Introduction to Repeated Games with Perfect Monitoring

김 진 옥*

(Kim, Jin-Ock)

Contents

I. Introduction

Game theory is the study of multi-person decision making in strategic situations. A crucial feature of many strategic situations is that people interact repeatedly over time, not just once. For example, Korean Airline and American Airline compete for business every day, principals try to induce agents' full effort, and suppliers and

* Associate Professor, Department of Economics, College of Economics and Commerce, Cheju National University.

buyers make deals repeatedly, and so on. People may behave quite differently toward those with whom they would expect to have a long term relationship, compared to their behavior toward those whom they expect to have a short-term interactions. To capture this notion of ongoing interactions, the repeated game model is devised. In reality, the strategic situations players face change over time. The basic model of repeated play abstracts from this situations, and focuses just on the effect of repetition. In this introductory paper, we confined ourselves to the study of repeated play when players' past moves are perfectly monitored. We organize this paper as follows. To motivate our understanding on repeated play, we review some examples in section 2. In section 3, we introduce a general model of repeated play. In section 4, we present the well known Folk Theorems, and sketched the proof of those theorems.

Lastly, we make some comments on the variations and strengthenings of Perfect monitoring folk theorems.

Ⅱ. Some Examples

2.1 Example 1: Prisoner's Dilemma Game

To motivate the study of repeated games, we begin with a well-known example, the repeated Prisoner's Dilemma game in Table 1.

Table 1

Player 2

		C	D
Player 1	C	1,1	-1,1
	D	2,1	0,0

Here, C is a cooperative move, and D is a move of defection. The unique Nash equilibrium is (D,D) when the game is played just once. As we know, this Nash equilibrium is Pareto inferior, while (C,C) is a Pareto superior. A pareto superior outcome (C,C) may be voluntarily elicited when the game is infinitely repeated. Let's consider that players 1 and 2 expect to play the game repeatedly at time t=0,1,2,.... For example, suppose that the number of times that the game will be played is a random variable, which is unknown to the players until the game ends, and that the random stopping time is a geometric distribution with expected value 100: the probability of play continuing for exactly k rounds is $(0.99)^{k-1} \times 0.01$. In this repeated game, if both players do the generous moves(C,C) forever, then each player will get an expected total future payoff of 500: $\sum_{k=1}^{\infty} (0.99^{k-1})(0.01)5k=500$. On the contrary, if both players play the noncooperative moves(D,D), then each player will get an expected future payoff of 100: $\sum_{k=1}^{\infty} (0.99^{k-1})(0.01)k =100$. We will present the following proposition with a loose sketch of proof.

Proposition 1 A pareto optimal outcome(C,C) may be elicited in every period if the players follow "grim trigger" strategies:

I. Play C in every period unless someone plays D, in which case go to II.

II. Play D forever.

Proof. If both players follow these strategies, then, at any time in the game each player will get an expected total future payoff of $500(=\sum_{k=1}^{\infty}(0.99^{k-1})(0.01)5k)$, so long as no one has deviated. But, If either player I deviated from these strategies and choose defection(D) on particular day, then her expected total future payoff from this day onward would be $105(=6+\sum_{k=2}^{\infty}(0.99^{k-1})(0.01)k)$. What if at some time t, D has already been played. Then, if player j will play C instead of playing D, her expected total future payoff is 99(=0+$\sum_{k=2}^{\infty}(0.99^{k-1})(0.01)k$), which is less than her expected total future 100 if she doesn't deviate from playing D. So D is definitely optimal.

2.2 Example 2: Repeating the Game Twice.

The second example we consider is the following game, represented as in Table 2.

Table 2

Player 2

		Left	Middle	Right
Player 1	Top	0,0	3,4	6,0
	Middle	4,3	0,0	0,0
	Bottom	0,6	0,0	5,5

We repeated the above game twice with a discount factor δ. Player i's payoff for the entire repeated game is:

$$U_i(\{a^1,a^2\}) = U_i(a^1) + \delta U_i(a^2)$$

where a^1=period 1's strategy profile, a^2=period 2's strategy profile.

If we play the game just once, the Nash Equilibria will be {Middle ,Left}, {Top, Middle} and {(3/7)Top +(4/7)Middle, (3/7)Left+(4/7)Middle}. In this twice repeated game, we can construct the following equilibrium strategy(I and II) for δ >7/9:

I. Play {Bottom ,Right} in the first period, and {Middle ,Left} in the second period.

II. If either player deviates in the first period, play the mixed strategy game equilibrium in the second period.

Cheating in the first period incurs a penalty of at least 9/7 in the second period. The present net gain of deviating from equilibrium is 1, which is greater than a penalty of 9/7 in the second period if $\delta > 7/9$. In this finitely repeated game, a stage NE is required in the last period.

Ⅲ. A General Model of Infinitely Repeated Games

The following defines an infinitely repeated simultaneous move games.

- Players: a finite set $I = \{1,2,...,i\}$.
- Let G be a stage strategic form game with action sets $A_1...,A_I$, payoff functions $U_i : A \rightarrow R$, where $A = A_1 \times ... \times A_I$.
- Let G^∞ be the infinitely repeated version of G played at t=0,1,2,... where players discount at δ and observe all previous actions.
- A history is $H^t = \{a^0,...,a^{t-1}\}$, where $a^i \in A$ for i=0,...t-1.
- A strategy is $s_{it} : H^t \rightarrow A_i$, Here A_i could be the set of mixed actions.
- Average payoffs for I are:

  $U_i(s_i, s_{-i}) = (1-\delta) \Sigma_{t=0}^{\infty} \delta^t U_i(a_i, a_{-i})$, here $a_i \in A_i$, $a_{-i} \in A_{-i}$.

The question we are very interested in is what possible average payoff could be obtained from different equilibria in the game when discount factor is sufficiently large. That is, what might be happened in equilibrium when players are patient. We present some facts and a definition of min-max payoff before the wellknown general feasibility theorem called as Folk Theorem:

Fact 1 (Feasibility) If payoffs vector $v=(v_1,v_2,...v_I)$ is in a Nash equilibrium, then v is an element of convex hull of V, where V=convex hull$\{(x_1,x_2,...,x_I): \exists\ a \in A$ with $U_i(a)=x_i\}$.

Definition 1 Player i's min-max payoff is defined as follows

$$\upsilon_i=\min_{\sigma_{-i}}\max_{\sigma_i}U_i(\sigma_i,\sigma_{-i})$$

Fact 2 (Individual Rationality) In any Nash equilibrium, player i must receive at least min-max payoff υ_i.

Proof. In a static equilibrium α, α_i is a best response to α_{-i}, which implies that $U_i(\alpha_i,\alpha_{-i})$ is no less than the min-max payoff υ_i. Nower consider a Nash equilibrium $\sigma(\sigma_i,\sigma_{-i})$ in a repeated game. Then let σ_i' be the strategy of playing a static best-response to σ_{-i} in each period. Then (σ_i',σ_{-i}) will give player I a payoff of min-max value υ_i. Thus playing σ_i must give at least this much.

Ⅳ. The General Feasibility Theorem

The general intuition that we take from in the preceding examples is that almost any feasible payoff allocation above the min-max security level can be realized in an equilibrium of the repeated game when players are sufficiently patient. That is, the feasible payoff allocation in equilibria of a standard repeated game may generally coincide with the payoff that are feasible in the corresponding one-round game with contracts. In the game theory literature, these feasibility theorems have been referred to as Folk Theorem because some weak feasibility theorems are understood and believed in an oral folk tradition before any rigorous statements were published. We first state the Nash Folk theorem with a brief sketch of proof.

Theorem 1 (Nash Folk Theorem) If payoffs vector $v=(v_1,v_2,...v_I)$ is feasible and strictly individually rational, then there exists a $\delta_1<1$ such that for all $\delta>\delta_1$, there is a Nash Equilibrium of $G^{\infty}(\delta)$ with average payoffs v.

Proof. Assume that there exists a profile $a=(a_1,a_2,...,a_I)$ such that $U_i(a)=v_i$ for all i. Let m_{-i} be the min-max strategy profile of players other than player i holding him to his min-max payoff υ_i; and write m_i for player i's best response to m_{-i}. Now construct the following equilibrium:

I. Play a profile $a=(a_1,a_2,...,a_I)$ as long as no one deviates.

II. If some player j deviates, players other than j plays min-max strategy profile m_{-j} against player j thereafter.

If any player i plays this strategy, he gets v_i. If he deviates in some period t, the most average payoff v_d that player i could get is:

$$(1-\delta)[v_i+\delta v_i+...+\delta^{t-2}v_i+\delta^{t}v_i^{max}+\delta^{t+1}\upsilon_i+\delta^{t+2}\upsilon_i+...]$$

where v_i^{max}=supremum$_a$ $U_i(a)$.

Following the suggested strategy will be optimal if $v_i>v_d$, which implies the following inequality.

$$[\delta/(1-\delta)](v_i-\upsilon_i)\geq(v_i^{max}-\upsilon_i)$$

As $\delta \to 1$, the ratio $\delta/(1-\delta) \to \infty$, so we simply pick $\delta_1 = \max_i((v_i^{max} - \upsilon_i)/(v_i - \upsilon_i)$.
Q.E.D.

This Nash folk theorem says that any feasible payoff vector can be supported as a Nash equilibrium when players are sufficiently patient. But one caveat in applying Nash equilibrium as a solution concept remains. That is because this solution concept requires to specify an implausible punishment scheme for players who are not punished. For example, if player j participate in the punishment scheme against player i who deviates, then he should follow the min-max strategy profile m_{-i} cooperatively, which will hurt him a lot. Fudenberg and Maskin (1986) devised the following folk theorem to avoid this problem[1]:

Theorem 2 (Folk Theorem)

Let V^* be the set of feasible and strictly individually rational payoffs. Assume that dim $V^*=I$. Then any payoff vector $(v_1,v_2,...v_I) \in V^*$, there exists a δ_1, such that for any $\delta > \delta_1$, there is a subgame perfect equilibrium of $G^\infty(\delta)$ with average payoffs $(v_1,v_2,...v_I)$.

Proof. Firstly fix a payoff vector $(v_1,v_2,...v_I) \in V^*$. For convenience, we assume that there exists some profile $a=(a_1,a_2,...,a_I) \in A$, such that $U_i(a)=v_i$ for all i. To prove the above theorem, we construct a subgame perfect equilibrium(SPE) of achieving a fixed payoff vector $v=(v_1,v_2,...v_I)$. The key to the proof is find payoffs that allows us to reward all agents except i who are participating in the punishment phase when i deviates and has to be min-maxed for some length of time. To construct a SPE, we set up the following things.

1) Rubinstein (1979) proved a general folk theorem for subgame perfect equilibria of standard repeated games with the overtaking criterion. With discounted average payoff criterion, Fudenberg and Maskin (1986) proved a general feasibility theorem for subgame-perfect equilibria of standard repeated games.

- Choose $v' \in \text{Interior}(V^*)$ such that $v_i' < v_i$ for all i.
- Choose T such that $\max_a U_i(a) + T\upsilon_i < \min_a U_i(a) + Tv_i'$
- Choose $\varepsilon > 0$ such that for each i, $v^i(\varepsilon) = (v_1' + \varepsilon, \ldots, v_{i-1}' + \varepsilon, v_i', v_{i+1}' + \varepsilon, \ldots, v_I' + \varepsilon)$ which is supported by the profile $a^i \in A$.
- Let m^i be the profile that min-maxes i, so that $U_i(m^i) = \upsilon_i$.

We construct the following strategies for i=1,2,...,I, which is a SPE .

Phase I. Each player i plays a_i so long as no one deviates from $a \in A$. If some player j alone deviates, go to punishment phase II_j.

Phase II_j. Play m^j for T periods, then go to pahse III_j if no one deviates. If k deviates. restart II_k.

Phase III_j. Play the reward profile a^j for players other than j so long as no one deviates. If k deviates, go to II_k.

Note that these strategies involve both punishments (the stick) and rewards(the carrot). Now we are in a position that these strategies are a SPE using the one-shot deviation principle. To do this, We will check for each of different subgames.

Subgame I. Compare i's payoff playing the strategy with his payoff of deviating:

player i follows the suggested strategy : $(1-\delta)[v_i + \delta v_i + \ldots] = v_i$

player i deviates : $(1-\delta)[v_i^{max} + \delta\,\upsilon_i + \ldots + \delta^T \upsilon_i + \delta^{T+1} v_i' + \ldots]$

We can pick some δ such that $v_i > (1-\delta)[v_i^{max} + \delta\,\upsilon_i + \ldots + \delta^T \upsilon_i + \delta^{T+1} v_i' + \ldots]$.

Subgame II_i. (suppose that there are $T' \leq T$ periods left); this is the punishment phase against player i.

player i follows strategy : $(1-\delta^{T})\upsilon_i + \delta^{T}v_i'$

player I deviates : $(1-\delta)\upsilon_i + \delta[(1-\delta^{T})\upsilon_i + \delta^{T}v_i']$

Subgame IIj. (suppose that there are $T' \leq T$ periods left); this is the punishment phase against j.

palyer i follows the suggested strategy : $(1-\delta^{T'})U_i(m^j) + \delta^{T}(v_i'+\varepsilon)$

player i deviates : $(1-\delta)v_i^{max} + \delta(1-\delta^{T})\upsilon_i + \delta^{T+1}v_i'$

Subgame III_i, $III_{j,:}$ we will consider player i's payoff.

player i follows strategy : v_i'

player i deviates : $(1-\delta)v_i^{max} + \delta(1-\delta^{T})\upsilon_i + \delta^{T+1}v_i'$

We have presented both the least payoff player i could get if he follows the strategy and the most payoff he could get if he deviates. We can pick some δ_1 such that for $\delta > \delta_1$, it is best for him to follow the strategy for every subgame rather than deviating.

Q.E.D.

Notice that this perfect folk theorem requires an assumption that dim V^*=I. This assumption implies that any player i can be singled out when he deviates from equilibrium. The following example shows that no individual can be punished without punishing everybody:

Example: There are three players; Player 1 is a column player, Player 2 is row player, and player 3 chooses the matrix.

matrix A

	A	B
A	1,1,1	0,0,0
B	0,0,0	0,0,0

matrix B

	A	B
A	0,0,0	0,0,0
B	0,0,0	1,1,1

In this game the min-max level is zero 0. To min-max some player i, the other two players j and k need to mis-coordinate . The set of feasible and individually rational payoffs is: $V^{*}=\{(v, v, v):v\in[0,1]\}$.

Claim. there is no SPE of G^{∞} with average payoff less than 1/4.

Proof. Fix δ, and let x = infimum{v : (v, v, v) is a SPE payoff}. The first thing we show is that if (v, v, v) is an SPE, then $v \geq (1-\delta)(1/4)+\delta x$. To show this, we denote that $(\sigma_1, \sigma_2, \sigma_3)$ is the first period mixed strategies used in a SPE with payoff v. Then there must exist either two players with $\sigma_i(A) \geq 1/2$ or two players $\sigma_i(B) \geq 1/2$. Take the former case, and suppose that $\sigma_1(A), \sigma_2(A) \geq 1/2$.

Suppose that player 3 plays A in the first period and then follows his equilibrium strategy. Then his payoff from this will be at least $(1-\delta)(1/4)+\delta x$, given that continuous play will be an SPE. Since this deviation is unprofitable, the initial claim holds. That is, $x = \text{infimum}_{(v \text{ is SPE})}\, v \geq (1-\delta)(1/4)+\delta x$, which implies that $x \geq 1/4$. Basically, the problem is that no individual can be punished for deviating without punishing everyone. So there is no way of reward the punishers. Q.E.D.

V. Some comments

We conclude this introductory paper of repeated play with perfect monitoring. Of course, there are ongoing researches on "repeated play with imperfect public monitoring", and on repeated play with imperfect private monitoring". Here we willconfine ourselves to making a brief comments on several variations and strengthenings of the perfect monitoring folk theorem. Dutta and Smith (1994) relax the assumption of dim V^{*}=I (full dimensionality). They show that it is always possible to single out individuals for punishment in case that no two players have payoffs that are transformations of each other. Benoit and Krishna (1986) prove a folk theorem for finitely repeated game. Clearly, we know that this can't be done in the prisoner's dilemma where backward induction says that (D, D) will be played in every period. In this finitely repeated game, for folk theorem to hold, the stage game must have multiple Nash equilibria to allow for rewards and punishments towards the end of the game. Kreps, Milgrom, Roberts, and Wilson (1982) considered a finitely repeated Prisoner's dilemma game with the sum of payoffs criterion. In their model, they assumed that there is an initial doubt of player 2's rationality. That is, there is a small probability ε such that player 2 is not really rational but is instead a machine that always plays tit-for-tat. In that case, some generous moves can be elicited before the last some rounds. Kandori (1992) and Ellison (1994) show the folk theorems for games with overlapping generations and for games where players face a new opponent randomly drawn form the population in each period respectively. There are also a folk theorem for games where some players are long run, and others have myopic horizon.

References

Abreu, D. (1988), "Towards a theory of discounted repeated games," *Econometrica* 56: 383-396.

Benoit, J.P. and V. Krishna. (1986), "Finitely Repeted Games," *Econometrica* 53: 890=904.

Ellison, G. (1994), "Cooperating in the Repeated Prisoners' Dilemma with Anonymous Randon Matching," *Review of Economic Studies* 61, 567-588.

Fudenberg, D. and E. Maskin. (1986), "The Folk Theorem with Discounting and with Incomplete Information," *Econometrica* 54: 533-556.

Kanndori, M. (1992), "Social Norms and Community Enforcement," *Review of Econmomic Studies* 59, 63-80.

Kreps, D., P. Milgrom, J.Roberts, and R. Wilson. (1982), "Rational Cooperation in the Finitely Repeated Priosoner's Dilemma." *Journal of Economic Theory* 27, 245-252.