Stevens Institute of Technology

Game Theory Notes II*

Author: Ruming L†

January 4, 2023

*Reference here

†rliu38@stevens.edu

Contents

1 Repeated games **3**

1.1 Mutistage games 3

1.2 Conditional strategies 3

1.3 Time value of the payoff 4

1.4 Subgame-Perfect Nash equilibrium 5

1.5 Revenge or cooperation 8

1.6 One-stage deviation principle 9

1.7 Finitely repeated games 10

1.8 Infinitely repeated games 10

1.9 Grim-trigger strategies 11

1.10 The forlk theorem 13

1 Repeated games

1.1 Mutistage games

We define a **multistage game** as a finite sequence of normal-form stage-games, in which each stage-game is an independent, well-defined game of complete but imperfect information (a simultaneous-move game).

Remark. *Several keywords in this definition.*

- *Both players understand the sequence game will have an end.*
- *Independent means the payoff in each round is independent. It doesn't mean the strategies in each round are independent, e.g. player can make actions based on previous game results.*
- *Imperfect information comes from the simultaneous move in each round, but all players can see the strategies in this round after the payoff.*

1.2 Conditional strategies

If all players have bad memory, they can't recall what have happened in the past. It's obvious that they will play Nash equilibrium in each round. Then the repeated playing doesn't introduce any new thing. What if players can recall everything happened before? So we need to make a definition so that players can make their strategies based on the history.

Definition 1. History: We adopt the convention that each game is played in a distinct period, so the game 1 is played at period $t = 1$, game 2 in period $t = 2$, and so on, up until period $t = T$, which will be the last stage in the game.[1] We define the history h_t of period t

[1] It is possible to consider an infinite sequence of games, but we restrict attention to a finite sequence of games in this chapter. In the later chapter we will consider such infinite sequences, in which the stage-games are always the same game that repeats itself again and again.

as:

$$h_t = (a_1, a_2, ...a_{t-1}),$$

where a_τ is the action profile played at period τ.

Now with the defnition of history, we are able to define the conditional strategies of each player based on the past game results.

Definition 2. Consider a finitely multi-stages game. Let H_t denote the set of all possible histories of period t, $h_t \in H_t$ where $1 \leq t \leq T$. A **pure strategy** for player i at stage t is a mapping $a_i(t) : H_t \to A_i(t)$[2]. Similarly a **behavioral strategy** of player i at stage t is $\sigma_i(t) : H_t \to \Delta S_i(t)$.

That means every players in each period t can make decisions based on the history of the past games H_t.

1.3 Time value of the payoff

How should we evaluate the total payoffs from a sequence of outcomes in each of the sequentially played stage-games? We adopt the well-defined notion of present value.

In particular at any period t we will collapse any sequence of payments from the games played starting at period t onward into a single value that can be evaluated at period t . To do this we take the payoffs derived from each of the individual games and add them up with the added assumption that future payoffs are discounted at a rate $0 \leq \delta \leq 1$. A higher discount factor δ means that the players are more patient and will care more about future payoffs. An extreme case of impatience occurs when $\delta = 0$, which means that only the payoffs of the current stage-game matter.

Definition 3. Consider a multistage game in which there are T stage-games played in each of the periods $1, 2, ..., T$. Let v_i^t be player i's payoff from the anticipated outcome in the stage-game played in

[2] $A_i(t)$ is the actions set of player i in game t

		Player 2	
		m	f
Player 1	M	$(4,4)$	$(-1,5)$
	F	$(5,-1)$	$(1,1)$

Table 1: Prisoner dilemma

period t. We denote by v_i the total payoff of player i from playing the sequence of games in the multistage game and define it as

$$v_i = v_i^1 + \delta v_i^2 + \delta^2 v_i^3 + ...\delta^{T-1} v_i^T = \sum_{t=1}^{T} \delta^{t-1} v_i^t.$$

1.4 Subgame-Perfect Nash equilibrium

With above special definitions in muti-stage games, we are trying to form the game in strategic form (matrix). A natural question is how large the dimension of the matrix will be?

We use a two-stage game that called the Prisoner-Revenge game as an example. Suppose that in a first period labeled $t = 1$ two players from different local neighborhoods play the familiar Prisoner's Dilemma with pure actions mum and fink (uppercase for player 1 and lowercase for player 2) and with the payoff matrix in table 1.

Now imagine that after this game is over the same two players play the following Revenge Game in a second period $t = 2$. Each player can choose whether to join his local neighborhood gang or remain a "loner." If the players remain loners then they go their separate ways, regardless of the outcome of the first game, and obtain second-period payoffs of 0 each. If both join gangs, then because of the nature of neighborhood gangs they will fight each other and suffer a loss, causing each to receive a payoff of 3. Finally if player i joins a gang while player j remains a loner then the loner suffers dearly since he has no gang to defend him, thus receiving a payoff of -4, while the new gang member suffers much less and receives a payoff of -1. The Revenge Game is described by the following matrix:

		Player 2	
		l	g
Player 1	L	$(0,0)$	$(-4,-1)$
	G	$(-1,-4)$	$(-3,-3)$

Table 2: Prisoner revenge

How do we use matrix to represent the two-stage game? Recall in the Spence's education signalling game, we use the strategy **"if player 1 take education, we do this, if player 1 does not take education we do this..."** for player 2. Similarly, with the help of definition 2, we can use the technique **"if history is h_t, we play this, if history is h'_t, we do this..."** to represent the matrix.

Formally, we can define a strategy for player i in the multistage Prisoner-Revenge Game as:

$$s_i = (s_i^1, s_i^2(Mm), s_i^2(Mf), s_i^2(Fm), s_i^2(Ff)),$$

here s_i^1 is what player i will do in the first-stage Prisoner's Dilemma game and $s_i^2(ab)$ is what player i will do in the second-stage Revenge Game. Each element can take two values. So easily, we know the matrix is a 32×32 dimension matrix with 1024 pure-strategy combinations.

With the well-defined matrix, we can easily solve the NE for the whole game. Then the next step is to do equilibrium refinement.

Because every period games are strategic games with simultaneous move. This will lead some information sets are non-singleton. And recall the extensive-form games with imperfect information, we usually use the concept subgame-perfect to induce "reasonable" and refined equilibrium. The procedure is almost the same with the matrix representation procedure above, but instead of focusing on the whole game, we focus on those proper subgames. Let's define the subgame perfect first:

Definition 4. A profile of pure strategies $(s_1^*, s_2^*, ..., s_n^*)$, $s_i : H \to S_i$ for all $i \in N$, is a subgame-perfect equilibrium if the restriction of

$(s_1^*, s_2^*, ..., s_n^*)$ is a NE in every subgame. That is, for any history of the game h_t, the continuation play dictated by $(s_1^*, s_2^*, ..., s_n^*)$ is a NE. (similarly for behaviral strategies.)

Theorem 1. Consider a multistage game with T stages, and let σ^{t*} be a Nash equilibrium strategy profile for the t-th stage-game. There exists a subgame-perfect equilibrium in the multistage game in which the equilibrium path coincides with the path generated by $\sigma^{1*}, \sigma^{2*}, ..., \sigma^{T*}$.

Remark. *The intuitive proof is following the backward induction logic: At final stage T, the previous payoff is determined, what you can do is do the best response in this stage, so playing σ^{T*} is subgame perfect in the subgame at stage T. At stage $T-1$, the previous payoff is determined, and you know in the next stage T, equilibrium will be σ^{T*} which is independent with the outcomes of stage $T-1$, what you can do is do best response in current stage $T-1$. This implies that $\sigma^{T-1,*}$ followed by σ^{T*} in any subgame at stage T is a Nash equilibrium in any subgame starting at time $T-1$. Using backward induction, this argument implies that at any stage t, current play does not affect future play, and this inductive argument for a finite number of stages in turn implies that in any subgame these strategies constitute a Nash equilibrium.* ■

Theorem 2. If σ^* is a Nash equilibrium of the multistage game consisting of stage-games $G_1, G_2, ..., G_T$ then the restriction of σ^* to the stage-game in period T must be a Nash equilibrium of that stage-game.

The proposition states that for any finite stage-game of length T, in the last stage-game G_T the players must play a Nash equilibrium of that stage-game.The reasoning is simple: because there is no future that can depend on the actions taken at stage T, the only thing that determines what is best for any player at that stage is to play a best response to what he believes the other players are doing at that stage. This insight implies another important fact:

Theorem 3. If a finite multistage game consists of stage-games that each have a unique Nash equilibrium then the multistage game has a unique subgame-perfect equilibrium.

Remark. *This theorem comes from Theorem 1 with unique NE in each stage.*

1.5 Revenge or cooperation

In theorem 3, we have seen that in a finite muti-stage game, each stage has unique NE, then everything is same with the game without repeating. In some cases, is it possible that strategies which are not NE can be played in the early stage? The answer is yes, this will be possible only when is there are multiple NE in the later stages.

We now use the Prisoner-Revenge Game as an illustration of what players can do when the stage-game played at the end of a multistage game has more than one Nash equilibrium. As we all know, players play (M, m) to cooperate is Pareto optimal but not NE, in any one-time prisoner dilemma game, the cooperation can never be reached. But interestingly, if there is a revenge game in period 2, and notice that the revenge game has 2 NE, (L, l) and (G, g). Now consider below strategies:

- **Player 1** Play M in stage 1. In stage 2 play L if (M, m) was played in stage 1, otherwise play G in stage 2. $s_1^* = (s_1^1, s_1^2(Mm), s_1^2(Mf), s_1^2(Fm), s_1^2(Ff)) = (M, L, G, G, G)$.
- **Player 2** Play m in stage 1. In stage 2 play l if (M, m) was played in stage 1, otherwise play g in stage 2. $s_2^* = (s_2^1, s_2^2(Mm), s_2^2(Mf), s_2^2(Fm), s_2^2(Ff)) = (m, l, g, g, g)$.

We are left to check if this pair of strategies is a subgame-perfect equilibrium. It is easy to see that in each of the subgames beginning at stage 2 the players are playing a Nash equilibrium (either both play (L, l) following (M, m) or both play (G, g) following other histories of play). Hence to check that this is a subgame-perfect equilibrium we need to check that players would not want to deviate from mum in the first stage of the game. In other words, is mum a best response in period 1, given what each player believes about his opponent, and given the continuation play in the second-period stage-games? Consider player 1 and observe that

$$v_1(M, s_2) = 4 + 0 \times \delta, v_1(F, s_1) = 5 + (-3) \times \delta$$

which implies that M is a best response for player 1 if $\delta \geq 1/3$. (Note: this is also the condition leads m is the BR for player 2.)

Why this interesting subgame perfect equilibrium exists? Two requirements lead the equilibrium:

1. There must be at least two distinct equilibrium in the second stage: a "stick" and a "carrot".

2. The discount factor has to be large enough for the difference in payoffs between the “stick” and the “carrot” to have enough impact in the first stage of the game.

In the revenge game, (G, g) is a "sticker" and (L, l) is a "carrot". If everyone cooperates, they get a "carrot". Of course, if the "sticker" is not too bad or players are impatient(discount factor is small), the "sticker" will keep two players stay in cooperation. That will guarantee the existence of such interesting subgame perfect equlibrium.

1.6 One-stage deviation principle

As an illustration of multistage games, the Prisoner-Revenge Game was not hard to analyze. In particular, to check that a certain proposed strategy profile was a subgame- perfect equilibrium, we needed only to check that no player wanted to deviate in the first stage, because we constructed our candidates for subgame-perfect equilibria to consist of Nash equilibria in each of the second-stage subgames. It might seem that if we had, say, a five-stage game it would be more complicated to check that a profile of strategies constitutes a subgame-perfect equilibrium, and that in particular we would have to check that the players do not want to deviate at any single stage-game. That is, maybe they could gain by combining several deviations in separate stages of the game. This is where the **one-stage deviation principle** comes in to simplify what seems like a rather daunting task.

Definition 5. A strategy σ_i is **one-stage unimprovable** if there is no information set h_i, action $a \in A_i(h_i)$ and corresponding strategy σ_i^{a,h_i} such that $v_i(\sigma_i^{a,h_i}, h_i) > v_i(\sigma_i, h_i)$.

Theorem 4. A one-stage unimprovable strategy must be optimal.

Remark. *We can use this theorem to prove some strategies are subgame perfect. We only need to show that for every player i, given σ^*_{-i}, player i does not have a single information set from which he would want to deviate.*

1.7 Finitely repeated games

In the previous chapter on mutistage games, we found that the behavior that can be supported need not be a static best response in the early stage-game because the conditional strategies in later stage-games can act as a powerful incentive scheme to help players resist the short-run temptation of deviating from the proposed path of play.

A special case of multistage games has received considerable attention over the years: the case of repeated games. A repeated game is simply a multistage game in which the same stage-game is being played at every stage.

Triviallyall definitions, theorems still work fine in **finitely repeated games**.

1.8 Infinitely repeated games

Why infinitely repeated games happen? Consider a monopolist public firm, the firm's goal is to live forever, it may play entry-battle game round by round, that's a possible situation we need to define an infinitely repeated games. Because the period is infinite now, we need to modify some definitions below:

Definition 6. Given the discount factor $0 < \delta < 1$, the **present value** of an infinite sequence of payoffs $\{v_i^t\}_{t=1}^{\infty}$ for player i is $v_i = \sum_{t=1}^{\infty} \delta^{t-1} v_i^t$.

Definition 7. Given $\delta < 1$, the **modified present value (average payoff**[3]**)** of an infinite sequence of payoffs $\{v_i^t\}_{t=1}^{\infty}$ is

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

This is a renormalization of utility that doesn't change player i's ranking of any two infinite sequences of payoffs. The reason we time $1-\delta$ is that $(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} c = c(1-\delta)\frac{1}{1-\delta} = c$. That is, the average payoff of an infinite fixed sequence of some value c is itself equal to c. As we will soon see, using average payoffs will help simplify the analysis of subgame-perfect equilibria in infinitely repeated games. This approach will also help us identify the set of payoffs that can be supported by strategies that form a subgame-perfect equilibrium of a repeated game.

1.9 Grim-trigger strategies

Some general propositions will be omitted (e.g. he repeated-game strategy for each player i to be the history-independent Nash strategy is a subgame-perfect equilibrium...) in this chapter.

Now consider a repeated prisoners dilemma game in table 3. Absolutely, playing fink unconditionally in every period by each player is a subgame-perfect equilibrium.

But this equilibrium is not interesting, let's consider below "carrot" or "sticker" strategies:

[3] The reason we call it average payoff is that this formular can be regarded as $\frac{\sum_{t=1}^{\infty} \delta^{t-1} v_i^t}{\infty}$.

		Player 2	
		m	f
Player 1	M	$(4,4)$	$(-1,5)$
	F	$(5,-1)$	$(1,1)$

Table 3: Prisoner dilemma

1. In the first stage play $s_1^1 = M$. For any stage $t > 1$, play $s_1^t(h_{t-1}) = M$ if and only if the history h_{t-1} is a sequence that consists only of (M, m). Otherwise, play $s_1^t(h_{t-1}) = F$.

2. In the first stage play $s_1^1 = m$. For any stage $t > 1$, play $s_1^t(h_{t-1}) = m$ if and only if the history h_{t-1} is a sequence that consists only of (M, m). Otherwise, play $s_1^t(h_{t-1}) = f$.

These strategies are commonly referred to as grim-trigger strategies because they include a natural trigger: once someone deviates from mum, this is the trigger that causes the players to revert their behavior to fink forever, resulting in a very grim future. More generally the idea behind grim-trigger strategies is as follows. If the stage-game has a Nash equilibrium, and we are trying to support subgame-perfect equilibrium behavior that results in outcomes that are better than any Nash equilibrium, we use the established subgame-perfect equilibrium of playing the grim static Nash outcome forever to support the more desirable outcomes that are not supported by static best responses being played. Hence we use the grim trigger to provide incentives for the players to stick to behavior for which short-run temptations to deviate are present.

Let's confirm that this strategy is subgame perfect by one-stage deviation principle introduced before. Past can not be changed, the payoff of both player from following the strategy and not deviating at period t is

$$v_i^* = 4 + \delta 4 + \delta^2 4 + ... = 4 + \delta \frac{4}{1-\delta}.$$

If the player deviates from mum and chooses fink instead, then he gets 5 at period t instead of 4 in the immediate stage of deviation, followed by his continuation payoff, which is an infinite sequence of 1s.

$$v_i^* = 5 + \delta 1 + \delta^2 1 + ... = 5 + \delta \frac{1}{1-\delta}.$$

Obviously, if $\delta \geq 1/4$, for any $t \in [1, \infty)$, both players will not deviate at period t. Which means the strategy profile is a subgame perfect equilibrium.

Once again, as with multistage games, we see the value of patience. If the players are sufficiently patient, so that the future carries a fair amount of weight in their preferences, then there is a reward-and-punishment strategy that will allow them to cooperate forever. The loss in continuation payoffs will more than offset the gains from immediate defection, and this will keep the players on the cooperative path of play.

1.10 The forlk theorem

The general idea of this theorem is that: Any "reasonable" achievable payoff can be realized in a subgame perfect equilibrium of the repeated game, if players are patient enough ($\delta \to 1$).

Remark. *The achievable payoff is the convex hull of $V = (V_1, V_2)$ in figure 1, in other words, is the combination of all possible payoffs. The "reasonable" means that $v_i > v_i^*, \forall i \in N$. Otherwise some player will prefer to play Nash strategy. Finally, with a high enough discount factor, we can guarantee that wedge sufficiently large to deter short-run gains from deviation, no matter how large they are.*

The trigger strategy to achieve average payoff $(2, 2)$ in figure 1 is that players cooperate to play (F, m) in odd period and play (M, f) in even period. If anyone defects, they choose the "sticker" strategy (F, f) forever.

Intuitively, average payoff $(2, 2)$ is better than NE $(1, 1)$, and players are very patient (they care future a lot), so two players will cooperate to achieve this average payoff.

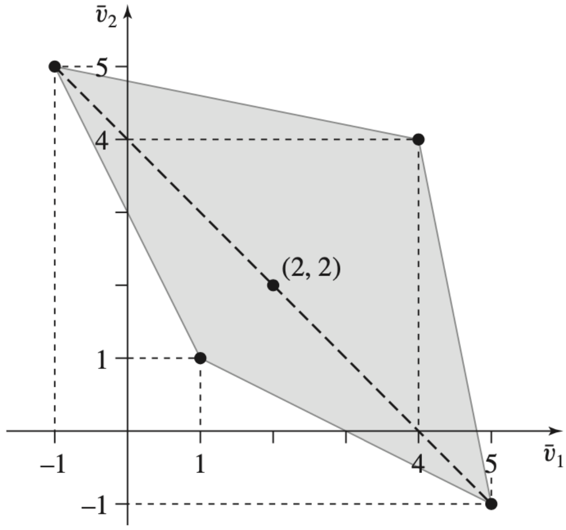

FIGURE 10.5 Constructing the folk theorem for the repeated Prisoner's Dilemma game.

Figure 1: The convex hull of average payoffs for the Prisoner's Dilemma game.