JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 33, No. 1, JANUARY 1981

TECHNICAL NOTE

Characterizing Properties of the Value Function of Stochastic Games

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

Communicated by Y. C. Ho

Abstract. In the present note, the axiomatic characterization of the value function of two-person, zero-sum games in normal form by Vilkas and Tijs is extended to the value function of discounted, two-person, zero-sum stochastic games. The characterizing axioms can be indicated by the following terms: objectivity, monotony, and sufficiency for both players; or sufficiency for one of the players and symmetry. Also, a characterization without using the monotony axiom is given.

Key Words. Stochastic games, values of stochastic games, characterization of the value function.

1. Introduction

We consider infinite stage β-discounted, stochastic, two-person, zero-sum games with finite state space and finite action spaces. Such a game Γ can be characterized by a five-tuple

$$\langle S, \{A_i(k)\colon k \in S, i \in \{1, 2\}\}, r, P, \beta\rangle,$$

where $S = \{1, 2, \ldots, N\}$ is the *state space*, $A_1(k) = \{1, 2, \ldots, m_k\}$ is the *pure action space for player 1* in state k, $A_2(k) = \{1, 2, \ldots, n_k\}$ is the *pure action space for player 2* in state k, and r is the real-valued *reward function*, defined on the set of triples

$$T = \{(k, i, j) : k \in S, i \in A_1(k), j \in A_2(k)\};$$

[1] Senior Lecturer, University of Nijmegen, Nijmegen, Holland.
[2] Research Associate, Mathematical Center, Amsterdam, Holland.

0022-3239/81/0100-0145$03.00/0

further,

$$P = \{p(l \mid k, i, j) : l \in S, (k, i, j) \in T\}$$

is the *set of transition probabilities*, and $\beta \in (0, 1)$ is the *discount factor*.

In Ref. 1, Shapley proved that such a game Γ possesses a value $V(\Gamma) \in \mathbb{R}^N$ and that both players possess optimal stationary strategies. Furthermore, a stationary strategy $\pi_i = (\pi_{i1}, \ldots, \pi_{iN})$ is optimal for player i iff, for each $k \in S$, the mixed action π_{ik} in state k is an optimal action for player i in the matrix game

$$\left[r(k, i, j) + \beta \sum_{l \in S} p(l \mid k, i, j) V(\Gamma)(l) \right]_{i=1, j=1}^{m_k\ n_k}$$

whose value is $V(\Gamma)(k)$. For a survey on stochastic games, we refer to Ref. 2. In the following, we denote by $G(N, \beta)$ the family of β-discounted stochastic games with N states and finite action spaces for both players.

In Ref. 3, Vilkas characterized the value function for matrix games in an axiomatic way. Tijs (Ref. 4) extended this idea to the characterization of the value function on the class of all determined, two-person, zero-sum games in normal form. The purpose of this note, which is a shortened version of Ref. 5, is to give an axiomatic description of the value function $V : G(N, \beta) \to \mathbb{R}^N$; where, for each $\Gamma \in G(N, \beta)$, $V(\Gamma)$ is the value of the game Γ.

2. Characterizing Properties of the Value Function for Discounted Stochastic Games

We start with some definitions.

Definition 2.1. Let $\Gamma \in G(N, \beta)$, $i \in A_1(k)$; and let $\bar{\pi}_{1k}$ be a mixed action for player 1 in state k with $\bar{\pi}_{1k}(i) = 0$. We call action i *superfluous* (in view of action $\bar{\pi}_{1k}$) if, for each $j \in A_2(k)$,

$$r(k, i, j) + \beta \sum_{l \in S} p(l \mid k, i, j) V(\Gamma)(l) \leq r(k, \bar{\pi}_{1k}, j) + \beta \sum_{l \in S} p(l \mid k, \bar{\pi}_{1k}, j) V(\Gamma)(l).$$

An action $i \in A_1(k)$ is called *weakly superfluous* if, for each mixed action π_{1k} of player 1 in state k, there exists a mixed action $\tilde{\pi}_{1k}$ such that $\tilde{\pi}_{1k}(i) = 0$ and

$$\inf_{j \in A_2(k)} r(k, \pi_{1k}, j) + \beta \sum_{l \in S} p(l \mid k, \pi_{1k}, j) V(\Gamma)(l) \leq \inf_{j \in A_2(k)} r(k, \tilde{\pi}_{1k}, j) + \beta \sum_{l \in S} p(l \mid k, \tilde{\pi}_{1k}, j) V(\Gamma)(l).$$

It is obvious what is meant by superfluous and weakly superfluous actions for player 2. Now, we state our main result.

Theorem 2.1. A function $f: G(N, \beta) \to \mathbb{R}^N$ equals the value function iff f obeys the following axioms.

Axiom A1. *Objectivity.* If $\Gamma \in G(N, \beta)$ is such that, for state $k \in S$, we have

$$m_k = n_k = 1 \quad \text{and} \quad p(k \mid k, 1, 1) = 1,$$

then

$$f(\Gamma)(k) = (1-\beta)^{-1} r(k, 1, 1).$$

Axiom A2. *Monotony.* If all game parameters of Γ' and Γ'' in $G(N, \beta)$ are the same, except the reward functions r' and r'', and if $r' \leq r''$, then

$$f(\Gamma') \leq f(\Gamma'').$$

Axiom A3.i. *Sufficiency for player i, i = 1, 2.* If $\Gamma' \in G(N, \beta)$ results from $\Gamma \in G(N, \beta)$ by deleting a superfluous action of player i, then

$$f(\Gamma') = f(\Gamma).$$

Proof. It is left to the reader to show that V satisfies the axioms A1, A2, A3.1, and A3.2 (see Ref. 5, pp. 6–7). The proof of the "if" part of the theorem proceeds in two steps.

(i) First, we look at games $\Gamma \in G(N, \beta)$, which have a state k and $(i_0, j_0) \in A_1(k) \times A_2(k)$, such that

$$p(k \mid k, i, j) = 1, \quad \text{if } i = i_0 \text{ or } j = j_0,$$

and such that

$$\inf_{j \in A_2(k)} r(k, i_0, j) = r(k, i_0, j_0) = \sup_{i \in A_1(k)} r(k, i, j_0). \tag{1}$$

We want to show that

$$f(\Gamma)(k) = V(\Gamma)(k).$$

Obviously,

$$V(\Gamma)(k) = (1-\beta)^{-1} r(k, i_0, j_0).$$

Take a large real number M and look at $\Gamma', \Gamma'' \in G(N, \beta)$, which differ from Γ only in the reward functions r' and r'' as follows:

$$r'(l, i, j) = r''(l, i, j) = r(l, i, j), \quad \text{if } l = k \text{ and } i = i_0 \text{ or } j = j_0,$$

$$r'(l, i, j) = r(l, i, j) - M,\ r''(l, i, j) = r(l, i, j) + M, \quad \text{otherwise.}$$

It is obvious that, in the game Γ', for M large enough, each $i \in A_1(k) - \{i_0\}$ is superfluous in view of action i_0,

$$V(\Gamma)(k) = V(\Gamma')(k).$$

By Axiom A3.1, all actions $i \neq i_0$ may be deleted, without disturbing the f-value, resulting in a game with only one action i_0 for player 1 in state k.

In the new game, by Axiom A3.2 and (1), we can delete each action $j \neq j_0$ in state k. This results in a game $\tilde{\Gamma}$, with

$$f(\Gamma')(k) = f(\tilde{\Gamma})(k) = (1-\beta)^{-1} r(k, i_0, j_0),$$

by Axiom A1. So,

$$f(\Gamma')(k) = V(\Gamma)(k).$$

Analogously,

$$f(\Gamma'')(k) = V(\Gamma)(k).$$

But then,

$$f(\Gamma)(k) = V(\Gamma)(k),$$

because, by Axiom A2,

$$f(\Gamma') \leq f(\Gamma) \leq f(\Gamma'').$$

(ii) Now, take an arbitrary $\Gamma \in G(N, \beta)$, let $k \in S$. Consider the game $\Gamma^k \in G(N, \beta)$, which is constructed from Γ, by adding in state k an action i_0 for player 1 and an action j_0 for player 2 and by extending r and P of Γ such that

$$r(k, i, j) = (1-\beta) V(\Gamma)(k)$$

and

$$p(k \mid k, i, j) = 1, \qquad \text{if } i = i_0 \text{ or } j = j_0.$$

Clearly,

$$V(\Gamma^k) = V(\Gamma).$$

Furthermore, Γ^k is a game with state k of the type treated in the first step of the proof. Hence,

$$f(\Gamma^k)(k) = (1-\beta)^{-1} r(k, i_0, j_0) = V(\Gamma^k)(k) = V(\Gamma)(k).$$

It is easy to prove that, in the game Γ^k, i_0 is superfluous for player 1 (in view of $\hat{\pi}_{1k}$, if $\hat{\pi}_{1k}$ is the kth coordinate of an optimal stationary strategy $\hat{\pi}_1$ in the game Γ) and that also j_0 is superfluous for player 2 in Γ^k. But then, in view of Axioms A3.1 and A3.2,

$$V(\Gamma)(k) = f(\Gamma^k)(k) = f(\Gamma)(k).$$

This completes the proof of the theorem.

We note that the four axioms in Theorem 2.1 are independent (see Ref. 5, pp. 10–12). Another interesting property of the value function is the symmetry property

$$V(\Gamma^T) = -V(\Gamma),$$

where Γ^T is the stochastic game that we obtain from Γ by interchanging the names of the players. Now, Axiom A3.2 follows from Axiom A3.1 and Axiom A4 below. This implies that we have the following alternative characterization of the value function.

Theorem 2.2. A function $f: G(N, \beta) \to \mathbb{R}^N$ equals the value function iff f obeys Axioms A1, A2, A3.1, and Axiom A4 given below.

Axiom A4. $f(\Gamma^T) = -f(\Gamma)$, for all $\Gamma \in G(N, \beta)$.

Now, we give another characterization of the value function, in which the monotony property no longer plays a role.

Theorem 2.3. A function $f: G(N, \beta) \to \mathbb{R}^N$ equals the value function iff f satisfies Axiom A1 and Axiom A3.*iw* given below.

Axiom A3.iw. Weak Sufficiency for Player i, i = 1, 2. If $\Gamma' \in G(N, \beta)$ results from $\Gamma \in G(N, \beta)$ by deleting a weakly superfluous action i of player i, then

$$f(\Gamma') = f(\Gamma).$$

Proof. In the proof of Theorem 2.1, the only place where the monotony axiom is used is in the first step, where Γ is compared with Γ' and Γ''. From the last two games, superfluous actions could be deleted. But now we no longer need games Γ' and Γ'', because, directly in game Γ with a state k with a saddle point as in the proof of Theorem 2.1, at once all actions of player 1 unequal to i_0 can be deleted successively, as they are all weakly superfluous in view of action i_0. The remainder of the proof proceeds analogously as the proof of Theorem 2.1.

For an axiomatic characterization of the value function for other classes of dynamic games, we refer the reader to Ref. 5, pp. 14–16.

References

1. SHAPLEY, L. S., *Stochastic Games*, Proceedings of the National Academy of Sciences of the USA, Vol. 39, pp. 1095–1100, 1953.
2. PARTHASARATHY, T., and STERN, M., *Markov Games, A Survey*, University of Illinois, Chicago, Illinois, 1976.

3. VILKAS, E. I., *Axiomatic Definition of the Value of a Matrix Game*, Theory of Probability and Its Applications, Vol. 8, pp. 304–307, 1963.
4. TIJS, S. H., *A Characterization of the Value of Zero-Sum Two-Person Games*, University of Nijmegen, Nijmegen, Holland, Department of Mathematics, Report No. 7632, 1976; Naval Research Logistics Quarterly, Vol. 28, 1981.
5. TIJS, S. H., and VRIEZE, O. J., *Characterizing Properties of the Value Function of Stochastic Games*, Mathematical Center, Amsterdam, Holland, Report No. BW 110/79, 1979.