# ECE700.07: Game Theory with Engineering Applications

## Lecture 6: Repeated Games

Seyed Majid Zahedi

UNIVERSITY OF WATERLOO

# Outline

- Finitely and infinitely repeated games w/ and w/o perfect monitoring
- Trigger strategies
- Folk theorems

- Readings:
  - MAS Sec. 6.1, GT Sec. 5.1 and 5.5

# Finitely Repeated Games (with Perfect Monitoring)

- In repeated games, *stage game* $G$ is played by same agents for R rounds
  - Agents discount utilities by discount factor $0 \leq \delta \leq 1$
  - Game is denoted by $G^R(\delta)$
- At each round, outcomes of all past rounds are observed by all agents
- Agents' overall utility is sum of discounted utilities at each round
- Given sequence of utilities $u_i^{(1)}, \dots, u_i^{(R)}$

$$u_i = \sum_{r=1}^{R} \delta^{r-1} u_i^{(r)}$$

- In general, strategies at each round could depend on history of play
  - Memory-less (also called *stationary*) strategies are special cases

# Example: Finitely-Repeated Prisoners' Dilemma

- Suppose that *Prisoners' Dilemma* is played in R (< ∞) rounds

| Prisoner 1 \ Prisoner 2 | Stay Silent | Confess |
|---|---|---|
| Stay Silent | (-1, -1) | (-3, 0) |
| Confess | (0, -3) | (-2, -2) |

- What is SPE of this game?
  - We can use backward induction
  - Starting from last round, (C, C) is dominant strategy
  - Regardless of history, (C, C) is dominant strategy at each round
  - There exists unique SPE which is (C, C) at each round

# SPE in Finitely Repeated Games

[Theorem]

- If stage game $G$ has unique pure strategy equilibrium $s^*$, then $G^R(\delta)$ has unique SPE in which $s^{(r)} = s^*$ for all $r = 1, \dots, R$, regardless of history

[Proof]

- By backward induction, at round $R$, we have $s^{(R)} = s^*$
- Given this, then we have $s^{(R-1)} = s^*$, and continuing inductively, $s^{(r)} = s^*$ for all $r = 1, \dots, R$, regardless of history

# Infinitely Repeated Games

- Infinitely repeated play of $G$ with discount factor $\delta$ is denoted by $G^{\infty}(\delta)$
- Agents' utility is average of *discounted utilities* at each round
- For $\delta < 1$, given sequence of utilities $u_i^{(1)}, \ldots, u_i^{(\infty)}$

$$u_i = (1 - \delta) \sum_{r=1}^{\infty} \delta^{r-1} u_i^{(r)}$$

- For $\delta = 1$, given sequence of utilities $u_i^{(1)}, \ldots, u_i^{(\infty)}$

$$u_i = \lim_{R \to \infty} \frac{\sum_{r=1}^{R} u_i^{(r)}}{R}$$

# Trigger Strategies (TS)

- Agents get punished if they deviate from agreed profile
- In non-forgiving TS (or grim TS), punishment continues forever

$$s_i^{(t)} = \begin{cases} s_i^* & \text{if } s^{(r)} = s^*, \forall r < t \\ s_i^j & \text{if } s_j^{(r)} \neq s_j^*, \exists r < t \end{cases}$$

- Here, $s^*$ is agreed profile, and $s_i^j$ is punishment strategy of $i$ for agent $j$
- Single deviation by $j$ trigers agent $i$ to switch to $s_i^j$, forever

# Trigger Strategies in Repeated Prisoners' Dilemma

| Prisoner 2 / Prisoner 1 | Stay Silent | Confess |
|---|---|---|
| Stay Silent | (-1, -1) | (-3, 0) |
| Confess | (0, -3) | (-2, -2) |

- Suppose both agents use following trigger strategy
  - Play S unless someone has ever played C in past
  - Play C forever if someone has played C in past
- Under what conditions is this SPE?
- We use one-stage deviation principle
- Step 1: (S is best response to S)
  - Utility from S: $-(1-\delta)(1+\delta+\delta^2+\cdots) = -(1-\delta)/(1-\delta) = -1$
  - Utility from C: $-(1-\delta)(0+2\delta+2\delta^2+\cdots) = -2\delta(1-\delta)/(1-\delta) = -2\delta$
  - S is better than C if $\delta \geq 1/2$
- Step 2: (C is best response to C)
  - Other agents will always play C, thus C is best response

# Remarks

- Cooperation is equilibrium, but so are many other strategy profiles
- If $s^*$ is NE of $G$, then "*each agent plays* $s_i^*$" is SPE of $G^R(\delta)$
  - Future play of other agents is independent of how each agent plays
  - Optimal play is to maximize current utility, *i.e.*, play static best response
- Sets of equilibria for finite and infinite horizon versions can be different
  - Multiplicity of equilibria in repeated prisoner's dilemma only occurs at $R = \infty$
  - For any finite $R$ (thus for $R \to \infty$), repeated prisoners' dilemma has unique SPE

# TS in Finitely Repeated Games

- If $G$ has multiple equilibria, then $G^R(\delta)$ does not have unique SPE
- Consider following example

| Agent 1 \ Agent 2 | x | y | z |
|---|---|---|---|
| x | (3, 3) | (0, 4) | (-2, 0) |
| y | (4, 0) | (1, 1) | (-2, 0) |
| z | (0, -2) | (0, -2) | (-1, -1) |

- Stage game has two pure NE: (y, y) and (z, z)
- Socially optimal outcome, (x, x), is not equilibrium
- In twice repeated play, we can support (x, x) in first round

# TS in Finitely Repeated Games (cont.)

- TS strategy
  - Play x in first round
  - Play y in second round if opponent played x; otherwise, play z
- We can use one-shot deviation principle
  - For simplicity, suppose $\delta = 1$
  - Playing x first and y next leads to utility of 4
  - Playing y first triggers opponent to play z next, which leads to utility 3
  - Deviation is not profitable!

# Repetition Can Lead to Bad Outcomes

- Consider this game

| Agent 1 \ Agent 2 | x | y | z |
|---|---|---|---|
| x | (2, 2) | (2, 1) | (0, 0) |
| y | (1, 2) | (1, 1) | (-1, 0) |
| z | (0, 0) | (0, -1) | (-1, -1) |

- Strategy x strictly dominates y and z for both agents
- Unique Nash equilibrium of stage game is (x, x)
- If $\delta \geq 1/2$, this game has SPE in which (y, y) is played in every round
- It is supported by slightly more complicated strategy than grim trigger
  - I. Play y in every round unless someone deviates, then go to II
  - II. Play z. If no one deviates go to I. If someone deviates stay in II

# Feasible and Individually Rational Utilities

- $V$ = *Convex hull* of $\{v \in \mathbb{R}^{|\mathcal{I}|} \mid \text{there exists } s \in S \text{ such that } u(s) = v\}$
  - Utility in repeated game is just a weighted average of utilities in stage game
- Note that $V \neq \{v \in \mathbb{R}^{|\mathcal{I}|} \mid \text{there exists } \sigma \in \Sigma \text{ such that } u(\sigma) = v\}$
- Recall minmax value of agent $i$

$$\underline{v}_i = \min_{\sigma_{-i}} \max_{s_i} u_i(s_i, \sigma_{-i})$$

- Also recall minmax strategy against $i$

$$m^i_{-i} = \arg\min_{\sigma_{-i}} \max_{s_i} u_i(s_i, \sigma_{-i})$$

- Utility vector $v \in \mathbb{R}^{|\mathcal{I}|}$ is strictly individually rational if $v_i > \underline{v}_i, \ \forall i$

# Example

| Agent 1 \ Agent 2 | L | R |
|---|---|---|
| U | (-2, 2) | (1, -2) |
| M | (1, -2) | (-2, 2) |
| D | (0, 1) | (0, 1) |

- What is minimax value of agent 1?
  - Let $q$ denote probability that agent 2 chooses L
  - $\underline{v}_1 = \min_{0 \leq q \leq 1} \max \{1 - 3q, -2 + 3q, 0\} = 0$
  - $m_2^1 \in [1/3, 2/3]$
- What is minimax value of agent 2?
  - Let $p$ and $q$ denote probabilities that agent 1 chooses U and M, respectively
  - $\underline{v}_2 = \min_{0 \leq p, q \leq 1} \max \{p - 3q + 1, -3p + q + 1\} = 0$
  - $m_1^2 = (1/2, 1/2, 0)$

# Minmax Utility Lower Bounds

[Theorem]

- If $\sigma^*$ is NE of $G$, then $u_i(\sigma^*) \geq \underline{v}_i$
- If $\sigma^*$ is NE of $G^\infty(\delta)$, then $u_i(\sigma^*) \geq \underline{v}_i$

[Proof]

- Agent $i$ can always guarantee herself $\underline{v}_i$ in stage game and also in each round of repeated game, meaning that she can always achieve at least this utility against even most adversarial opponents

# Nash Folk Theorem

[Nash Folk Theorem]

- If $v$ is feasible and strictly individually rational, then there exists $\underline{\delta} < 1$, such that for all $\delta > \underline{\delta}$, $G^{\infty}(\delta)$ has NE with utilities $v$

[Proof]

- Suppose for simplicity that there exists pure strategy profile $s^*$ such that $u_i(s^*) = v_i$ (otherwise, proof is more involved)
- Consider following grim trigger strategy for agent $i$
  - Play $s_i$ as long as no one deviates
  - If some agent $j$ deviates, then play $m_i^j$ thereafter
  - If $i$ plays $s$, her utility is $v_i$

# Proof of Nash Folk Theorem

- We can use one-shot deviation principle
- Suppose $i$ deviates from $s$ in round $r$
- Define $\bar{v}_i = \max_{s_i} u_i(s_i, s^*_{-i})$
- We have

$$u_i \leq (1-\delta)(v_i + \delta v_i + \cdots + \delta^{r-1} v_i + \delta^r \bar{v}_i + \delta^{r+1} \underline{v}_i + \delta^{r+2} \underline{v}_i + \dots)$$

- Following $s^*$ will be optimal if

$$\begin{aligned} v_i &\geq (1-\delta^r) v_i + \delta^r (1-\delta) \bar{v}_i + \delta^{r+1} \underline{v}_i \\ &= v_i - \delta^r (v_i - \bar{v}_i + \delta(\bar{v}_i - \underline{v}_i)) \end{aligned}$$

- This means, $s^*$ is NE of $G^\infty(\delta)$ if

$$\delta \geq \underline{\delta} = \max_i \frac{\bar{v}_i - v_i}{\bar{v}_i - \underline{v}_i}$$

# Problems with Nash Folk Theorem

- Any utility can be achieved when agents are patient enough
- NE involves non-forgiving TS which may be costly for punishers
- NE may include non-credible threats
- NE may not be subgame perfect
- Example:
  - Unique NE in this game is (D, L)
  - Minmax values are given by $\underline{v}_1 = 0$ and $\underline{v}_2 = 1$
  - Minmax strategy against agent 1 requires agent 2 to play R
  - R is strictly dominated by L for agent 2

| Agent 1 \ Agent 2 | L | R |
|---|---|---|
| U | (6, 6) | (0, -100) |
| D | (7, 1) | (0, -100) |

# Subgame Perfect Folk Theorem

[Theorem]

- Let $s^*$ be NE of stage game $G$ with utilities $v^*$
- For any feasible utility $v$ with $v_i > v_i^*, \forall i \in \mathcal{I}$, there exists $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, $G^\infty(\delta)$ has SPE with utilities $v$

[Proof]

- Simply construct non-forgiving TS with punishment by static NE
- Punishments are therefore subgame perfect
- For $\delta$ sufficiently close to 1, it is better for each agent $i$ to obtain $v_i$ rather than deviate and get $v_i^*$ forever thereafter
- This shows that any utility higher than NE utilities can be sustained as SPE

# Repeated Games with Imperfect Monitoring

- At each round, all agents observe some public outcome, which is correlated with stage game actions
- Let $y \in Y$ denote publicly observed outcome
- Each strategy profile $s$ induces probability distribution over $y$
- Let $\pi(y, s)$ denote probability distribution of $y$ under action profile $s$
- Public information at round $r$ is $h^{(\mathrm{r})} = \left(y^{(1)}, \dots, y^{(\mathrm{r}-1)}\right)$
- Strategy of agent $i$ is sequence of maps $s_i^{(r)}: h^{(r)} \rightarrow S_i$

# Repeated Games with Imperfect Monitoring (cont.)

- Each agent's utility depends only on her own action and public outcome
  - Dependence on actions of others is through their effect on distribution of $y$
- Agent $i$'s realized utility at round $r$ is

$$u_i(s_i^{(r)}, y^{(r)})$$

- Given strategy profile $s^{(r)}$, agent $i$'s expected utility is

$$u_i(s^{(r)}) = \sum_{y \in Y} \pi(y^{(r)}, s^{(r)}) u_i(s_i^{(r)}, y^{(r)})$$

- Agent $i$'s average discounted utility for sequence $\{s^t\}$ is

$$u_i = (1 - \delta) \sum_{r=1}^{\infty} \delta^{r-1} u_i(s^{(r)})$$

# Example: Cournot Competition with Noisy Demand

**[Green and Porter, Noncooperative Collusion under Imperfect Price Information, 1984]**

- Firms set production levels $q_1^{(\mathrm{r})}, \ldots, q_n^{(\mathrm{r})}$ privately at round $r$
- Firms do not observe each other's output levels
- Market demand is stochastic
- Market price depends on total production and market demand
- Low price could be due to high production or low demand
- Firms utility depends on their own production and market price

# Simpler Example: Noisy Prisoner's Dilemma

- Prisoners don't observe each others actions, they only observe signal $y$
  - $u_1(S, y) = 1 + y \qquad u_1(C, y) = 4 + y$
  - $u_2(S, y) = 1 + y \qquad u_2(C, y) = 4 + y$
- Signal $y$ is defined by continuous random variable $X$ with $E[X] = 0$
  - If $s = (S, S)$, then $y = X$
  - If $s = (S, C)$ or $(C, S)$, then $y = X - 2$
  - If $s = (C, C)$, then $y = X - 4$
- Normal form stage game is

| Prisoner 1 \ Prisoner 2 | Stay Silent | Confess |
|---|---|---|
| Stay Silent | (1+X, 1+X) | (-1+X, 2+X) |
| Confess | (2+X, -1+X) | (X, X) |

# Trigger-Price Strategy

- Consider following trigger strategy for noisy Prisoner's Dilemma
  - (I) - Play (S, S) until $y \leq y^*$, then go to (II)
  - (II) - Play (C, C) for $R$ rounds, then go back to (I)
- Note that punishment uses NE of stage game
- We can choose $y^*$ and $R$ such that this strategy profile is SPE
- We use one-shot deviation principle
- Deviation in (II) is obviously not beneficial
- In (I), if agents do not deviate, their expected utility is

$$v = (1-\delta)\Big((1+0) + F(y^*)\delta^{(R+1)}v + \big(1 - F(y^*)\big)\delta v\Big)$$

$$\Rightarrow \quad v = \frac{1-\delta}{1 - \delta(1-\delta)\big(1 - F(y^*)(1-\delta^R)\big)}$$

# Trigger-Price Strategy (cont.)

- If some agent deviates in (I), then her expected utility is

$$v_d = (1-\delta)\Big((2+0) + F(y^*+2)\delta^{(R+1)}v + \big(1 - F(y^*+2)\big)\delta v\Big)$$

- Deviation provides immediate utility, but increases probability of entering (II)
- To have SPE, we mush have $v \geq v_d$ which means

$$v \geq \frac{2(1-\delta)}{1-\delta(1-\delta)\big(1-F(y^*+2)(1-\delta^R)\big)}$$

$$\Rightarrow 2F(y^*+2) - F(y^*) \leq \frac{1-\delta(1-\delta)}{\delta(1-\delta)(1-\delta^R)}$$

- Any $R$ and $y^*$ that satisfy this constraint construct SPE
- Best trigger-price strategy could be found if we maximize $v$ subject to this constraint

# Questions?

# Acknowledgement

- This lecture is a slightly modified version of one prepared by
  - Asu Ozdaglar [MIT 6.254]