CUR 412: Game Theory and its Applications, Lecture 14

Prof. Ronaldo CARPIO

June 7, 2016

Announcements

- HW5 is due next week.
- The final exam will be on June 20, 4-6 PM, in Boxue 507.
- Final will be closed-book and covers the second half of the course. Old finals and solutions are on the course website.
- Bring a non-programmable calculator.

Review of Last Lecture

- A *repeated game* is a situation where a game is repeatedly played for several (possibly infinite) periods.
- We assume that agents value payoffs in the *present*, more than payoffs in the *future* by a constant *discount factor*, δ.
- δ is between 0 and 1; an agent with a lower δ is said to be more *impatient*, i.e. places lesser weight on the future.
- The overall value of a sequence of payoffs $w^1, w^2, w^3, \ldots$ is given by the *discounted average*:

$$(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}w^t$$

- A strategy for a repeated game must specify an action for every possible history.

Strategies in a Repeated Game

- As we've seen before, a strategy in an extensive game needs to specifies an action after every history in which it is the player's turn to move.
- In repeated games, all players move after every history, so a strategy must specify a player's action after any possible history.
- In an infinitely repeated game, this could potentially require specifying actions for all possible histories of any length.
- We can simplify things by only looking at a special class of strategies, in which actions can depend only a finite subset of the past history.
- In other words, the strategy has a limited "memory", and can only "remember" a finite number of past moves.
- Here are some examples of this kind of strategy:

Grim Trigger

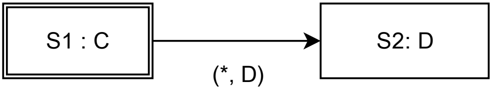
- **Grim Trigger**: This is a simple strategy that always plays C until the other player plays D; then it "punishes" the other player by always playing D.
- Formally, we define this strategy as:

$$s_i(\varnothing) = C$$

$$s_i(a^1, \ldots, a^t) = \begin{cases} C & \text{if } (a^1, \ldots, a^t) = (C, C, \ldots, C) \\ D & \text{otherwise} \end{cases}$$

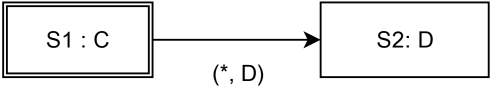
- The first part of the definition, $s_i(\varnothing) = C$, specifies what to do at the beginning of the game.
- The second part specifies what to do for any finite history.
- If the other player has never played D, then play C; if the other player has played D at any point in time, then play D.

Grim Trigger



- We can graphically represent strategies using a *state diagram*.
- Each box in this diagram represents a possible *state* of the strategy; here, there are two states: $S1, S2$.
- At the beginning of the game, the strategy starts out in the box with double edges, $S1$. The last word in the box, C, specifies what action to play in this state.

Grim Trigger



- Then, depending on what the outcome of the game in this period is, the strategy will either move to another state, or remain in the same state.
- The arrow labeled $(*, D)$ specifies when to *transition* to another state: if the outcome of the game is $(*, D)$, move to state $S2$.
- The $*$ means that any action of the first player, together with D played by the second player, will trigger this transition. If the outcome does *not* match $(*, D)$, then the strategy will remain in state $S1$.
- Once in state $S2$, the specified action is always D, and there are no more transition arrows out of this state, which means that the strategy will remain in this state forever (therefore, play D forever).

Always Cooperate

- **Always Cooperate**: This is one of the simplest strategies; it plays C after any history. Formally, it is defined as:

$$s_i(\varnothing) = C$$

$$s_i(a^1, ..., a^t) = C$$

- In the state diagram, the strategy begins in state $S1$, and remains there, always playing C.

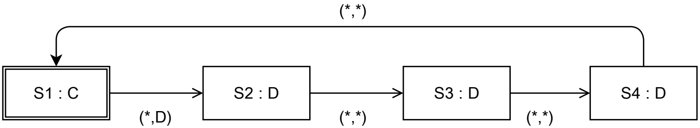
Always Defect

- **Always Defect**: Likewise, this strategy is defined as:

$$s_i(\varnothing) = D$$

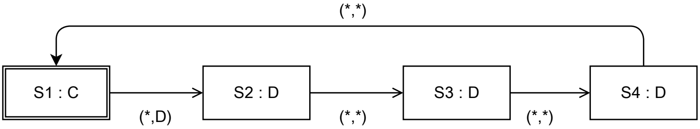
$$s_i(a^1, \ldots, a^t) = D$$

Punish for 3 periods



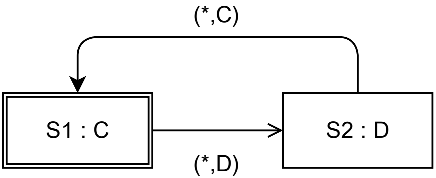
- Here's a an example of a more complicated strategy.
- **Punish for 3 periods**: This strategy plays C until the other player plays D, at which point this strategy will play D for 3 consecutive periods. Then, this strategy will "forget" the past and go back to its original state.
- This is complicated to define as a function, but relatively simple as a diagram.
- This strategy begins at state $S1$, which plays C.

Punish for 3 periods



- If the other player plays D:
  - then this strategy transitions to state $S2$, which plays D once;
  - then to state $S3$, which plays D once;
  - then to state $S4$, which plays D once;
  - then transitions back to the original state $S1$.
- Thus, this strategy will punish D by playing D 3 times, after which it will play C again (even if the other player previously played D in the last period of punishment).

Tit-for-Tat



- **Tit-for-Tat**: This strategy has an intuitive interpretation: do whatever the other player did previously. We can define it as:

$$s_i(\varnothing) = C$$

$$s_i(a^1, \ldots, a^t) = \begin{cases} C & \text{if } a^t = C \\ D & \text{if } a^t = D \end{cases}$$

Finitely repeated Prisoner's Dilemma

- In a one-shot Prisoner's Dilemma, the NE is when both players *Defect*.
- Can a *finitely repeated* Prisoner's Dilemma sustain a different NE?
- Suppose Player 1's strategy is s_1 and Player 2's strategy is s_2.
- Let t denote the last period in which the outcome is *not* (D, D) (and therefore the outcome in all periods after t is (D, D)).
- Suppose that Player 1 chose C in this period (we could also assume it was Player 2).
- We claim that Player 1 can deviate and get a higher payoff.

Finitely repeated Prisoner's Dilemma

- Let s_1' be any strategy such that the strategy profile (s_1', s_2) results in exactly the same history as (s_1, s_2), except that Player 1 chooses D in period t.
- This must increase Player 1's payoff in period t, while his payoff in periods *after* t cannot be worse, since Player 1 is already playing D in every period after t (by assumption).
- Therefore, the *outcome* in every NE is that (D, D) is played in every period.
- The *strategies* chosen by each player may specify playing C in response to some history, but those histories will never actually occur.
- Outcomes and histories that do not occur in equilibrium are said to be *off the equilibrium path*.

Finitely repeated Prisoner's Dilemma

- What about SPNE? This is easier to prove: in the last subgame, the only NE is (D, D) regardless of the previous history. Going back one step, the only NE is (D, D), and so on, until we reach the beginning of the game. Therefore, the only SPNE is when both players' strategies is to play (D, D) after every history.
- Punishment cannot be sustained in the finitely repeated Prisoner's Dilemma because in the last period, there is no way to deter *Defect*.
- However, in an *infinitely* repeated game, there is *always* the possibility of punishment in the future.

NE of Repeated PD: Always D

- Suppose both players play *Always Defect*: they play D after any history.
- The sequence of outcomes will be $(D, D), (D, D), \ldots$
- The sequene of payoffs will be $(1, 1), (1, 1), \ldots$
- By our construction of the discounted average, this gives a discounted sum of 1 to both players.
- Is this a Nash equilibrium? Suppose Player 1 deviates to any strategy that does *not* result in the outcome sequence (D, D) in every period.
- In some period, Player 1 will get a payoff of 0 instead of 1.
- This must decrease Player 1's discounted sum, so there is no incentive to deviate.

NE of Repeated PD: Grim Trigger

- Recall the *Grim Trigger* strategy: Play C until the other player plays D, then punish by playing D forever.
- Suppose both players play *Grim Trigger*.
  - In the first period, both players C.
  - In the second period, no one has played D, so both players play C.
  - Same for period 3, 4, 5...
- The sequence of outcomes is: $(C, C), (C, C), \ldots$
- The sequence of payoffs for both players is $2, 2, \ldots$ with a discounted average of

$$(1-\delta)(2+\delta 2+\delta^2 2+\ldots) = (1-\delta)2\sum_{t=0}^{\infty}\delta^t = 2$$

NE of Repeated PD: Grim Trigger

- Now, suppose Player 2 deviates by playing some other strategy that actually results in a different sequence of outcomes.
- For the sequence of outcomes to be different, Player 2 must play D at least once.
- Then Player 1 will play D forever starting at $t+1$, Player 2's best response to this is to also play D forever starting at $t+1$.
- Player 2's sequence of payoffs starting at period t is $(3, 1, 1, 1...)$

$$(1-\delta)(3+\delta+\delta^2+\delta^3+...) = (1-\delta)(3+\frac{\delta}{1-\delta})$$

$$= 3(1-\delta)+\delta$$

- This deviation will give a higher payoff *Grim Trigger* (or any other strategy that results in an outcome where (C, C) is always played) if and only if:

$$3(1-\delta)+\delta > 2 \rightarrow \delta < \frac{1}{2}$$

NE of Repeated PD: Grim Trigger

$$3(1-\delta)+\delta > 2 \rightarrow \delta < \frac{1}{2}$$

- Therefore, if $\delta \geq \frac{1}{2}$, both players playing *Grim Trigger* is a Nash equilibrium.
- And in general, one player playing Grim Trigger and the other playing any strategy that results in (C, C) every period is a Nash equilibrium.
- Note what the condition on δ implies: if players are *patient* enough, i.e. they place a high enough value on future payoffs, then the threat of punishment is enough to deter *Defect* in the present.
- If players have a sufficiently low discount factor δ, then the short-term gain of playing D outweighs the long-term gain of avoiding punishment.

NE: Tit-for-Tat

- Recall $Tit - for - Tat$: play C at the beginning of the game, then play whatever the other player chose in the previous round.
- If both players play this strategy, the outcome will be (C, C) each period, with a payoff stream of $(2, 2, 2...)$.
- Suppose Player 2 plays another strategy that plays D at time t. Player 1 will therefore play D in $t + 1$.
- Player 2 can either:
  - revert to C at $t + 1$, in which case we are back in our original situation, or
  - play D in $t + 1$, which guarantees in Player 1 playing D again in $t + 2$.
- If one deviation in the original situation is optimal, then repeated deviation must also be optimal, since the game reverts back to the original situation after a single deviation.

NE: Tit-for-Tat

- In short, Player 2 can deviate in two ways:
  - Play a strategy that alternates between C and D. The outcomes will alternate between (C, D) and (D, C). This gives a payoff stream of $(3, 0, 3, 0, \ldots)$ with a discounted average of

$$(1-\delta)\frac{3}{1-\delta^2} = \frac{3}{1+\delta}$$

  - Play a strategy that plays D in each period. The outcomes will be (D, D) in each period starting from $t+1$. This gives a payoff stream of $(3, 1, 1, 1, \ldots)$ with a discounted average of

$$3(1-\delta) + \delta = 3 - 2\delta$$

NE: Tit-for-Tat

- Comparing the discounted averages, both players playing $Tit - for - Tat$ can be a NE when:

$$2 \geq \frac{3}{1+\delta} \text{ and } 2 \geq 3 - 2\delta$$

- which are both satisfied if $\delta \geq \frac{1}{2}$.

- So far, we have shown that some strategies, when played by both players, can be a Nash equilibrium.
- In general, there are an infinite number of possible strategies for repeated games, and therefore an infinite number of possible ways in which a player can deviate from a NE.
- This makes it difficult to prove whether any given pair of strategies is a NE.
- In contrast, when we consider subgame perfect NE, there are only a limited number of ways in which a player can deviate, which makes it much easier to show if a given pair of strategies is a SPNE.

Ch 14.9: Subgame Perfect Equilibria and the One-Deviation Property

- We know that Nash equilibria of an extensive form game may include threats that are not credible.
- This is particularly important in repeated games, since the threat of punishment is the only way to deter a player from playing *Defect*.
- If punishment is not credible, then there will be no reason to expect anything but (D, D).
- The concept of subgame perfect equilibrium dealt with this issue by requiring that strategies also be Nash equilibria in all subgames.
- However, this is not easy to check in an infinite game.
- The good news is that we can take advantage of a result that gives a much simpler condition that is easier to check.

The One-Deviation Property

- Suppose we have a strategy profile s. s satisfies the **one-deviation property** if no player can increase his payoff in any subgame through a *one-shot deviation*:
  - deviating from s in the *first* period of the subgame, then *reverting* back to his strategy in s for the rest of the game.
- In a *finite* horizon game, a strategy profile is a SPNE if and only if it satisfies the one-deviation property.
- In an *infinite* horizon game where the discount factor is less than 1, a strategy profile is a SPNE if and only if it satisfies the one-deviation property.
- This is also called the *one-shot deviation principle*.

The One-Deviation Property

- The basic idea is that any change in strategy can be broken down into a sequence of one-period changes.
- If there is no one-period change that can increase payoffs, then there is no change in strategy that can increase payoffs.
- A full proof is somewhat technically advanced; however, if you are interested, you can search for "one-shot deviation principle".
- For repeated games, this means that in order to check whether some strategy profile s is a SPNE, we must check all possible one-shot deviations for all players.
- for every possible history, compare the payoffs to adhering to s, versus deviating for one period, then reverting back to s.
- Let's look at some of the strategies we've seen, and check if it is a SPNE when both players play them.

Grim Trigger

- *Grim Trigger* is a strategy that punishes defection, so it seems like it should be possible to be part of a SPNE.
- We will show that it is *not* a SPNE when both players play Grim Trigger.
- Suppose the previous history ended with (C, D).
- We don't specify how this occurred; in fact, if both players adhere to Grim Trigger, then the outcome will be (C, C) in every period. However, both players always have the choice of playing D, so it is one of the possible subgames.

Grim Trigger

- First, calculate the payoffs from *not* deviating. If both players adhere to *Grim Trigger*:
  - outcome path will be: $(D, C), (D, D), (D, D), \ldots$ (Player 1 will start punishing in the first period, Player 2 will start in the second period)
  - payoffs will be: $(3, 0), (1, 1), (1, 1), \ldots$
  - Player 1's discounted average:

$$(1-\delta)(3+\delta+\delta^2+\ldots) = (1-\delta)(3+\frac{\delta}{1-\delta}) = 3-2\delta$$

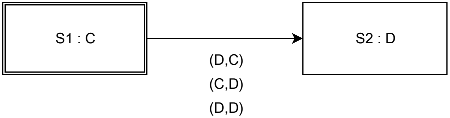
  - Player 2's discounted average:

$$(1-\delta)(0+\delta+\delta^2+\ldots) = \delta$$

Grim Trigger

- Now, suppose Player 2 deviates by choosing D in all periods (note that this is not a one-shot deviation).
  - outcome path: $(D, D), (D, D), (D, D), \ldots$
  - payoffs: $(1, 1), (1, 1), \ldots$
  - Both players' discounted average is 1.
- By our assumptions on $\delta, 1 > \delta$.
- Therefore, Player 2 has an incentive to deviate, given a history ending with (C, D), so this violates the requirement for a SPNE.

Modified Grim Trigger



- However, there is a modified version of Grim Trigger that is a SPNE when both players play it.
- *Grim Trigger* moves into the punishment state if the *other* player plays D.
- In contrast, *Modified Grim Trigger* moves into the punishment state if *either* player plays D.

Modified Grim Trigger

- Now, let's use the One-Deviation Property to show that it is a SPNE when both players play *Modified Grim Trigger*.
- We need to check all possible possible histories for the case where a one-shot deviation increases a player's payoff.
- This is vastly simplified because the strategy distinguishes between two types of histories:
  - D has never been played by either player (therefore, play C)
  - D has been played by either player (therefore, play D)
- For each of these types of histories, we will check the conditions that satisfy the one-deviation princple.

Case 1: D has not been played

- Case 1: Suppose D has never been played by either player.
- This is the case if the history is empty (i.e. at the beginning of the game), or if only (C, C) has been played in all periods.
- First, let's calculate the payoffs from *not* deviating.
  - outcome path: $(C, C), (C, C), \ldots$
  - payoffs: $(2, 2), (2, 2), \ldots$
  - Both players' discounted average is 2.

Case 1: D has not been played

- Now, consider the payoffs from a one-shot deviation. Suppose Player 1 does the one-shot deviation and plays D in the first period, then reverts back to *Modified Grim Trigger*.
  - outcome path: $(D, C), (D, D), (D, D), \ldots$
  - payoffs: $(3, 0), (1, 1), (1, 1), \ldots$
  - Player 1's discounted average is $3 - 2\delta$, same as shown above for Grim Trigger.
- Therefore, this one-shot deviation will give a higher payoff if $3 - 2\delta > 2$, or if $\delta < \frac{1}{2}$.
- The same is true if Player 2 does the one-shot deviation.

Case 2: D has been played

- Case 2: D has been played in the past, by either player.
- Then, *Modified Grim Trigger* will switch to the punishment state. If the players do *not* deviate:
  - outcome path: $(D, D), (D, D), ...$
  - payoffs: $(1, 1), (1, 1), ...$
  - Both players' discounted average is 1.

Case 2: D has been played

- Suppose Player 1 does a one-shot deviation by playing C.
  - outcome path: $(C, D), (D, D), (D, D), \ldots$
  - payoffs: $(0, 3), (1, 1), (1, 1), \ldots$
  - Player 1's discounted average is $(1 - \delta)(\delta + \delta^2 + \ldots) = \delta$
- This one-shot deviation will not give a higher payoff for any value of δ.
- The same is true if it is Player 2 that considers a one-shot deviation.

Modified Grim Trigger

- The requirements for the two cases we've seen are:
  - Case 1: D has never been played. A one-shot deviation is profitable if $\delta < \frac{1}{2}$.
  - Case 2: D has been played in the past, by either player. A one-shot deviation is never profitable.
- Combining the requirements for both cases, a profitable (i.e. higher payoff) one-shot deviation does not exist if $\delta \geq \frac{1}{2}$, and therefore the strategy profile is a SPNE.

Tit for Tat

- Recall: *Tit – for – Tat* plays C at the beginning of the game, then plays whatever the other player did in the previous period.
- Therefore, in terms of determining the outcome, we need only look at the previous round's actions. There are four possibilities:
    1. History ends in (C, D)
    2. History ends in (D, C)
    3. History ends in (C, C)
    4. History ends in (D, D)
- Let's examine each case in turn. We will find the conditions on δ that ensure a profitable one-shot deviation does not exist.
- We need to check both players' possible one-shot deviations.

Case 1: Previous outcome was (C, D)

- Suppose the history ends with (C, D).
- If both players do *not* deviate:
  - outcome path: $(D, C), (C, D), (D, C), \ldots$. (both players will alternate between C and D forever)
  - payoffs: $(3, 0), (0, 3), (3, 0), \ldots$
  - Player 1's discounted average:

$$(1-\delta)(3+0+3\delta^2+0+3\delta^4+\ldots) = (1-\delta)\frac{3}{1-\delta^2} = \frac{3}{1+\delta}$$

  - Player 2's discounted average:

$$(1-\delta)(0+3\delta+0+3\delta^3+0+\ldots) = (1-\delta)\frac{3\delta}{1-\delta^2} = \frac{3\delta}{1+\delta}$$

Case 1: Previous outcome was (C, D)

- Suppose Player 1 does a one-shot deviation by playing C in the first period.
  - outcome path: $(C, C), (C, C), \ldots$
  - payoffs: $(2, 2), (2, 2), \ldots$
  - Player 1's discounted average is 2.
- Therefore, this one-shot deviation will give a higher payoff if $2 > \frac{3}{1+\delta}$, or if $\delta > \frac{1}{2}$.
- Now, suppose Player 2 does a one-shot deviation by playing D in the first period:
  - outcome path: $(D, D), (D, D), \ldots$
  - payoffs: $(1, 1), (1, 1), \ldots$
  - Player 2's discounted average is 1.
- Therefore, this one-shot deviation will give a higher payoff if $1 > \frac{3\delta}{1+\delta}$, or if $\delta < \frac{1}{2}$.
- Combining the two conditions, a profitable one-shot deviation does not exist if $\delta = \frac{1}{2}$.

Case 2: Previous outcome was (D, C)

- Suppose the previous outcome was (D, C).
- This is the same situation as the previous case, with the players reversed.
- Therefore, the result is the same: a profitable one-shot deviation does not exist if $\delta = \frac{1}{2}$.

Case 3: Previous outcome was (C, C)

- Suppose the previous outcome was (C, C).
- If both players do *not* deviate:
  - outcome path: $(C, C), (C, C)...$
  - payoffs: $(2, 2), (2, 2), (2, 2), ...$
  - Both players' discounted average: 2

Case 3: Previous outcome was (C, C)

- Suppose Player 1 does a one-shot deviation by playing D in the first round.
  - outcome path: $(D, C), (C, D), (D, C), (C, D), \ldots$
  - payoffs: $(3, 0), (0, 3), (3, 0), \ldots$
  - Player 1's discounted average: $\frac{3}{1+\delta}$ (shown in the first type of history).
- The one-shot deviation will be profitable if $\frac{3}{1+\delta} > 2$, or if $\delta < \frac{1}{2}$.
- The same is true if it is Player 2 that is considering a one-shot deviation.
- Therefore, a profitable one-shot deviation does not exist if $\delta \geq \frac{1}{2}$.

Case 4: Previous outcome was (D, D)

- Suppose the previous outcome was (D, D).
- If both players do *not* deviate:
  - outcome path: $(D, D), (D, D)...$
  - payoffs: $(1, 1), (1, 1), (1, 1), ...$
  - Both players' discounted average: 1

Case 4: Previous outcome was (D, D)

- Suppose Player 1 does a one-shot deviation by playing C in the first round.
  - outcome path: $(C, D), (D, C), (C, D), \ldots$
  - payoffs: $(0, 3), (3, 0), (0, 3), (3, 0), \ldots$
  - Player 1's discounted average: $\frac{3\delta}{1+\delta}$ (see Case 1 for the formula for the discounted avg)
- The one-shot deviation will be profitable if $\frac{3\delta}{1+\delta} > 1$, or if $\delta > \frac{1}{2}$.
- The same is true if it is Player 2 that is considering a one-shot deviation.
- A profitable one-shot deviation will not exist if $\delta \leq \frac{1}{2}$.

Tit for Tat

- Summarizing the results for each case:
  - Case 1 (previous outcome was (C, D)): no profitable deviation exists if $\delta = \frac{1}{2}$.
  - Case 2 (previous outcome was (D, C)): no profitable deviation exists if $\delta = \frac{1}{2}$.
  - Case 3 (previous outcome was (C, C)): no profitable deviation exists if $\delta \geq \frac{1}{2}$.
  - Case 4 (previous outcome was (D, D)): no profitable deviation exists if $\delta \leq \frac{1}{2}$.
- Combining all the conditions, no profitable deviation exists if and only if $\delta = \frac{1}{2}$.
- This result depends on the particular payoffs of the game in each period.
- If we change the payoffs, the range of δ that supports a SPNE may change.

Tacit Collusion Among Firms

- One important application of repeated-game equilibria is the study of *tacit collusion* among firms.
- Usually, it is considered illegal for competitors to fix prices (i.e. agree not to compete on price, and keep prices high).
- If executives from rival firms are seen to meet and discuss prices, this is grounds for a lawsuit.
- However, even if they do not communicate, firms may collude, simply by observing the past history of prices.
- "Price wars" can be seen as times when firms seek to punish each other.

- Consider this Cournot duopoly problem with two firms.
- Each firm has cost of production $c_i(q_i) = 10q_i$.
- Market demand is given by $P = 100 - Q$.
- The two firms repeatedly play the Cournot duopoly game in time periods $t = 1, 2, 3, \dots$ with discount factor δ.
- The NE of the Cournot game in a single period is $q_1 = q_2 = 30$, $p = 40$, and profits for each firm are 900.
- If there was a single monopolist firm, the optimal $q = 45$, $p = 55$, profits $= 2025$.
- Suppose in each period, each firm can choose to *Collude*, in which case they produce $q_i = 22.5$, half the monopoly quantity.
- Or, they can choose to *Defect*, in which case they maximize their own profit, given the quantity of the other firm.

- If both firms choose *Defect*, they choose the NE quantities ($q_i = 30$), which results in a profit of 900 for each firm.
- If both firms choose *Collude*, they choose half the monopoly quantities ($q_i = 22.5$), which results in a profit of 1012.5 for each firm.
- If one chooses *Collude* while the other chooses *Defect*, the firm that *Collude*s chooses $q_i = 22.5$, while the firm that *Defect*s chooses the best response to that, which is $q_j = 33.75$.
- The firm that chose *Collude* gets a profit of 759.375, while the firm that chose *Defect* gets a profit of 1139.06.
- Note that this situation is a Prisoner's Dilemma.
- We can then use the one-shot deviation principle to show that a given strategy profile (e.g. both players play Modified Grim Trigger) is a SPNE, for a certain value of the discount factor δ.

- The one-shot deviation principle is not limited to repeated Prisoner's Dilemma; it can be used in any repeated game situation to prove that a given strategy profile is or is not a SPNE.

Announcements

- HW5 is due next week.
- The final exam will be on June 20, 4-6 PM, in Boxue 507.
- Final will be closed-book and covers the second half of the course. Old finals and solutions are on the course website.
- Bring a non-programmable calculator.