Discounted, Positive, and Noncooperative Stochastic Games

By T. PARTHASARATHY

Indian Statistical Institute, Calcutta, and University of Illinois, Chicago

Abstract: In this paper, we consider the stochastic games of SHAPLEY and prove under certain conditions the stochastic game has a value and both players have optimal strategies. We also prove a similar result for noncooperative stochastic games.

1. Introduction

A stochastic game is determined by the following objects: S, $A(s)$, $B(s)$, q, r. Here, S is a nonempty Borel subset of a Polish space, the set of states of the system; $A(s)$ is a nonempty Borel subset of a Polish space, the set of actions available to player I at state s; $B(s)$ is a nonempty Borel subset of a Polish space, the set of actions available to player II at state s. We shall assume throughout that $A(s) \subseteq A$ and $B(s) \subseteq B$ for every $s \in S$ where A and B are Polish spaces. q associates Borel measurably with each triple $(s,a,b) \in S \times A \times B$ a probability measure on the Borel subsets of S; r, the reward function, is a bounded measurable function on $S \times A \times B$. Periodically (say, once a day) players I and II observe the current state s of the system and choose actions $a(s) \in A(s)$ and $b(s) \in B(s)$ respectively; the choice of the actions is made with full knowledge of the history of the system as it has evolved to the present. As a consequence of the actions chosen by the players, two things happen: player II pays player I $r(s,a(s),b(s))$ units of money, and the system moves to a new state s' according to the distribution $q(\cdot/s,a(s),b(s))$. (In the noncooperative stochastic game there will be two bounded measurable functions r_1 and r_2 defined on $S \times A \times B$ in which case players I and II will receive $r_1(s,a(s),b(s))$ and $r_2(s,a(s),b(s))$ units of money.) Then the whole process is repeated from the new state s'. The problem, then, is to maximize player I's expected income as the game proceeds over the infinite future and to minimize player II's expected loss. (In the noncooperative case both will try to maximize their expected pay-offs.)

We shall assume unless otherwise stated S, A and B to be compact metric. Let P_A and P_B denote the space of all probability distributions on A and B respectively. It is well-known that P_A and P_B are compact and metrizable in the

[1]) A.M.S. (1969) subject Classification: Primary 9070, 9072. Secondary 9072.

Key words: Discounted, positive and noncooperative stochastic games, optimal stationary strategies.

weak topology. We shall also assume $A(s) \subseteq A$, $B(s) \subseteq B$ to be non-empty and compact for each $s \in S$. We, further, suppose the following point to set mappings F and G are continuous:

$$F : s \to P_{A(s)}$$
$$G : s \to P_{B(s)} .$$

In other words, $\{s : F(s) \cap K \neq \emptyset\}$ and $\{s : F(s) \subseteq K\}$ are open in S whenever K is open in P_A.

A strategy Π for player I is a sequence $(\Pi_1, \Pi_2, \ldots)$ where Π_n specifies the action to be chosen by player I on the n^{th} day by associating Borel measurably with each history $h = (s_1, a_1(s_1), b_1(s_1), \ldots, s_{n-1}, a_{n-1}(s_{n-1}), b_{n-1}(s_{n-1}), s_n)$ of the system a probability distribution $\Pi_n(\cdot/h) \in P_{A(s_n)}$. A strategy Π is said to be stationary if there is a Borel map f from S to P_A such that $f(s) \in P_{A(s)}$ and $\Pi_n = f$ for each $n \geq 1$: and in this case Π is denoted by f^∞. Similarly strategies and stationary strategies are defined for II.

Let β be any fixed number satisfying $0 \leq \beta < 1$. A pair (Π, Γ) of strategies for players I and II associates with each initial state s an n^{th}-day expected gain $r_n(\Pi, \Gamma)(s)$ for player I and a total expected discounted gain for player I.

$$I(\Pi, \Gamma)(s) = \sum_{n=1}^{\infty} \beta^{n-1} r_n(\Pi, \Gamma)(s) .$$

Positive stochastic games are those where $r(s,a,b) \geq 0$ and $\beta = 1$. (In the non-cooperative case we will have a pair of total discounted expected reward functions corresponding to the two functions r_1 and r_2. In this case we will write $I_1(\Pi, \Gamma)(s)$ and $I_2(\Pi, \Gamma)(s)$.)

A strategy Π^* is optimal for player I if $\inf_{\Gamma} \sup_{\Pi} I(\Pi, \Gamma)(s) \leq I(\Pi^*, \Gamma)(s)$ for every Γ and $s \in S$: a strategy Γ^* is optimal for player II if $\sup_{\Pi} \inf_{\Gamma} I(\Pi, \Gamma)(s) \geq I(\Pi, \Gamma^*)(s)$ for every Π and $s \in S$. We shall say that the stochastic game has a value if $\sup_{\Pi} \inf_{\Gamma} I(\Pi, \Gamma)(s) = \inf_{\Gamma} \sup_{\Pi} I(\Pi, \Gamma)(s)$ for every $s \in S$. In case the stochastic game has a value, the quantity $\sup_{\Pi} \inf_{\Gamma} I(\Pi, \Gamma)(s)$ as a function on S, is called the value function. (In the noncooperative case, call (Π^*, Γ^*) an equilibrium pair if

$$I_1(\Pi^*, \Gamma^*)(s) \geq I_1(\Pi, \Gamma^*)(s) \;\forall\, \Pi \quad \text{and} \quad s$$

and

$$I_2(\Pi^*, \Gamma^*)(s) \geq I_2(\Pi^*, \Gamma)(s) \;\forall\, \Gamma \quad \text{and} \quad s.$$

The stochastic game problem was first formulated by SHAPLEY [1953] who took S, A, and B to be finite, assumed that play would terminate in a finite number of states with probability one, and considered only what we have called stationary strategies. Incidentally, the restriction that play should terminate in a finite number of stages is to keep the total expected gain for player I finite; we have bypassed this difficulty by introducing a discount factor. GILLETTE [1957] and HOFFMAN and KARP [1966] have investigated the non-terminating case of a stochastic game by taking the average gain of player I per play as the pay-off. TAKAHASHI [1962] considered the case when S is finite but A and B are inifinite and $A(s) \equiv A$, $B(s) \equiv B$. In MAITRA and PARTHASARATHY [1970] this problem was solved under suitable assumptions when S, $A(s) \equiv A$, $B(s) \equiv B$ are compact metric spaces. ROGERS [1969] and SOBEL [1969] were the first to consider noncooperative stochastic games and showed the existence of equilibrium strategies when S, A and B are finite. The main purpose of this paper is to get extensions of the known results in discounted and noncooperative stochastic games.

Section 2 is somewhat expository in nature where we prove a known selection theorem. In section 3 we prove two theorems on discounted stochastic games while in section 4 a theorem, perhaps known, is proved for positive stochastic games. Section 5 is concerned with noncooperative (discounted) stochastic games. The last section contains a few remarks and an open problem.

2. Selection Theorem

In order to state the selection theorem, we need some preliminaries. Let Y denote a compact metric space and 2^Y denote the collection of nonempty closed subsets of Y. We introduce a metric d on 2^Y – the Hausdorff metric – as follows. For any $E, F \in 2^Y$,

$$d(E,F) = \max\{\sup_{x\in E} d'(x,F), \sup_{y\in F} d'(y,E)\}$$

where d' is the metric on Y and $d'(x,G) = \inf_{x'\in G} d'(x,x')$ for $G \subseteq Y$. It is well known that $(2^Y, d)$ is compact metric – for a proof see KURATOWSKI [1966].

Definition:

Let X be a complete separable metric space. A map $D: X \to 2^Y$ is called upper semicontinuous in the sense of KURATOWSKI if x_n, $x \in X$, $y_n \in D(x_n)$. $x_n \to x$, $y_n \to y$ as $n \to \infty$ implies $y \in D(x)$.

Remark 2.1:

One can easily prove the following fact: Let $D(x)$ be a nonempty subset of Y, $x \in X$. Then $\{(x,y): y \in D(x)\}$ is closed if and only if $D(x)$ is closed for every x and $D: X \to 2^Y$ is upper semicontinuous. Next we state a lemma and prove it.

Lemma 2.1: HINDERER [1970]*:*

We make the following assumptions.

(i) Y is compact metric.

(ii) X is a complete separable metric space.

(iii) $D(x)$ is a nonempty subset of Y for $x \in X$, such that $K = \{(x,y): y \in D(x)\}$ is closed, that is, $D: X \to Y$ is upper semicontinuous.

(iv) $w: K \to R$ (= set of all real numbers) is upper semicontinuous, that is, a_n, $a \in K$ and $a_n \to a \Rightarrow \limsup w(a_n) \leq w(a)$.

Let $W(x) = \{y : w(x,y) = \sup_{y' \in D(x)} w(x,y')\}$. Then $W: X \to 2^Y$ is measurable (that is, $\{x : W(x) \cap G \neq \emptyset\}$ is Borel in X for every $G \in 2^Y$).

Proof:

Let $t(x) = \sup_{y \in D(x)} w(x,y)$. It is not hard to check that t is upper semicontinuous. Let

$$T = \{(x,r) : t(x) \geq r\}$$

where $r \in R$. Obviously T is nonempty. In particular, $(x, t(x)) \in T$, for every $x \in X$. T is also closed. In fact, consider a sequence of points $(x_n, r_n) \in T$ converging to (x,r). Then there are points $y_n \in D(x_n)$ such that $w(x_n, y_n) \geq r_n$. Since Y is compact, there exists a subsequence $\{y_{n_k}\}$ converging to some $y \in Y$. Then $(x_{n_k}, y_{n_k}) \to (x,y) \in K$ since K is closed. It follows that $r = \lim r_n \leq \limsup w(x_{n_k}, y_{n_k}) \leq w(x,y)$, hence $(x,r) \in T$.

Define the map $\tilde{W}: T \to 2^Y$, by means of $\tilde{W}(x,r) = \{y : y \in D(x) \text{ and } w(x,y) \geq r\}$. Obviously $\tilde{W}$ is always nonempty and closed, since w is upper semicontinuous. Now we shall show that $\tilde{W}$ is upper semicontinuous. Let (x_n, r_n), $(x,r) \in T$; $y_n \in \tilde{W}(x_n, r_n)$ and $(x_n, r_n) \to (x,r)$ and $y_n \to y$. We have to show y belongs to $\tilde{W}(x,r)$. We have $r = \lim r_n \leq \limsup w(x_n, y_n) \leq w(x,y)$, hence $y \in \tilde{W}(x,r)$. It follows that $\tilde{W}$ is measurable. The set T is closed, hence measurable extension of $\tilde{W}$ (which we continue to denote by $\tilde{W}$) from T to $X \times R$. The map $t: X \to R$ is upper semicontinuous, hence measurable. It follows that the map $x \to (x, t(x))$ from X to $X \times R$ is measurable, hence also $x \to \tilde{W}(x, t(x)) = W(x)$ is measurable. The proof of the lemma is complete.

Now we are ready to prove the following selection theorem (see pp. 113–115 in HINDERER [1970]).

Theorem 2.1:

Under the assumptions of lemma 2.1 there exists a measurable map $f: X \to Y$ such that

$$f(x) \in D(x), \quad x \in X \quad \text{and}$$

$$w(x, f(x)) = \sup_{y \in D(x)} w(x,y), \quad x \in X.$$

Proof:

It is well-known that there exists a sequence of (v_n) continuous maps $v_n: Y \to R$ which separate points in Y; that is, for any pair y, $y' \in Y$ there exists some n such that $v_n(y) \neq v_n(y')$. For every n, let us derive a map $V_n: 2^Y \to 2^Y$ by $V_n(K) = y : y \in K$, $v_n(y) = \sup v_n(y')$. MAITRA [1968] has shown that V_n is measurable. Let W be as in lemma 2.1. Define $F_0 = W$, $F_n = V_n \circ F_{n-1}$ for every n. It follows from lemma 2.1 that $F_n: X \to 2^Y$ is measurable. For any x the sequence of set $F_n(x)$ is decreasing, since $V_n(K) \subset K$. Since Y is compact, the set $F(x) = \bigcap_{n=1}^{\infty} F_n(x)$ is not empty. Fix $x \in X$ and consider two points y, $y' \in F(x)$, hence y, $y \in F_n(x)$ for every n. Hence $v_n(y) = \max_{z \in F_{n-1}} v_n(z) = v_n(y')$ for all n. Since the sequence $\{v_n\}$ is separating points, we have $y = y'$, that is, $F(x)$ is a singleton, say $F(x) = \{f(x)\}$. Moreover, we have $f(x) \in F_0(x) = W(x)$, consequently $f(x) \in D(x)$ and $w(x, f(x)) = \max_{y' \in D(x)} w(x, y'), x \in X$.

It is well-known that for any $x \in X$ the sequence of sets $F_n(x)$, being a decreasing sequence of closed sets converges in the metric 2^Y to $F(x)$. The limit of a convergent sequence of measurable maps into a metric space is a measurable map. Hence $F: X \to 2^Y$ is measurable. For any closed subset B of Y we have $\{x : f(x) \in B\} = \{x : F(x) \in \{B\}\}$, and the latter is a Borel set in X, since F is measurable and since the singleton $\{B\}$ is a Borel set in the metric space 2^Y. The σ-algebra of Borel sets in Y is generated by the system of closed subsets. It follows that $f: X \to Y$ is measurable and the proof is complete.

We will close this section with the statement of one more selection theorem due to OLECH [1965] which we need in the sequel.

Theorem 2.2:

Let X be a Borel subset of a Polish space and let Y be a compact subset of R^n ($= n$ dimensional Euclidean space). Let $w: X \times Y \to R^1$. Suppose w satisfies that (i) $w(x, y)$ is Borel measurable in x for each $y \in Y$ and (ii) $w(x, y)$ is continuous in y for each $x \in X$. Let $D: X \to 2^Y$ be a measurable map. Then there is a measurable function f from $X \to Y$ such that $f(x) \in D(x)$ and $w(x, f(x)) = \sup_{y \in D(x)} w(x, y)$. For a proof of this theorem refer to OLECH [1965].

3. Discounted Stochastic Games

In this section we shall prove two theorems on discounted stochastic games. We need the following lemma.

Lemma 3.1:

Let X and Y be two topological spaces. Let $\emptyset: X \to R$ be continuous. Let Γ be a compact valued function from X to Y, that is, for each x, $\Gamma(x) \subseteq Y$ and compact. Suppose Γ is continuous, that is, $\{x : \Gamma(x) \cap G \neq \emptyset\}$ and $\{x : \Gamma(x) \cap G\}$ are open in X whenever G is open. Then the numerical function M defined by

$$M(x) = \max \{\emptyset(x, y) : y \in \Gamma(x)\}$$

is continuous.

Proof:

Suppose that $x_0 \in X$ and let Y_0 be such that

$$y_0 \in \Gamma(x_0); \emptyset(x_0, y_0) \geq M(x_0) - \varepsilon .$$

There exist neighborhoods $U(x_0)$ and $V(y_0)$ such that

$$(x, y) \in U(x_0) \times V(y_0) \Rightarrow \emptyset(x, y) \geq \emptyset(x_0, y_0) - \varepsilon \geq M(x_0) - 2\varepsilon$$

and there exists a neighborhood $U'(x_0)$ such that

$$x \in U'(x_0) \Rightarrow \Gamma(x) \cap V(y_0) \neq \emptyset .$$

Therefore $x \in U(x_0) \cap U'(x_0) \Rightarrow M(x) \geq M(x_0) - 2\varepsilon$ and hence M is lower semicontinuous. We will complete the proof by showing that M is also upper semicontinuous. Suppose that $x_0 \in X$; to each y in $\Gamma(x_0)$ there correspond neighborhoods $U_y(x_0)$ and $V(y)$ such that

$$(x,z) \in U_g(x_0) \times V(y) \Rightarrow \emptyset(x,z) \leq \emptyset(x_0,y) + \varepsilon .$$

Since $\Gamma(x)$ is compact, it can be covered by a finite number of neighborhoods of the form $V(y)$, say $V(y_1), V(y_2), \ldots, V(y_n)$. Putting $U'(x_0) = \bigcap_{i=1}^{n} U_{y_i}(x_0)$ and $V(\Gamma x_0) = \bigcup_{i=1}^{n} V(y_i)$, we have

$$x \in U'(x_0), y \in V(\Gamma x_0) \Rightarrow \emptyset(x,y) \leq \max_i \emptyset(x_0,y_i) + \varepsilon \leq M(x_0) + \varepsilon .$$

Moreover there exists a neighborhood $U(x_0)$ such that

$$x \in U(x_0) \Rightarrow \Gamma(x) \subseteq V(\Gamma(x_0))$$

and so $x \in U(x_0) \cap U'(x_0) \Rightarrow M(x) = \max_{y \in \Gamma x} \emptyset(x,y) \leq M(x_0) + \varepsilon$ which shows M is also upper semicontinuous and this terminates the proof.

Theorem 3.1:

Let S, A, B be compact metric spaces. Let $A(s) \subseteq A$ and $B(s) \subseteq B$ be compact every $\in S$. Let r be a continuous real-valued function on $S \times A \times B$. Suppose, whenever $(s_n, a_n, b_n) \to (s_0, a_0, b_0)$ in $S \times A \times B$, $q(\cdot/s_n, a_n, b_n)$ converges weakly to $q(\cdot/s_0, a_0, b_0)$. Moreover assume that the following set valued functions

$$F : S \to 2^{P_A} \quad \text{defined by} \quad F(s) = P_{A(s)}$$

and

$$G : S \to 2^{P_A} \quad \text{defined by} \quad G(s) = P_{B(s)}$$

are continuous, where P_A stand for the set of all probability distributions on A etc. Then, the stochastic game has a value, the value function is continuous, and players I and II have optimal stationary strategies.

Remark 3.1:

This theorem is a variant of theorem 4.1 in [MAITRA et al. 1970] where we assumed $A(s) \equiv A$ and $B(s) \equiv B$ for every $s \in S$, as can be seen from the following remark:

Remark 3.2:

A referee points out that it is enough to define the function r only for $s \in S$. $a \in A(s)$ and $b \in B(s)$. We can always extend r to $S \times A \times B$ as follows:

$$r(s,a,b) = \begin{cases} r(s,a,b) & \text{if } a \in A(s),\ b \in B(s) \\ r(s,f(s),b) & \text{if } a \notin A(s),\ b \in B(s) \\ r(s,a,g(s)) & \text{if } a \in A(s),\ b \notin B(s) \\ r(s,f(s),g(s)) & \text{if } a \notin A(s),\ b \notin B(s) \end{cases}$$

where f and g are Borel functions with $f(s) \in A(s)$ and $g(s) \in B(s)$. The existence of such functions follows from the assumptions of theorem 3.1. We can do the same for q also. The problem is essentially reduced to the case $A(s) \equiv A, B(s) \equiv B$.

Proof of theorem 3.1:

For each $w \in C(S)$ (= space of real-valued continuous functions on S) defines Tw as follows.

$$(Tw)(s) = \min_{P_{B(s)}} \max_{P_{A(s)}} [r(s,\mu,\lambda) + \beta \int w(\cdot) dq(\cdot/s,\mu,\lambda)]$$

where $r(s,\mu,\lambda) = \iint r(s,a,b) d\mu(a) d\lambda(b)$ and

$$q(\cdot/s,\mu,\lambda) = \iint q(\cdot/s,a,b) d\mu(a) d\lambda(b).$$

From lemma 3.1, it is not hard to check that $Tw \in C(S)$. Since $r(s,\mu,\lambda) + \beta w(\cdot) d(\cdot/s,\mu,\lambda)$ is continuous on $P_{A(s)} \times P_{B(s)}$ for every fixed s and $P_{A(s)}$, $P_{B(s)}$ are compact convex, it follows from Sion's minimax theorem see PARTHASARATHY and RAGHAVAN [1971] that

$$\begin{aligned}(Tw)(s) &= \min_{P_{B(s)}} \max_{P_{A(s)}} [r(s,\mu,\lambda) + \beta \int w(\cdot) dq(\cdot/s,\mu,\lambda)] \\ &= \max_{P_{A(s)}} \max_{P_{B(s)}} [r(s,\mu,\lambda) + \beta \int w(\cdot) dq(\cdot/s,\mu,\lambda)] .\end{aligned}$$

Plainly T is a contraction mapping on $C(S)$ since $0 \leq \beta \leq 1$. Since $C(S)$, when equipped with the supremum, is a complete metric space, T has a unique fixed point in $C(S)$, by virtue of the Banach fixed point theorem. Let w^* be the fixed point of T. Then it follows from theorem 2, that there exist Borel maps f^* and g^* from S to P_A and P_B, respectively, such that, for every $s \in S$, $f^*(s) \in P_{A(s)}$, $g^*(s) \in P_{B(s)}$ and

$$\begin{aligned}w^*(s) &= \min_{P_{B(s)}} [r(s,f^*(s),\lambda) + \beta \int w^*(\cdot) dq(\cdot/s,f^*(s),\lambda)] \\ &= \max_{P_{A(s)}} [r(s,\mu,g^*(s) + \beta \int w^*(\cdot) dq(\cdot/s,\mu,g^*(s))] \\ &= r(s,f^*(s),g^*(s)) + \beta w^*(\cdot) dq(\cdot/s,f^*(s),g^*(s)] .\end{aligned}$$

From the above equation it follows (see Lemma 4.1 in MAITRA and PARTHASARATHY [1970]) $w^* = I(f^{*(\infty)},g^{*(\infty)})$. In view of this the above equation can be written as

$$\begin{aligned}I(f^{*(\infty)},g^{*(\infty)})(s) &= \max [r(s,\mu g^*(s)) + \beta \int I(f^{*(\infty)},g^{*(\infty)})(\cdot) dq(\cdot/s,\mu,g^*(s))] \\ &= \min [r(s,f^*(s),\lambda)) + \beta I(f^{*(\infty)} g^{*(\infty)})(\cdot) dq(\cdot/s,f^*(s),\lambda)] .\end{aligned}$$

It follows from a Theorem of BLACKWELL (HINDERER [1970]) that,

$$I(f^{*(\infty)},g^{*(\infty)})(s) = \sup I(\Pi,g^{*(\infty)})(s) = \inf_{\Gamma} I(f^{*(\infty)},\Gamma)(s)$$

for every $s \in S$. Consequently, $I(f^{*(\infty)}, g^{*(\infty)})(s) \geq \inf_{\Gamma} \sup_{\Pi} I(\Pi,\Gamma)(s)$. On the other hand, $I(f^{*(\infty)}, g^{*(\infty)})(s) \leq \sup_{\Pi} \inf_{\Gamma} I(\Pi,\Gamma)(s)$. Hence $\inf_{\Gamma} \sup_{\Pi} I(\Pi,\Gamma)(s) = I(f^{*(\infty)}, g^{*(\infty)})(s) = \sup_{\Pi} \inf_{\Gamma} I(\Pi,\Gamma)(s)$. This proves that the stochastic game has a value, that the value function is $I(f^{*(\infty)}, g^{*(\infty)})(s) = w^*(s)$ and continuous and that $f^{*(\infty)}, g^{*(\infty)}$ are optimal strategies for players I and II respectively. This terminates the proof. We shall state our next theorem.

Theorem 3.2:

Let S be a complete separable metric space. Let A and B be finite sets. Suppose $r(s,a,b)$ is a bounded measurable function in s and $q(\cdot/s,a,b)$ is also measurable in s. Then the discounted stochastic game has a value and the value function is measurable and the two players have optimal stationary strategies.

Remark 3.2:

This theorem includes theorem 1 in PARTHASARATHY [1971].

Proof:

Clearly $r(s,\mu,\lambda)$ and $q(\cdot/s,\mu,\lambda)$ are continuous in (μ,λ) for every s. We shall assume without any loss of generality $A(s) \equiv A$ and $B(s) \equiv B$. Denote by $M(s)$ the family of all bounded Borel functions on S. With each $w \in M(s)$, define $Tw \in M(s)$ as follows:

$$Tw(s) = \min \max \left[r(s,\mu,\lambda) + \beta \int w(\cdot) dq(\cdot/s,\mu,\lambda) \right].$$

Plainly T is a contraction mapping on $M(s)$. Let w^* be the unique fixed point of T that is,

$$\begin{aligned} w^*(s) &= \min \max \left[r(s,\mu,\lambda) + \beta \int w^*(\cdot) dq(\cdot/s,\mu,\lambda) \right] \\ &= \max \min \left[r(s,\mu,\lambda) + \beta \int w^*(\cdot) dq(\cdot/s,\mu,\lambda) \right]. \end{aligned}$$

It follows from theorem 2.2 that there exist Borel maps f^* and g^* from S to P_A and P_B respectively, such that for every $s \in S$,

$$\begin{aligned} w^*(s) &= \min \left[r(s,f^*(s),\lambda) + \beta \int w^*(\cdot) dq(\cdot/s,f^*(s),\lambda) \right] \\ &= \max \left[r(s,\mu,g^*(s)) + \beta \int w^*(\cdot) dq(\cdot/s,\mu,g^*(s)) \right] \\ &= r(s,f^*(s),g^*(s) + \beta \int w^*(\cdot) dq(\cdot/s,f^*(s),g^*(s)). \end{aligned}$$

The rest of the proof is similar to theorem 3.1 and we omit the details.

Remark 3.3:

Here the value function is Borel measurable and need not be continuous as in the previous theorem, for we are not having any continuity assumption on $r(s,a,b)$ as a function on S.

4. Positive Stochastic Games

In this section we are considering positive stochastic games, that is, we shall assume $r(s,a,b)$ is non-negative and $\beta = 1$. We prove the following theorem.

Theorem 4.1:

Let S, A and B are finite sets. (For simplicity we assume $A(s) \equiv A$ and $B(s) \equiv B$. Let $r(s,a,b) \geq 0$ and $\beta = 1$. Suppose there exists a positive constant K independent of Π, Γ and s such that $I(\Pi,\Gamma)(s) \leq K$ for all Π, Γ and s. Then the positive stochastic game has a value and the two players have optimal stationary strategies.

Remark 4.1:

The referee points out that any positive stochastic game with the above assumptions is simply a terminating stochastic sequence of finite games – see SHAPLEY [1953], or a discounted stochastic game.

Remark 4.2:

Suppose S is countable, A and B are finite. Then in the discounted case the two players have optimal stationary strategies – but in the positive case the players need not have optimal stationary strategies. We will give an example in the last section to demonstrate this.

Proof of theorem 4.1:

Perhaps this theorem is known but we give a proof. Let $0 < \beta < 1$. Let $v_\beta(s)$ be the value of the discounted stochastic game. Since r is non-negative, $v_\beta(s)$ is an increasing function of β. Let $v^*(s) = \lim_{\beta \uparrow 1} v_\beta(s)$. Since $I(\Pi,\Gamma)(s) \leq K$, the limit exists. Hence we have

$$\begin{aligned} v^*(s) &= \min_{P_B} \max_{P_A} \left[r(s,\mu,\lambda) + \int v^*(\cdot)dq(\cdot/s,\mu,\lambda)\right] \\ &= \max_{P_A} \min_{P_B} \left[r(s,\mu,\lambda) + \int v^*(\cdot)dq(\cdot/s,\mu,\lambda)\right]. \end{aligned}$$

Here one can replace the integral sign by summation sign since S is finite. For every s, let $f^*(s)$ and $g^*(s)$ be any pair of optimal strategies for the two players for the finite game whose payoff is given by

$$r(s,a,b) + \int v^*(\cdot)dq(\cdot/s,a,b).$$

That is

$$\begin{aligned} v^*(s) &= \max_{P_A} \left[r(s,\mu,g^*(s)) + \int v^*(\cdot)dq(\cdot/s,\mu,g^*(s))\right] \\ &= \min_{P_B} \left[r(s,f^*(s),\lambda) + \int v^*(\cdot)dq(\cdot/s,f^*(s),\lambda)\right] \\ &= r(s,f^*(s),g^*(s)) + \int v^*(\cdot)dq(\cdot/s,f^*(s),g^*(s)). \end{aligned}$$

At this stage we need the following lemma.

Lemma 4.1:

Let $r(f^*,g^*)$ be a vector whose s^{th} coordinate is given by $r(s,f^*(s),g^*(s))$. Let $Q(f^*,g^*)$ stand for the finite stochastic matrix whose (s,s')-th element is given by $q('/s,f^*(s),g^*(s))$. If $r(f^*,g^*) + Q(f^*,g^*)v^* = v^*$, then

$$I(f^{*(\infty)},g^{*(\infty)})(s) = v^*(s) \quad \text{for every} \quad s.$$

Proof:

Since $r(f^*,g^*) + Q(f^*,g^*)v^* = v^*$, it follows

$$\sum_{K=1}^{n-1} Q^K(f^*,g^*)r(f^*,g^*) + Q^n(f^*,g^*)v^* = v^*$$

where $Q^K = Q\,Q\ldots$ (K times). To complete the proof of the lemma it is enough if we show $Q^n(f^*,g^*)v^* \to 0$ as in $n \to \infty$. Recall S is finite. If $j \in S$ is a transient state (see FELLER [1950]) then $q^n_{ij} \to 0$ for all i where q^n_{ij} is the (ij)-th element in $Q^n(f^*,g^*)$. If j is a recurrent state, $v^*(j) = 0$; otherwise the total expected pay-off will remain unbounded. Hence,

$$(Q^n(f,g)v)_i = \sum_j q^{(n)}_{ij} v^*_j = \sum_{j'} q^{(n)}_{ij'} v^*_{j'}$$

(Here summation is taken over all j' which are transient states.) But the last expression obviously goes to zero for every i, since $q^{(n)}_{ij'} \to 0$ whenever j' is transient and since S is finite. Thus the proof of the lemma is complete.

Now we continue the proof of theorem 4.1. From the lemma it follows that $v^*(s) = I(f^{*(\infty)}, g^{*(\infty)})(s)$. It is not hard to check that $f^{*(\infty)}$ and $g^{*(\infty)}$ are optimal strategies and that $v^*(s)$ is the value of the positive stochastic game. Thus the proof is complete.

Remark 4.3:

Lemma 4.1 is false if S is not finite.

5. Noncooperative Stochastic Games

Non-cooperative stochastic games were first studied by ROGERS [1969] and SOBEL [1969] who proved the existence of equilibrium strategies for the two players under the assumptions S, A and B are finite and that the play terminates with probability one. We shall prove a similar result when S is countable and A and B are finite. Precisely we prove the following theorem.

Theorem 5.1:

Let S be countable and A,B are finite. (For simplicity we shall assume $A(s) \equiv A$, $B(s) \equiv B$ for every s.) Let r_1 and r_2 be bounded measurable functions on $S \times A \times B$. Then in the discounted noncooperative stochastic game the two players have equilibrium stationary strategies.

Proof:

Let $f : S \to P_A$ and $g : S \to P_B$. Let v_g and u_f be the unique fixed points of the operators T_1 and T_2 on $M(s)$ defined as follows. For every $w \in M(s)$

$$(T_1 w)(s) = \max_{P_A} \left[r(s,\mu,g(s)) + \beta \int w(\cdot)\,dq(\cdot/s,\mu,g(s))\right]$$

and

$$(T_2 w)(s) = \max_{P_B} \left[r_2(s,f(s),\lambda) + \beta \int w(\cdot)\,dq(\cdot/s,f(s),\lambda)\right].$$

Define $G: P_A^S \times P_A^S \to P_A^S \times P_B^S$ as follows. For every $f,g \in P_A^S \times P_B^S$, $G(f,g) = \{(f',g'): u_f(s) = r_2(s,f(s),g'(s)) + \beta \int u_f(\cdot)dq(\cdot/s,f(s),g'(s))$ and $v_g(s) = r_1(s,f'(s),g(s)) + \beta \int v_g(\cdot)dq(\cdot/s,f'(s),g(s))$ for each $s \in S\}$. Clearly $G(f,g)$ is nonempty, compact and convex for every $(f,g) \in P_A^S \times P_B^S$. We shall now prove that the set valued function G is upper semicontinuous, that is $y^n \in G(f^n,g^n)$, $y^n \to y^0, (f^n,g^n) \to (f^0,g^0) \Rightarrow y^0 \in G(f^0,g^0)$. For this we need the following lemma.

Lemma 5.1:

Let S be countable, A and B be finite. Let u_n be the fixed point of the operator T_n on $M(s)$ defined as above corresponding to the function f_n. Suppose $u_n(s) \to u(s)$ and $f_n(s) \to f(s)$ for every $s \in S$. Then u is the fixed point of the operator T associated with f.

Proof:

$u_n(s) = \max[r_2(s,f_n(s),\lambda) + \beta u_n(\cdot)dq((\cdot/s,f_n(s),\lambda))$. Let $f_n(s) = (\xi_1^n(s),\xi_2^n(s),\ldots,\xi_k^n(s))$ and $f(s) = (\xi_1(s),\ldots,\xi_k(s))$. Since $u_n(s) \to u(s)$ and $u_n's$ are uniformly bounded

$$\int u_n(\cdot)dq(\cdot/s,a_i,b_j) \to \int u(\cdot)dq(\cdot/s,a_i,b_j)$$

for every $a_i \in A$ and $b_j \in B$ and for $s \in S$.

$$\begin{aligned}
&\left|\int u_n(\cdot)dq(\cdot/s,f_n(s),\lambda) - \int u(\cdot)dq(\cdot/s,f_0(s),\lambda)\right| \\
&\le \left|\int u_n(\cdot)dq(\cdot/s,f_n(s),\lambda) - \int u(\cdot)dq(\cdot/s,f_n(s),\lambda)\right| \\
&+ \left|\int u(\cdot)dq(\cdot/s,f_n(s),\lambda) - \int u(\cdot)dq(\cdot/s,f(s),\lambda)\right| \\
&\le \sum_i \xi_i^n(s)\left|\int (u_n(\cdot) - u(\cdot))dq(\cdot/s,a_i,\lambda)\right| \\
&+ \sum_i |\xi_i^n(s) - \xi(s)| \int |u(\cdot)| dq(\cdot/s,a_i,\lambda).
\end{aligned}$$

Here both the terms tend to zero uniformly in λ for every fixed s as $n \to \infty$. From this it follows

$$u(s) = \max_{P_B}(r_2(s,f(s),\lambda) + \beta \int u(\cdot)dq(\cdot/s,f(s),\lambda)).$$

Since the fixed point is unique, u is the fixed point of the operator T associated with f. We now continue the proof of the theorem. From the lemma, it is not hard to check that G is upper semicontinuous. Hence we can conclude from Kakutani's fixed point theorem that there exist (f^*,g^*) such that $(f^*,g^*) \in G(f^*,g^*)$, that is,

$$\begin{aligned}
u_{f^*}(s) &= \max_{P_B}[r_2(s,f^*(s),\lambda) + \beta \int u_{f^*}(\cdot)dp(\cdot/s,f^*(s),\lambda)] \\
&= r_2(s,f^*(s),g^*(s)) + \beta \int u_{f^*}(\cdot)dq(\cdot/s,f^*(s),g^*(s))
\end{aligned}$$

and

$$\begin{aligned}
v_{g^*}(s) &= \max_{P_A}[r_1(s,\mu,g^*(s) + \beta \int v_{g^*}(\cdot)dq(\cdot/s,\mu,g^*(s)) \\
&= r_1(s,f^*(s),g^*(s)) + \beta \int v_{g^*}(\cdot)dq(\cdot/s,f^*(s),g^*(s)).
\end{aligned}$$

From these equations it is not hard to check that

$$u_{f^*}(s) = I_2(f^{*(\infty)}, g^{*(\infty)})(s) = \sup_{\Gamma} I_2(f^{*(\infty)}, \Gamma)(s)$$

and

$$v_{g^*}(s) = I_1(f^{*(\infty)}, g^{*(\infty)})(s) = \sup_{\Pi} I_1(\Pi, g^{*(\infty)})(s).$$

That is, $(f^{*(\infty)}, g^{*(\infty)})$ is a pair of equilibrium strategies for the two players. This terminates the proof of the theorem.

Remark 5.1:

When $r_2 = -r_1$ for every s, a and b then we are in the set up of stochastic games considered in section 3.

Remark 5.2:

Theorem 5.1 is a slight extension of [ROGERS 1969] and [SOBEL 1969]. The proofs in [ROGERS 1969] and [SOBEL 1969] with some modifications, apply to theorem 5.1. Alternatively, one could build a proof around Derman's "Markovian Sequential Control processes – denumerable State space", Jour. Math. Analy and Appl. Vol. 10, 295–302, 1965. The author is grateful to one of the referees for pointing out this reference.

6. Miscellaneous Remarks

We now present an example to show that theorem 4.1 is false when S is countable and compact and A and B are finite.

Example:

Let $S = \{1/2, 2/3, 3/4, \ldots\} \cup \{1\} \cup \{t\}$ and $A = \{0,1\}$ $B = \{0\}$. Let $r(s,0,0\} = s$ and $r(s,1,0) \equiv 0$ and t is a terminal state where you do not receive any reward. The transition probabilities are given as follows

$$q(\cdot/n/n+1, 1, 0) = \delta(n+1/n+2)$$

$$q(\cdot/n/n+1, 0, 0) = \delta(t)$$

$$q(\cdot/1, 1, 0) = \delta(1)$$

$$q(\cdot/1, 0, 0) = \delta(t).$$

This example is actually a maximizing one-person game (dynamic programming). Plainly the value of the game $v^*(s) \equiv 1$ for $s \neq t$ and $v^*(t) = 0$; it is not hard to check that there is no optimal stationary strategy. The same example shows that lemma 4.1 is false. Let $f^*(s) \equiv 1$. Then it follows

$$r(s, f^*(s), 0) + \int v^*(\cdot) dq(\cdot/s, f^*(s), 0) = v^*(s)$$

for every s. Obviously $I(f^{*(\infty)}, \{0\})(s) \equiv 0$ for every s but $v^*(s) \equiv 1$ for every $s \neq t$; that is,

$$v^*(s) \neq I(f^{*(\infty)}, \{0\})(s) \quad \text{for any} \quad s \neq t.$$

We conjecture that theorem 5.1 will be true with the following assumptions: S, A, B are compact metric and $A(s) \equiv A$, $B(s) \equiv B$ for every $s \in S$. Suppose

$r_1(s,a,b)$, $r_2(s,a,b)$ and $(q(\cdot/s,a,b)$ are continuous in $S \times A \times B$. In the proof of theorem 5.1, measurability question does not arise since S is assumed to be countable. The conjecture is true in the discounted case – see theorem 3.1.

Acknowledgment

I wish to express my sincere thanks to Professor ROBERT AUMANN (who suggested theorem 3.2), Professor ASHOK MAITRA and the referees for several useful suggestions. My thanks are also due to the editor for some useful comments.

References

BLACKWELL, D.: Discounted dynamic programming, Ann. of Math. Stat. **36**, 1965.

FELLER, W.: An introduction to probability theory and its applications, Vol. 1, 1950 (second edition), John Wiley and Sons, Inc.

GILLETTE, D.: Stochastic games with zero-stop probabilities, Ann. of Math. Studies **39**, 1957, Princeton University Press, Princeton, New Jersey.

HINDERER, K.: Foundations of Non-stationary dynamic programming with discrete time parameter, Lecture notes in operations research and Mathematical systems **33**, 1970, Springer-Verlag.

HOFFMAN, A. J., and R. M. KARP: On non-terminating stochastic games, Management Science **12**, 1966.

KURATOWSKI, K.: Topology, Vol. 1, 1966. Academic Press, New York and London.

MAITRA, A.: Discounted dynamic programming on compact metric spaces, Sankhya, Series A, **30**, 1968.

MAITRA, A., and T. PARTHASARATHY: On stochastic games, Journ. Opti. Theory and its Appl. **5**, 1970.

MAITRA, A., and T. PARTHASARATHY: On stochastic games II, Journ. Opti. Theory and its Appl. **7**, 1971.

OLECH, C.: A note concerning set valued measurable functions, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astron. Phys. **13**, 1965.

PARTHASARATHY, T.: Discounted and positive stochastic games, Bull. Amer. Math. Soc. **77**, 1971.

PARTHASARATHY, T., and T. E. S. RAGHAVAN: Some topics in two-person games 1971, American Elsevier Publishing Company, New York.

ROGERS, P. D.: Non-zero sum stochastic games; Ph. D. thesis submitted to the University of California, Berkeley 1969.

SHAPLEY, L. S.: Stochastic games, Proc. National Acad. Sci. U.S.A. **39**, 1953.

SOBEL, M. J.: Non-cooperative stochastic games, Report No. 21, Department of Administrative Sciences, Yale University 1969.

TAKAHASHI, M.: Stochastic games with infinitely many strategies, Journal of Sciences, Hiroshima University, **26**, 1962.