Repeated Games I: Perfect Monitoring

Jonathan Levin

May 2006

A crucial feature of many strategic situations is that players interact repeatedly over time, not just once. For instance, American Airlines and United compete for business every day, bosses try to motivate workers on an ongoing basis, suppliers and buyers make deals repeatedly, nations engage in ongoing trade, and so on. The repeated game model is perhaps the simplest model that captures this notion of ongoing interaction. Of course, in all these examples, there is a strong argument to be made that the game itself changes over time. The basic repeated game model abstracts from this issue, and focuses just on the effect of repetition.

1 Some Examples

1.1 Example 1: Cooperation in the Prisoners' Dilemma

Probably the best-known repeated game argument is that ongoing interaction can explain why people might behave cooperatively when it is against their self-interest in the short run. The classic example is the repeated prisoners' dilemma.

	C	D
C	$1,1$	$-1,2$
D	$2,-1$	$0,0$

The unique Nash equilibrium if the game is played once is (D,D).

Suppose that players 1 and 2 play the game repeatedly at time $t = 0,1,2...$ and that i's payoff for the entire repeated game is:

$$u_i\left(\{a^1,a^2,....\}\right) = (1-\delta)\sum_{t=0}^{\infty}\delta^t g_i(a_i^t,a_{-i}^t)$$

where $\delta\in[0,1)$. The fact that $\delta<1$ means that players discount the future — a dollar tomorrow is less than a dollar today. The overall payoffs are multiplied by $(1-\delta)$ to get a per-period average payoff for the game (note that this makes the repeated game payoff comparable to the stage game payoffs).

Proposition 1 *If $\delta \geq \frac{1}{2}$, the repeated prisoners' dilemma game has a subgame perfect equilibrium in which (C, C) is played in every period.*

Proof. Suppose the players use "grim trigger" strategies:

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

II. Play D forever.

To check that these strategies form a subgame perfect equilibrium if $\delta \geq \frac{1}{2}$, we need to verify that there is no single period where i can make a profitable deviation (this is the "one-stage deviation principle" — for details, see Fudenberg and Tirole, p. 108–110).

Suppose up to time t, D has never been played. Then i's payoffs looking forward are:

$$\begin{array}{lcl} \text{Play } C & \Rightarrow & (1-\delta)\left[1+\delta+\delta^2+...\right]=1 \\ \text{Play } D & \Rightarrow & (1-\delta)\left[2+\delta\cdot 0+\delta^2\cdot 0+...\right]=(1-\delta)2 \end{array}$$

so if $\delta \geq \frac{1}{2}$, C is optimal.

Suppose that at time t, D has already been played. Then j will play D and no matter what will continue to play D, so i's payoffs are:

$$\begin{array}{lcl} \text{Play } C & \Rightarrow & (1-\delta)\left[-1+\delta\cdot 0+\delta^2\cdot 0+...\right]=(1-\delta)(-1) \\ \text{Play } D & \Rightarrow & (1-\delta)\left[0+\delta\cdot 0+\delta^2\cdot 0+...\right]=0 \end{array}$$

So D is definitely optimal. *Q.E.D.*

1.2 Example 2: Use of Non-Nash Reversion

A second 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	$0,0$	$0,-1$	$-1,-1$

In this game, A is strictly dominant, and the unique Nash Equilibrium is (A, A).

Proposition 2 *If $\delta \geq \frac{1}{2}$, this game has a subgame perfect equilibrium in which (B, B) is played in each period.*

Proof. Here, we construct slightly more complicated strategies than grim trigger.

I. Play B in every period unless someone deviates, in which case go to II.

II. Play C. If no one deviates, go to I. If someone deviates, stay in II.

These strategies have what Abreu (1988) calls a "stick" (threatening to play C if someone deviates from B) *and* a "carrot" (promising to go back to B if everyone carries out the C punishment). Let's check that this is a SPE.

Suppose no one deviated at $t-1$, so players should play B at time t (i.e. they're in phase I):

$$\begin{aligned} \text{Play } B \quad &\Rightarrow \quad (1-\delta)\left[1+\delta+\delta^2+\delta^3+...\right] = 1 \\ \text{Best Dev. } (A) \quad &\Rightarrow \quad (1-\delta)\left[2+\delta(-1)+\delta^2+\delta^3+...\right] = 1+(1-\delta)(1-2\delta) \end{aligned}$$

so it's optimal to play B if $\delta \geq \frac{1}{2}$.

Suppose someone deviated at $t-1$, so players should play C at time t (i.e. they're in phase II)

$$\begin{aligned} \text{Play } C \quad &\Rightarrow \quad (1-\delta)\left[-1+\delta+\delta^2+\delta^3+...\right] = 1-(1-\delta)(2) \\ \text{Best Dev. } (A) \quad &\Rightarrow \quad (1-\delta)\left[0+\delta(-1)+\delta^2+\delta^3+...\right] = 1-(1-\delta)(1+2\delta) \end{aligned}$$

so it's optimal to play C if $\delta \geq \frac{1}{2}$. *Q.E.D.*

2 A General Model

- Let G be a normal form game with action spaces $A_1, ..., A_I$, payoff functions $g_i : A \to \mathbb{R}$, where $A = A_1 \times ... \times A_I$.
- Let $G^\infty(\delta)$ 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 $H^t = \{(a_1^0, ..., a_I^0), ..., (a_1^{t-1}, ..., a_I^{t-1})\}$.
- A *strategy* is $s_{it} : H^t \to A_i$.

- Average payoffs for i are:

$$u_i(s_i, s_{-i}) = (1-\delta)\sum_{t=0}^{\infty}\delta^t g_i(a_i, a_{-i}).$$

We now investigate what average payoffs could result from different equilibria when δ is near 1. That is, what can happen in equilibrium when players are very patient?

Fact 1 (Feasibility) If $(v_1, ..., v_I)$ are average payoffs in a Nash equilibrium, then

$$(v_1, ..., v_I) \in Conv\{(x_1, ..., x_I) : \exists (a_1, ..., a_I) \text{ with } g_i(a) = x_i \text{ for all } i\}$$

Definition 1 *Player i's min-max payoff is*

$$\underline{v}_i = \min_{\sigma_{-i}} \max_{\sigma_i} g_i(\sigma_i, \sigma_{-i})$$

Fact 2 (Individual Rationality) In any Nash equilibrium, player i must receive at least $\underline{v}_i$.

Proof. Suppose (σ_i, σ_{-i}) is a Nash equilibrium. 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 i a payoff of at least $\underline{v}_i$, and playing σ_i must give at least this much. *Q.E.D.*

3 The Folk Theorem

The first result is the (Nash) folk theorem which states that any feasible and strictly individually rational payoff vector can be achieved as a *Nash equilibrium* of the repeated game, provided players are sufficienly patient.

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

Proof. Assume there exists a profile $a = (a_1, ..., a_I)$ such that $g_i(a) = v_i$ for all i (I'll comment on this assumption later.) Let m_{-i}^i denote the strategy profile of players other than i that holds i to at most $\underline{v}_i$ and write m_i^i for i's best-response to m_{-i}^i.

Now consider the following strategies:

I. Play $(a_1, ..., a_I)$ as long as no one deviates.

II. If some player j deviates, play m_i^j thereafter.

If i plays this strategy, he gets v_i. If he deviates in some period t, then if $\overline{v}_i = \sup_a g_i(a)$, the *most* that i could get is:

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

Following the suggested strategy will be optimal if:

$$\frac{\delta}{1-\delta}\left(v_i - \underline{v}_i\right) \geq \left(\overline{v}_i - v_i\right)$$

As $\delta \to 1$, the ratio $\frac{\delta}{1-\delta} \to \infty$, so simply pick $\underline{\delta} = \max_i \left(\overline{v}_i - v_i\right) / \left(\overline{v}_i - \underline{v}_i\right)$. *Q.E.D.*

This Nash folk theorem says that essentially anything goes as a Nash equilibrium when players are sufficiently patient. Of course, we should be a little bit cautious about using Nash equilibrium as our solution concept since it might specify punishment behavior that is implausible. For example, consider the game

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

.

The Folk Theorem says that $(6, 6)$ is possible as a Nash equilibrium payoff, but the strategies suggested in the proof require the column player to play R in every period following a deviation. While this will hurt Row, it will hurt Column a lot — it seems unreasonable to expect her to carry ought the threat.

What we'd like to do is get $(6, 6)$, or more generally, the whole set of feasible and individually rational payoff vectors as subgame perfect equilibrium payoffs. The Fudenberg and Maskin (1986) folk theorem says that this possible.

Theorem 2 *(Folk Theorem) Let V^* be the set of feasible and strictly individually rational payoffs. Assume that $\dim V^* = I$. Then for any $(v_1, ..., v_I) \in V^*$, there exists a $\underline{\delta} < 1$, such that for any $\delta > \underline{\delta}$, there is a subgame perfect equilibrium of $G^\infty(\delta)$ with average payoffs $(v_1, ..., v_I)$.*

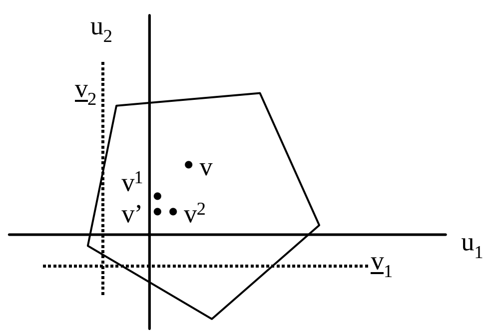
Proof. Fixing a payoff vector $(v_1, ..., v_I) \in V^*$, we construct a SPE that achieves it. For convenience, let's assume that there is some profile

$(a_1, ..., a_I)$ such that $g_i(a) = v_i$ for all i. The key to the proof is find payoffs that allow us to "reward" all agents $j \neq i$ in the event that i deviates and has to be min-maxed for some length of time.

- Choose $v' \in Int(V^*)$ such that $v'_i < v_i$ for all i.
- Choose T such that $\max_a g_i(a) + T\underline{v}_i < \min_a g_i(a) + Tv'_i$
- Choose $\varepsilon > 0$ such that for each i,

$$v^i(\varepsilon) = \left(v'_1 + \varepsilon, ..., v'_{i-1} + \varepsilon, v'_i, v'_{i+1} + \varepsilon, ..., v'_I + \varepsilon\right).$$

- Let a^i be the profile with $g(a^i) = v^i(\varepsilon)$
- Let m^i be the profile that min-maxes i, so $g_i(m^i) = \underline{v}^i$.



Consider the following strategies for $i = 1, 2, ..., I$.

I. Play a_i so long as no player deviates from $(a_1, ..., a_I)$. If j alone deviates, go to II$_j$. (If two or more players simultaneously deviate, play stays in I.)

II$_j$. Play m_i^j for T periods, then go to III$_j$ if no one deviates. If k deviates, re-start II$_k$.

III$_j$. Play a_i^j so long as no one deviates. If k deviates, go to II$_k$.

Note that strategies involve both punishments (the stick) and rewards (the carrot). Let's check that they are indeed a subgame perfect equilibrium using the one-shot deviation principle. We need to check for each of the different subgames.

Subgame I. Consider i's payoff to playing the strategy and deviating:

$$\begin{aligned} i \text{ follows strategy} &: \quad (1-\delta)\left[v_i + \delta v_i + ...\right] = v_i \\ i \text{ deviates} &: \quad (1-\delta)\left[\overline{v}_i + \delta \underline{v}_i + ... + \delta^T \underline{v}_i + \delta^{T+1} v_i' + ...\right] \end{aligned}$$

Subgame II$_i$. (suppose there are $T' \leq T$ periods left)

$$\begin{aligned} i \text{ follows strategy} &: \quad (1-\delta^{T'})\underline{v}_i + \delta^{T'} v_i' \\ i \text{ deviates} &: \quad (1-\delta)\overline{v}_i + \delta(1-\delta^T)\underline{v}_i + \delta^{T+1} v_i' \end{aligned}$$

Subgame II$_j$. (suppose there are $T' \leq T$ periods left)

$$\begin{aligned} i \text{ follows strategy} &: \quad (1-\delta^{T'})g_i(m^j) + \delta^{T'}(v_i' + \varepsilon) \\ i \text{ deviates} &: \quad (1-\delta)\overline{v}_i + \delta(1-\delta^T)\underline{v}_i + \delta^{T+1} v_i' \end{aligned}$$

Subgame III$_i$,III$_j$. Consider i's payoff to playing the strateg and deviating:

$$\begin{aligned} i \text{ follows strategy} &: \quad v_i' \\ i \text{ deviates} &: \quad (1-\delta)\overline{v}_i + \delta(1-\delta^T)\underline{v}_i + \delta^{T+1} v_i' \end{aligned}$$

The payoffs here are the least i could get if he follows the strategy and the most he could get if he deviates. A small amount of algebra shows that for $\delta \approx 1$, it is best not to deviate. *Q.E.D.*

Remark 1 *The equilibrium constructed in the above proof involves both the "stick" (Phase II) and the "carrot" (Phase III). Often, however, only the stick is necessary. The carrot phase is needed only if the parties punishing in Phase II get less than their min-max payoffs.*

Note that the (perfect) Folk Theorem requires an extra (though relatively mild) assumption, namely that $\dim V^* = I$. The assumption ensures that each player i can, in the event of a deviation, be singled out for punishment. It rules out special games like the one in the following example.

Example 1 Consider the following game with three players: 1 chooses Row, 2 chooses column, and 3 chooses matrix.

$\underline{A}$

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

$\underline{B}$

	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. To min-max i, j and k just need to mis-coordinate. The set of feasible and individually rational payoffs is:

$$V^* = \{(v,v,v) : v \in [0,1]\}$$

Claim. For any $\delta \in (0,1)$, there is no SPE of $G^\infty(\delta)$ with average payoff less than $\frac{1}{4}$.

Proof. Fix δ, and let $x = \inf\{v : (v,v,v) \text{ is a SPE payoff}\}$. The first step of the proof is to show that it (v,v,v) is an SPE payoff then:

$$v \geq (1-\delta)\frac{1}{4} + \delta x.$$

To see this, let $(\sigma_1, \sigma_2, \sigma_3)$ denote the first period mixtures used in a SPE with payoff v. Then there must exist either two players with $\sigma_i(A) \geq \frac{1}{2}$ or two players with $\sigma_i(B) \geq \frac{1}{2}$. Assume the former, and suppose $\sigma_1(A), \sigma_2(A) \geq \frac{1}{2}$.

Suppose 3 plays A in the first period and then follows his equilibrium strategy. His payoff from this will be *at least* $(1-\delta)\frac{1}{4} + \delta x$ — since $\sigma_1(A), \sigma_2(A) \geq \frac{1}{2}$, he gets at least $\frac{1}{4}$ in the first period, and over all future periods he must average at least x, given that continuation play will be an SPE. Since this deviation is unprofitable, the initial claim holds.

But now we're essentially done, since

$$x = \inf_{v \text{ is } SPE} v \geq (1-\delta)\frac{1}{4} + \delta x \qquad \Longrightarrow \qquad x \geq \frac{1}{4}.$$

The problem is that no individual can be punished for deviating without punishing everyone, so there is no way to "reward" the punishers. *Q.E.D.*

4 Comments

There are many variations and strengthenings of the perfect monitoring folk theorem.

1. The assumption that $\dim V^* = I$ (full-dimensionality) can be relaxed. Abreu, Dutta and Smith (1994) show that what is really needed is that no two players have payoffs that are affine transformations of each other — in this case it is always possible to single out individuals for punishment.

2. The assumption of strict individual rationality can also be relaxed.

3. There are some subtle issues involving randomization. The proof above assumes that there are profiles a, a^i, m^i to achieve the given payoff vectors in each period, and that deviations from these profiles are observable. There are several ways to justify this. The simplest is to assume players can carry out a public coin toss (public randomization) before each period, and that mixed strategies are observable. Randomization can also be replaced by a deterministic variation in play over time (which is more tricky).

4. Benoit and Krishna (1986) prove a folk theorem for finitely repeated games. Clearly this can't be done in the prisoners' dilemma where backward induction says that (D, D) will be played in each period. The stage game must have multiple nash equilibria to allow for rewards and punishments towards the end of the game.

5. There are also folk theorems for games where some players are long-run (infinite-horizon) and others are short-run (myopic), for games with overlapping generations of players, and for games where players face a new opponent randomly drawn from the population in each period (Kandori, 1992; Ellison, 1994).

References

[1] Abreu, D. (1988) "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica.*

[2] Abreu, D., P. Dutta and L. Smith (1994) "The Folk Theorem for Repeated Games: A NEU Condition," *Econometrica*, 62(4), 939-948.

[3] Benoit, J.P. and V. Krishna (1986) “Finitely Repeated Games,” *Econometrica.*

[4] Ellison, G. (1994) “Cooperation in the Repeated Prisoners’ Dilemma with Anonymous Random Matching,” *Review of Economic Studies.*

[5] Fudenberg, D. and E. Maskin (1986) “The Folk Theorem with Discounting and with Incomplete Information,” *Econometrica.*

[6] Kandori, M. (1992) “Social Norms and Community Enforcement,” *Review of Economic Studies.*