Lecture 19: Infinitely Repeated Games

Mauricio Romero

Introduction to Infinitely Repeated Games

Introduction to Infinitely Repeated Games

- One of the features of **finitely** repeated games was that if the stage game had a **unique** Nash equilibrium, then the only subgame perfect Nash equilibrium was the repetition of that unique stage game Nash equilibrium

- One of the features of **finitely** repeated games was that if the stage game had a **unique** Nash equilibrium, then the only subgame perfect Nash equilibrium was the repetition of that unique stage game Nash equilibrium

- This happened because there was a last period from which we could induct backwards (and there was a domino effect!)

- One of the features of **finitely** repeated games was that if the stage game had a **unique** Nash equilibrium, then the only subgame perfect Nash equilibrium was the repetition of that unique stage game Nash equilibrium

- This happened because there was a last period from which we could induct backwards (and there was a domino effect!)

- When the game is instead **infinitely** repeated, this argument no longer applies since there is no such thing as a last period

- Lets first define what an infinitely repeated game is

- Lets first define what an infinitely repeated game is
- We start with a stage game whose utilities are given by $u_1, u_2, \ldots, u_n$

- Lets first define what an infinitely repeated game is
- We start with a stage game whose utilities are given by $u_1, u_2, \ldots, u_n$
- Each player i has an action set A_i

- Lets first define what an infinitely repeated game is
- We start with a stage game whose utilities are given by $u_1, u_2, \ldots, u_n$
- Each player i has an action set A_i
- In each period $t = 0, 1, 2, \ldots$, players simultaneously choose an action $a_i \in A_i$ and the chosen action profile $(a_1, a_2, \ldots, a_n)$ is observed by all players

- Lets first define what an infinitely repeated game is
- We start with a stage game whose utilities are given by $u_1, u_2, \ldots, u_n$
- Each player i has an action set A_i
- In each period $t = 0, 1, 2, \ldots$, players simultaneously choose an action $a_i \in A_i$ and the chosen action profile $(a_1, a_2, \ldots, a_n)$ is observed by all players
- Then play moves to period $t + 1$ and the game continues in the same manner.

- It is impossible to draw the extensive form of this infinitely repeated game

- It is impossible to draw the extensive form of this infinitely repeated game

- Each information set of each player i associated with a finitely repeated game corresponded to a history of action profiles chosen in the past

- It is impossible to draw the extensive form of this infinitely repeated game

- Each information set of each player *i* associated with a finitely repeated game corresponded to a history of action profiles chosen in the past

- We can represent each information set of player *i* by a history:

$$h^0 = (\emptyset), h^1 = (a^0 := (a_1^0, \ldots, a_n^0)), \ldots, h^t = (a^0, a^1, \ldots, a^{t-1})$$

- It is impossible to draw the extensive form of this infinitely repeated game
- Each information set of each player i associated with a finitely repeated game corresponded to a history of action profiles chosen in the past
- We can represent each information set of player i by a history:

$$h^0 = (\emptyset), h^1 = (a^0 := (a_1^0, \ldots, a_n^0)), \ldots, h^t = (a^0, a^1, \ldots, a^{t-1})$$

- We denote the set of all histories at time t as H^t

Prisoner's Dilemma

	C_2	D_2
C_1	$1, 1$	$-1, 2$
D_1	$2, -1$	$0, 0$

- For example, if the stage game is the prisoner's dilemma, at period 1, there are 4 possible histories:

$$\{(C_1, C_2), (C_1, D_2), (D_1, C_2), (D_1, D_2)\} = H^1.$$

- For example, if the stage game is the prisoner's dilemma, at period 1, there are 4 possible histories:

$$\{(C_1, C_2), (C_1, D_2), (D_1, C_2), (D_1, D_2)\} = H^1.$$

- For time t, H^t consists of 4^t possible histories

- For example, if the stage game is the prisoner's dilemma, at period 1, there are 4 possible histories:

$$\{(C_1, C_2), (C_1, D_2), (D_1, C_2), (D_1, D_2)\} = H^1.$$

- For time t, H^t consists of 4^t possible histories

- This means that there is a **one-to-one** mapping between all possible histories and the information sets if we actually wrote out the whole extensive form game tree

- For example, if the stage game is the prisoner's dilemma, at period 1, there are 4 possible histories:

$$\{(C_1, C_2), (C_1, D_2), (D_1, C_2), (D_1, D_2)\} = H^1.$$

- For time t, H^t consists of 4^t possible histories

- This means that there is a **one-to-one** mapping between all possible histories and the information sets if we actually wrote out the whole extensive form game tree

- As a result, we can think of each $h^t \in H^t$ as representing a particular information set for each player i in each time t

- What is a strategy in an infinitely repeated game?

- What is a strategy in an infinitely repeated game?

- It is simply a prescription of what player i would do at every information set or history

- What is a strategy in an infinitely repeated game?

- It is simply a prescription of what player i would do at every information set or history

- Therefore, it is a function that describes:

$$s_i : \bigcup_{t \geq 0} H^t \rightarrow A_i.$$

- What is a strategy in an infinitely repeated game?

- It is simply a prescription of what player i would do at every information set or history

- Therefore, it is a function that describes:

$$s_i : \bigcup_{t \geq 0} H^t \to A_i.$$

- Intuitively, s_i describes exactly what player i would do at every possible history h^t, where $s_i(h^t)$ describes what player i would do at history h^t

- For example in the infinitely repeated prisoner's dilemma, the strategy $s_i(h^t) = C_i$ for all h^t and all t is the strategy in which player i always plays C_i regardless of the history

- For example in the infinitely repeated prisoner's dilemma, the strategy $s_i(h^t) = C_i$ for all h^t and all t is the strategy in which player i always plays C_i regardless of the history

- There can be more complicated strategies such as the following:

$$s_i(h^t) = \begin{cases} C_i & \text{if } t = 0 \text{ or } h^t = (C, C, \dots, C), \\ D_i & \text{otherwise.} \end{cases}$$

- For example in the infinitely repeated prisoner's dilemma, the strategy $s_i(h^t) = C_i$ for all h^t and all t is the strategy in which player i always plays C_i regardless of the history

- There can be more complicated strategies such as the following:

$$s_i(h^t) = \begin{cases} C_i & \text{if } t = 0 \text{ or } h^t = (C, C, \ldots, C), \\ D_i & \text{otherwise.} \end{cases}$$

- The above is called a **grim trigger strategy**

- How are payoffs determined in the repeated game?

- How are payoffs determined in the repeated game?

- Suppose the *strategies* $s_1, \ldots, s_n$ are played which lead to the infinite sequence of action profiles:

$$a^0, a^1, \ldots, a^t, a^{t+1}, \ldots.$$

- How are payoffs determined in the repeated game?

- Suppose the *strategies* $s_1, \ldots, s_n$ are played which lead to the infinite sequence of action profiles:

$$a^0, a^1, \ldots, a^t, a^{t+1}, \ldots.$$

- Then the payoff of player i in this repeated game is given by:

$$sum_{t=0}^{\infty} \delta^t u_i(a^t).$$

- How are payoffs determined in the repeated game?
- Suppose the *strategies* $s_1, \ldots, s_n$ are played which lead to the infinite sequence of action profiles:

$$a^0, a^1, \ldots, a^t, a^{t+1}, \ldots.$$

- Then the payoff of player i in this repeated game is given by:

$$sum_{t=0}^{\infty} \delta^t u_i(a^t).$$

- Intuitively, the contribution to payoff of time t action profile a^t is discounted by δ^t

- It may be unreasonable to think about an infinitely repeated game

- It may be unreasonable to think about an infinitely repeated game
- However the discount factor instead could be interpreted by the probability of the game/relationship ending at any point in time.

- It may be unreasonable to think about an infinitely repeated game
- However the discount factor instead could be interpreted by the probability of the game/relationship ending at any point in time.
- Thus, an infinitely repeated game does not necessarily represent a scenario in which there are an infinite number of periods, but rather a relationship which ends in finite time with probability one, but in which the time at which the relationship ends is uncertain

- Lets see some examples of how to compute payoffs in the repeated game

- Lets see some examples of how to compute payoffs in the repeated game
- Consider first the strategy profile in which $s_i(h^t) = C_i$ for all $i = 1, 2$ and all h^t.

- Lets see some examples of how to compute payoffs in the repeated game
- Consider first the strategy profile in which $s_i(h^t) = C_i$ for all $i = 1, 2$ and all h^t.
- In this case, the payoff of player 1 in this repeated game is given by:

$$\sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta}$$

- Lets see some examples of how to compute payoffs in the repeated game
- Consider first the strategy profile in which $s_i(h^t) = C_i$ for all $i = 1, 2$ and all h^t.
- In this case, the payoff of player 1 in this repeated game is given by:

$$\sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta}$$

- What about in the grim trigger strategy profile?

- Lets see some examples of how to compute payoffs in the repeated game
- Consider first the strategy profile in which $s_i(h^t) = C_i$ for all $i = 1, 2$ and all h^t.
- In this case, the payoff of player 1 in this repeated game is given by:

$$\sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta}$$

- What about in the grim trigger strategy profile?
- In that case, if all players play the grim trigger strategy profile, the sequence of actions that arise is again $(C, C, \ldots)$

- Lets see some examples of how to compute payoffs in the repeated game
- Consider first the strategy profile in which $s_i(h^t) = C_i$ for all $i = 1, 2$ and all h^t.
- In this case, the payoff of player 1 in this repeated game is given by:
$$\sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta}$$
- What about in the grim trigger strategy profile?
- In that case, if all players play the grim trigger strategy profile, the sequence of actions that arise is again $(C, C, \ldots)$
- Thus the payoffs of all players is again $\frac{1}{1-\delta}$.

- How about a more complicated strategy profile?

- How about a more complicated strategy profile?

- Suppose that $s_i(h^0) = (C_1, D_2)$ and the strategy profile says to do exactly what the opponent did in the previous period

- How about a more complicated strategy profile?
- Suppose that $s_i(h^0) = (C_1, D_2)$ and the strategy profile says to do exactly what the opponent did in the previous period
- Then if both players play these strategies, then the sequence of actions that arise is:

$$(C_1, D_2), (D_1, C_2), (C_1, D_2), \ldots$$

- How about a more complicated strategy profile?
- Suppose that $s_i(h^0) = (C_1, D_2)$ and the strategy profile says to do exactly what the opponent did in the previous period
- Then if both players play these strategies, then the sequence of actions that arise is:

$$(C_1, D_2), (D_1, C_2), (C_1, D_2), \dots$$

- Then the payoff to player 1 in this game is given by:

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

- Sometimes its more convenient to represent payoffs in terms of a flow payoff today and a continuation value from tomorrow on

- Sometimes its more convenient to represent payoffs in terms of a flow payoff today and a continuation value from tomorrow on
- At history $h^t = (a^0, \ldots, a^{t-1})$, if the players play according to the strategy profile s, then play unfolds in the following manner:

$$(a^0, a^1, \ldots, a^{t-1}, a^t(s_i, s_{-i} \mid h^t), a^{t+1}(s_i, s_{-i} \mid h^t) \ldots),$$

  where $a^\tau(s_i, s_{-i} \mid h^t)$ denotes the action profile that will be played at time $\tau \geq t$ if players indeed play the strategy profile (s_i, s_{-i}) after history h^t

- Sometimes its more convenient to represent payoffs in terms of a flow payoff today and a continuation value from tomorrow on
- At history $h^t = (a^0, \ldots, a^{t-1})$, if the players play according to the strategy profile s, then play unfolds in the following manner:

$$(a^0, a^1, \ldots, a^{t-1}, a^t(s_i, s_{-i} \mid h^t), a^{t+1}(s_i, s_{-i} \mid h^t) \ldots),$$

  where $a^\tau(s_i, s_{-i} \mid h^t)$ denotes the action profile that will be played at time $\tau \geq t$ if players indeed play the strategy profile (s_i, s_{-i}) after history h^t
- Then we can define the following payoff to the strategy profile (s_i, s_{-i}) conditional on the history h^t:

$$U_i(s_i, s_{-i} \mid h^t) = \sum_{\tau=0}^{t-1} \delta^\tau u_i(a^\tau) + \sum_{\tau=t}^{\infty} \delta^\tau u_i(a^\tau(s_i, s_{-i} \mid h^t)).$$

- Since the payoffs from time $0, 1, \ldots, t-1$ do not matter for the incentives after time t (since these payoffs are sunk), we delete them and normalize by dividing through by δ^t

- Since the payoffs from time $0, 1, \ldots, t-1$ do not matter for the incentives after time t (since these payoffs are sunk), we delete them and normalize by dividing through by δ^t
- We then obtain:

$$W_i(s_i, s_{-i} \mid h^t) := \sum_{\tau=t}^{\infty} \delta^{\tau-t} u_i(a^\tau(s_i, s_{-i} \mid h^t))$$

- Since the payoffs from time $0, 1, \ldots, t-1$ do not matter for the incentives after time t (since these payoffs are sunk), we delete them and normalize by dividing through by δ^t
- We then obtain:

$$W_i(s_i, s_{-i} \mid h^t) := \sum_{\tau=t}^{\infty} \delta^{\tau-t} u_i(a^\tau(s_i, s_{-i} \mid h^t))$$

- This is what is called the **continuation value** of the strategy profile (s_i, s_{-i}) at history h^t

- Since the payoffs from time $0, 1, \ldots, t-1$ do not matter for the incentives after time t (since these payoffs are sunk), we delete them and normalize by dividing through by δ^t
- We then obtain:

$$W_i(s_i, s_{-i} \mid h^t) := \sum_{\tau=t}^{\infty} \delta^{\tau-t} u_i(a^\tau(s_i, s_{-i} \mid h^t))$$

- This is what is called the **continuation value** of the strategy profile (s_i, s_{-i}) at history h^t
- The subgame following a time-t history h^t essentially is equivalent to another infinitely repeated game

- Since the payoffs from time $0, 1, \ldots, t-1$ do not matter for the incentives after time t (since these payoffs are sunk), we delete them and normalize by dividing through by δ^t
- We then obtain:

$$W_i(s_i, s_{-i} \mid h^t) := \sum_{\tau=t}^{\infty} \delta^{\tau-t} u_i(a^\tau(s_i, s_{-i} \mid h^t))$$

- This is what is called the **continuation value** of the strategy profile (s_i, s_{-i}) at history h^t
- The subgame following a time-t history h^t essentially is equivalent to another infinitely repeated game
- The value $W_i(s_i, s_{-i} \mid h^t)$ represents the value that i accrues in this subgame, following history h^t, when players play according to h^t, viewing payoffs from time t perspective (as if time t is time 0)

- We can represent the payoff $U_i(s_i, s_{-i} \mid h^t)$ using continuation values:

$$U_i(s_i, s_{-i} \mid h^t) = \sum_{\tau=0}^{t-1} \delta^\tau u_i(a^\tau) + \delta^t W_i(s_i, s_{-i} \mid h^t).$$

- We can represent the payoff $U_i(s_i, s_{-i} \mid h^t)$ using continuation values:

$$U_i(s_i, s_{-i} \mid h^t) = \sum_{\tau=0}^{t-1} \delta^\tau u_i(a^\tau) + \delta^t W_i(s_i, s_{-i} \mid h^t).$$

- $W_i(s_i, s_{-i} \mid h^t)$ can also be decomposed as follows:

$$W_i(s_i, s_{-i} \mid h^t) = u_i(s_i(h^t), s_{-i}(h^t)) + \delta W_i(s_i, s_{-i} \mid (h^t, s_i(h^t), s_{-i}(h^t$$

- What is a subgame perfect Nash equilibrium in an infinitely repeated game?

- What is a subgame perfect Nash equilibrium in an infinitely repeated game?
- It is exactly the same idea as in the finitely repeated game or more generally extensive form games

- What is a subgame perfect Nash equilibrium in an infinitely repeated game?

- It is exactly the same idea as in the finitely repeated game or more generally extensive form games

- That is a strategy profile $s = (s_1, \ldots, s_n)$ is a subgame perfect game Nash equilibrium if and only if s is a Nash equilibrium in every subgame of the repeated game.

- This is however still a bit mysterious, so how do we make it a bit more transparent?

- This is however still a bit mysterious, so how do we make it a bit more transparent?
- First notice that a particular subgame corresponds to an infinitely repeated game that starts after a certain history h^t

- This is however still a bit mysterious, so how do we make it a bit more transparent?
- First notice that a particular subgame corresponds to an infinitely repeated game that starts after a certain history h^t
- Furthermore the fact that s is a Nash equilibrium after the history means that after every history $h^t = (a^0, \ldots, a^{t-1})$, s_i is a best response against s_{-i} at such a history:

$$U_i(s_i, s_{-i} \mid h^t) = \max_{s_i'} U_i(s_i', s_{-i} \mid h^t).$$

- This is however still a bit mysterious, so how do we make it a bit more transparent?
- First notice that a particular subgame corresponds to an infinitely repeated game that starts after a certain history h^t
- Furthermore the fact that s is a Nash equilibrium after the history means that after every history $h^t = (a^0, \ldots, a^{t-1})$, s_i is a best response against s_{-i} at such a history:

$$U_i(s_i, s_{-i} \mid h^t) = \max_{s_i'} U_i(s_i', s_{-i} \mid h^t).$$

- Rewriting the above we get that for all s_i',

$$\sum_{\tau=0}^{t-1} \delta^\tau u_i(a^\tau) + \delta^t W_i(s_i, s_{-i} \mid h^t)$$
$$\geq \sum_{\tau=0}^{t-1} \delta^\tau u_i(a^\tau) + \delta^t W_i(s_i', s_{-i} \mid h^t)$$

- Can be simplified slightly by noticing that the payoffs from time 0 until $t-1$ do not matter in the above maximization problem

- Can be simplified slightly by noticing that the payoffs from time 0 until $t-1$ do not matter in the above maximization problem

- Furthermore we can divide all utilities by δ^t to realize that:

$$W_i(s_i, s_{-i} \mid h^t) = \max_{s_i'} W_i(s_i', s_{-i} \mid h^t).$$

- Can be simplified slightly by noticing that the payoffs from time 0 until $t-1$ do not matter in the above maximization problem
- Furthermore we can divide all utilities by δ^t to realize that:

$$W_i(s_i, s_{-i} \mid h^t) = \max_{s_i'} W_i(s_i', s_{-i} \mid h^t).$$

- The above is still a bit complicated since checking that a strategy s_i is a best response against s_{-i} may be quite difficult since there are infinitely many pure strategies s_i' that player i could potentially deviate to

- Can be simplified slightly by noticing that the payoffs from time 0 until $t-1$ do not matter in the above maximization problem

- Furthermore we can divide all utilities by δ^t to realize that:

$$W_i(s_i, s_{-i} \mid h^t) = \max_{s_i'} W_i(s_i', s_{-i} \mid h^t).$$

- The above is still a bit complicated since checking that a strategy s_i is a best response against s_{-i} may be quite difficult since there are infinitely many pure strategies s_i' that player i could potentially deviate to

- However, the following proposition makes the check quite simple

Theorem (One-stage deviation principle)

s is a subgame perfect Nash equilibrium (SPNE) if and only if for all times t, each history h^t, and each player i,

$$u_i(s_i(h^t), s_{-i}(h^t)) + \delta W_i(s_i, s_{-i} \mid (h^t, s_i(h^t), s_{-i}(h^t)))$$
$$= \max_{a_i' \in A_i} u_i(a_i', s_{-i}(h^t)) + \delta W_i(s_i, s_{-i} \mid (h^t, a_i', s_{-i}(h^t))).$$

- In words the above states that if s is a subgame perfect Nash equilibrium if and only if at every time t, and every history and every player i, player i cannot profit by deviating just at time t and following the strategy s_i' from time $t+1$ on

- In words the above states that if s is a subgame perfect Nash equilibrium if and only if at every time t, and every history and every player i, player i cannot profit by deviating just at time t and following the strategy s_i' from time $t+1$ on

- This is extremely useful since we only need to check that s_i is optimal against all possible one-stage deviations rather than having to check that it is optimal against all s_i'.

- In words the above states that if s is a subgame perfect Nash equilibrium if and only if at every time t, and every history and every player i, player i cannot profit by deviating just at time t and following the strategy s_i' from time $t+1$ on

- This is extremely useful since we only need to check that s_i is optimal against all possible one-stage deviations rather than having to check that it is optimal against all s_i'.

- We will now put this into practice to analyze subgame perfect Nash equilibria of infinitely repeated games

- Lets go back to the infinitely repeated prisoner's dilemma

- Lets go back to the infinitely repeated prisoner's dilemma
- What is an example of a subgame perfect Nash equilibrium?

- Lets go back to the infinitely repeated prisoner's dilemma
- What is an example of a subgame perfect Nash equilibrium?
- One kind of equilibrium should be straightforward: each player plays D_1 and D_2 always at all information sets

- Lets go back to the infinitely repeated prisoner's dilemma
- What is an example of a subgame perfect Nash equilibrium?
- One kind of equilibrium should be straightforward: each player plays D_1 and D_2 always at all information sets
- Why is this a SPNE?

- Lets go back to the infinitely repeated prisoner's dilemma
- What is an example of a subgame perfect Nash equilibrium?
- One kind of equilibrium should be straightforward: each player plays D_1 and D_2 always at all information sets
- Why is this a SPNE?
- We can use the one-stage deviation principle

Prisoner's Dilemma

	C_2	D_2
C_1	$1, 1$	$-1, 2$
D_1	$2, -1$	$0, 0$

- Under this strategy profile s_1^*, s_2^*, for all histories h^t,

$$W_1(s_1^*, s_2^* \mid h^t) = W_2(s_1^*, s_2^* \mid h^t) = 0.$$

- Under this strategy profile s_1^*, s_2^*, for all histories h^t,

$$W_1(s_1^*, s_2^* \mid h^t) = W_2(s_1^*, s_2^* \mid h^t) = 0.$$

- Thus, for all histories h^t,

$$\underbrace{u_i(D_i, D_{-i})}_{0} + \delta \underbrace{W_i(s_1^*, s_2^* \mid h^t)}_{0} > \underbrace{u_i(C_i, D_{-i})}_{-1} + \delta \underbrace{W_i(s_1^*, s_2^* \mid h^t)}_{0}$$

- Under this strategy profile s_1^*, s_2^*, for all histories h^t,

$$W_1(s_1^*, s_2^* \mid h^t) = W_2(s_1^*, s_2^* \mid h^t) = 0.$$

- Thus, for all histories h^t,

$$\underbrace{u_i(D_i, D_{-i})}_{0} + \delta \underbrace{W_i(s_1^*, s_2^* \mid h^t)}_{0} > \underbrace{u_i(C_i, D_{-i})}_{-1} + \delta \underbrace{W_i(s_1^*, s_2^* \mid h^t)}_{0}$$

- Thus, (s_1^*, s_2^*) is a SPNE

In fact this is not specific to the prisoner's dilemma as we show below:

Theorem

Let a^* be a Nash equilibrium of the stage game. Then the strategy profile s^* in which all players i play a_i^* at all information sets is a SPNE for any $\delta \in [0, 1)$.

Proof

- We use the one-stage deviation principle again

Proof

- We use the one-stage deviation principle again
- We need to show that for every t and all h^t, and all i,

$$\begin{aligned} & u_i(s_i^*(h^t), s_{-i}^*(h^t)) + \delta W_i(s_i^*, s_{-i}^* \mid (h^t, s_i^*(h^t), s_{-i}^*(h^t))) \\ &= \max_{a_i' \in A_i} u_i(a_i', s_{-i}^*(h^t)) + \delta W_i(s_i^*, s_{-i}^* \mid (h^t, a_i', s_{-i}^*(h^t))). \end{aligned}$$

Proof

- We use the one-stage deviation principle again
- We need to show that for every t and all h^t, and all i,

$$u_i(s_i^*(h^t), s_{-i}^*(h^t)) + \delta W_i(s_i^*, s_{-i}^* \mid (h^t, s_i^*(h^t), s_{-i}^*(h^t)))$$
$$= \max_{a_i' \in A_i} u_i(a_i', s_{-i}^*(h^t)) + \delta W_i(s_i^*, s_{-i}^* \mid (h^t, a_i', s_{-i}^*(h^t))).$$

- Note that

$$W_i(s_i^*, s_{-i}^* \mid (h^t, s_i(h^t), s_{-i}(h^t)))$$
$$= W_i(s_i, s_{-i} \mid (h^t, a_i', s_{-i}(h^t))) = u_i(a^*)$$

Proof

- We use the one-stage deviation principle again
- We need to show that for every t and all h^t, and all i,

$$u_i(s_i^*(h^t), s_{-i}^*(h^t)) + \delta W_i(s_i^*, s_{-i}^* \mid (h^t, s_i^*(h^t), s_{-i}^*(h^t)))$$
$$= \max_{a_i' \in A_i} u_i(a_i', s_{-i}^*(h^t)) + \delta W_i(s_i^*, s_{-i}^* \mid (h^t, a_i', s_{-i}^*(h^t))).$$

- Note that

$$W_i(s_i^*, s_{-i}^* \mid (h^t, s_i(h^t), s_{-i}(h^t)))$$
$$= W_i(s_i, s_{-i} \mid (h^t, a_i', s_{-i}(h^t))) = u_i(a^*)$$

- So we just need to show that

$$u_i(a_i^*, a_{-i}^*) = \max_{a_i' \in A_i} u_i(a_i', a_{-i}^*),$$

Proof

- We use the one-stage deviation principle again
- We need to show that for every t and all h^t, and all i,

$$u_i(s_i^*(h^t), s_{-i}^*(h^t)) + \delta W_i(s_i^*, s_{-i}^* \mid (h^t, s_i^*(h^t), s_{-i}^*(h^t)))$$
$$= \max_{a_i' \in A_i} u_i(a_i', s_{-i}^*(h^t)) + \delta W_i(s_i^*, s_{-i}^* \mid (h^t, a_i', s_{-i}^*(h^t))).$$

- Note that

$$W_i(s_i^*, s_{-i}^* \mid (h^t, s_i(h^t), s_{-i}(h^t)))$$
$$= W_i(s_i, s_{-i} \mid (h^t, a_i', s_{-i}(h^t))) = u_i(a^*)$$

- So we just need to show that

$$u_i(a_i^*, a_{-i}^*) = \max_{a_i' \in A_i} u_i(a_i', a_{-i}^*),$$

- This is true by assumption that a^* is a Nash equilibrium of the stage game

- What other kinds of SPNE are there?

- What other kinds of SPNE are there?

- In finitely repeated games, this was the only SPNE with prisoner's dilemma since the stage game had a unique Nash equilibrium

- What other kinds of SPNE are there?

- In finitely repeated games, this was the only SPNE with prisoner's dilemma since the stage game had a unique Nash equilibrium

- When the repeated game is infinitely repeated, this is no longer true

- Consider for example the grim trigger strategy profile that we discussed earlier. Each player plays the following strategy:

$$s_i^*(h^t) = \begin{cases} C_i & \text{if } h^t = (C, C, \ldots, C) \\ D_i & \text{if } h^t \neq (C, C, \ldots, C). \end{cases}$$

- Consider for example the grim trigger strategy profile that we discussed earlier. Each player plays the following strategy:

$$s_i^*(h^t) = \begin{cases} C_i & \text{if } h^t = (C, C, \ldots, C) \\ D_i & \text{if } h^t \neq (C, C, \ldots, C). \end{cases}$$

- We will show that if δ is sufficiently high, so that the players are sufficiently patient, the strategy profile of grim trigger strategies is indeed a SPNE

- Consider for example the grim trigger strategy profile that we discussed earlier. Each player plays the following strategy:

$$s_i^*(h^t) = \begin{cases} C_i & \text{if } h^t = (C, C, \ldots, C) \\ D_i & \text{if } h^t \neq (C, C, \ldots, C). \end{cases}$$

- We will show that if δ is sufficiently high, so that the players are sufficiently patient, the strategy profile of grim trigger strategies is indeed a SPNE

- The *equilibrium path of play* for this SPNE is for players to play C in every period

- How do we show that the above is indeed an SPNE?

- How do we show that the above is indeed an SPNE?

- We use the one-stage deviation principle again

- How do we show that the above is indeed an SPNE?

- We use the one-stage deviation principle again

- We need to check the one-stage deviation principle at every history h^t.

Case 1:

- Suppose first that $h^t \neq (C, C, \ldots, C)$

Case 1:

- Suppose first that $h^t \neq (C, C, \ldots, C)$
- Players are each suppose to play D_i

Case 1:

- Suppose first that $h^t \neq (C, C, \ldots, C)$
- Players are each suppose to play D_i
- Thus, we need to check that

$$u_i(D_i, D_{-i}) + \delta W_i(s^* \mid (h^t, D)) \\ \geq u_i(C_i, D_{-i}) + \delta W_i(s^* \mid (h^t, (C_i, D_{-i})))$$

Case 1:

- Suppose first that $h^t \neq (C, C, \ldots, C)$
- Players are each suppose to play D_i
- Thus, we need to check that

$$\begin{aligned} &u_i(D_i, D_{-i}) + \delta W_i(s^* \mid (h^t, D)) \\ &\geq u_i(C_i, D_{-i}) + \delta W_i(s^* \mid (h^t, (C_i, D_{-i}))) \end{aligned}$$

- But since $h^t \neq (C, C, \ldots, C)$,
$W_i(s^* \mid (h^t, D)) = W_i(s^* \mid (h^t, (C_i, D_{-i}))) = u_i(D_i, D_{-i})$.

Case 1:

- Suppose first that $h^t \neq (C, C, \ldots, C)$
- Players are each suppose to play D_i
- Thus, we need to check that

$$u_i(D_i, D_{-i}) + \delta W_i(s^* \mid (h^t, D)) \geq u_i(C_i, D_{-i}) + \delta W_i(s^* \mid (h^t, (C_i, D_{-i})))$$

- But since $h^t \neq (C, C, \ldots, C)$,
  $W_i(s^* \mid (h^t, D)) = W_i(s^* \mid (h^t, (C_i, D_{-i}))) = u_i(D_i, D_{-i})$.
- So the above inequality is satisfied if and only if

$$u_i(D_i, D_{-i}) \geq u_i(C_i, D_{-i}).$$

Case 1:

- Suppose first that $h^t \neq (C, C, \ldots, C)$
- Players are each suppose to play D_i
- Thus, we need to check that

$$u_i(D_i, D_{-i}) + \delta W_i(s^* \mid (h^t, D))$$
$$\geq u_i(C_i, D_{-i}) + \delta W_i(s^* \mid (h^t, (C_i, D_{-i})))$$

- But since $h^t \neq (C, C, \ldots, C)$,
$W_i(s^* \mid (h^t, D)) = W_i(s^* \mid (h^t, (C_i, D_{-i}))) = u_i(D_i, D_{-i})$.
- So the above inequality is satisfied if and only if

$$u_i(D_i, D_{-i}) \geq u_i(C_i, D_{-i}).$$

- But this is satisfied since D is a Nash equilibrium of the stage game

Case 2:

- Suppose instead that $h^t = (C, C, \ldots, C)$

Case 2:

- Suppose instead that $h^t = (C, C, \ldots, C)$
- Players are both supposed to play C_i

Case 2:

- Suppose instead that $h^t = (C, C, \ldots, C)$
- Players are both supposed to play C_i
- Thus, we need to check that

$$\begin{aligned} & u_i(C_i, C_{-i}) + \delta W_i(s^* \mid (h^t, C)) \\ & \geq u_i(D_i, C_{-i}) + \delta W_i(s^* \mid (h^t, (D_i, C_{-i}))). \end{aligned}$$

Case 2:

- Suppose instead that $h^t = (C, C, \ldots, C)$
- Players are both supposed to play C_i
- Thus, we need to check that

$$u_i(C_i, C_{-i}) + \delta W_i(s^* \mid (h^t, C)) \\ \geq u_i(D_i, C_{-i}) + \delta W_i(s^* \mid (h^t, (D_i, C_{-i}))).$$

- In this case,

$$W_i(s^* \mid (h^t, C)) = u_i(C_i, C_{-i}) \\ = 1, W_i(s^* \mid (h^t, (D_i, C_{-i}))) = u_i(D) = 0.$$

Case 2:

- Suppose instead that $h^t = (C, C, \dots, C)$
- Players are both supposed to play C_i
- Thus, we need to check that

$$u_i(C_i, C_{-i}) + \delta W_i(s^* \mid (h^t, C)) \\ \geq u_i(D_i, C_{-i}) + \delta W_i(s^* \mid (h^t, (D_i, C_{-i}))).$$

- In this case,

$$W_i(s^* \mid (h^t, C)) = u_i(C_i, C_{-i}) \\ = 1, W_i(s^* \mid (h^t, (D_i, C_{-i}))) = u_i(D) = 0.$$

- Therefore, the above is satisfied if and only if

$$1 + \delta \geq 2 \iff \delta \geq 1/2.$$

Case 2:

- Suppose instead that $h^t = (C, C, \dots, C)$
- Players are both supposed to play C_i
- Thus, we need to check that

$$u_i(C_i, C_{-i}) + \delta W_i(s^* \mid (h^t, C)) \\ \geq u_i(D_i, C_{-i}) + \delta W_i(s^* \mid (h^t, (D_i, C_{-i}))).$$

- In this case,

$$W_i(s^* \mid (h^t, C)) = u_i(C_i, C_{-i}) \\ = 1, W_i(s^* \mid (h^t, (D_i, C_{-i}))) = u_i(D) = 0.$$

- Therefore, the above is satisfied if and only if

$$1 + \delta \geq 2 \iff \delta \geq 1/2.$$

- Thus the grim trigger strategy profile s^* is a SPNE if and only if $\delta \geq 1/2$.

- The above findings that SPNE may involve the repetition of action profile that is not a stage game NE is not specific to just the infinitely repeated prisoner's dilemma as the following theorem demonstrates.

Theorem (Folk theorem)

Suppose that a^* is a Nash equilibrium of the stage game. Suppose that $\hat{a}$ is an action profile of the Nash equilibrium such that

$$u_1(\hat{a}) > u_1(a^*), \ldots, u_n(\hat{a}) > u_n(a^*).$$

Then there is some $\delta^* < 1$ such that whenever $\delta > \delta^*$, there is a SPNE in which on the equilibrium path of play, all players play $\hat{a}$ in every period.

Proof

- The proof is essentially the same argument as that in verification of the grim trigger strategy of the prisoner's dilemma

Proof

- The proof is essentially the same argument as that in verification of the grim trigger strategy of the prisoner's dilemma
- Consider the following strategy profile:

$$s_i(h^t) = \begin{cases} \hat{a}_i & \text{if } t = 0 \text{ or } h^t = (\hat{a}, \hat{a}, \ldots, \hat{a}) \\ a_i^* & \text{if } t > 0 \text{ and } h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a}). \end{cases}$$

Proof

- The proof is essentially the same argument as that in verification of the grim trigger strategy of the prisoner's dilemma
- Consider the following strategy profile:

$$s_i(h^t) = \begin{cases} \hat{a}_i & \text{if } t = 0 \text{ or } h^t = (\hat{a}, \hat{a}, \ldots, \hat{a}) \\ a_i^* & \text{if } t > 0 \text{ and } h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a}). \end{cases}$$

- When is the above a SPNE? Let $M_i = \max_{a_i' \in A_i} u_i(a_i', \hat{a}_{-i})$

Proof

- The proof is essentially the same argument as that in verification of the grim trigger strategy of the prisoner's dilemma

- Consider the following strategy profile:

$$s_i(h^t) = \begin{cases} \hat{a}_i & \text{if } t = 0 \text{ or } h^t = (\hat{a}, \hat{a}, \ldots, \hat{a}) \\ a_i^* & \text{if } t > 0 \text{ and } h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a}). \end{cases}$$

- When is the above a SPNE? Let $M_i = \max_{a_i' \in A_i} u_i(a_i', \hat{a}_{-i})$

- We again have two cases

Proof

- **Case 1:** $t > 0$ and $h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a})$

Proof

- **Case 1:** $t > 0$ and $h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play a^* forever

Proof

- **Case 1:** $t > 0$ and $h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play a^* forever
- To check that this is optimal, we need to check that for all a_i',

$$u_i(a^*)+\delta W_i(s \mid (h^t, a^*)) \geq u_i(a_i', a_{-i}^*)+\delta W_i(s \mid (h^t, (a_i', a_{-i}^*))).$$

Proof

- **Case 1:** $t > 0$ and $h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play a^* forever
- To check that this is optimal, we need to check that for all a_i',

$$u_i(a^*) + \delta W_i(s \mid (h^t, a^*)) \geq u_i(a_i', a_{-i}^*) + \delta W_i(s \mid (h^t, (a_i', a_{-i}^*))).$$

- But since from time $t + 1$ on, all players a^* forever afterwards regardless of what happens in period t

Proof

- **Case 1:** $t > 0$ and $h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play a^* forever
- To check that this is optimal, we need to check that for all a_i',

$$u_i(a^*) + \delta W_i(s \mid (h^t, a^*)) \geq u_i(a_i', a_{-i}^*) + \delta W_i(s \mid (h^t, (a_i', a_{-i}^*))).$$

- But since from time $t+1$ on, all players a^* forever afterwards regardless of what happens in period t
- $W_i(s \mid (h^t, a^*)) = W_i(s \mid (h^t, (a_i', a_{-i}^*)) = u_i(a^*)$

Proof

- **Case 1:** $t > 0$ and $h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play a^* forever
- To check that this is optimal, we need to check that for all a_i',

$$u_i(a^*) + \delta W_i(s \mid (h^t, a^*)) \geq u_i(a_i', a_{-i}^*) + \delta W_i(s \mid (h^t, (a_i', a_{-i}^*))).$$

- But since from time $t+1$ on, all players a^* forever afterwards regardless of what happens in period t
- $W_i(s \mid (h^t, a^*)) = W_i(s \mid (h^t, (a_i', a_{-i}^*)) = u_i(a^*)$
- Thus, the above holds if and only if for all a_i',

$$u_i(a^*) \geq u_i(a_i', a_{-i}^*).$$

Proof

- **Case 1:** $t > 0$ and $h^t \neq (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play a^* forever
- To check that this is optimal, we need to check that for all a_i',

$$u_i(a^*)+\delta W_i(s \mid (h^t, a^*)) \geq u_i(a_i', a_{-i}^*)+\delta W_i(s \mid (h^t, (a_i', a_{-i}^*))).$$

- But since from time $t + 1$ on, all players a^* forever afterwards regardless of what happens in period t
- $W_i(s \mid (h^t, a^*)) = W_i(s \mid (h^t, (a_i', a_{-i}^*)) = u_i(a^*)$
- Thus, the above holds if and only if for all a_i',

$$u_i(a^*) \geq u_i(a_i', a_{-i}^*).$$

- This satisfied by the assumption that a^* is a stage game NE.

Proof

- **Case 2:** $t = 0$ or $h^t = (\hat{a}, \hat{a}, \ldots, \hat{a})$

Proof

- **Case 2:** $t = 0$ or $h^t = (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play $\hat{a}$ in period t

Proof

- **Case 2:** $t = 0$ or $h^t = (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play $\hat{a}$ in period t
- To check that this is optimal,

$$u_i(\hat{a}) + \delta W_i(s \mid (h^t, \hat{a})) \geq u_i(a_i', \hat{a}_{-i}) + \delta W_i(s \mid (h^t, (a_i', \hat{a}_{-i}))).$$

Proof

- **Case 2:** $t = 0$ or $h^t = (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play $\hat{a}$ in period t
- To check that this is optimal,

$$u_i(\hat{a}) + \delta W_i(s \mid (h^t, \hat{a})) \geq u_i(a_i', \hat{a}_{-i}) + \delta W_i(s \mid (h^t, (a_i', \hat{a}_{-i}))).$$

- But $W_i(s \mid (h^t, \hat{a})) = u_i(\hat{a})$ and $W_i(s \mid (h^t, a_i', \hat{a}_{-i})) = u_i(a^*)$

Proof

- **Case 2:** $t = 0$ or $h^t = (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play $\hat{a}$ in period t
- To check that this is optimal,

$$u_i(\hat{a}) + \delta W_i(s \mid (h^t, \hat{a})) \geq u_i(a_i', \hat{a}_{-i}) + \delta W_i(s \mid (h^t, (a_i', \hat{a}_{-i}))).$$

- But $W_i(s \mid (h^t, \hat{a})) = u_i(\hat{a})$ and $W_i(s \mid (h^t, a_i', \hat{a}_{-i})) = u_i(a^*)$
- Thus the above is satisfied if and only if for all a_i',

$$u_i(\hat{a}) \geq u_i(a_i', \hat{a}_{-i}) + \delta u_i(a^*).$$

Proof

- **Case 2:** $t = 0$ or $h^t = (\hat{a}, \hat{a}, \dots, \hat{a})$
- Then all players are supposed to play $\hat{a}$ in period t
- To check that this is optimal,

$$u_i(\hat{a}) + \delta W_i(s \mid (h^t, \hat{a})) \geq u_i(a_i', \hat{a}_{-i}) + \delta W_i(s \mid (h^t, (a_i', \hat{a}_{-i}))).$$

- But $W_i(s \mid (h^t, \hat{a})) = u_i(\hat{a})$ and $W_i(s \mid (h^t, a_i', \hat{a}_{-i})) = u_i(a^*)$
- Thus the above is satisfied if and only if for all a_i',

$$u_i(\hat{a}) \geq u_i(a_i', \hat{a}_{-i}) + \delta u_i(a^*).$$

- The above is satisfied if

$$u_i(\hat{a}) \geq M_i + \delta u_i(a^*) \iff \delta \geq \frac{M_i - u_i(\hat{a})}{(u_i(\hat{a}) - u_i(a^*)) + (M_i - u_i(\hat{a}))}.$$

Proof

- **Case 2:** $t = 0$ or $h^t = (\hat{a}, \hat{a}, \ldots, \hat{a})$
- Then all players are supposed to play $\hat{a}$ in period t
- To check that this is optimal,

$$u_i(\hat{a}) + \delta W_i(s \mid (h^t, \hat{a})) \geq u_i(a_i', \hat{a}_{-i}) + \delta W_i(s \mid (h^t, (a_i', \hat{a}_{-i}))).$$

- But $W_i(s \mid (h^t, \hat{a})) = u_i(\hat{a})$ and $W_i(s \mid (h^t, a_i', \hat{a}_{-i})) = u_i(a^*)$
- Thus the above is satisfied if and only if for all a_i',

$$u_i(\hat{a}) \geq u_i(a_i', \hat{a}_{-i}) + \delta u_i(a^*).$$

- The above is satisfied if

$$u_i(\hat{a}) \geq M_i + \delta u_i(a^*) \iff \delta \geq \frac{M_i - u_i(\hat{a})}{(u_i(\hat{a}) - u_i(a^*)) + (M_i - u_i(\hat{a}))}.$$

- Setting $\delta^* = \max_{i=1}^{n} \frac{M_i - u_i(\hat{a})}{(u_i(\hat{a}) - u_i(a^*)) + (M_i - u_i(\hat{a}))}$ concludes the proof

- The above have the simple feature that once a deviation occurs, players punish each other forever by playing the stage game Nash equilibrium forever after

- The above have the simple feature that once a deviation occurs, players punish each other forever by playing the stage game Nash equilibrium forever after
- There are many other SPNE than those that we have just discussed

- The above have the simple feature that once a deviation occurs, players punish each other forever by playing the stage game Nash equilibrium forever after

- There are many other SPNE than those that we have just discussed

- In fact as δ becomes large, the number of SPNE explodes to infinity