6.207/14.15: Networks
Lecture 15: Repeated Games and Cooperation

Daron Acemoglu and Asu Ozdaglar
MIT

November 2, 2009

Outline

- The problem of cooperation
- Finitely-repeated prisoner's dilemma
- Infinitely-repeated games and cooperation
- Folk theorems
- Cooperation in finitely-repeated games
- Social preferences

- Reading:
- Osborne, Chapters 14 and 15.

Prisoners' Dilemma

- How to sustain cooperation in the society?
- Recall the **prisoners' dilemma**, which is the canonical game for understanding incentives for defecting instead of operating.

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- Recall that the strategy profile (D, D) is the unique NE. In fact, D strictly dominates C and thus (D, D) is the dominant equilibrium.
- In society, we have many situations of this form, but we often observe some amount of cooperation.
- Why?

Repeated Games

- In many strategic situations, players interact repeatedly over time.
- Perhaps repetition of the same game might foster cooperation.
- By repeated games we refer to a situation in which the same stage game (strategic form game) is played at each date for some duration of T periods.
- Such games are also sometimes called "supergames".
- Key new concept: **discounting**.
- We will imagine that future payoffs are discounted and are thus less valuable (e.g., money and the future is less valuable than money now because of positive interest rates; consumption in the future is less valuable than consumption now because of *time preference*).

Discounting

- We will model time preferences by assuming that future payoffs are discounted proportionately ("*exponentially*") at some rate $\delta \in [0, 1)$, called the **discount rate**.
- For example, in a two-period game with stage payoffs given by u^1 and u^2, overall payoffs will be

$$U = u^1 + \delta u^2.$$

- With the interest rate interpretation, we would have

$$\delta = \frac{1}{1+r},$$

where r is the interest rate.

Mathematical Model

- More formally, imagine that I players are playing a strategic form game $G = \langle \mathcal{I}, (A_i)_{i \in \mathcal{I}}, (u_i)_{i \in \mathcal{I}} \rangle$ for T periods. At each period, the outcomes of all past periods are observed by all players.
- Let us start with the case in which T is finite, but we will be particularly interested in the case in which $T = \infty$.
- Here A_i denotes the set of actions at each stage, and

$$u_i : \ A \to \mathbb{R},$$

  where $A = A_1 \times \cdots \times A_I$.
- That is, $u_i \left(a_i^t, a_{-i}^t \right)$ is the state payoff to player i when action profile $a^t = \left(a_i^t, a_{-i}^t \right)$ is played.

Mathematical Model (continued)

- We use the notation $\mathbf{a} = \{a^t\}_{t=0}^{T}$ to denote the sequence of action profiles. We could also define $\boldsymbol{\sigma} = \{\sigma^t\}_{t=0}^{T}$ to be the profile of mixed strategies.
- The payoff to player i in the repeated game

$$U(\mathbf{a}) = \sum_{t=0}^{T} \delta^t u_i(a_i^t, a_{-i}^t)$$

  where $\delta \in [0, 1)$.
- We denote the T-period repeated game with discount factor δ by $G^T(\delta)$.

Finitely-Repeated Prisoners' Dilemma

- Recall

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- What happens if this game was played $T < \infty$ times?
- We first need to decide what the equilibrium notion is. Natural choice, **subgame perfect Nash equilibrium (SPE).**
- Recall: SPE $\Longleftrightarrow$ backward induction.
- Therefore, start in the last period, at time T. What will happen?

Finitely-Repeated Prisoners' Dilemma (continued)

- In the last period, "defect" is a dominant strategy regardless of the history of the game. So the subgame starting at T has a dominant strategy equilibrium: (D, D).
- Then move to stage $T - 1$. By backward induction, we know that at T, no matter what, the play will be (D, D). Then given this, the subgame starting at $T - 1$ (again regardless of history) also has a dominant strategy equilibrium.
- With this argument, we have that there exists a unique SPE: (D, D) at each date.
- In fact, this is a special case of a more general result.

Equilibria of Finitely-Repeated Games

Theorem

Consider repeated game $G^T(\delta)$ for $T < \infty$. Suppose that the stage game G has a unique pure strategy equilibrium a^*. Then G^T has a unique SPE. In this unique SPE, $a^t = a^*$ for each $t = 0, 1, ..., T$ regardless of history.

Proof: The proof has exactly the same logic as the prisoners' dilemma example. By backward induction, at date T, we will have that (regardless of history) $a^T = a^*$. Given this, then we have $a^{T-1} = a^*$, and continuing inductively, $a^t = a^*$ for each $t = 0, 1, ..., T$ regardless of history.

Infinitely-Repeated Games

- Now consider the **infinitely-repeated game** G^∞.
- The notation $\mathbf{a} = \{a^t\}_{t=0}^\infty$ now denotes the (infinite) sequence of action profiles.
- The payoff to player i is then

$$U(\mathbf{a}) = (1-\delta)\sum_{t=0}^{\infty} \delta^t u_i(a_i^t, a_{-i}^t)$$

  where, again, $\delta \in [0,1)$.
- Note: this summation is well defined because $\delta < 1$.
- The term in front is introduced as a normalization, so that utility remains bounded even when $\delta \to 1$.

Trigger Strategies

- In infinitely-repeated games we can consider trigger strategies.
- A trigger strategy essentially threatens other players with a "worse," *punishment*, action if they deviate from an implicitly agreed action profile.
- A **non-forgiving trigger strategy** (or *grim trigger strategy*) s would involve this punishment *forever* after a single deviation.
- A non-forgiving trigger strategy (for player i) takes the following form:

$$a_i^t = \begin{cases} \bar{a}_i & \text{if } a^\tau = \bar{a} \text{ for all } \tau < t \\ \underline{a}_i & \text{if } a^\tau \neq \bar{a} \text{ for some } \tau < t \end{cases}$$

- Here if $\bar{a}$ is the implicitly agreed action profile and $\underline{a}_i$ is the punishment action.
- This strategy is non-forgiving since a single deviation from $\bar{a}$ induces player i to switch to $\underline{a}_i$ forever.

Cooperation with Trigger Strategies in the Repeated Prisoners' Dilemma

- Recall

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- Suppose both players use the following non-forgiving trigger strategy s^*:
  - Play C in every period unless someone has ever played D in the past
  - Play D forever if someone has played D in the past.
- We next show that the preceding strategy is an SPE if $\delta \geq 1/2$.

Cooperation with Trigger Strategies in the Repeated Prisoners' Dilemma

- Step 1: cooperation is best response to cooperation.
  - Suppose that there has so far been no D. Then given s^* being played by the other player, the payoffs to cooperation and defection are:

$$\text{Payoff from } C : (1-\delta)[1+\delta+\delta^2+\cdots] = (1-\delta)\times\frac{1}{1-\delta} = 1$$
$$\text{Payoff from } D : (1-\delta)[2+0+0+\cdots] = 2(1-\delta)$$

  - Cooperation better if $2(1-\delta) \geq 1$.
  - This shows that for $\delta \geq 1/2$, deviation to defection is not profitable.

Cooperation with Trigger Strategies in the Repeated Prisoners' Dilemma (continued)

- Step 2: defection is best response to defection.
  - Suppose that there has been some D in the past, then according to s^*, the other player will always play D. Against this, D is a best response.
- This argument is true in every subgame, so s^* is a subgame perfect equilibrium.
- **Note:** cooperating in every period would be a best response for a player against s^*. But unless that player herself also plays s^*, her opponent would not cooperate. Thus SPE requires both players to use s^*.

Multiplicity of Equilibria

- Cooperation is an equilibrium, but so are many other strategy profiles.
- Multiplicity of equilibria endemic in repeated games.
- Note that this multiplicity only occurs at $T = \infty$.
- In particular, for any finite T (and thus by implication for $T \to \infty$), prisoners' dilemma has a unique SPE.
- Why? The set of Nash equilibria is an upper hemi-continuous correspondence in parameters. It is not necessarily lower hemi-continuous.

Repetition Can Lead to Bad Outcomes

- The following example shows that repeated play can lead to *worse* outcomes than in the one shot game:

	A	B	C
A	$2, 2$	$2, 1$	$0, 0$
B	$1, 2$	$1, 1$	$-1, 0$
C	$0, 0$	$0, -1$	$-1, -1$

- For the game defined above, the action A strictly dominates both B and C for both players; therefore the unique Nash equilibrium of the stage game is (A, A).
- If $\delta \geq 1/2$, this game has an SPE in which (B, B) is played in every period.
- It is supported by the trigger strategy: Play B in every period unless someone deviates, and play C if there is any deviation.
- It can be verified that for $\delta \geq 1/2$, (B, B) is an SPE.

Folk Theorems

- In fact, it has long been a "folk theorem" that one can support cooperation in repeated prisoners' dilemma, and other "non-one-stage"equilibrium outcomes in infinitely-repeated games with sufficiently high discount factors.
- These results are referred to as "folk theorems" since they were believe to be true before they were formally proved.
- Here we will see a relatively strong version of these folk theorems.

Feasible Payoffs

- Consider stage game $G = \langle \mathcal{I}, (A_i)_{i \in \mathcal{I}}, (u_i)_{i \in \mathcal{I}} \rangle$ and infinitely-repeated game $G^\infty(\delta)$.
- Let us introduce the *Set of feasible payoffs*:

$$V = \text{ Conv}\{v \in \mathbb{R}^I \mid \text{there exists } a \in A \text{ such that } u(a) = v\}.$$

- That is, V is the convex hull of all I- dimensional vectors that can be obtained by some action profile. Convexity here is obtained by *public randomization*.
- **Note:** V is not equal to $\{v \in \mathbb{R}^I \mid \text{there exists } \sigma \in \Sigma \text{ such that } u(\sigma) = v\}$, where Σ is the set of mixed strategy profiles in the stage game.

Minmax Payoffs

- *Minmax payoff of player i*: the lowest payoff that player i's opponent can hold him to:

$$\underline{v}_i = \min_{a_{-i}} \left[\max_{a_i} u_i(a_i, a_{-i}) \right]$$
$$= \max_{a_i} \left[\min_{a_{-i}} u_i(a_i, a_{-i}) \right].$$

- The player can never receive less than this amount.
- Minmax strategy profile against i:

$$m^i_{-i} = \arg\min_{a_{-i}} \left[\max_{a_i} u_i(a_i, a_{-i}) \right]$$

Example

- Consider

	L	R
U	$-2, -2$	$1, -2$
M	$1, -1$	$-2, 2$
D	$0, 1$	$0, 1$

- To compute $\underline{v}_1$, let q denote the probability that player 2 chooses action L.
- Then player 1's payoffs for playing different actions are given by:

$$U \to 1 - 3q$$
$$M \to -2 + 3q$$
$$D \to 0$$

Example

- Therefore, we have

$$\underline{v}_1 = \min_{0 \le q \le 1} \left[\max\{1 - 3q, -2 + 3q, 0\} \right] = 0,$$

  and $m_2^1 \in [\frac{1}{3}, \frac{2}{3}]$.
- Similarly, one can show that: $\underline{v}_2 = 0$, and $m_1^2 = (1/2, 1/2, 0)$ is the unique minimax profile.

Minmax Payoff Lower Bounds

Theorem

- *Let σ be a (possibly mixed) Nash equilibrium of G and $u_i(\sigma)$ be the payoff to player i in equilibrium σ. Then*

$$u_i(\sigma) \geq \underline{v}_i.$$

- *Let $\boldsymbol{\sigma}$ be a (possibly mixed) Nash equilibrium of $G^\infty(\delta)$ and $U_i(\boldsymbol{\sigma})$ be the payoff to player i in equilibrium σ. Then*

$$U_i(\boldsymbol{\sigma}) \geq \underline{v}_i.$$

Proof: Player i can always guarantee herself $\underline{v}_i = \min_{a_{-i}} [\max_{a_i} \; u_i(a_i, a_{-i})]$ in the stage game and also in each stage of the repeated game, since $\underline{v}_i = \max_{a_i} \left[\min_{a_{-i}} \; u_i(a_i, a_{-i})\right]$, meaning that she can always achieve at least this payoff against even the most adversarial strategies.

Folk Theorems

Definition

A payoff vector $\mathbf{v} \in \mathbb{R}^I$ *is strictly individually rational if* $v_i > \underline{v}_i$ *for all* i.

Theorem

(Nash Folk Theorem) *If* $(v_1, \ldots, v_I)$ *is feasible and strictly individually rational, then there exists some* $\underline{\delta} < 1$ *such that for all* $\delta > \underline{\delta}$, *there is a Nash equilibrium of* $G^\infty(\delta)$ *with payoffs* $(v_1, \cdots, v_I)$.

Proof

Proof:

- Suppose for simplicity that there exists an action profile $a = (a_1, \cdots, a_I)$ s.t. $u_i(a) = v$ [otherwise, we have to consider mixed strategies, which is a little more involved].
- Let m^i_{-i} these the minimax strategy of opponents of i and m^i_i be i's best response to m^i_{-i}.
- Now consider the following grim trigger strategy.
- For player i: Play $(a_1, \cdots, a_I)$ as long as no one deviates. If some player deviates, then play m^j_i thereafter.
- We next check if player i can gain by deviating form this strategy profile. If i plays the strategy, his payoff is v_i.

Proof (continued)

- If i deviates from the strategy in some period t, then denoting $\overline{v_i} = \max_a u_i(a)$, the most that player i could get is given by:

$$(1-\delta)\left[v_i + \delta v_i + \cdots + \delta^{t-1} v_i + \delta^t \overline{v}_i + \delta^{t+1} \underline{v}_i + \delta^{t+2} \underline{v}_i + \cdots\right].$$

- Hence, following the suggested strategy will be optimal if

$$\frac{v_i}{1-\delta} \geq \frac{1-\delta^t}{1-\delta} v_i + \delta^t \overline{v}_i + \frac{\delta^{t+1}}{1-\delta} \underline{v}_i,$$

thus if

$$\begin{aligned} v_i &\geq \left(1-\delta^t\right) v_i + \delta^t \left(1-\delta\right) \overline{v}_i + \delta^{t+1} \underline{v}_i \\ &= v_i - \delta^t [v_i - (1-\delta)\overline{v}_i - \delta \underline{v}_i]. \end{aligned}$$

- The expression in the bracket is non-negative for any

$$\delta \geq \underline{\delta} \equiv \max_i \frac{\overline{v}_i - v_i}{\overline{v}_i - \underline{v}_i}.$$

- This completes the proof.

Problems with Nash Folk Theorem

- The Nash folk theorem states that essentially any payoff can be obtained as a Nash Equilibrium when players are patient enough.
- However, the corresponding strategies involve this non-forgiving punishments, which may be very costly for the punisher to carry out (i.e., they represent non-credible threats).
- This implies that the strategies used may not be subgame perfect. The next example illustrates this fact.

	L (q)	R $(1-q)$
U	$6, 6$	$0, -100$
D	$7, 1$	$0, -100$

- The unique NE in this game is (D, L). It can also be seen that the minmax payoffs are given by

$$\underline{v}_1 = 0, \qquad \underline{v}_2 = 1,$$

and the minmax strategy profile of player 2 is to play R.

Problems with the Nash Folk Theorem (continued)

- Nash Folk Theorem says that (6,6) is possible as a Nash equilibrium payoff of the repeated game, but the strategies suggested in the proof require player 2 to play R in every period following a deviation.
- While this will hurt player 1, it will hurt player 2 a lot, it seems unreasonable to expect her to carry out the threat.
- Our next step is to get the payoff $(6, 6)$ in the above example, or more generally, the set of feasible and strictly individually rational payoffs as subgame perfect equilibria payoffs of the repeated game.

Subgame Perfect Folk Theorem

- The first subgame perfect folk theorem shows that any payoff above the static Nash payoffs can be sustained as a subgame perfect equilibrium of the repeated game.

Theorem

(Friedman) *Let a^{NE} be a static equilibrium of the stage game with payoffs e^{NE}. For any feasible payoff v with $v_i > e_i^{NE}$ for all $i \in \mathcal{I}$, there exists some $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there exists a subgame perfect equilibrium of $G^{\infty}(\delta)$ with payoffs v.*

Proof: Simply construct the non-forgiving trigger strategies with punishment by the static Nash Equilibrium. Punishments are therefore subgame perfect. For δ sufficiently close to 1, it is better for each player i to obtain v_i rather than deviate get a high deviation payoff for one period, and then obtain e_i^{NE} forever thereafter.

Subgame Perfect Folk Theorem (continued)

Theorem

(Fudenberg and Maskin) *Assume that the dimension of the set V of feasible payoffs is equal to the number of players I. Then, for any $v \in V$ with $v_i > \underline{v}_i$ for all i, there exists a discount factor $\underline{\delta} < 1$ such that for all $\delta \geq \underline{\delta}$, there is a subgame perfect equilibrium of $G^\infty(\delta)$ with payoffs v.*

- The proof of this theorem is more difficult, but the idea is to use the assumption on the dimension of V to ensure that each player i can be *singled out* for punishment in the event of a deviation, and then use rewards and punishments for other players to ensure that the deviator can be held down to her minmax payoff.

Cooperation in Finitely-Repeated Games

- We saw that finitely-repeated games with unique stage equilibrium do not allow corporation or any other outcome than the repetition of this unique equilibrium.
- But this is no longer the case when there are multiple equilibria in the stage game.
- Consider the following example

	A	B	C
A	$3, 3$	$0, 4$	$-2, 0$
B	$4, 0$	$1, 1$	$-2, 0$
C	$0, -2$	$0, -2$	$-1, -1$

- The stage game has two pure Nash equilibria (B, B) and (C, C). The most cooperative outcome, (A, A), is not an equilibrium.
- **Main result in example:** in the twice repeated version of this game, we can support (A, A) in the first period.

Cooperation in Finitely-Repeated Games (continued)

- Idea: use the threat of switching to (C, C) in order to support (A, A) in the first period and (B, B) in the second.
- Suppose, for simplicity, no discounting.
- If we can support (A, A) in the first period and (B, B) in the second, then each player will receive a payoff of 4.
- If a player deviates and plays B in the first period, then in the second period the opponent will play C, and thus her best response will be C as well, giving her -1. Thus total payoff will be 3. Therefore, deviation is not profitable.

How Do People Play Repeated Games?

- In lab experiments, there is more cooperation in prisoners' dilemma games than predicted by theory.
- More interestingly, cooperation increases as the game is repeated, even if there is only finite rounds of repetition.
- Why?
- Most likely, in labs, people are confronted with a payoff matrix of the form:

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- Entries are monetary payoffs. But we should really have people's **full payoffs**.
- These may differ because of social preferences.

Social Preferences

- Types of social preferences:
  - **Altruism**: people receive utility from being nice to others.
  - **Fairness:** people receive utility from being fair to others.
  - **Vindictiveness:** people like to punish those deviating from "fairness" or other accepted norms of behavior.
- All of these types of social preferences seem to play some role in experimental results.

14.15J / 6.207J Networks
Fall 2009