A Folk Theorem for Stochastic Games with Infrequent State Changes*

(Very Preliminary)

Marcin Pęski**

Thomas Wiseman***

April 2011

Abstract

Fudenberg and Yamamoto (forthcoming) and Hörner, Sugaya, Takahashi, and Vieille (forthcoming) study dynamic stochastic games with finite states, and give conditions on monitoring under which a folk theorem holds as players become very patient (so that players discount vanishingly little both the time until the next period and the expected time until the next state transition). Here, we consider the case what happens as the length of a period shrinks, but players' rate of time discounting remains fixed. Now the discounting between periods shrinks to zero in the limit, but the discounting of the expected time until a state transition does not. Our main result is a folk theorem that holds under Fudenberg, Levine, and Maskin's (1994) monitoring conditions. Unlike FY and HSTV, we do not require that the stochastic game be irreducible.

---

* We thank Drew Fudenberg and Johannes Hörner for helpful discussions

** Department of Economics, University of Texas at Austin, Austin, TX 78712; phone: (512)475-6843, fax: (512)471-3510, email: mpeski@gmail.com.

*** Department of Economics, University of Texas at Austin, Austin, TX 78712; phone: (512)475-8516, fax: (512)471-3510, email: wiseman@austin.utexas.edu.

1. Introduction

T

Literature review.

Examples.

2. Model

There are N expected-utility maximizing players playing an infinite-horizon stochastic game. The time between periods is given by $\Delta > 0$, and all players discount the future at rate $r > 0$ (so that the per-period discount rate is $e^{-r\Delta} \equiv \delta$.) There is a finite set S of K payoff-relevant states of the world. In each different state s, the stage game $G(s)$ has different payoffs but the same set of action profiles $A = A_1 \times A_2 \times \ldots \times A_N$, where A_i is the (finite) set of actions for player i. (The assumption that the set of actions is independent of the state is without loss of generality – if, for example, a player has fewer actions in one state than in another, we can add a "dummy" action identical to an existing action.) In each period t, the state s is publicly observed. Then each player i chooses an action $a_i \in A_i$ and observes a public signal y drawn from the distribution $\rho(a, s)$ with finite support Y, where a is the profile of actions of all players. Player i's payoff when his action is a_i and signal y is realized in state s is $u_i(a_i, y, s)$; let $g_i(a, s)$ be player i's expected payoff when action profile a is played in state s:

$$g_i(a, s) \equiv \sum_{y \in Y} [\rho(a,s)(y)] u_i(a_i, y, s) .$$

Denote by $g(a, s)$ the vector of expected payoffs for each player. At the end of period in which when action profile a is played in state s, the probability that the state changes to state $s' \neq s$ is equal to $\Delta \hat{\gamma}(s'; a, s)$: the constant rate $\hat{\gamma}(s'; a, s)$ multiplied by the period length Δ. That is, the transition probability per unit of *time* is fixed, so (for small Δ) the transition probability per *period* is proportional to the length of the period. With the remaining probability $1 - \Delta\hat{\gamma}(a,s)$, where

$$\hat{\gamma}(a,s) \equiv \sum_{s'\neq s} \hat{\gamma}(s';a,s),$$

the state does not change. Note that when the period length Δ is close to zero, then the discount rate per period δ is approximately equal to $1 - r\Delta$, and so the probability of transition from state s to state s' given action profile a in each period, $\Delta\hat{\gamma}(s';a,s)$, is approximately equal to $(1-\delta)\frac{1}{r}\hat{\gamma}(s';a,s)$. Since we will focus on the limiting case as $\Delta \to 0$, for notational simplicity we will re-specify the per-period transition probabilities as $(1-\delta)\gamma(s';a,s)$, where

$$\gamma(s';a,s) \equiv \tfrac{1}{r}\hat{\gamma}(s';a,s).$$

Similarly, we re-specify the probability that no transition occurs as $1-(1-\delta)\gamma(a,s)$, where

$$\gamma(a,s) \equiv \sum_{s'\neq s} \gamma(s';a,s).$$

We will then consider the limit as $\delta \to 1$.

All these definitions extend in a natural way to mixed actions. This structure is common knowledge. We assume that a public randomization device is available to the players.

The set of *public histories* in period t is equal to $H_t \equiv Y^{t-1} \times S^t$, with element $h_t = (s_1, y_1, \ldots, y_{t-1}, s_t)$, where s_t denotes the state at the beginning of period t and y_t denotes the public signal realized at the end of period t. Player i's *private history* in period t is $h_t^i = (s_1, y_1, a_{i,1}, \ldots, y_{t-1}, a_{i,t-1}, s_t)$, where $a_{i,t}$ is player i's action in period t; $H_t \equiv (Y \times A_i)^{t-1} \times$

S^t is the set of such private histories. Define $H \equiv \bigcup_t H_t$ and $H^i \equiv \bigcup_t H_t^i$. A strategy for player i is a mapping α_i : $H^i \to \Delta A_i$. A public strategy for player i is a mapping σ_i : $H \to \Delta A_i$. Let Σ_i and Σ_i^P, respectively, denote the set of strategies and the set of public strategies for player i; let Σ and Σ^P denote the sets of strategy profiles and of public strategy profiles, respectively. Given a profile of strategies $\alpha \in \Sigma$ and an initial state $s \in S$, player i's expected payoff in the dynamic game, $v^\delta(a,s)$, is given by

$$v^\delta(a,s) = (1-\delta)E\sum_{t=1}^{\infty}\delta^{t-1}g(a_t,s_t),$$

where the expectation is taken with respect to the distribution over actions and states induced by the strategy α and initial state s. For each public strategy $\sigma \in \Sigma^P$ and public history $h \in H$, the continuation payoffs $v^\delta(\sigma,h)$ are calculated in the usual way.

Let $u = (u_s)_{s\in S} \in \mathbf{R}^{N\times K}$ specify the vector of continuation payoffs as a function of next period's state. For each such u, current state s, and action profile a, we define a vector of payoffs

$$\psi^\delta(a,s,u) \equiv \frac{1}{1+\delta\gamma(a,s)\left[g(a,s)+\delta\sum_{s'\neq s}\gamma(s';a,s)u_{s'}\right]}.$$

We refer to $\psi^\delta(a,s,u)$ as the *pseudo-instantaneous payoff* (or just the pseudo-payoff) from playing action profile a in state s, given continuation payoffs u. To motivate the name, we note that the expected payoff in this situation is given by

$$\begin{aligned}
&(1-\delta)g(a,s)+\delta\sum_{s'\neq s}(1-\delta)\gamma(s';a,s)u_{s'}+\delta[1-(1-\delta)\,((a,s)]u_s\\
&=(1-\delta)(1+\delta\gamma(a,s))\psi^\delta(a,s,u)+[1-(1-\delta)(1+\delta\gamma(a,s))]u_s\\
&=\beta(a,s)\psi^\delta(a,s,u)+[1-\beta(a,s)]u_s\ ,
\end{aligned}$$

where $\beta(a,s) \equiv (1-\delta)(1+\delta\gamma(a,s))$.

Thus, we can represent the expected payoff as a convex combination of the pseudo-instantaneous payoff and the continuation payoff if no state transition occurs. This definition will be useful later.

2.1 Feasible and individually rational payoffs

Define the (convex hull of) the set of feasible payoffs in initial state *s*, $\hat{V}^{\delta}(s)$, as

$$\hat{V}^{\delta}(s) \equiv \text{co}\left\{v^{\delta}(\sigma,s) : \sigma \in \Sigma^{P}\right\}.$$

Note that, in contrast to the setting in FY and HSTV, the set of feasible payoffs $\hat{V}^{\delta}(s)$ varies with the state even in the limit as δ approaches 1: the payoffs to the stage game vary with the state, and both the discount rate and the rate of transition between states are fixed per unit of time as δ grows.

For each player *i* and state *s*, we define player *i*'s *public minmax payoff* in state *s*, $m_i^{\delta}(s)$, as

$$m_i^{\delta}(s) \equiv \min_{\sigma_{-i} \in \Sigma_{-i}^{P}} \max_{\sigma_i \in \Sigma_i^{P}} v^{\delta}((\sigma_i, \sigma_{-i}), s) .$$

Next, define the (convex hull of the) set of feasible and individually rational payoffs in initial state *s*, $V^{\delta}(s)$, as

$$V^{\delta}(s) \equiv \text{co}\left\{v^{\delta}(\sigma,s) : \sigma \in \Sigma^{P} \text{ and } v_i^{\delta}(\sigma,h) \ge m_i^{\delta}(s(h)) \ \forall i \in N, \forall h \in H\right\}$$

Note that $V^{\delta}(s)$ typically is not equivalent to the set $\left\{v \in \hat{V}^{\delta}(s) : v_i \geq m_i^{\delta}(s) \ \forall i \in N\right\}$.

EXAMPLE.

Finally, let $V_{\varepsilon}^{\delta}(s) \subseteq V^{\delta}(s)$ denote the set of payoffs that are feasible and that after every history yield each player a payoff at least ε greater than his minmax payoff:

$$V_{\varepsilon}^{\delta}(s) \equiv \text{co}\left\{v^{\delta}(\sigma, s) : \sigma \in \Sigma^{P} \text{ and } v_i^{\delta}(\sigma, h) \geq m_i^{\delta}(s(h)) + \varepsilon \ \forall i \in N, \forall h \in H\right\}.$$

Later we will consider conditions under which $V_{\varepsilon}^{\delta}(s)$ converges to $V^{\delta}(s)$ as ε shrinks to 0. (Note that even in a standard repeated game, with only one state, that convergence may not occur – for example, if the profile of minmax profiles is Pareto optimal.)

2.2 Identifiability

Definitions of pairwise full rank and individual full rank are the same as in FLM.

Identifiability Condition: The identifiability conditions of FLM's Theorem 6.2 are satisfied in every state.

3. Results

The *uniform interior condition* requires that there exists some positive d such that in each state s there is a payoff vector in the interior of $V_{\varepsilon}^{\delta}(s)$ that lies at a distance of at least r from the boundary of $V_{\varepsilon}^{\delta}(s)$. For any vector $v \in \mathbf{R}^N$ and scalar $d > 0$, let $B(v, d)$ denote the closed ball centered at v with radius d.

Uniform Interior Condition: Let $\varepsilon > 0$ be given. For high enough values of δ, there exists $d > 0$ and $\underline{v}_s \in V_{\varepsilon}^{\delta}(s)$ for each state s such that $B(v_s, d) \subseteq V_{\varepsilon}^{\delta}(s)$.

With that definition, we can state our main result, a folk theorem:

Theorem 1. Let $\varepsilon > 0$ be given. Suppose that the Uniform Interior Condition holds for ε, and that the Identifiability Condition holds. Then there exists $\delta^* < 1$ such that the following holds: for any initial state s, any $\hat{v} \in V_\varepsilon^\delta(s)$, and any $\delta \geq \delta^*$, there exists a perfect public equilibrium strategy profile σ such that $v^\delta(\sigma, s) = \hat{v}$.

The proof is based on the techniques in the proof of FLM's folk theorem for games with imperfect public monitoring. That proof shows that any smooth set of payoffs W strictly in the interior of the feasible and individually rational set can be attained in equilibrium – a key step is to show that any payoff on the boundary of W can be achieved as the weighted average of a stage-game payoff in the current period that lies outside W (thus the requirement that W is strictly in the interior of the feasible set) and expected continuation payoffs that lie in W. Here, we want to do something similar, with pseudo-instantaneous payoffs taking the place of the stage-game payoffs. We have to do some work to ensure that there is a pseudo-payoff outside $V_\varepsilon^\delta(s)$ in each direction (and, of course, for every state). As a building block, we considers strategies that, after each period, with positive probability switch permanently to a strategy that yields payoffs in the middle of $V_\varepsilon^\delta(s)$. Later (in Lemma 3), we will show that any payoff in $V_\varepsilon^\delta(s)$ can be attained with such strategies.

3.1 η-Modified Strategies

Given ε, suppose that the Uniform Interior Condition holds, and for each state s fix a payoff vector v_s as in the statement of that condition. Let σ^s denote a strategy that yields $\underline{v}_s$ in state s (and at least $m_i^\delta(s(h)) + \varepsilon$ for all players after every history, by the definition of $V_\varepsilon^\delta(s)$.)

Fix $\eta \in (0, 1)$. For any public strategy profile σ, construct the η-*modified* strategy $\sigma^{\eta}(\sigma)$ in the following way: there are $K + 1$ regimes, regime 0 and regime s for each state s. Play begins in regime 0. After any period in which play was in regime 0, the state was s, and the state did not change, the regime switches to s with probability $\eta(1 - \delta)$, and otherwise remains in regime 0. After any period in which play was in regime 0, the state was s, and the state transitions to $s' \neq s$, the regime switches to s' with probability η, and otherwise remains in regime 0. Regime s is absorbing for all s. In regime 0, strategy $\sigma^{\eta}(\sigma)$ specifies exactly the same play as strategy σ. If play switches to regime s after period t, then in each period $t' > t$, after t'-period history $(h_t, h_{t'})$, strategy $\sigma^{\eta}(\sigma)$ follows the prescription of strategy $\sigma^{s}(h_{t'})$. That is, the continuation payoff after the switch to regime s is $\underline{v}_s$.

Next, define the set of individually rational payoffs achievable using η-modified strategies, starting from state s:

$$V^{\delta,\eta}(s) \equiv \text{co}\left\{v^{\delta}(\sigma^{\eta}(\sigma), s) : \sigma \in \Sigma^{P} \text{ and } v_i^{\delta}(\sigma^{\eta}(\sigma), h) \geq m_i^{\delta}(s(h)) \ \forall i \in N, \forall h \in H\right\}.$$

Our first lemma establishes that for any state s and any payoff $v \in V^{\delta,\eta}(s)$, we can pick any direction and find an action profile a and continuation payoffs in $V^{\delta,\eta}(s')$ (where s' is the state in the next period) such that the pseudo-payoff from a is at least $\delta\eta d$ farther in that direction than any payoff in $V^{\delta,\eta}(s)$.

Lemma 2. Suppose that the Uniform Interior Condition holds. Choose any state s and any vector $\lambda \in \mathbf{R}^N$ such that $\|\lambda\| = 1$. Then there exist an action profile a and continuation payoffs $(u_{s'})_{s' \in S}$, with $u_{s'} \in V^{\delta,\eta}(s')$ for all states s', such that

$$\lambda \cdot \psi^{\delta}(a, s, u) \geq \delta\eta d + \max_{v \in V^{\delta,\eta}(s)} \lambda \cdot v.$$

Proof of Lemma 2: Pick $v^* \in \max_{v \in V^{\delta,\eta}(s)} \lambda \cdot v$ and corresponding strategy profiles σ and $\sigma^\eta(\sigma)$ such that $v^\delta(\sigma^\eta(\sigma), s) =$ v^*, and $v_i^\delta(\sigma^\eta(\sigma), h) \geq m_i^\delta(s(h))$ for each player i and public history h. For simplicity, suppose that σ is a pure strategy. Let $u_{s'}(y)$ denote the continuation payoff in state s' under σ when the realization of the public signal is y, and let $u_{s'}(a) \equiv \sum_y \rho(a,s')(y) u_{s'}(y)$ be the expected value of $u_{s'}(y)$ when a is played. (Note that these are also the continuation payoffs under $\sigma^\eta(\sigma)$ if play stays in regime 0.)

We can write v^* as

$$v^* = (1-\delta) g(a,s) + \delta \sum_{s' \neq s} (1-\delta)\gamma(s';a,s)\left[(1-\eta)u_{s'}(a) + \eta \underline{v}_{s'}\right]$$
$$+\delta[1-(1-\delta)\gamma(a,s)]\left[(1-\eta(1-\delta))u_s(a) + \eta(1-\delta)\underline{v}_s\right]$$

$$= \beta(a,s)\psi^\delta(a,s,(1-\eta)u(a) + \eta \underline{v}) + [1-\beta(a,s)]\left[(1-\eta(1-\delta))u_s(a) + \eta(1-\delta)\underline{v}_s\right],$$

where for each state s', $\underline{v}_{s'}$ is the payoff vector in the interior of $V_\varepsilon^\delta(s)$ described in the Uniform Interior Condition, and $\underline{v} \equiv (\underline{v}_{s'})_{s' \in S}$ is the vector of those payoffs. Thus,

$$\begin{aligned} &\beta(a,s)\lambda \cdot \left[\psi^\delta(a,s,(1-\eta)u(a) + \eta \underline{v}) - v^*\right] \\ &= [1-\beta(a,s)]\left[(1-\eta(1-\delta))\lambda \cdot \left[v^* - u_s(a)\right] + \eta(1-\delta)\lambda \cdot \left[v^* - \underline{v}_s\right]\right] \end{aligned} \quad \textbf{(1)}$$

For each state $s' \neq s$, choose $u_{s'}^* \in \max_{v \in V^{\delta,\eta}(s')} \lambda \cdot v$. Then since $u_{s'}(a) \in V^{\delta,\eta}(s')$,

$$\lambda \cdot \left[u_{s'}^* - u_{s'}(a)\right] \geq 0, \text{ and } \lambda \cdot \left[u_{s'}^* - \underline{v}_{s'}\right] \geq d \text{ by the Uniform Interior Condition.}$$

Thus, letting $u^* \equiv (u_{s'}^*)_{s' \in S}$, we have

$$\lambda \cdot \left[\psi^{\delta}(a,s,u^*) - \psi^{\delta}(a,s,(1-\eta)u(a)+\eta\underline{v}) \right]$$

$$\geq \frac{\delta\gamma(a,s)}{1+\delta\gamma(a,s)} \eta\lambda \cdot \left[u^*_{s'} - \underline{v}_{s'} \right] \geq \frac{\delta\gamma(a,s)}{1+\delta\gamma(a,s)} \eta d \; .$$

Similarly, $\lambda \cdot \left[v^* - u_s(a) \right] \geq 0$, and $\lambda \cdot \left[v^* - \underline{v}_s \right] \geq d$, so we have

$$\lambda \cdot \left[\psi^{\delta}(a,s,u^*) - v^* \right]$$

$$= \lambda \cdot \left[\psi^{\delta}(a,s,u^*) - \psi^{\delta}(a,s,(1-\eta)u(a)+\eta\underline{v}) \right] + \lambda \cdot \left[\psi^{\delta}(a,s,(1-\eta)u(a)+\eta\underline{v}) - v^* \right]$$

$$\geq \frac{\delta\gamma(a,s)}{1+\delta\gamma(a,s)} \eta d + \lambda \cdot \left[\psi^{\delta}(a,s,(1-\eta)u(a)+\eta\underline{v}) - v^* \right]$$

$$\geq \frac{\delta\gamma(a,s)}{1+\delta\gamma(a,s)} \eta d + \frac{(1-\delta)(1-\beta(a,s))}{\beta(a,s)} \eta d \quad \text{(using Equation 1)}$$

$$= \frac{\delta\gamma(a,s) + 1 - (1-\delta)(1+\delta\gamma(a,s))}{1+\delta\gamma(a,s)} \eta d = \eta d$$

Q.E.D.

PICTURES.

Our next lemma shows that for any state s and any $\varepsilon > 0$, the set $V_{\varepsilon}^{\delta}(s)$ is contained in the set $V^{\delta,\eta}(s)$ of payoffs achieved through η-modified strategies, if the probability η of a regime change is small enough.

Lemma 3. Fix a state $s \in S$ and $\varepsilon > 0$. Then there exists $\overline{\eta}$ such that $V_{\varepsilon}^{\delta}(s) \subseteq V^{\delta,\eta}(s)$ for all $\eta \leq \overline{\eta}$.

Proof of Lemma 3: First, define M as the highest absolute value of the expected stage-game payoff to any player from any action in any state: $M \equiv \max_{i,a,s} g_i(a, s)$. Similarly, let $\Gamma \equiv \max_{,a,s} \gamma(a,s)$ be the highest possible total transition rate away from any state.

Note that for any strategy σ, the probability in any period, under the η-modified strategy $\sigma^{\eta}(\sigma)$, that play switches from regime 0 to some regime s is η if a state transition occurs and $\eta(1-\delta)$ otherwise. That probability is bounded above by

$$\Gamma(1-\delta)\eta + [1-\Gamma(1-\delta)]\eta(1-\delta) = (1-\delta)\eta(1+\delta\Gamma) < (1-\delta)\eta(1+\Gamma).$$

Thus,

$$\left| v^{\delta}(\sigma,s) - v^{\delta}(\sigma^{\eta}(\sigma),s) \right| \le M \sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)\eta(1+\Gamma)\big(1-(1-\delta)\eta(1+\Gamma)\big)^{t-1}$$

$$= M(1-\delta)\eta(1+\Gamma)\frac{1}{1-\delta\big(1-(1-\delta)\eta(1+\Gamma)\big)}$$

$$= M\eta(1+\Gamma)\frac{1}{1+\delta\eta(1+\Gamma)} \le M\eta(1+\Gamma).$$

It follows that for $\eta \le \frac{\varepsilon}{M(1+\Gamma)}$, $V_{\varepsilon}^{\delta}(s) \subseteq V^{\delta,\eta}(s)$ for all s. *Q.E.D.*

Now we can prove Theorem 1.

Proof of Theorem 1: Follows FLM's proof of Theorem 6.2.

4. Summary and Discussion

This paper

References

Fudenberg, D., Levine, D., and Maskin, E. (1994). "The Folk Theorem with Imperfect Public Information," *Econometrica*, **62**, pp. 997-1039.

Fudenberg, D. and Yamamoto, Y. (forthcoming). "The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring," *Journal of Economic Theory*.

Hörner, J., Sugaya, T., Takahashi, S., and Vieille, N. (forthcoming). "Recursive Methods in Discounted Stochastic Games: An Algorithm for $\delta \to 1$ and a Folk Theorem," *Econometrica*.