 \times 3 + 50 + 3) + (50 + 3 + 10 \times 3) = 166$$

**Remark:** Equilibrium between links

Consider the following case of two links from $A$ to $B$. The cost on the upper route is always equal to 1, while the cost on the lower route is proportional to the unit of travelers $i \in [0, 1]$ on the link. In the beginning, if there are 10% of people choosing the upper route, they notice that the time cost on the lower route is $0.9 < 1$, and will transfer to the lower route. The equilibrium will be achieved until the cost of the two links are identical, which means everyone will choose the lower route.

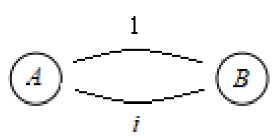


Figure 2: Basic case

## 2.2 Augmented Network

Now reconsider the problem with one link added from C to D as in Fig 1.(b). There are now three routes available $\{BCE, BDE, BCDE\}$. Suppose the unit of traffic on each route is $x_1, x_2, x_3$ accordingly, the Nash solution in this new game is

$$\text{Time on each route: } \begin{cases} t_{BCE} = 10(x_1 + x_3) + 50x_1 \\ t_{BDE} = 10(x_2 + x_3) + 50x_2 \\ t_{BCDE} = 10(x_1 + x_3) + 10(x_2 + x_3) + 10 + x_3 \end{cases}$$

$$t_{BCE} = t_{BDE} = t_{BCDE}$$
$$x_1 + x_2 + x_3 = 6$$

$$\implies \text{Equilibrium: } x_1 = x_2 = x_3 = 2$$
$$\text{Total traveling time: } T = 92 \times 3 = 276$$

Braess' paradox is the observation that adding one or more roads to a road network can end up impeding overall traffic flow through it. Apparently, in this case, adding the number of routes has led to a worse result in total traveling time. Note that travelers acting selfishly without considering overall performance is the cause of the paradox.

# 3 Correlated Equilibrium

## 3.1 Game of Chicken (Road Intersection Game)

Game of chicken can be thought of as two players walking towards each other in front of an intersection. A player can choose to go (G) and pass through the intersection or yield (Y) and stop. Both players are maximizers. The payoff matrix is given by

| | G | Y |
|---|---|---|
| G | (-10,-10) | (5,0) |
| Y | (0,5) | (-1,-1) |

There are two pure strategy Nash Equilibrium: $(G, Y)$ and $(Y, G)$.

There is also a Nash Equilibrium in mixed strategies. Suppose the probabilities of choosing different actions are $P_1 : (y, 1-y)$, $P_2 : (z, 1-z)$. The mixed strategy can be obtained by applying best response.

$$BR_1(z) = \begin{cases} [0,1] & z = \frac{3}{8} \\ 0 & z > \frac{3}{8} \\ 1 & z < \frac{3}{8} \end{cases} \qquad BR_2(y) = \begin{cases} [0,1] & y = \frac{3}{8} \\ 0 & y > \frac{3}{8} \\ 1 & y < \frac{3}{8} \end{cases}$$

The mixed strategy Nash Equilibrium is $P_1 : (\frac{3}{8}, \frac{5}{8})$, $P_2 : (\frac{3}{8}, \frac{5}{8})$ The four possible action profiles have the following probability under this equilibrium.

| | G | Y |
|---|---|---|
| G | 9/64 | 15/64 |
| Y | 15/64 | 25/64 |

If we interpret this Nash Equilibrium as a recommendation, it doesn't make sense in the way it suggests the players get negative utilities (collision in this case) with probability $\frac{9}{64}$. A better equilibrium which is fair without danger of collision could be the following

| | G | Y |
|---|---|---|
| G | 0 | 1/2 |
| Y | 1/2 | 0 |

However, the independent strategies of the players cannot lead to this solution. The above outcome requires that players are somehow correlated with each other.

### 3.2 Correlated Equilibrium

**Definition** *Let $\pi$ be a probability distribution over $X_1, X_2, \ldots, X_n$. The correlated mixed strategy $\pi$ is a* ***correlated equilibrium*** *if*

$$\sum_{a_{-i}} \pi(a_i, a_{-i}) U_i(a_i, a_{-i}) \geq \sum_{a_{-i}} \pi(a_i, a_{-i}) U_i(a_i', a_{-i}) \qquad \forall i \in \{1, 2, \ldots, n\}, \forall a_i, a_i' \in X_i \tag{1}$$

**Example** Consider the game of chiken in the previous example. Assign probability to each action profiles. The correlated equilibrium can be solved using (1).

| | G | Y |
|---|---|---|
| G | $\pi_{11}$ | $\pi_{12}$ |
| Y | $\pi_{21}$ | $\pi_{22}$ |

(1) Player 1

$$\sum_{a_2} \pi(a_1, a_2)(U_1(a_1, a_2) - U_1(a_1', a_2)) \geq 0$$

$a_1 = G$:

$$\begin{aligned} \pi(G,G)(U_1(G,G) - U_1(Y,G)) + \pi(G,Y)(U_1(G,Y) - U_1(Y,Y)) \geq 0 \\ \Longrightarrow -10\pi_{11} + 6\pi_{12} \geq 0 \end{aligned}$$

$a_1 = Y$:

$$\begin{aligned} \pi(Y,G)(U_1(Y,G) - U_1(G,G)) + \pi(Y,Y)(U_1(Y,Y) - U_1(G,Y)) \geq 0 \\ \Longrightarrow 10\pi_{21} - 6\pi_{22} \geq 0 \end{aligned}$$

(2) Player 2 follows the same logic

$$\sum_{a_1} \pi(a_1, a_2)(U_2(a_1, a_2) - U_2(a_1, a_2')) \geq 0$$

We can get a set of constraints over the probability that leads to a set of correlated equilibrium.

$$\begin{cases} 10\pi_{11} - 6\pi_{12} \leq 0 \\ -10\pi_{21} + 6\pi_{22} \leq 0 \\ 10\pi_{11} - 6\pi_{21} \leq 0 \\ -10\pi_{12} + 6\pi_{22} \leq 0 \\ \pi_{11} + \pi_{12} + \pi_{21} + \pi_{22} = 1 \\ \pi_{ij} \geq 0 & i, j = \{1, 2\} \end{cases}$$

**Remark**

- Correlated equilibrium is a set of equilibrium.
- Solution examples:

| | $\pi_{11}$ | $\pi_{12}$ | $\pi_{21}$ | $\pi_{22}$ |
|---|---|---|---|---|
| 1 | 0 | 0.55 | 0.4 | 0.05 |
| 2 | 0 | 0.5 | 0.5 | 0 |
| 3 | 0 | 0.7 | 0.3 | 0 |
| 4 | 9/64 | 15/64 | 15/64 | 25/64 |
| 5 | 0 | 1 | 0 | 0 |
| 6 | 0 | 0 | 1 | 0 |

  Notice that pure strategies (5)(6) and mixed strategy (4) are in the set of correlated equilibrium.

- Solution is equivalent to feasibility of Linear Programming, could be solved using LP techniques.
- "Obedience". Players have the incentive to obey correlated equilibrium.
- "Signals". The recommendation may not be direct. It could be a signal.

**Fact** Nash equilibrium always exists in mixed strategies, thus there always exists correlated equilibrium. We have $NE \subseteq CE$.

# 4 No-Regret Learning [3] (Converge to a set of CE)

Consider the strategic form of the game: $\langle N, (S_i)_{i\in N}, (u_i)_{i\in N}\rangle$, where $N$ represents the players, $S$ is the collections of action space, and $U$ is the collections of utility functions. The game is played repeatedly through time $t = 1, 2, \ldots$.

## 4.1 Learning Process

(1) At time t+1

Given a shared history of plays $h_t = (s^\tau)_{\tau=1}^t$, where $s^\tau = \{s_1^\tau, s_2^\tau, \ldots, s_N^\tau\}$ is the collection of action profiles. Each player $i \in N$, choose action $s_i^{t+1} \in S_i$ according to "some" probability distribution.

(2) Player $i$: $s_i = \{k, j\} \in S_i$.

If player $i$ replace action $j$ with action $k$ at time $\tau$ (others remain the same).

**Payoff**

$$W_i^\tau(j,k) := \begin{cases} u_i(k, s_{-i}^\tau) & \text{if } s_i^\tau = j \text{, change it to } k \\ u_i(s_i^\tau) & \text{otherwise, remain the same} \end{cases} \tag{2}$$

**Accumulated chagne**

$$D_i^t(j,k) := \frac{1}{t}\sum_{\tau=1}^{t} W_i^{\tau}(j,k) - \frac{1}{t}\sum_{\tau=1}^{t} u_i(s^{\tau}) \tag{3}$$

Every time player $i$ plays j, replace it with action $k$. This value $D$ represents the difference between the new accumulated average payoff and the original payoff up to time $t$. We can multiply a discounting factor $\alpha^{\tau}$ in each term to give weights on history.

**Regret**

$$R_i^t(j,k) := \max\{D_i^t(j,k), 0\} \tag{4}$$

When $D_i^t(j,k)$ is negative, there is no regret. When $D_i^t(j,k)$ is positive, we can choose action $k$ instead of $j$ with probability proportional to regret. Let $j \in S_i$ be the strategy last choosen by player $i$, i.e. $j = s_i^t$. Then the probability distribution used by $i$ at time $t+1$ should be

$$\begin{cases} P_i^{t+1}(k) = \frac{1}{\mu} R_i^t(j,k) & \text{for all } k \neq j \\ P_i^{t+1}(j) = 1 - \sum_{k \in S_i, k \neq j} P_i^{t+1}(k) \end{cases} \tag{5}$$

where $\mu$ is a normalization factor.

## 4.2 Results

For every $t$, let $z_t \in \Delta(S)$ be the empirical distribution of the N-tuples of strategies played up to time $t$. That is, for every $s \in S$,

$$z_t(s) := \frac{1}{t}|\{\tau \leq t : s_{\tau} = s\}| \tag{6}$$

Here $|Q|$ is the number of elements of a finite set Q. (6) is the relative frequency that the N-tuple s has been played in the first $t$ periods.

**Theorem[3]** *If every player plays according to the adaptive procedure in Sec.4.1, then the empirical distributions of play $z_t$ converge almost surely as $t \to \infty$ to the set of correlated equilibrium distributions of the game.*

**Remark**

- Correlation in this learning process comes from the fact that players share the same history. The shared history correlates the players' action choice even when the action picking is independent by themselves.
- Correlated equilibrium can also be interpreted in "Extensive Games".

# 5 Extensive Game

The games we've seen so far are strategic form games (normal-form games) with matrix representation. This kind of form does not capture the sequential nature of the play. The sequence of action picking doesn't matter (e.g. *sealed bid auction*). If we want to capture the sequence of actions, we should look at the extensive form games with the tree structure.

**Example: Model of Entry** Suppose player $A$ is a big company. Player $B$ is trying to decide whether he should enter or not enter $A$'s company. If $B$ chooses to enter, $A$ would decide whether he should fight with $B$ or accommodate $B$. The payoff is shown in Fig. 3. In this game, player $A$ can observe what $B$ has done, and then choose action.

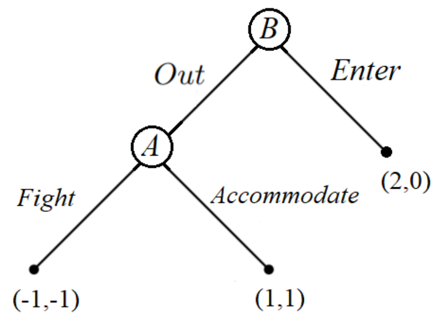


Figure 3: Model of Entry

**Example: Chicken Game** In Fig 4, we add a dotted line to represent that player 2 cannot distinguish two states. In other words, player 2 does not know that player 1 has done. This game then has the same information structure as the original chicken game in matrix form as in Sec. 3.1.

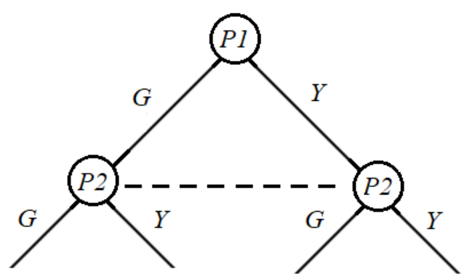


Figure 4: Chicken Game

## References


[1] Basar T, Olsder GJ. "Dynamic noncooperative game theory." Vol. 23. Siam, 1999.

[2] Aumann, R. J., "Subjectivity and correlation in randomized strategies." *Journal of mathematical Economics*, 1(1), 67-96, 1974

[3] Sergiu Hart  Andreu Mas-Colell, "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica*, vol. 68(5), pages 1127-1150, 2000.