International Journal of Game Theory (1989) 18: 189–194

Existence of Stationary Equilibrium Strategies in Non-zero Sum Discounted Stochastic Games With Uncountable State Space and State-Independent Transitions

By T. Parthasarathy and S. Sinha[1]

Abstract: Non-zero sum discounted stochastic games with uncountable state space and state independent transitions have stationary equilibrium strategies.

Key words: Stochastic games, uncountable state space, state independent transitions, stationary equilibrium strategies.

1 Introduction

A non-zero sum N-person stochastic game is determined by $S, A_i, r_i, q; i = 1, 2, \ldots, N$. Here $S = [0, 1]$ is the state space, $A_i = \{1, 2, \ldots, k_i\}$, finite sets describing actions available to player i, $r_i(s, .)$, the immediate pay-off to player i, which depends on the current state s and the actions of the players. As a consequence of the actions chosen by the N-players:

(a) Player i receives $r_i(s,.)$.

(b) The system moves to a new state s' according to the transitions law $q(s'/s, .)$.

(c) The whole process is repeated from the new state s'.

The game is played over the infinite future. Every player wants to maximize his accumulated rewards, the rewards are accumulated with a discount factor $\beta \in [0, 1)$, that is, on the n-th day, the pay-off to player i is $\beta^{n-1} r_i^n(s, .)$, where r_i^n stands for the reward on the n-th day.

A strategy π_i for player i is a sequence $(f_1, f_2, \ldots, f_k, \ldots)$ where f_k specifies the actions to be chosen on the k-th day, depending on the past history h_{k-1}. Note that f_k is a probability distribution on A_i given h_{k-1} and $f_k(E|h_{k-1})$ is measurable in

[1] T. Parthasarathy and S. Sinha, Indian Statistical Institute, New Delhi-110016, India.

h_{k-1} for every Borel set E. A strategy π_i is called stationary if there exists a Borel map $f : S \to P_{A_i}$ such that $f_k \equiv f$ for all k, where P_{A_i} stands for the space of probability vectors on A_i.

Let $(\pi_i : 1 \leqslant i \leqslant N)$ be an N-tuple of strategies for the N players. Associated with each initial state s, the total discounted expected pay-off to player i is denoted by $I_i(\pi_1, \pi_2, \ldots, \pi_N)(s) = \sum_{n=1}^{\infty} \beta^{n-1} r_i^n(\pi_1, \ldots, \pi_N)(s)$. An N-tuple $(\pi_i^* : 1 \leqslant i \leqslant N)$ is an equilibrium in the sense of Nash if

$$I_i(\pi_1^*, \ldots, \pi_i^*, \ldots \pi_N^*)(s) \geqslant I_i(\pi_1^*, \ldots, \pi_{i-1}^*, \pi_i, \pi_{i+1}^*, \ldots, \pi_N^*)(s)$$

for all $s \in S$, for all π_i and for all $i = 1, 2, \ldots, N$. The law of transition $q(ds'|s, .)$ is called state independent (denoted by SIT) if $q(ds'|s_1, .) = q(ds'|s_2, .)$ for all $s_1, s_2 \in S$. In other words, transition probability is independent of the initial state. For notational convenience, if the transition is SIT, we denote it by $q(ds'|.)$ only.

Our problem is to find a Nash equilibrium stationary strategy for every player under SIT assumption.

For related results of this problem, we refer to [1, 5, 6, 7, 8, 9, 10, 11].

Assumptions

(i) r_i is bounded by some constant K for all $i = 1, \ldots, N$.

(ii) r_i is continuous $S \times A_1 \times \ldots \times A_N$ for all $i = 1, 2, \ldots, N$.

(iii) There exists a fixed nonatomic measure μ such that $q(.|.) \ll \mu$, that is transition probabilities are absolutely continuous with respect to μ.

(iv) We shall use the expected pay-off $I_i(\pi_1, \ldots, \pi_N)(s) = (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} r_i^n(\pi_1, \ldots \pi_N)(s)$ so that I_i is bounded by K for all i.

2 Main Theorem

Under assumptions (i), (ii), (iii), and (iv), a SIT stochastic game under discounted pay-off has stationary Nash equilibrium strategies.

In order to prove this theorem we need some preliminaries. Denote by F the set of all possible measurable pay-off functions, each component being bounded by K. That is,

$$F = \{f : S \to R^n : f = (f_1, f_2, \ldots, f_N) \text{ and } |f_k(s)| \leqslant K$$
$$\text{for all } k = 1, 2, \ldots n \text{ and } s \in S\}.$$

Given $f \in F$, $s \in [0, 1]$, define an N-person non-zero sum finite game $G_f(s)$, where the pay-off to the k-th player is given by

$$(1-\beta) r_k(s, i, j, \ldots) + \beta \int f_k(s') dq(s' | i, j, \ldots), \quad k = 1, 2, \ldots, N.$$

For every s, this finite game has a Nash equilibrium point. Let

$$N_f(s) = \{h(s) = (h_1(s), \ldots, h_N(s)) : h_k(s) \text{ is a Nash equilibrium pay-off to the } k\text{-th player where } k = 1, 2, \ldots N\}.$$

Define $\bar{N}_f(s)$ = convex-hull of $N_f(s)$ for each s. Let N_f and $\bar{N}_f$ be all measurable selections from $N_f(s)$ and $\bar{N}_f(s)$ respectively. One can show that $N_f \neq \phi$ as well as $\bar{N}_f \neq \phi$ using [3].

We shall endow F as well as the space of stationary strategies with weak* ($=w^*$) topology induced by the fixed non-atomic measure μ. We first prove the following basic lemma.

Lemma 1: The correspondence $\Psi : F \to 2^F$ with $\Psi(f) = \bar{N}_f$ is closed, convex-valued and w^*-upper semi-continuous.

Proof: Clearly $\bar{N}_f$ is convex-valued since $\bar{N}_f(s)$ is convex for each s.

Let $f_k \in F$, $f_k \xrightarrow{w^*} f$, $\phi_k \in \bar{N}_{f_k}$ for each k and $\phi_k \xrightarrow{w^*} \phi$. To show Ψ is upper semi-continuous it is enough to prove that $\phi \in \bar{N}_f$. In fact it is enough to show that $\phi(s) \in \bar{N}_f(s)$ almost everywhere, since it is not going to matter if we change the value of a function on a set of measure zero.

Clearly some subsequence of convex combinations of ϕ_k converges to ϕ almost everywhere. We continue to call the subsequence as ϕ_k without loss of generality. Since $f_k \to f$ in w^*-topology, $G_{f_k}(s) \to G_f(s)$ for every s. Hence $\overline{\lim}\, N_{f_k}(s) \subset N_f(s)$ or $\overline{\lim}\, \bar{N}_{f_k}(s) \subset \bar{N}_f(s)$ for every s. This means given $\epsilon > 0$, there is a k_0 such that for all $k \geqslant k_0$, $\bar{N}_{f_k}(s) \subset \epsilon$-neighbourhood of $\bar{N}_f(s)$, in other words, $\phi_k(s) \in \epsilon$-neighbourhood of $\bar{N}_f(s)$. That is $\phi(s) \in \epsilon$-neighbourhood of $\bar{N}_f(s)$. Since ϵ is arbitrary $\phi(s) \in \bar{N}_f(s)$

almost everywhere. Thus $\phi \in \bar{N}_f$ which proves Ψ is upper semicontinuous. The same proof also shows that Ψ_f is closed for each f. This terminates the proof of the lemma.

Lemma 2: There exists a $g_0 \in F$ such that $g_0 \in \bar{N}_{g_0}$.

Proof: Observe F is compact and metrizable in the w^*-topology. Now use Lemma 1 to invoke Glicksberg's fixed point theorem which yields the desired result.

Note that $g_0 \in \bar{N}_{g_0}$ almost everywhere, modify g_0 on a null set to make it measurable and $g_0 \in \bar{N}_{g_0}$ everywhere. This terminates the proof of Lemma 2.

Lemma 3: There exists a g^* such that $g^* \in N_{g^*}$ under the SIT assumption.

Proof: By Lemma 2, there is a $g_0 \in \bar{N}_{g_0}$. The idea is to replace g_0 by a g^* such that (1) $g^* \in N_{g_0}$ and (2) $N_{g^*} = N_{g_0}$.

For every s consider the game $G_{g_0}(s)$. Here the pay-off to player k is $(1-\beta)r_k(s, i, j, \ldots) + \beta \int g_0(.)dq(.|i, j, \ldots)$. Since $g_0(s) \in \bar{N}_{g_0}(s) =$ convexhull of $N_{g_0}(s)$, it follows that $\int g_0(.)dq(.|i, j, \ldots) \in \text{Conv} \int N_{g_0} h d\mu$, where h is the density of $q(.|i, j, \ldots)$ with respect to μ. By Liapunov, one can conclude that there is a selection g^* from N_{g_0} such that $\int g_0(.)dq(.|i, j, \ldots) = \int g^* dq(.|i, j, \ldots)$. Observe that finite game $G_{g_0}(s)$ does not change if we replace g_0 by g^* and consequently, $N_{g^*}(s) = N_{g_0}(s)$ for every s. Hence $g^* \in N_{g^*}$ everywhere. This terminates the proof of Lemma 3.

Lemma 3 precisely means that for every s, $g^*(s)$ is an equilibrium pay-off to the finite game

$$(1-\beta)r_k(s, i, j, \ldots) + \beta \int g^*(.)dq(.|i, j, \ldots).$$

Next we show that g^* is an equilibrium pay-off to the stochastic game under the assumptions (i) through (iv) and SIT. We also show the existence of stationary equilibrium strategies for the N-players corresponding to the Nash pay-off g^*.

Proof of the Main Theorem

From Lemma 3, we have a $g^* \in N_{g^*}$ everywhere. Fix the finite game $G_{g^*}(s)$ throughout the rest of the argument. Consider the following set E

$E = \{(s, p, \mu_1, \mu_2, \ldots \mu_N)$: for each s, p is a feasible Nash pay-off and $\mu_1, \mu_2, \ldots, \mu_N$ Nash equilibrium points that induce p corresponding to the finite game $G_{g^*}(s)\}$.

One can check that E is a Borel set and its (s, p) sections are nonempty and compact. Hence from Kunugui-Novikoff's theorem [2] we can find a Borel map σ^* such that $(s, p, \sigma^*(s, p)) \in E$. Existence of σ^* can also be seen from [4]. In other words $\sigma^*(s, p)$ is an equilibrium point yielding a Nash pay-off p at state s. Since $g^*(s) \in N_{g^*}(s)$, $\sigma^*(s, g^*(s))$ yields the required stationary Nash equilibrium points for the stochastic game under consideration. This terminates the proof of the main theorem.

Remark: In [6] we show that the present result can be strengthened by dropping the nonatomicity assumption on μ. For more general results see [5]. However we are unable to show the existence of stationary equilibria in general. This problem appears to be a very difficult open problem.

Acknowledgements: Our thanks are due to Professor J. F. Mertens, an anonymous referee and an associate editor for several useful comments.

References

1. Himmelberg CJ, Parthasarathy T, Raghavan TES, van Vleck FS (1976) Existence of p-equilibrium and optimal stationary strategies in stochastic games. Proc Amer Math Soc 60:245–251
2. Himmelberg CJ, Parthasarathy T, van Vleck FS (1976) Optimal plans for dynamic programming problems. Math OR 1:390–394
3. Kuratowski K, Ryll-Nardzewski C (1965) A general theorem on selectors. Bull Acad Poland Sci Ser Math Astronom Phys 13:397–403
4. Mertens JF (1987) A measurable "measurable choice" theorem. CORE Discussion Paper No. 8749
5. Mertens JF, Parthasarathy T (1987) Equilibria for discounted stochastic games. CORE Discussion Paper No. 8750
6. Mertens JF, Parthasarathy T (1988) Non-zero sum stochastic games. Preprint, Indian Statistical Institute

7. Nowak AS (1985) Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space. JOTA 54:591–602
8. Parthasarathy T (1982) Existence of equilibrium stationary strategies in discounted stochastic games. Sankhya, Series A 44:114–127
9. Rieder O (1979) Equilibrium plans for non-zero sum Markov games. In: Moeschlin O, Pallaschke D (eds) Game theory and related topics. North Holland, Amsterdam, Holland
10. Sobel MJ (1973) Continuous stochastic games. Jour Appl Prob 10:497–504
11. Whitt W (1980) Representation and approximation of non-cooperative sequential games. SIAM Jour on Control and Optimization 18:33–48

Received June 1987