JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 31, No. 4, AUGUST 1980

Relations Between Game Parameters, Value, and Optimal Strategy Spaces in Stochastic Games and Construction of Games with Given Solution

O. J. VRIEZE[1] AND S. H. TIJS[2]

Communicated by Y. C. Ho

Abstract. Results of Bohnenblust, Karlin, and Shapley and results of Shapley and Snow, concerning solutions of matrix games, are extended to the class of discounted stochastic games. Prior to these extensions, relations between the game parameters, value, and optimal stationary strategy spaces are established. Then, the inverse problem of constructing stochastic games, given the solution, is considered.

Key Words. Discounted stochastic games, characterization of solutions, game construction with prescribed solution.

1. Introduction

This paper considers relations between the game parameters and the solution of two-person, zero-sum stochastic games. As a result, this leads one to investigate to what extent certain game parameters can be constructed, when the other game parameters are given and when the solution is prescribed. To be specific, for the discounted stochastic game, we tackle two problems in Section 4. First of all, we consider the problem of the construction of a suitable reward function, given the other parameters and the solution; secondly, we consider the construction of a suitable transition probability map, given the other parameters and the solution. As a preamble, the construction of matrix games with prescribed solutions, due to Bohnenblust, Karlin, and Shapley, is reviewed and extended slightly in Section 3.

[1] Research Associate, Department of Operations Research, Mathematical Centre, Amsterdam, Holland.
[2] Senior Lecturer, Department of Mathematics, Catholic University, Nijmegen, Holland.

0022-3239/80/0800-0501$03.00/0

2. Preliminaries

In this paper, we consider two-person zero-sum games with finite state and action spaces. A game will be denoted by $\langle G_1, G_2, \ldots, G_n \rangle$, where $G_1, G_2, \ldots, G_n$ are the constituent game parameters. These parameters include, for example, state space, action spaces, reward function, and transition probability map. The nonconstituent game parameters prescribe the strategy spaces from which the players may take a strategy, and also the manner in which, for a pair of strategies of both players, the corresponding payoff is calculated. So, when player 1 uses strategy π_1 and player 2 uses strategy π_2, then the game parameters determine a payoff, say $f(\pi_1, \pi_2)$, which is paid to player 1 by player 2. Note that $f(\pi_1, \pi_2)$ is a vector for stochastic games, where each coordinate corresponds with one of the states (the initial state). The game is said to have a value, if

$$\sup_{\pi_1} \inf_{\pi_2} f(\pi_1, \pi_2) = \inf_{\pi_2} \sup_{\pi_1} f(\pi_1, \pi_2),$$

componentwise. The value of a game will be denoted by $\operatorname{val}(\langle G_1, \ldots, G_n \rangle)$. A strategy π_1 will be called an optimal strategy for player 1, when

$$\inf_{\pi_2} f(\pi_1, \pi_2) \geq \operatorname{val}(\langle G_1, \ldots, G_n \rangle),$$

and a strategy π_2 is optimal for player 2, if

$$\sup_{\pi_1} f(\pi_1, \pi_2) \leq \operatorname{val}(\langle G_1, \ldots, G_n \rangle).$$

The set of optimal strategies for player i, $i = 1, 2$, is denoted by O_i. In general, a game may or may not have a value, or it may have a value, but no optimal strategies for one or both players. If we speak of the solution of a game, we mean the value together with the sets of optimal strategies.

3. Review of Some Results for Matrix Games

In this section, we recall some results of Bohnenblust, Karlin, and Shapley (Ref. 1 and Ref. 2, Chapter 3) and give a slight extension to be used later on. The fundamental simplices of mixed strategies for player 1 and player 2 in matrix games with m rows and n columns will be denoted by X_m and Y_n, respectively. We assign to a convex polyhedron X in X_m three integers $d_1(X)$, $t_1(X)$, $u_1(X)$, as follows:

(i) $d_1(X) - 1$ is the dimension of X;

(ii) $d_1(X)+t_1(X)$ is the number of elements in the carrier $C(x)$ of X,

$$C(x)=\{i \in \{1, 2, \ldots, m\}: x(i)>0, \text{ for some } x \in X\};$$

(iii) $u_1(X)$ is the number of unnatural faces of X, (i.e., faces which are not entirely contained in the relative boundary of X_m).

We assign to a convex polyhedron Y in Y_n, in a similar way, integers $d_2(Y)$, $t_2(Y)$, $u_2(Y)$.

Definition 3.1. We say that the pair (X, Y), where X is a polyhedron in X_m and Y is a polyhedron in Y_n, possesses the BKS properties (Bohnenblust, Karlin, and Shapley properties), if

(i) $t_1(X)=t_2(Y)$, dimension relation;
(ii) $m \geq d_1(X)+t_1(X)+u_2(Y)$, $n \geq d_2(Y)+t_2(Y)+u_1(X)$.

Important, for the construction of matrix games with prescribed optimal strategies, is the beautiful result of Bohenblust, Karlin, and Shapley (Ref. 1), given in the next lemma. A proof can also be found in Karlin (Ref. 2, Chapter 3).

Lemma 3.1. (i) If O_1 and O_2 are the sets of optimal mixed strategies for an $m \times n$ matrix game, then the pair (O_1, O_2) possesses the BSK properties.

(ii) If a pair (O_1, O_2) of convex polyhedra in $X_m \times Y_n$ possesses the BKS properties, and if V is a real number, then an $m \times n$ matrix game can be constructed, such that V is the value and O_1 and O_2 are the sets of optimal strategies for player 1 and player 2, respectively.

In the following, we need a slight extension, which is given in the following corollary.

Corollary 3.1. Let $V \in \mathbb{R}$ and (O_1, O_2) be a pair of polyhedra in $X_m \times Y_n$, such that the BKS properties hold, and let $\epsilon > 0$. Then, there exists an $m \times n$ matrix game with value V and optimal strategy sets O_1 and O_2, such that the distance of each entry in the matrix to V is at most ϵ.

Proof. In view of Lemma 3.1, there is an $m \times n$ matrix $[a_{ij}]$ with solution V, O_1, O_2. Now, let $[b_{ij}]$ be the $m \times n$ matrix with

$$b_{ij} = V + s^{-1}(a_{ij} - V), \qquad \text{for } i = 1, \ldots, m, \qquad j = 1, \ldots, n,$$
$$\text{where } s \in (0, \infty).$$

Then, this matrix has the desired properties, if s is a sufficient large real number. □

For the convenience of the reader, we now give a sketch of the construction procedure presented in Ref. 1. Henceforth, (O_1, O_2) is a fixed pair of polyhedrons in $X_m \times X_n$, satisfying the BKS properties. So, we abbreviate $d_i(O_i)$, $u_i(O_i)$ to d_i, u_i; and we abbreviate $t_1(O_1) = t_2(O_2)$ to t. Without loss of generality, we suppose that

$$m = d_1 + t + u_2, \qquad n = d_2 + t + u_1,$$

and that the carrier $C(O_i)$ of O_i, $i = 1, 2$, is equal to $\{1, 2, \ldots, d_i + t)$. Let $X(d_1, d_1 + t + u_2)$ be a $d_1 \times (d_1 + t + u_2)$ matrix in which the rows form a system of d_1 linearly independent elements of O_1. And again, without loss of generality, we may suppose that the $d_1 \times d_1$ submatrix $X(d_1, d_1)$ is nonsingular (rearrange otherwise the coordinates in the carrier). Analogously, let $Y(d_2 + t + u_1, d_2)$ be a matrix, in which the d_2 columns are linearly independent elements of O_2 and where $Y(d_2, d_2)$ is nonsingular. Now, the construction of a matrix $A(m, n)$ with optimal strategy sets O_1 and O_2 proceeds in six steps, where a partition of $A(m, n)$ plays a role in nine still unknown submatrices, namely,

$$A(m, n) = \begin{bmatrix} A(d_1, d_2) & A(d_1, t) & A(d_1, u_1) \\ A(t, d_2) & A(t, t) & A(t, u_1) \\ A(u_2, d_2) & A(u_2, t) & A(u_2, u_1) \end{bmatrix}.$$

Here, $A(k, l)$ is a $k \times l$ matrix, $k, l \in \mathbb{N}$.

The following partitions correspond with the one above:

$$X(d_1, d_1 + t + u_2) = [X(d_1, d_1), X(d_1, t), X(d_1, u_2)],$$

$$Y(d_2 + t + u_1, d_2) = \begin{bmatrix} Y(d_2, d_2) \\ Y(t, d_2) \\ Y(u_1, d_2) \end{bmatrix}.$$

Now, we describe the steps of the construction.

(C1) Choose an arbitrary, nonsingular $t \times t$ matrix $A(t, t)$, and take a real number V unequal to zero, which will become the value of the matrix game that we are going to construct. In the following, $V(k, l)$ is a $k \times l$ matrix with a V in each entry.

(C2) $A(d_1, t)$ has to satisfy

$$V(d_1, t) = X(d_1, d_1)A(d_1, t) + X(d_1, t)A(t, t) + 0.$$

Hence, we define

$$A(d_1, t) = X(d_1, d_1)^{-1}(V(d_1, t) - X(d_1, t)A(t, t)).$$

(C3) Similarly, we define

$$A(t, d_2) = (V(t, d_2) - A(t, t)Y(t, d_2))(Y(d_2, d_2))^{-1}.$$

(C4) Now, it is necessary for $A(d_1, d_2)$ to satisfy

$$X(d_1, d_1)A(d_1, d_2) + X(d_1, t)A(t, d_2) + 0 = V(d_1, d_2).$$

So, $A(d_1, d_2)$ needs to be

$$A(d_1, d_2) = (X(d_1, d_1))^{-1}(V(d_1, d_2) - X(d_1, t)A(t, d_2)).$$

In view of (C3), we define

$$A(d_1, d_2) = (X(d_1, d_1))^{-1}(V(d_1, d_2) - X(d_1, t)(V(t, d_2) - A(t, t)Y(t, d_2))(Y(d_2, d_2))^{-1}).$$

(C5) The dimension of an unnatural face F_j, $j = 1, \dots, u_1$, of O_1 is at most $d_1 - 2$, and F_j has a carrier lying in $\{1, 2, \dots, d_1 + t\}$. Hence, there exists a supporting hyperplane of O_1 through F_j and the origin, with a normal of the form

$$b_j = (b_{lj}, \dots, b_{d_1+t,j}, 0, \dots, 0),$$

such that

$$\pi_1 b_j^T = 0, \qquad \text{for each } \pi_1 \in F_j,$$

$$\pi_1 b_j^T > 0, \qquad \text{for } \pi_1 \in O_1 \backslash F_j.$$

Choose now as the $(d_2 + t + j)$th column of $A(m, n)$ the vector $b_j^T + V(d_1 + t + u_2, 1)$.

(C6) Similarly, unnatural faces of O_2 give rise to the last u_2 rows of $A(m, n)$ of the form $c_i + V(1, d_2 + t + u_1)$, $i = 1, \dots, u_2$.

The matrix $A(m, n)$ constructed in this way has V, O_1, O_2 as its solution.

4. Structure of the Optimal Strategy Sets in Discounted Stochastic Games and Construction Problems

A discounted two-person, zero-sum stochastic game can be identified by a sixtuple

$$\Gamma = \langle S, \{A_1(k): k \in S\}, \{A_2(k): k \in S\}, r, P, \beta \rangle,$$

where

(i) $S = \{1, 2, \dots, N\}$;

(ii) $A_1(k)$ and $A_2(k)$ are nonempty finite sets for each $k \in S$;

(iii) r is a real-valued function defined on the set of triples $T = \{(k, a_1, a_2): k \in S, a_1 \in A_1(k), a_2 \in A_2(k)\}$;

(iv) P is a map from T into the set $\mathcal{P}(S)$ of probability measures on S;

(v) $\beta \in (0, 1)$.

S will be called the state space, $A_i(k)$ the set of pure actions of player i, $i = 1, 2$, in state k, r the reward function for player 1, $-r$ the reward function for player 2, P the transition probability map, and β the discount factor. Such a stochastic game corresponds to a dynamic system, where the dynamic behavior and the rewards are influenced in the following way by the players at discrete points in time (called stages), say $t = 0, 1, 2, \ldots$.

At each stage t, the players observe the current state of the system. Then, they have to select, independently of one another, an action. If, at stage t, the system is in state k, and if player 1 selects action $a_1 \in A_1(k)$ and player 2 selects action $a_2 \in A_2(k)$, then two things happen:

(a) player 1 obtains an immediate reward $r(k, a_1, a_2)$ from player 2;

(b) the system moves with probability $P(k, a_1, a_2)\{l\}$, which we denote by $p(l|k, a_1, a_2)$ from now on, to state $l \in S$, which will be observed at the next stage $t+1$.

Furthermore, one supposes that a reward r, paid at time t, has worth $\beta^t r$ at time 0 (called the discounted reward) and that player 1 (player 2) wants to maximize (minimize) the total discounted expected reward.

For a two-person game in normal form with finite, pure strategy spaces A_1 and A_2, and payoff function $k: A_1 \times A_2 \to \mathbb{R}$, the value of the mixed extension will be denoted by $\mathrm{val}_{A_1 \times A_2}[k(\cdot, \cdot)]$ and is also called the value of the matrix game $[k(\cdot, \cdot)]$. A mixed strategy for player i in such a game is a probability measure on A_i, $i = 1, 2$.

A stationary strategy for player i in the stochastic game Γ is an N-tuple

$$\pi_i = (\pi_i(1), \pi_i(2), \ldots, \pi_i(N)),$$

where $\pi_i(k)$ is a probability measure on $A_i(k)$, $k \in S$. Playing π_i means that, each time $t \in \{0, 1, 2, \ldots\}$ the system is in state $k \in S$, player i chooses his action according to the probability measure $\pi_i(k)$. In the following, we restrict our attention to the sets of stationary strategies for both players. The next lemma is a well-known result in the theory of stochastic games, and a proof of it can be found in several papers, the first of which was a paper of Shapley (Ref. 3).

Lemma 4.1. The stochastic game Γ has a value. The vector

$$V := (V(1), V(2), \ldots, V(N))$$

is the value iff

$$V(k) = \operatorname*{val}_{A_1(k)\times A_2(k)} \left[r(k,\cdot,\cdot) + \beta \sum_{l=1}^{N} p(l|k,\cdot,\cdot)V(l) \right], \qquad k = 1, 2, \ldots N.$$

A stationary strategy $(\pi_i(1), \ldots, \pi_i(N))$ is optimal for player i in the game Γ iff, for each $k \in S$, the probability measure $\pi_i(k)$ on $A_i(k)$ is an optimal mixed strategy in the matrix game

$$\left[r(k,\cdot,\cdot) + \beta \sum_{l=1}^{N} p(l|k,\cdot,\cdot)V(l) \right].$$

From Lemma 4.1 and Lemma 3.1, we obtain the following corollary.

Corollary 4.1. The set of optimal stationary strategies O_i for player i, $i = 1, 2$, in the discounted stochastic game is equal to the Cartesian product

$$\mathop{\times}_{k=1}^{N} O_i(k),$$

where $O_i(k)$ is the set of optimal mixed strategies for player i in the matrix game

$$\left[r(k,\cdot,\cdot) + \beta \sum_{l=1}^{N} p(l|k,\cdot,\cdot)V(l) \right].$$

For each $k \in S$, the pair $(O_1(k), O_2(k))$ has the BKS properties.

Now, we want to sketch an extension to discounted stochastic games of a result by Shapley and Snow in Ref. 4, concerning the characterization of extreme optimal mixed strategies for matrix games (Ref. 2, pp. 44–51). Roughly speaking, Shapley and Snow showed that, for each pair (π_1, π_2) of extreme strategies for a matrix game A, there exists a square submatrix M of A, such that π_1 and π_2 can be calculated explicitly in terms of M by solving systems of linear equalities with coefficient matrix M (Ref. 2, Theorems 2.4.1 and 2.4.3). This also suggested a method of obtaining all the extreme optimal strategies of a matrix game (Ref. 2, pp. 50–51).

Let us now look at our stochastic game Γ. Note that $\pi_i \in O_i$ is an extreme point of O_i iff $\pi_i(k)$ is an extreme point of $O_i(k)$ for each $k \in S$, $i = 1, 2$. But then it is obvious that we can extend the above results for matrix games as follows.

Theorem 4.1. Let Γ, O_1, O_2 be as above.

Let π_1 and π_2 be extreme points of O_1 and O_2. Then, there exists a stochastic subgame

$$\langle S, \{B_1(k): k \in S\}, \{B_2(k): k \in S\}, \tilde{r}, \tilde{P}, \beta \rangle$$

of Γ, where

(i) for all $k \in S$, the sets $B_1(k)$ and $B_2(k)$ are subsets of $A_1(k)$ and $A_2(k)$, respectively, with an equal number of elements;

(ii) $\tilde{r}$ and $\tilde{P}$ are restrictions of the maps r and P to the set of triples

$$\tilde{T} = \{(k, a_1, a_2): k \in S, a_1 \in B_1(k), a_2 \in B_2(k)\},$$

and such that $\pi_1(k)$ and $\pi_2(k)$ can be calculated in the Shapley–Snow manner, where the square matrix game

$$\left[\tilde{r}(k, \cdot, \cdot) + \beta \sum_{l=1}^{N} \tilde{p}(l|k, \cdot, \cdot) V(l)\right],$$

with pure strategy sets $B_1(k)$ and $B_2(k)$, plays the role of M.

This theorem also suggests a method of finding the finite number of extreme optimal strategies of O_1 and O_2, when the value of the game is known, by looking at the finite number of stochastic subgames in which, at each state, both players have an equal number of pure actions.

Now, we fix S, $\{A_1(k): k \in S\}$, $\{A_2(k): k \in S\}$, and analyze two problems.

Problem 4.1. Let V, P, β, and $O_i(1), \ldots, O_i(N)$, $i = 1, 2$, be given, where $O_i(k)$ is, for each $k \in S$, a convex polyhedron in the set of probability measures on $A_i(k)$. Construct a function r such that the following properties hold:

(P1) $V = \text{val}(\langle S, \{A_1(k): k \in S\}, \{A_2(k): k \in S\}, r, P, \beta\rangle)$;

(P2) $\times_{k=1}^{N} O_i(k)$ is the set of optimal stationary strategies for player i.

Solution of Problem 4.1. If there is an r with the specific desired properties, then $(O_1(k), O_2(k))$ has the BKS properties for each $k \in S$, by Corollary 4.1. Conversely, if $(O_1(k), O_2(k))$ has, for each $k \in S$, the BKS properties, then we can construct (in view of Lemma 3.1) a function $b(k, \cdot, \cdot)$ on $A_1(k) \times A_2(k)$, such that

$$\text{val}[b(k, \cdot, \cdot)] = V(k)$$

and such that $O_1(k)$ and $O_2(k)$ are the optimal strategy spaces of the matrix game $[b(k, \cdot, \cdot)]$. Now, for

$$k \in S, \qquad i \in A(k), \qquad j \in B(k),$$

let

$$r(k, i, j) = b(k, i, j) - \beta \sum_{l=1}^{N} p(l|k, i, j) V(l).$$

Then, with the aid of Lemma 4.1 and Corollary 4.1, we may conclude that r is a function, such that Properties (P1) and (P2) hold. Hence, we have proved the following theorem.

Theorem 4.2. Problem 4.1 can be solved iff, for each $k \in S$, the pair $(O_1(k), O_2(k))$ possesses the BKS properties.

Problem 4.2. Let V, r, β, and $O_i(1), \ldots, O_i(N)$, $i = 1, 2$, be given, where $O_i(k)$ is, for each $k \in S$, a convex polyhedron as in Problem 4.1. Construct a map P such that Properties (P1) and (P2) hold.

Discussion of Problem 4.2. Similarly as for Problem 4.1, it is necessary that, for each $k \in S$, the pair $(O_1(k), O_2(k))$ has the BKS properties. But now, r, β, V cannot be chosen independently, as the following theorem shows.

Theorem 4.3. Necessary and sufficient conditions for the existence of a map P, such that Property (P1) holds, are given by the following system of inequalities:

$$m \leq \beta^{-1}(V(k) - w(k)) \leq M, \qquad \text{for each } k \in S, \tag{1}$$

where

$$m = \min_{k \in S} V(k), \qquad M = \max_{k \in S} V(k), \qquad w(k) = \mathrm{val}[r(k, \cdot, \cdot)].$$

Proof. (a) First, suppose that there exists a map P such that Property (P1) holds. Let $(\pi_1^*(1), \pi_1^*(2), \ldots, \pi_1^*(N))$ be an optimal stationary strategy for player 1 in the stochastic game, and let $\tilde{\pi}_2(k)$ be an optimal strategy for player 2 in the matrix game $[r(k, \cdot, \cdot)]$. Then, in view of Lemma 4.1, we have

$$V(k) \leq r(k, \pi_1^*(k), \tilde{\pi}_2(k)) + \beta \sum_{l=1}^{N} p(l|k, \pi_1^*(k), \tilde{\pi}_2(k)) V(l)$$

$$\leq r(k, \pi_1^*(k), \tilde{\pi}_2(k)) + \beta M \leq w(k) + \beta M.$$

Note that $r(k, \pi_1(k), \pi_2(k))$ is the expected immediate reward when the players use the mixed strategies $\pi_1(k)$ and $\pi_2(k)$, respectively; similarly, $p(l|k, \pi_1(k), \pi_2(k))$ is defined.

Analogously, one can prove that

$$V(k) \geq w(k) + \beta m.$$

Hence, (1) holds.

(b) Now, suppose that (1) holds. Take

$$s_* \in S, \qquad \text{with } V(s_*) = m,$$
$$s^* \in S, \qquad \text{with } V(s^*) = M.$$

In view of (1), there exists an $\alpha_k \in [0, 1]$, such that

$$\beta^{-1}(V(k) - w(k)) = \alpha_k m + (1 - \alpha_k)M = \alpha_k V(s_*) + (1 - \alpha_k) V(s^*).$$

Now, choose the map P in such a way that

$$p(s_*|k, i, j) = \alpha_k,$$
$$p(s^*|k, i, j) = 1 - \alpha_k,$$
$$p(l|k, i, j) = 0, \qquad \text{if } l \neq s_*, s^*.$$

Then,

$$\mathrm{val}[r(k, \cdot, \cdot) + \beta \sum_{l=1}^{N} p(l|k, \cdot, \cdot) V(l)] = \mathrm{val}[r(k, \cdot, \cdot) + V(k) - w(k)] = V(k),$$

for each $k \in S$. Thus, by Lemma 4.1, we can conclude that Property (P1) holds. □

From the first two inequalities in part (a) of the proof of Theorem 4.3, we immediately infer the following theorem.

Theorem 4.4. A necessary condition for the existence of a map P, such that a strategy π_1^* is optimal for player 1, is that

$$V(k) \leq \min_{\pi_2(k)} r(k, \pi_1^*(k), \pi_2(k)) + \beta M, \qquad \text{for each } k \in S,$$

and that a strategy π_2^* is optimal for player 2 is that

$$V(k) \geq \max_{\pi_1(k)} r(k, \pi_1(k), \pi_2^*(k)) + \beta m, \qquad \text{for each } k \in S.$$

In the following theorem, we give sufficient conditions for solving Problem 4.2. Here, m and M are as in Theorem 4.3.

Theorem 4.5. If, for each $k \in S$, the pair $(O_1(k), O_2(k))$ has the BKS properties and if, for each $k \in S$ and each $(i, j) \in A_1(k) \times A_2(k)$,

$$r(k, i, j) + \beta m < V(k) < r(k, i, j) + \beta M, \tag{2}$$

then there is a solution map P for Problem 4.2.

Proof. First, we construct, for each $k \in S$, a matrix game $[b(k, \cdot, \cdot)]$ with pure strategy spaces $A_1(k)$ and $A_2(k)$, such that

(i) $\text{val}[b(k, \cdot, \cdot)] = V(k)$;
(ii) $O_1(k)$ and $O_2(k)$ are the optimal strategy sets for both players;
(iii) for each $k \in S$ and each $i \in A_1(k)$ and $j \in A_2(k)$, we have

$$r(k, i, j) + \beta m \leq b(k, i, j) \leq r(k, i, j) + \beta M. \tag{3}$$

This is possible in view of (2) and Corollary 3.1. Let $s_*, s^* \in S$, be such that

$$V(s_*) = m, \qquad V(s^*) = M.$$

By (3), for each (k, i, j), there exists an $\alpha \in [0, 1]$ such that

$$\beta^{-1}(b(k, i, j) - r(k, i, j)) = \alpha m + (1 - \alpha)M = \alpha V(s_*) + (1 - \alpha) V(s^*).$$

Now, define

$$p(s_* | k, i, j) = \alpha,$$
$$p(s^* | k, i, j) = 1 - \alpha,$$
$$p(l | k, i, j) = 0, \qquad \text{if } l \neq s_*, s^*.$$

Obviously, in such a way we obtain a map P, which is a solution of Problem 4.2. □

In the rest of this section, we restrict our attention to Problem 4.2 for the important subclass of stochastic games where, at each state, one of the players is a dummy. Recall that a player i in the game Γ is called a dummy at state $k \in S$, if the player has at state k only one action available. We note that discounted Markov decision problems can be seen as such games, where at each state player 2 is a dummy.

Furthermore, note that, for stochastic games with a dummy at each state, there exist pure optimal actions for each player. If we denote by $Z_i(k)$ the set of pure optimal strategies of player i in the matrix game

$$[r(k, \cdot, \cdot) + \beta \sum_{l \in S} p(l | k, \cdot, \cdot) V(l)],$$

then $O_i(k)$ is completely determined by $Z_i(k)$, because $O_i(k)$ is the set of probability measures μ on $A_i(k)$ with

$$\mu(Z_i(k)) = 1.$$

Let

$$m = \min_{k \in S} V(k), \qquad M = \max_{k \in S} V(k), \qquad .$$

and suppose that player 2 is a dummy at state k in the stochastic game Γ. Then, the kth equation of the system of value-determining equations (see Lemma 4.1) becomes

$$V(k) = \max_{i \in A_1(k)} r(k, i, a_2) + \beta \sum_{l \in S} p(l|k, i, a_2) V(l),$$

where

$$A_2(k) = \{a_2\}.$$

Moreover,

$$Z_1(k) = \left\{ i^* \in A_1(k): r(k, i^*, a_2) + \beta \sum_{l \in S} p(l|k, i^*, a_2) V(l) = V(k) \right\}.$$

Hence, we may conclude that the following properties hold.

$I_k(2)$. If $i \in Z_1(k)$, then $\beta^{-1}(V(k) - r(k, i, a_2)) \in [m, M]$;

$J_k(2)$. If $i \notin Z_1(k)$, then $\beta^{-1}(V(k) - r(k, i, a_2)) \in (m, \infty)$.

In the case player 1 is a dummy at state k, the kth equation of the system of value-determining equations becomes

$$V(k) = \min_{j \in A_2(k)} r(k, a_1, j) + \beta \sum_{l \in S} p(l|k, a_1, j) V(l),$$

where

$$A_1(k) = \{a_1\}.$$

Properties similar to $I_k(2)$ and $J_k(2)$ can be stated, denoted $I_k(1)$ and $J_k(1)$. The only difference is that, for $J_k(1)$, we have to replace (m, ∞) by $(-\infty, M)$.

Now, for stochastic games with a dummy at each state, Problem 4.2 can be reformulated as follows.

Problem 4.3. Given V, r, β, $\langle Z_1(k): k \in S \rangle$, $\langle Z_2(k): k \in S \rangle$, with $\phi \neq Z_1(k) \subset A_i(k)$, for each $k \in S$ and $i \in \{1, 2\}$, find a map P such that Property (P1) holds and also Property (P3) given below.

(P3) $\times_{k=1}^{N} Z_i(k)$ is the set of pure optimal stationary strategies for player i.

The solution is given in the following theorem.

Theorem 4.6. Let Γ be a stochastic game with a dummy at each state. Let m and M be as above. Then, Problem 4.3 is solvable iff Properties $I_k(i)$ and $J_k(i)$ hold for each $k \in S$, where player i is a dummy, $i = 1, 2$.

Proof. The *only if* part of the theorem is already shown above. So, it remains to prove the *if* part, i.e., to choose a suitable map P when Properties

$I_k(i)$ and $J_k(i)$ hold. Fix $k \in S$. Let

$$s_* \in S, \qquad \text{such that } V(s_*) = m,$$

$$s^* \in S, \qquad \text{such that } V(s^*) = M.$$

First, let $(a_1, a_2) \in Z_1(k) \times Z_2(k)$. Then, it follows, from Properties $I_k(1)$ and $I_k(2)$ (one of them holds), that there exists an $\alpha(k, a_1, a_2) \in [0, 1]$, such that

$$\beta^{-1}(V(k) - r(k, a_1, a_2)) = \alpha(k, a_1, a_2) V(s_*) + (1 - \alpha(k, a_1, a_2)) V(s^*).$$

For such (a_1, a_2), define

$$p(s_* | k, a_1, a_2) = \alpha(k, a_1, a_2),$$

$$p(s^* | k, a_1, a_2) = 1 - \alpha(k, a_1, a_2),$$

$$p(l | k, a_1, a_2) = 0, \qquad \text{for } l \neq s_*, s^*.$$

Now, let $(a_1, a_2) \notin Z_1(k) \times Z_2(k)$, and suppose that player $i \in \{1, 2\}$ is a dummy in state k. Then, take

$$p(s_* | k, a_1, a_2) = i - 1,$$

$$p(s^* | k, a_1, a_2) = 2 - i,$$

$$p(l | k, a_1, a_2) = 0, \qquad \text{for } l \neq s_*, s^*.$$

Now, it is easy to show that the map P, fixed in the above way, has the desired properties. □

In a forthcoming paper, we will look at the two construction problems for undiscounted stochastic games with the average reward criterion.

References

1. BOHNENBLUST, H. F., KARLIN, S., and SHAPLEY, L. S., *Solutions of Discrete Two-Person Games*, Annals of Mathematics Studies, Vol. 24, pp. 51–72, 1950.
2. KARLIN, S., *Mathematical Methods and Theory in Games, Programming, and Economics*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1959.
3. SHAPLEY, L. S., *Stochastic Games*, Proceedings of the National Academy of Sciences, USA, Vol. 39, pp. 1095–1100, 1953.
4. SHAPLEY, L. S., and SNOW, R. N., *Basic Solutions of Discrete Games*, Annals of Mathematics Studies, Vol. 24, pp. 27–35, 1950.