

Stochastic Games with Infinitely Many Strategies

Masayuki TAKAHASHI

(Received September 20, 1962)

Introduction

A stochastic game is originated by L. S. Shapley [1]. It is a game consisting of a finite collection of positions among which two players 1 and 2 proceed by steps from position to position according to certain prescribed transition probabilities jointly controlled by them: there is assumed a finite number of positions $1, 2, \dots, N$, and at the position k , a game Γ_k is played, in which Player 1 can choose any strategy among the given m_k pure strategies and Player 2 can also choose any strategy among the given n_k pure strategies. We assume that, at the position k , the players 1 and 2 choose their i -th and j -th alternatives respectively. We also assume that the game stops with the probability $p_{k_0}^{i,j} > 0$ and the game moves to another position l with the probability $p_{kl}^{i,j}$. Thus the game may not be bounded in length. Player 1 takes the gain $g_k^{i,j}$ from Player 2 whenever the pair i, j of pure strategies is chosen at the position k . In Shapley's stochastic game, both players use the so-called stationary strategies, namely at the position k , whenever and by whatever route the position may be reached, the probability distributions of choosing pure strategies are specified. And payments accumulate throughout the course of the play. Let $\vec{\Gamma}_k$ denote the infinite game begun with Γ_k . With the aid of dummy games, Shapley gave a method of finding the value of the stochastic game which is the collection $\vec{\Gamma} = \{\vec{\Gamma}_k, k=1, 2, \dots, N\}$.

As stated above, L. S. Shapley has assumed that at each position there are only finite numbers of pure strategies from among which each player can choose one. Generalizations of his theory to infinite sets of alternatives seem yet to be obtained although he has promised in [1] to discuss them in another place. It seems interesting to generalize his theory to infinite sets of pure strategies or to an infinite number of positions.

After giving preliminary remarks in Section 1, we proceed, in Section 2, to the definition of the stochastic game, at each position of which each player may choose any one out of infinite pure strategies. With suitably imposed conditions on pay-offs and transition probabilities, we show that the stochastic games thus defined are strictly determined (Theorem 1 below). Some considerations are given centering around the ϵ -optimal strategies. The proof of the theorem 1 is carried out with the aid of the dummy games associated with the original stochastic game (Lemmas 3, 4).

In final Section 3, we concern ourselves with the stochastic games with

infinite positions. From this, we can generalize Shapley's stochastic game to the case where non stationary strategies are adopted by both players. However, it is to be noticed that the value of the game itself is irrelevant whether the stationary strategies are adopted or not.

§1. Preliminaries

Let A (resp. B) be an abstract set (not necessarily finite) called a strategic set of Player 1 (resp. Player 2) or merely a strategic set. Any element of A or B is called a pure strategy. A game is defined to be a quintuplet $(A, B, K, \mathfrak{M}, \mathfrak{N})$, where K is a real valued function, called the pay-off of the game, defined on the product set $A \times B$, and $\mathfrak{M}, \mathfrak{N}$ are the mixed strategic sets of the players 1 and 2 respectively which will be defined later. Player 1 (resp. Player 2) chooses a pure strategy a (resp. b) from A (resp. B), each choice being made independently of the other. Then Player 1 gets $K(a, b)$ and Player 2 — $K(a, b)$. Clearly Player 1 wishes to maximize $K(a, b)$ and Player 2 wishes to minimize $K(a, b)$. In this paper, we always assume that $K(a, b)$ is a bounded function. Now we define the intrinsic distance of any two points a_1 and a_2 of A by

$$\delta^K(a_1, a_2) = \sup_{b \in B} |K(a_1, b) - K(a_2, b)|.$$

Similarly the intrinsic distance of any two elements b_1 and b_2 of B is defined by

$$\delta^K(b_1, b_2) = \sup_{a \in A} |K(a, b_1) - K(a, b_2)|.$$

The space A (resp. B) with this distance is a pseudo metric space which we shall denote by a pair (A, δ^K) (resp. (B, δ^K)) and we shall call it a strategy space of Player 1 (resp. Player 2). Let \mathfrak{A} (resp. \mathfrak{B}) be a σ -algebra of subsets of A (resp. B) such that \mathfrak{A} (resp. \mathfrak{B}) contains one point set $\{a\}$ for any $a \in A$ (resp. $\{b\}$ for any $b \in B$). Let μ (resp. ν) be any probability distribution defined on \mathfrak{A} (resp. \mathfrak{B}). (A, \mathfrak{A}, μ) (resp. (B, \mathfrak{B}, ν)) stands for a probability space. Let \mathfrak{M} (resp. \mathfrak{N}) be the set of all such μ (resp. ν). Any $\mu \in \mathfrak{M}$ (resp. $\nu \in \mathfrak{N}$) is called a mixed strategy of Player 1 (resp. Player 2). Let \mathfrak{C} be the smallest σ -algebra of subsets of $A \times B$ which contains the Cartesian product $\mathfrak{A} \times \mathfrak{B}$. We assume that $K(a, b)$ is \mathfrak{C} -measurable over $A \times B$. When Player 1 (resp. Player 2) selects μ (resp. ν), then the expected value of the pay-off $K(a, b)$ is defined by

$$K(\mu, \nu) = \iint_{B \times A} K(a, b) d\mu(a) d\nu(b).$$

It is clear that $K(a, b) = K(\mu_a, \nu_b)$, where μ_a (resp. ν_b) is a probability measure which assigns the probability measure 1 to the point a (resp. b). We shall also write $K(a, \nu)$ instead of $K(\mu_a, \nu)$, and $K(\mu, b)$ instead of $K(\mu, \nu_b)$. A. Wald [2]

has shown that, if one of the reduced strategy spaces, say A , obtained by identifying any two elements of A whose distance is zero, is precompact, and if the pay-off $K(a, b)$ is a bounded \mathbb{C} -measurable function as we assumed, then the reduced B is also precompact, and the game is strictly determined, i.e.

$$(1) \quad \sup_{\mu \in \mathfrak{M}} \inf_{\nu \in \mathfrak{N}} K(\mu, \nu) = \inf_{\nu \in \mathfrak{N}} \sup_{\mu \in \mathfrak{M}} K(\mu, \nu).$$

The common value of (1) is called the value of the game.

The following lemmas shall be needed for our later purpose.

LEMMA 1. Let $\Gamma_1 = (A, B, K, \mathfrak{M}, \mathfrak{N})$ and $\Gamma_2 = (A, B, H, \mathfrak{M}, \mathfrak{N})$ be two games. If the strategy spaces (A, δ^K) and (A, δ^H) are precompact, and c is an arbitrary constant, then (A, δ^{K+cH}) is also precompact.

PROOF. A space is precompact if and only if any sequence of A contains a Cauchy subsequence. Then the statement of our lemma is obvious from the inequality:

$$\delta^{K+cH}(a, a') \leq \delta^K(a, a') + |c| \delta^H(a, a') \quad \text{for any } a, a' \in A.$$

Using (1) and Lemma 1, we have

LEMMA 2. Let the strategy spaces (A, δ^K) and (A, δ^H) be precompact. Then the game $(A, B, K+cH, \mathfrak{M}, \mathfrak{N})$ is strictly determined.

§2. Stochastic Games with Infinitely Many Pure Strategies

We begin with the definition of the stochastic game. Suppose we are given N positions $1, 2, \dots, N$. To each position k we consider a game

$$\Gamma_k = (A_k, B_k, g_k, \mathfrak{M}_k, \mathfrak{N}_k),$$

which we call a component game of the stochastic game which will be defined below. Let Players 1 and 2 choose a pair $(a, b) \in A_k \times B_k$. Then the transition probabilities $p_{kl}(a, b)$ and the stop probability $p_{k0}(a, b)$ are given. Here $p_{kl}(a, b)$ denotes the probability with which the game Γ_k moves to the next game Γ_l when both players choose the pair (a, b) , and $p_{k0}(a, b)$ denotes the probability with which the game stops at this position k when both players choose the pair (a, b) . We assume that $p_{kl}(a, b), p_{k0}(a, b)$ are bounded and \mathbb{C}_k -measurable, so that we can consider the games $(A_k, B_k, p_{kl}, \mathfrak{M}_k, \mathfrak{N}_k)$ and $(A_k, B_k, p_{k0}, \mathfrak{M}_k, \mathfrak{N}_k)$ according to the definition given in §1. The stochastic game Γ is defined as the collection of all $\Gamma_k, p_{kl},$ and p_{k0} for $k, l=1, 2, \dots, N$. In the following we assume that

- (i) the pay-offs $g_k(a, b)$ are bounded and \mathbb{C}_k -measurable,
- (ii) the transition probabilities $p_{kl}(a, b)$ are \mathbb{C}_k -measurable for every $k, l,$ and $\inf_{k, a, b} p_{k0}(a, b) = p_0 > 0,$

- (iii) both players use stationary strategies,
- (iv) payments accumulate throughout the course of the play, and moreover we assume that
- (v) strategy spaces $(A_k, \delta^{g_k}), (A_k, \delta^{p_{kl}})$ are precompact for $k, l=1, 2, \dots, N$.

Now we shall define dummy games associated with Γ (note that Luce, and Raiffa used “truncated games” in [8]). We consider the infinite game $\vec{\Gamma}_k$ beginning with Γ_k . When Player 1 (resp. Player 2) selects mixed strategies μ_j (resp. ν_j) in the component game Γ_j , then the expected value $G_k(\vec{\mu}, \vec{\nu})$ of the gains of Player 1 is given by

$$(2) \quad G_k(\vec{\mu}, \vec{\nu}) = g_k(\mu_k, \nu_k) + \sum_{l=1}^N p_{kl}(\mu_k, \nu_k)g_l(\mu_l, \nu_l) \\ + \sum_{l=1}^N \sum_{l_1=1}^N p_{kl}(\mu_k, \nu_k)p_{ll_1}(\mu_l, \nu_l)g_{l_1}(\mu_{l_1}, \nu_{l_1}) + \dots$$

It is our main purpose to prove that the infinite games ($k=1, 2, \dots, N$) are strictly determined, i.e.

$$\sup_{\mu} \inf_{\nu} G_k(\vec{\mu}, \vec{\nu}) = \inf_{\nu} \sup_{\mu} G_k(\vec{\mu}, \vec{\nu}).$$

Now it is clear that $G_k(\vec{\mu}, \vec{\nu})$ is a solution of the simultaneous linear equations:

$$(3) \quad v_1 = g_1(\mu_1, \nu_1) + \sum_{l=1}^N p_{1l}(\mu_1, \nu_1)v_l \\ v_2 = g_2(\mu_2, \nu_2) + \sum_{l=1}^N p_{2l}(\mu_2, \nu_2)v_l \\ \dots \\ v_N = g_N(\mu_N, \nu_N) + \sum_{l=1}^N p_{Nl}(\mu_N, \nu_N)v_l$$

This consideration leads us to the following

DEFINITION 1. Let $h_k(a, b, \vec{v}) = g_k(a, b) + \sum_{l=1}^N p_{kl}(a, b)v_l$, \vec{v} being any fixed vector (v_1, v_2, \dots, v_N) . The game

$$(A_k, B_k, h_k(a, b, \vec{v}), \mathfrak{M}_k, \mathfrak{N}_k)$$

is called a dummy game, and denoted by $\Gamma_k^d(\vec{v})$.

It is to be noticed that the dummy game $\Gamma_k^d(\vec{v})$ is strictly determined by Lemma 2.

DEFINITION 2. A vector $\vec{v}^* = (v_1^*, v_2^*, \dots, v_N^*)$ is called a principal value vector of the dummy game if its components v_k^* satisfy the following conditions:

$$v_k^* = \sup_{\mu_k} \inf_{\nu_k} \int_{B_k} \int_{A_k} h_k(a, b, \vec{v}^*) d\mu_k(a) d\nu_k(b) = \inf_{\nu_k} \sup_{\mu_k} \int_{B_k} \int_{A_k} \dots,$$

where $h_k(a, b, \vec{v}^*) = g_k(a, b) + \sum_{l=1}^N p_{kl}(a, b)v_l^*$ for every k .

Now we shall show the existence of the principal value vector of the dummy games. Consider the map $T: \vec{v}^0 \rightarrow \vec{v}^1$ (called a value transformation) defined by

$$v_k^1 = \sup_{\mu_k} \inf_{\nu_k} \int_{B_k} \int_{A_k} h_k(a, b, \vec{v}^0) d\mu_k(a) d\nu_k(b)$$

for $k=1, 2, \dots, N$. Define the norm of \vec{v} by

$$\|\vec{v}\| = \max_k |v_k|.$$

Then we have

$$\begin{aligned} (4) \quad \|T\vec{w} - T\vec{v}\| &= \max_k |\text{value of } \Gamma_k^d(\vec{w}) - \text{value of } \Gamma_k^d(\vec{v})| \\ &\leq \max_k [\sup_{a,b} |h_k(a, b, \vec{w}) - h_k(a, b, \vec{v})|] \\ &= \sup_{k, a, b} \left| \sum_{l=1}^N p_{kl}(a, b) (w_l - v_l) \right| \\ &\leq \sup_{k, a, b} \left| \sum_{l=1}^N p_{kl}(a, b) \right| \max_l |w_l - v_l| \\ &= (1 - p_0) \|\vec{w} - \vec{v}\|. \end{aligned}$$

Then, by the principle of contraction (cf. e.g. [3]), there exists a unique \vec{v}^* which satisfies $T\vec{v}^* = \vec{v}^*$, which implies that the dummy game Γ_k^d has a unique principal value vector.

REMARK. From the inequality:

$$\|T^n \vec{v}^0 - \vec{v}^*\| \leq (1 - p_0) \|T^{n-1} \vec{v}^0 - \vec{v}^*\| \leq (1 - p_0)^n \|\vec{v}^0 - \vec{v}^*\|,$$

we see that the error estimates of approximate value vectors $T^n \vec{v}^0$ decrease with increasing n .

Next we shall introduce the notion of ϵ -optimal strategies (cf. e.g. [4]) of the dummy games, where ϵ is a non-negative number.

DEFINITION 3. Let \vec{v}^* be the principal value vector of the dummy games. For each k , any pair $(\mu_k^\epsilon, \nu_k^\epsilon) \in \mathfrak{M}_k \times \mathfrak{N}_k$ is said to be ϵ -optimal strategies of Players 1 and 2 of the dummy games at the position k , when

$$h_k(\mu_k^\epsilon, \nu_k, \vec{v}^*) \geq v_k^* - \epsilon \quad \text{for any } \nu_k \in \mathfrak{N}_k,$$

and

$$h_k(\mu_k, \nu_k^\epsilon, \vec{v}^*) \leq v_k^* + \epsilon \quad \text{for any } \mu_k \in \mathfrak{M}_k.$$

In case where $\epsilon = 0$, the ϵ -optimal strategy is called merely the optimal strategy.

Then we have

LEMMA 3. Any complete set of ϵ -optimal strategies of Players 1 and 2 of dummy games are ϵ/p_0 -optimal strategies of the original infinite games.

PROOF. Denote the expectation of gains of Player 1 of \vec{I}_k by $G_k(\vec{\mu}, \vec{\nu})$ as (2). Then we have

$$\begin{aligned} G_k(\vec{\mu}, \vec{\nu}) - v_k^* &= h_k(\mu_k, \nu_k, \vec{v}^*) - v_k^* + \sum_{l=1}^N p_{kl}(\mu_k, \nu_k) (v_l(\mu_k, \nu_k) - v_l^*) \\ &= h_k(\mu_k, \nu_k, \vec{v}^*) - v_k^* + \sum_{l=1}^N p_{kl}(\mu_k, \nu_k) (\mu_l, \nu_l, \vec{v}^*) - v_l^*) \\ &\quad + \sum_{l=1}^N \sum_{l_1=1}^N p_{kl}(\mu_k, \nu_k) p_{ll_1}(\mu_l, \nu_l) (h_{l_1}(\mu_{l_1}, \nu_{l_1}, \vec{v}^*) - v_{l_1}^*) + \dots \end{aligned}$$

As $h_k(\mu_k, \nu_k, \vec{v}^*) - v_k^* \leq \epsilon$, we have

$$G_k(\vec{\mu}, \vec{\nu}^\epsilon) - v_k^* \leq \epsilon + \epsilon \sum_{l=1}^N p_{kl} + \epsilon \sum_{l=1}^N \sum_{l_1=1}^N p_{kl} p_{ll_1} + \dots \leq \epsilon/p_0.$$

Namely

$$G_k(\vec{\mu}, \vec{\nu}^\epsilon) \leq v_k^* + \epsilon/p_0.$$

Similarly we can show that

$$G_k(\vec{\mu}^\epsilon, \vec{\nu}) \geq v_k^* - \epsilon/p_0,$$

and our lemma is proved.

LEMMA 4. For any given positive ϵ , there exist ϵ -optimal strategies of the dummy games.

PROOF. Let ϵ be any given positive number. We put $\epsilon' = \epsilon p_0 / 2p_0 + 1$. We divide A_k (resp. B_k) into non-empty measurable subsets $A_{k,1}, A_{k,2}, \dots, A_{k,m_k}$ (resp. $B_{k,1}, B_{k,2}, \dots, B_{k,n_k}$), where $A_{k,i}$ (resp. $B_{k,j}$) are smaller than ϵ' in diameter in the metric $\delta^{h_k(a,b,\vec{v}^*)}$. Let $\alpha_k = (a_{k,1}, a_{k,2}, \dots, a_{k,m_k})$ (resp. $\beta_k = (b_{k,1}, b_{k,2}, \dots, b_{k,n_k})$) denote the finite subset of A_k (resp. B_k) where $a_{k,i}$ (resp. $b_{k,j}$) is any point chosen from $A_{k,i}$ (resp. $B_{k,j}$). Let \mathfrak{M}'_k (resp. \mathfrak{N}'_k) be the set of probability measures concentrated on α_k (resp. β_k). For any $\mu \in \mathfrak{M}'_k$ (resp. $\nu \in \mathfrak{N}'_k$) we define $\bar{\mu} \in \mathfrak{M}'_k$ (resp. $\bar{\nu} \in \mathfrak{N}'_k$) as follows:

$$\bar{\mu}(a_{k,i}) = \mu(A_{k,i}) \quad (\text{resp. } \bar{\nu}(b_{k,j}) = \nu(B_{k,j})).$$

Then

$$\begin{aligned} (5) \quad |h_k(\mu, \nu, \vec{v}^*) - h_k(\bar{\mu}, \bar{\nu}, \vec{v}^*)| &\leq \sum_{i,j} \int_{A_{ki}} \int_{B_{kj}} |h_k(a, b, \vec{v}^*) - h_k(a_{ki}, b_{kj}, \vec{v}^*)| d\mu_k d\nu_k \\ &< \epsilon'. \end{aligned}$$

If we denote by T' the value transformation of

$$\Gamma_k^{\prime d} = (A_k, B_k, h_k(a, b, \bar{v}), \mathfrak{M}'_k, \mathfrak{N}'_k),$$

then it follows from (5) that

$$(6) \quad \|\bar{v}^* - T' \bar{v}^*\| < \epsilon'.$$

Let \bar{v}^* be the principal value vector of $\Gamma_k^{\prime d}$, $k=1, 2, \dots, N$. Then by (4), and (6) we have

$$\begin{aligned} \|T'^j \bar{v}^* - T'^{j+1} \bar{v}^*\| &\leq (1-p_0)^j \|\bar{v}^* - T' \bar{v}^*\| \\ &< \epsilon' (1-p_0)^j. \end{aligned}$$

Consequently

$$\begin{aligned} \|\bar{v}^* - \bar{v}'^*\| &= \lim_{n \rightarrow \infty} \|\bar{v}^* - T'^n \bar{v}^*\| \\ &< \epsilon' + \epsilon' (1-p_0) + \epsilon' (1-p_0)^2 + \dots \\ &= \epsilon' / p_0. \end{aligned}$$

According to von Neumann's theorem [5] for finite spaces, α_k and β_k are finite, there exist optimal strategies $\bar{\mu}^\infty, \bar{\nu}^\infty$ of the game $(A_k, B_k, h_k(a, b, \bar{v}^*), \mathfrak{M}'_k, \mathfrak{N}'_k)$, i.e.

$$(7) \quad h_k(\bar{\mu}^\infty, \bar{\nu}, \bar{v}^*) \geq v_k^* \geq v_k^* - \epsilon' / p_0, \quad \text{and}$$

$$(8) \quad h_k(\bar{\mu}, \bar{\nu}^\infty, \bar{v}^*) \leq v_k^* \leq v_k^* + \epsilon' / p_0.$$

Then for any $\nu \in \mathfrak{N}_k$, we have

$$(9) \quad \begin{aligned} h_k(\bar{\mu}^\infty, \nu, \bar{v}^*) &= h_k(\bar{\mu}^\infty, \bar{\nu}, \bar{v}^*) + h_k(\bar{\mu}^\infty, \nu - \bar{\nu}, \bar{v}^*) \\ &\quad + \sum_{l=1}^N (v_l^* - v_l'^*) p_{kl}(\bar{\mu}^\infty, \bar{\nu}), \quad \text{where} \end{aligned}$$

$$(10) \quad |h_k(\bar{\mu}, \nu - \bar{\nu}, \bar{v}^*)| < \epsilon', \quad \text{and}$$

$$(11) \quad \left| \sum_{l=1}^N (v_l^* - v_l'^*) p_{kl}(\bar{\mu}, \bar{\nu}) \right| < \epsilon'.$$

On account of (7), (9), (10), and (11), we have

$$\begin{aligned} h_k(\bar{\mu}^\infty, \nu, \bar{v}^*) &\geq v_k^* - \epsilon' - \epsilon' \\ &\geq v_k^* - \epsilon' / p_0 - 2\epsilon' = v_k^* - \epsilon. \end{aligned}$$

Similarly we can show that

$$h_k(\bar{\mu}, \bar{\nu}^\infty, \bar{v}^*) \leq v_k^* + \epsilon,$$

and our lemma is proved.

THEOREM 1. The stochastic game $\Gamma = \{\Gamma_k^{\vec{v}}, k=1, 2, \dots, N\}$ is strictly determined. The value of $\Gamma_k^{\vec{v}}$ is equal to v_k^* , the k -th component of the principal value

vector of the associated dummy games.

PROOF. By Lemmas 3 and 4, we see that there exist ϵ -optimal strategies $\vec{\mu}^\epsilon, \vec{\nu}^\epsilon$ of $\vec{\Gamma}_k$ for any positive ϵ , that is,

$$G_k(\vec{\mu}^\epsilon, \vec{\nu}) \geq v_k^* - \epsilon \quad \text{and} \quad G_k(\vec{\mu}, \vec{\nu}^\epsilon) \leq v_k^* + \epsilon.$$

Then

$$\inf_{\nu} G_k(\vec{\mu}^\epsilon, \vec{\nu}) \geq v_k^* - \epsilon.$$

As ϵ is any positive number, so we have

$$(12) \quad \sup_{\mu} \inf_{\nu} G_k(\vec{\mu}, \vec{\nu}) \geq v_k^*.$$

Similarly we have

$$(13) \quad \inf_{\nu} \sup_{\mu} G_k(\vec{\mu}, \vec{\nu}) \leq v_k^*.$$

On the other hand it is easy to see that

$$(14) \quad \sup_{\mu} \inf_{\nu} G_k(\vec{\mu}, \vec{\nu}) \leq \inf_{\nu} \sup_{\mu} G_k(\vec{\mu}, \vec{\nu}).$$

Therefore the inequalities (12), (13), together with the inequality (14) imply

$$v_k^* = \sup_{\mu} \inf_{\nu} G_k(\vec{\mu}, \vec{\nu}) = \inf_{\nu} \sup_{\mu} G_k(\vec{\mu}, \vec{\nu}),$$

which completes the proof of our theorem.

J. von Neumann has called a game to be fair if its value is zero [5]. According to this nomenclature, we shall say that a stochastic game is fair if its value is zero. As an immediate consequence of Theorem 1, we have

COROLLARY 1. If the games $\Gamma_k = (A_k, B_k, g_k, \mathfrak{M}_k, \mathfrak{N}_k)$ are fair for $k=1, 2, \dots, N$, then the stochastic game Γ is also fair however the transition probabilities p_{kl} may be chosen.

PROOF. Since the value of every Γ_k is zero, it is easy to see that the zero vector is the principal value vector of the associated dummy games. Therefore it follows from Theorem 1 that the value of $\vec{\Gamma}_k$ is zero. This completes the proof.

Let α_k (resp. β_k) be any finite subset of A_k (resp. B_k), $k=1, 2, \dots, N$. We denote by $K'(a', b)$ (resp. $K''(a, b')$, and resp. $K'''(a', b')$) the restriction of K on $\alpha_k \times B_k$ (resp. on $A_k \times \beta_k$, and resp. on $\alpha_k \times \beta_k$). Similar notations shall be applied to p_{kl} and p_{k0} and so on. Then we have the stochastic game

$$\Gamma' = (\Gamma'_k = (\alpha_k, B_k, g_k, \mathfrak{M}'_k, \mathfrak{N}_k), p'_{kl}, p'_{k0}, k=1, 2, \dots, N),$$

and so forth. Now we show

COROLLARY 2. For any positive ϵ , we can choose a finite subset α_k (resp. β_k) of A_k (resp. B_k) such that the value of the stochastic game Γ differs at most by ϵ from the values of Γ' , Γ'' , and Γ''' .

PROOF. It is sufficient by Theorem 1 to show the statements for the principal value vector of the associated dummy games. It was shown in the proof of Lemma 4 that we can choose α_k and β_k to satisfy the requisites of the statement. The proofs of the other cases may be carried out in a similar way.

COROLLARY 3. Suppose every B_k is finite, then for any positive ϵ we can choose a finite subset α_k of A_k such that the number of pure strategies contained in α_k does not exceed the number of pure strategies contained in B_k for every k , and that the values of both stochastic games Γ and Γ' differ at most by ϵ .

PROOF. By Corollary 2, we can choose a finite subset $\bar{\alpha}_k$ of A_k for every k such that the values of both stochastic games Γ and

$$\bar{\Gamma}' (= (\bar{\Gamma}'_k = \bar{\alpha}_k, B_k, \bar{g}'_k, \bar{\mathfrak{M}}'_k, \mathfrak{R}_k), \bar{p}'_{kl}, \bar{p}'_{k0}, k = 1, 2, \dots, N)$$

differ at most by ϵ . Denote the value vector of $\bar{\Gamma}'$ and the value transformation associated with $\bar{\Gamma}'$ by \bar{v}' and T' respectively, and denote a dummy game associated with the stochastic game $\bar{\Gamma}'$ by $\bar{\Gamma}'^d$. Suppose $\bar{\alpha}_k$ contain more points than B_k for $k=1, 2, \dots, p$ ($p \leq N$) but not for $k=p+1, \dots, N$, then since the strategic sets of both players are finite, we see that $\bar{\alpha}_k$ can be replaced by a subset α_k such that the values of $\bar{\Gamma}'^d_k$ and Γ'^d_k coincide, and α_k, B_k contain the same number of pure strategies for $k = 1, 2, \dots, p$ (cf. Theorem 7 of [7]). Put $\bar{\alpha}_k = \alpha_k$ for $k=p+1, \dots, N$. Then by denoting the value transformation associated Γ by T , we have

$$\bar{v}' = \bar{T}' \bar{v}' = T \bar{v}'.$$

Namely the values of $\bar{\Gamma}'$ and Γ' are equal. Since the values of Γ and $\bar{\Gamma}'$ differ at most by ϵ , so do the values of Γ and Γ' .

Up till now, in considering the strategy spaces, they have been given originally as abstract sets, in which the topologies were introduced by using functions $g_k, p_{kl}, k, l=1, 2, \dots, N$. Now, however, we shall assume that the sets $A_k, B_k, k=1, 2, \dots, N$ are from the outset compact topological spaces in the terminology of N. Bourbaki [6], and that g_k, p_{kl} are continuous functions on $A_k \times B_k$. It is easy to see that the reduced strategy spaces obtained from $(A_k, \delta^{g_k}), (B_k, \delta^{g_k})$ etc. are compact. Thus the results obtained in the preceding discussions can be applied to our case. Here \mathfrak{A}_k (resp. \mathfrak{B}_k) stands for the σ -algebra of Borel subsets of A_k (resp. B_k) for $k=1, 2, \dots, N$. Then we have

THEOREM 2. The stochastic game just mentioned is strictly determined, and has optimal strategies.

PROOF. By Theorem 1 the stochastic game is strictly determined. For any given positive number ϵ , we can choose ϵ -optimal strategies $\mu^\epsilon, \nu^\epsilon$ of the associated dummy games $\Gamma_k^d = (A_k, B_k, h_k(a, b, \vec{v}^*), \mathfrak{M}_k, \mathfrak{N}_k)$. Consider a U-net associated with $\{\mu^\epsilon\}$. Then it is known that it converges vaguely to a probability measure μ^0 . Then the inequality

$$h_k(\mu^\epsilon, \nu, \vec{v}^*) \geq v_k^* - \epsilon \quad \text{for any } \nu \in \mathfrak{N}_k$$

implies

$$h_k(\mu^0, \nu, \vec{v}^*) \geq v_k^* \quad \text{for any } \nu \in \mathfrak{N}_k.$$

Similar considerations for $\{\nu^\epsilon\}$ lead us to conclude that there exists a $\nu^0 \in \mathfrak{N}_k$ such that

$$h_k(\mu, \nu^0, \vec{v}^*) \leq v_k^* \quad \text{for any } \mu \in \mathfrak{M}_k.$$

The proof is completed.

COROLLARY 4. If every B_k consists of n_k pure strategies ($n_k < \infty$) for $k = 1, 2, \dots, N$. Then there exist finite strategic subsets α_k of A_k containing at most n_k pure strategies such that the value of the stochastic game Γ remains unchanged when A_k are replaced by α_k .

PROOF. Using Corollary 3 and Theorem 1, the proof can be carried out along a similar line as in the proof of Theorem 2.

§3. Stochastic Games with Infinite Positions

In the preceding discussions, as L. S. Shapley assumed, only the stationary strategies were considered in the stochastic games. Now we shall relax these restrictions on the mixed strategies which both players select. Then we shall see that the problem may be reduced to the stochastic games with countably infinite positions in which both players select stationary strategies as before. Suppose both players begin with the game Γ_1 , without any loss of generality, with mixed strategies $\mu_1^{(1)}, \nu_1^{(1)}$ respectively, then according to the transition probability $p_{1l}(a, b)$, depending on a choice of pure strategies $(a, b) \in A_1 \times B_1$, the game moves to Γ_l in the second trial. Both players then select mixed strategies $\mu_l^{(2)}, \nu_l^{(2)}$ respectively and so on. The expected value $G_1(\vec{\mu}, \vec{\nu})$ of the gains of Player 1 is given by

$$(15) \quad G_1(\vec{\mu}, \vec{\nu}) = g_1(\mu_1^{(1)}, \nu_1^{(1)}) + \sum_{l=1}^N p_{1l}(\mu_1^{(1)}, \nu_1^{(1)}) g_l(\mu_l^{(2)}, \nu_l^{(2)}) \\ + \sum_{j=1}^N \sum_{i=1}^N p_{1i}(\mu_1^{(1)}, \nu_1^{(1)}) p_{ij}(\mu_i^{(2)}, \nu_i^{(2)}) g_j(\mu_j^{(3)}, \nu_j^{(3)}) + \dots$$

The value of the game may be defined as

$$\sup_{\mu} \inf_{\nu} G_1(\vec{\mu}, \vec{\nu}) = \inf_{\nu} \sup_{\mu} G_1(\vec{\mu}, \vec{\nu})$$

provided this equation holds. The equation leads us to consider the sequence of games

$$(16) \quad \Gamma_1, \Gamma_2, \dots, \Gamma_N, \Gamma_1, \Gamma_2, \dots, \Gamma_N, \dots,$$

where Γ_l stands at the $(nN + 1 + l)$ -th position ($0 \leq l \leq N$) for every non-negative integer n . Taking (15), (16) into account, it is clear that we arrive at a special case of the stochastic games now defined below:

Suppose we are given countably infinite positions $1, 2, \dots$. To each position k there correspond the game

$$\Gamma_k = (A_k, B_k, g_k, \mathfrak{M}_k, \mathfrak{N}_k),$$

transition probabilities p_{kl} , and stop probabilities p_{k0} . Here we assume that

- (i) the pay-offs g_k are uniformly bounded and \mathfrak{C}_k -measurable,
- (ii) the transition probabilities p_{kl} are \mathfrak{C}_k -measurable for every k, l , and vanish identically except a finite number of l 's depending on k , and furthermore

$$\inf_{k,a,b} p_{k0}(a, b) = p_0 > 0,$$

- (iii) payments accumulate throughout the course of the play,
- (iv) strategy spaces $(A_k, \delta^{g_k}), (A_k, \delta^{p_{kl}})$ are precompact for $k, l, = 1, 2, \dots$.

If the game begins with the k -th position, then the expected value of the gains of Player 1 is given by

$$\begin{aligned} G_k(\vec{\mu}, \vec{\nu}) &= g_k(\mu_k^{(1)}, \nu_k^{(1)}) + \sum_{l=1}^N p_{kl}(\mu_k^{(1)}, \nu_k^{(1)}) g_l(\mu_l^{(2)}, \nu_l^{(2)}) \\ &+ \sum_{l=1}^N \sum_{j=1}^N p_{kl}(\mu_k^{(1)}, \nu_k^{(1)}) p_{lj}(\mu_l^{(2)}, \nu_l^{(2)}) g_j(\mu_j^{(3)}, \nu_j^{(3)}) + \dots \end{aligned}$$

Now it is clear that $G_k(\vec{\mu}, \vec{\nu})$ is a solution of the simultaneous linear equations of infinite number:

$$(17) \quad \begin{aligned} v_1 &= g_1(\mu_1, \nu_1) + \sum_{l=1}^N p_{1l}(\mu_1, \nu_1) v_l \\ v_2 &= g_2(\mu_2, \nu_2) + \sum_{l=1}^N p_{2l}(\mu_2, \nu_2) v_l \\ &\dots \end{aligned}$$

As in Section 2 we considered the dummy games, we can define the dummy game

$$(A_k, B_k, h_k(a, b, \vec{\nu}), \mathfrak{M}_k, \mathfrak{N}_k)$$

where

$$h_k(a, b, \vec{v}) = g_k(a, b) + \sum_{i=1}^N p_{ki}(a, b)v_i,$$

\vec{v} being any fixed vector (v_1, v_2, \dots) .

And $\vec{v}^* \in (l^\infty)$ is called a principal value vector. As in Section 2, we can show that there exists a unique principal value vector. As to the ϵ -optimal strategies we can go along the same line, and reach the same conclusion as Theorem 1.

Turning to the original problem, it is easy to see that the value of the stochastic game is irrelevant whether both players select stationary mixed strategies or not.

References

- [1] SHAPLEY, L. R., Stochastic games. Proc. Nat. Acad. Sci., U. S. A., 39 (1953), pp. 1095-1100.
- [2] WALD, A., Statistical Decision Functions. John Wiley and Sons, New York, 1950.
- [3] PETROWSKI, I. G., Vorlesungen über die Theorie der gewöhnlichen Differentialgleichungen. B. G. Teubner, Leipzig, 1954.
- [4] EVERETT, H., Recursive games. Contributions to the Theory of Games, III. Annals of Mathematics Studies, No. 39, Princeton, N. J., Princeton University Press, 1957, pp. 47-78.
- [5] VON NEUMANN, J., and O. MORGENSTERN, Theory of Games and Economic Behavior. Princeton University Press, Princeton, 1944.
- [6] BOURBAKI, N., Topologie générale. Hermann, Paris, 1961.
- [7] KAPLANSKY, I., A Contribution to von Neumann's Theory of Games. Ann. of Math., 46 (1945), pp. 474-479.
- [8] LUCE, R. DUNCAN and H. RAIFFA, Games and Decisions. John Wiley and Sons, New York, 1957.
- [9] KELLEY, J. L., General Topology. Van Nostrand, New York, 1955.