SIAM J. CONTROL AND OPTIMIZATION
Vol. 19, No. 5, September 1981
© 1981 Society for Industrial and Applied Mathematics
0363-0129/81/1905-0005 $01.00/0

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

EXISTENCE OF VALUE AND RANDOMIZED STRATEGIES IN ZERO-SUM DISCRETE-TIME STOCHASTIC DYNAMIC GAMES*

P. R. KUMAR† AND T. H. SHIAU†

Abstract. Two players with conflicting objectives are simultaneously controlling a discrete-time stochastic system. The goal of this paper is to analyze such zero-sum, discrete-time, stochastic systems when the two players are allowed to use randomized strategies.

Previous results have been restricted to systems with finite or compact state spaces. Such restrictions are usually untenable from the point of view of applications, since many applications frequently use either the integers or $\mathbb{R}^n$ as a state space. Our results are proved for complete, separable, metric spaces which are very useful for applications.

All previously known results emerge as special cases of our results. In addition, a variety of conjectures and open problems are resolved regarding the existence of a value function, its properties such as Borel measurability or continuity, and the existence for either or both players of optimal or ε-optimal stationary strategies.

1. Introduction. The goal of this paper is to analyze zero-sum discrete-time stochastic games where the two players are allowed to use randomized strategies (i.e., two players with conflicting objectives simultaneously controlling a stochastic system).

Starting with Shapley [1], many researchers, e.g., Everett [2], Maitra and Parthasarathy [3], [4] have treated such systems. However, all these treatments suffer from the fact that they restrict the state space to be either finite or compact. Many useful models of dynamic games however use a state space which is often the integers or $\mathbb{R}^n$—both of which fail to satisfy these restrictions. Even in some special cases which do satisfy these restrictions, earlier results are not applicable if the value function is not continuous. (We give such an example in § 7.) Also nonstationary (i.e., time-varying) systems, when made stationary by adjoining an additional state variable to count time, possess noncompact state spaces. The restriction of the state space to be either finite or compact is therefore untenable from the point of view of applications.

We provide in this paper a general theory of zero-sum discrete-time stochastic games which overcomes these restrictions. We consider two alternative models for such problems—a Borel model and a continuous model, both of which have complete, separable, metric state spaces and therefore include both integer state spaces and Euclidean state spaces. These two models are very useful in applications.

Our results include previously known results and also solve a number of open problems. The models and our results are stated in the next section.

2. Problem statement. We consider a system evolving in a state space X according to:

$$x_{k+1}=f(x_k, u_k, v_k, w_k). \tag{1}$$

Player I (the maximizer) controlling u wishes to maximize the expected cost

$$E\left[\sum_{k=0}^{\infty} \alpha^k c(x_k, u_k, v_k, x_{k+1})\right]. \tag{2}$$

At each time k, player I is allowed to choose u_k from some set $U_{x_k} \subset U$. Player II (the minimizer) controlling v wishes to minimize (2) by choosing, at each instant k, v_k from $V_{x_k} \subset V$. $\{w_k\}$ is a random disturbance.

* Received by the editors December 26, 1979, and in revised form September 6, 1980.

† Department of Mathematics, University of Maryland Baltimore County, 5401 Wilkens Ave., Baltimore, Maryland 21228. This research was supported by the U.S. Army Research Office.

We assume that X, U, and V are complete, separable, metric spaces. $\alpha \in (0, 1]$ is called the discount factor. By $P(A)$ we shall denote the set of all Borel probability measures on the complete, separable, metric space A. Instead of working with the state equation (1), we prefer to deal with the transition kernel, $q(B|x, u, v) = \text{Probability}(\{w: f(x, u, v, w) \in B\}|x, u, v)$, which we assume to be well defined for every Borel set $B \subset X$.

Since even elementary games need not possess a value, we need to impose additional conditions to make the problem meaningful. We impose two alternative sets of conditions—the Borel model and the continuous model. These two models are very useful from the point of view of applications.

Borel model.

(i) q is a Borel measurable stochastic kernel, i.e., $q(B|\cdot)$ is Borel measurable in the second argument for every fixed Borel subset $B \subset X$.

(ii) U_x and V_x are *finite* for each $x \in X$, and

$$
\begin{aligned}
&\Gamma_1 = \{(x, u): x \in X \text{ and } u \in U_x\} \subset X \times U,\\
&\Gamma_2 = \{(x, v): x \in X \text{ and } v \in V_x\} \subset X \times V,\\
&Q_1 = \{(x, T): x \in X \text{ and } T \in P(U_x)\} \subset X \times P(U),\\
&Q_2 = \{(x, R): x \in X \text{ and } R \in P(V_x)\} \subset X \times P(V)
\end{aligned}
\tag{3}
$$

are all Borel subsets of the corresponding spaces.

Continuous model.

(4)
(i) $U_x \equiv U$, $V_x \equiv V$ for all $x \in X$ where U and V are compact. (Note. This will be generalized in § 3 to allow U_x and V_x to depend on x as in the Borel model, but to preserve clarity of exposition, we prefer to state the simpler version first.)
(ii) c is continuous on $X \times U \times V \times X$.
(iii) $q: X \times U \times V \to P(X)$ is weakly continuous, i.e., continuous with respect to the weak topology on $P(X)$.

In both models we assume

$$0 \leqq c(x, u, v, y) \leqq \theta < \infty$$

(even this will be relaxed in § 3, Remark 1).

Players I and II are allowed to use randomized strategies. We define a randomized strategy for player I to be a sequence $F = \{F^0, F^1, F^2, F^3, \cdots\}$ where each $F^k = F^k(du_k|x_0, u_0, x_1, u_1, x_2, u_2, \cdots, x_k)$ is a Borel measurable stochastic kernel on U_{x_k}. Player I chooses u_k according to this probability distribution F^k which utilizes the past history $(x_0, u_0, x_1, u_1, x_2, u_2, \cdots, x_k)$ known to him in a Borel measurable way. A randomized strategy $F = \{F^0, F^1, F^2, \cdots\}$ is said to be Markovian if each F^k depends only on x_k, i.e., $F^k(du_k|x_k)$. A Markovian strategy is said to be stationary if all the F^k's are identical. The different types of randomized strategies for player II are defined similarly. For player i, the sets of randomized, Markovian and stationary strategies are denoted by D_i, M_i and S_i respectively.

The cases $0 < \alpha < 1$ and $\alpha = 1$ will be called the discounted and positive cases respectively and will be identified by the letters D and P. Similarly, B and C will denote the Borel model and the continuous model. A combination of letters such as BD will refer to the discounted cost case of the Borel model. When no letters are used, a result holds for all models and cases.

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

For every initial state x_0 and randomized strategy pair $(F, G) \in D_1 \times D_2$ adopted by the two players, the cost incurred is

$$J(x_0; F, G) := E_{F,G}\left[\sum_{k=0}^{\infty} \alpha^k c(x_k, u_k, v_k, x_{k+1}) \,|\, x_0\right],$$

where $E_{F,G}$ denotes the expectation under the probability measure induced on the future evolution of the system by (F, G) and the random disturbance. It is allowed to be $+\infty$, as are all functions throughout this paper.

Our main results are the following:

(i) For every $x_0 \in X$,

$$\inf_{G \in D_2} \sup_{F \in D_1} J(x_0; F, G) = \sup_{F \in D_1} \inf_{G \in D_2} J(x_0; F, G) =: J^*(x_0). \tag{5}$$

$J^*(\cdot)$ is called the value function (Theorem 2).

(ii) (6)

(B) $J^*(\cdot)$ is a Borel measurable function.

(CD) $J^*(\cdot)$ is a bounded continuous function.

(CP) $J^*(\cdot)$ is a lower-semicontinuous function (Lemma 3).

(iii) (D) $J^*(\cdot)$ is the *unique* solution of

$$\begin{aligned} J^*(x) &= \min_{R \in P(V_x)} \max_{T \in P(U_x)} \int_{V_x} \int_{U_x} \int_X [\alpha J^*(y) + c(x, u, v, y)] q(dy \,|\, x, u, v) T(du) R(dv) \\ &= \max_{T \in P(U_x)} \min_{R \in P(V_x)} \int_{U_x} \int_{V_x} \int_X [\alpha J^*(y) + c(x, y, v, y)] q(dy \,|\, x, u, v) R(dv) T(du) \end{aligned} \tag{7}$$

for every $x \in X$ (Theorem 1).

(P) $J^*(\cdot)$ satisfies

$$J^*(x) = \min_{R \in P(V_x)} \sup_{T \in P(U_x)} \int_{V_x} \int_{U_x} \int_X [J^*(y) + c(x, u, v, y)] q(dy|x, u, v) T(du) R(dv) \tag{8}$$

for every $x \in X$. Furthermore, if any nonnegative function $J(\cdot)$ satisfies

$$J(x) \geqq \inf_{R \in P(V_x)} \sup_{T \in P(U_x)} \int_{V_x} \int_{U_x} \int_X [J(y) + c(x, u, v, y)] q(dy \,|\, x, u, v) T(du) R(dv) \tag{9}$$

for every $x \in X$, then $J(x) \geqq J^*(x)$ for every $x \in X$ (Theorem 1).

(iv) If $G^* = \{G_0, G_0, G_0, \cdots\} \in S_2$ is such that $G_0(dv \,|\, x)$ attains the outer minimum in (7) or (8) then $J(x_0, F, G^*) \leqq J^*(x_0)$ for all $F \in D_1$ and all $x_0 \in X$. There always exists such a G^*. In words, player II has a stationary strategy G^* which is optimal irrespective of the initial state (Theorem 2).

(v) (D) If $F^* = \{F_0, F_0, F_0, \cdots\} \in S_1$ is such that $F_0(du \,|\, x)$ attains the outer maximum in (7), then $J(x_0; F^*, G) \geqq J^*(x_0)$ for all $G \in D_2$ and all $x_0 \in X$. There always exists such an F^*. In words, player I has a stationary strategy F^* which is optimal irrespective of the initial state (Theorem 3).

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(P) (a) For any probability measure λ on X and every $\varepsilon > 0$, there exists $F_{\lambda,\varepsilon} \in S_1$ such that

$$\lambda\left(\left\{x \in X: \inf_{G\in D_2} J(x; F_{\lambda,\varepsilon}, G) \geqq J^*(x) - \varepsilon \quad \text{if } J^*(x) < \infty, \atop \geqq \frac{1}{\varepsilon} \qquad\qquad \text{if } J^*(x) = \infty\right\}\right) \geqq 1 - \varepsilon.$$

(b) In particular, for every finite $S \subset X$, there exists an $F_{S,\varepsilon} \in S_1$ such that

$$\inf_{G\in D_2} J(x; F_{S,\varepsilon}, G) \geqq \begin{cases} J^*(x) - \varepsilon & \text{if } J^*(x) < \infty \text{ and } x \in S, \\ \dfrac{1}{\varepsilon} & \text{if } J^*(x) = \infty \text{ and } x \in S. \end{cases}$$

(c) If X is compact and $J^*(\cdot)$ is continuous, then for every $\varepsilon > 0$, player I has an ε-optimal stationary strategy; i.e., there exists $F_\varepsilon \in S_1$ such that (Theorem 3)

$$\inf_{G\in D_2} J(x; F_\varepsilon, G) \geqq J^*(x) - \varepsilon \quad \text{for all } x \in X.$$

Some comments about our results are in order.

(i) For the case where X is not necessarily compact (or finite) all the above mentioned results are new. In particular we call attention to our result (6) that $J^*(\cdot)$ is Borel measurable. This result is striking since it is even stronger than the previous *conjectures*. In [5, Open Problem 2, p. 253] it is conjectured that $J^*(\cdot)$ would be universally measurable. Our result goes beyond this conjecture and proves that $J^*(\cdot)$ is Borel measurable.

(ii) In the continuous model with undiscounted cost (CP), even when X is compact, our results are considerably stronger than earlier results [4]. In [4] an additional assumption regarding the equicontinuity of the family of value functions of the corresponding discounted games is made. A statement is also made [4, Remark 3.2] that the authors are unaware if the results hold when such an assumption is not made. Besides being a restrictive assumption which is not a priori verifiable, this assumption also results in a continuous value function. As evidenced by an example in § 7, there do exist simple games where $J^*(\cdot)$ is not continuous. We therefore eliminate this assumption. Our result stated in (6) is that $J^*(\cdot)$ is always lower-semicontinuous.

3. The truncated games. We start with a well-known result which we repeat here for convenience.

LEMMA 1. *Let $K: U \times V \to \mathbb{R}$ be continuous with U and V compact. Then*

$$\min_{R\in P(V)} \max_{T\in P(U)} \int_V \int_U K(u, v) T(du) R(dv) = \max_{T\in P(U)} \min_{R\in P(V)} \int_U \int_V K(u, v) R(dv) T(du).$$

Throughout this paper, by measurable we shall mean Borel measurable.

LEMMA 2.

(i) (B) *Let $J^n: X \to [0, M]$ be measurable. Then*

$$J^{n+1}(x) := \max_{T\in P(U_x)} \min_{v\in V_x} \int_{U_x} \int_X [\alpha J^n(y) + c(x, u, v, y)] q(dy | x, u, v) T(du)$$

$$\text{(10)} \qquad := \min_{R\in P(V_x)} \max_{u\in U_x} \int_{V_x} \int_X [\alpha J^n(y) + c(x, u, v, y)] q(dy | x, u, v) R(dv)$$

is well defined, measurable, nonnegative and bounded by $\alpha M + \theta$.

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(C) *Let* $J^n: X \to [0, M]$ *be continuous. Then* $J^{n+1}(\cdot)$ *defined by* (10) *is well defined, continuous, nonnegative and bounded by* $\alpha M + \theta$.

(ii) *There exist measurable stochastic kernels* $T^n(du|x)$ *and* $R^n(dv|x)$ *which achieve the outer maximum and outer minimum respectively in* (10) *for each* $x \in X$.

Proof. Let $K(x, u, v) := \int_X [\alpha J^n(y) + c(x, u, v, y)] q(dy|x, u, v)$.

(B) $K(\cdot)$ is measurable [6, Prop. 7.29] and $K(x, u, v) \leqq \alpha M + \theta$. For fixed x, U_x and V_x are finite; hence, from von Neumann [7], the two expressions in (10) are equal and therefore $J^{n+1}(\cdot)$ is well defined, nonnegative and bounded by $\alpha M + \theta$. To show it is measurable, let $l(x, T, v) = \int_U K(x, u, v) T(du)$. Then since l can be rewritten as $l(x, T, v) = \int_U \bar{K}(x, T, u, v) \varphi(du|x, T, v)$ where $\bar{K}(x, T, u, v) = K(x, u, v)$ and $\varphi(\cdot|x, T, v) \equiv T$ are both measurable, it follows from [6, Prop. 7.29] that l is measurable. To show that the map $(x, T) \mapsto \min_{v \in V_x} l(x, T, v)$ is measurable, let $D = \{(x, T, v): x \in X, \ T \in P(U), \ v \in V_x\}$. Then $D \subset X \times P(U) \times V$ is a Borel set since by assumption (3), Γ_2 is a Borel set. The (x, T)-section of D, $D_{(x,T)} = \{v \in V: (x, T, v) \in D\}$ is just V_x which is finite. From [8, Cor. 1] it follows that the map $m: (x, T) \mapsto \min_{v \in V_x} l(x, T, v)$ is measurable. For fixed x and $v \in V_x$, $l(x, \cdot, v)$ is linear on $P(U_x)$. Since U_x is finite, the set of all probability measures on U_x, i.e., $P(U_x)$ is a standard simplex in a finite dimensional space. Since a linear mapping defined on a subset of a finite dimensional space is continuous, it follows that $l(x, \cdot, v)$ is continuous on $P(U_x)$. $m(x, \cdot)$: $T \mapsto \min_{v \in V_x} l(x, T, v)$ is the minimum of a finite number of such functions and therefore $m(x, \cdot)$ is continuous on $P(U_x)$. A repeated application of [8, Cor. 1] to $-m$ then shows that $J^{n+1}(x) = \max_{T \in P(U_x)} m(x, T)$ is also measurable and that there exists a $T^n(du|x)$ maximizing (10) which is a Borel measurable stochastic kernel. A similar proof holds for $R_n(dv|x)$.

(C) From Lemma 1, the two expressions in (10) are equal and therefore $J^{n+1}(\cdot)$ is well defined, nonnegative and bounded by $\alpha M + \theta$. Define K and l as in the proof of case (B), immediately above. From [6, Prop. 7.30], l is continuous since it can be rewritten as in the proof of case (B) and $\bar{K}$, φ are continuous. By a repeated application, $J^{n+1}(\cdot)$ is also continuous. The existence of Borel-measurable stochastic kernels $T^n(du|x)$ *and* $R^n(dv|x)$ follows as in the proof of case (B).

We now define $J^0(x) \equiv 0$ for all $x \in X$. $J^0(\cdot)$ satisfies the condition of Lemma 2 and by induction it follows that $J^n(\cdot)$ inductively defined by (10) also satisfies the condition of Lemma 2. Moreover, since $J^1(x) \geqq J^0(x) \equiv 0$ for all $x \in X$, it follows by induction and (10) that $J^{n+1}(x) \geqq J^n(x)$ for all $x \in X$ and all n. Define

$$J^*(x) = \lim_{n \to \infty} J^n(x) \quad \text{for every } x \in X. \tag{11}$$

From Lemma 2, it easily follows that

(12)
- (BP) $J^*(\cdot)$ is a Borel measurable function.
- (CP) $J^*(\cdot)$ is a lower semicontinuous function (since it is the increasing limit of continuous functions).

LEMMA 3.

(13)
- (CD) $J^*(\cdot)$ *is a nonnegative continuous function bounded by* $\theta/(1-\alpha)$.
- (BD) $J^*(\cdot)$ *is a nonnegative Borel measurable function bounded by* $\theta/(1-\alpha)$.

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof.

(CD) Let $K_n(x) := \int_X [\alpha J^n(y) + c(x, u, v, y)] q(dy|x, u, v)$. Since

$$K_n(x, u, v) - K_{n-1}(x, u, v) = \int_X \alpha[J^n(y) - J^{n-1}(y)] q(dy|x, u, v)$$
$$\leqq \alpha \|J^n - J^{n-1}\|_\infty,$$

we obtain

$$\int_{U_x} K_n(x, u, v) T(du) \leqq \int_{U_x} K_{n-1}(x, u, v) T(du) + \alpha \|J^n - J^{n-1}\|_\infty.$$

Hence

$$\max_{T \in P(U_x)} \min_{v \in V_x} \int_{U_x} K_n(x, u, v) T(du)$$
$$\leqq \max_{T \in P(U_x)} \min_{v \in V_x} \int_{U_x} K_{n-1}(x, u, v) T(du) + \alpha \|J^n - J^{n-1}\|_\infty.$$

It follows that $J^{n+1}(x) - J^n(x) \leqq \alpha \|J^n - J^{n-1}\|_\infty$ for every $x \in X$, and hence $\|J^{n+1} - J^n\|_\infty \leqq \alpha \|J^n - J^{n-1}\|_\infty$ for all n. Therefore, $\|J^{n+p} - J^n\|_\infty \leqq (\alpha^{n+p-1} + \cdots + \alpha^n)\|J^1 - J^0\|_\infty = (\alpha^{n+p-1} + \cdots + \alpha^n)\|J^1\|_\infty$. Hence $J^n(\cdot)$ converges uniformly to $J^*(\cdot)$. Let $n = 0$, $p \to +\infty$ we have $\|J^*\|_\infty \leqq (1/(1-\alpha))\|J^1\|_\infty \leqq \theta/(1-\alpha)$, proving the lemma.

(BD) The proof is similar.

Remark 1. The condition (6) defining the Borel and the continuous models can be replaced by the weaker condition $\|J^1\|_\infty \leqq \theta < \infty$. All the results of this paper will continue to hold under this weaker condition.

We now wish to deal with the generalization of the continuous model (5) to allow U_x and V_x to depend on x. To ensure the equality of the two expressions in (10) we need the continuity of $J^n(\cdot)$. Before stating the most general version, we consider the following one-dimensional situation.

LEMMA 4. (C) *Let*

$$U_x = \{u : a_1(x) \leqq u \leqq b_1(x)\} \subset U \subset \mathbb{R},$$
$$V_x = \{v : a_2(x) \leqq v \leqq b_2(x)\} \subset V \subset \mathbb{R},$$

where a_i and b_i are continuous functions on X satisfying $a_i(x) \leqq b_i(x)$ for all $x \in X$ and $i = 1, 2$. Let X be locally compact. Then the results of Lemma 2 hold.

Proof. Given $x_0 \in X$, by the local compactness of X, there exists an open neighborhood $N(x_0)$ of x_0 with $\overline{N(x_0)}$ compact. Let

$$A_i = \min\{a_i(x) : x \in \overline{N(x_0)}\}, \qquad B_i = \max\{b_i(x) : x \in \overline{N(x_0)}\}.$$

Define $U(x_0) =: [A_1, B_1]$ and $V(x_0) = [A_2, B_2]$; then $U_x \subset U(x_0)$ and $V_x \subset V(x_0)$ for all $x \in N(x_0)$. Define $\varphi_1 : X \times U \times V \to \Gamma_1 \times V$ by

$$\varphi_1(x, u, v) = \begin{cases} (x, u, v), & \text{if } a_1(x) \leqq u \leqq b_1(x), \quad \text{i.e., } (x, u) \in \Gamma_1, \\ (x, a_1(x), v) & \text{if } u < a_1(x), \\ (x, b_1(x), v) & \text{if } u > b_1(x). \end{cases}$$

Also define $\varphi_2: \Gamma_1 \times V \to D := \{(x, u, v): u \in U_x, v \in V_x, x \in X\}$ by

$$\varphi_2(x, u, v) = \begin{cases} (x, u, v), & \text{if } a_2(x) \leqq v \leqq b_2(x), \\ (x, u, a_2(x)) & \text{if } v < a_2(x), \\ (x, u, b_2(x)) & \text{if } v > b_2(x). \end{cases}$$

Let $\varphi = \varphi_2 \circ \varphi_1$. Then since φ_1 and φ_2 are continuous, it follows that φ also is continuous. Furthermore, $\varphi(x, u, v) = (x, u'(x, u), v)$ for $v \in V_x$ and $\varphi(x, u, v) = (x, u, v'(x, v))$ for $u \in U_x$. Define $\hat{K}(x, u, v) = K(\varphi(x, u, v))$ where K is defined as in the proof of Lemma 2. For fixed $x \in \overline{N(x_0)}$, let $T^* \in P(U_x)$, $R^* \in P(V_x)$ be the saddle point of the static "game on the unit square" $K(x, u, v)$ on $U_x \times V_x$. For $v \in V_x$ and $u \in U(x_0)$, $\hat{K}(x, u, v) = K(\varphi(x, u, v)) = K(x, u', v)$ where $u' = u'(x, u) \in U_x$. Hence

$$\int_{V(x_0)} \hat{K}(x, u, v) R^*(dv) = \int_{V_x} \hat{K}(x, u, v) R^*(dv) = \int_{V_x} K(x, u', v) R^*(dv)$$

and, therefore,

$$\max_{u \in U(x_0)} \int_{V(x_0)} \hat{K}(x, u, v) R^*(dv) = \max_{u' \in U_x} \int_{V_x} K(x, u', v) R^*(dv) = J^{n+1}(x).$$

Hence

$$\min_{R \in P(V(x_0))} \max_{u \in U(x_0)} \int_{V(x_0)} \hat{K}(x, u, v) R(dv) \leqq J^{n+1}(x).$$

Similarly, we can prove

$$\max_{T \in P(U(x_0))} \min_{v \in V(x_0)} \int_{U(x_0)} \hat{K}(x, u, v) T(du) \geqq J^{n+1}(x).$$

Therefore

$$\min_{R \in P(V(x_0))} \max_{u \in U(x_0)} \int_{V(x_0)} \hat{K}(x, u, v) R(dv) = \min_{R \in P(V_x)} \max_{u \in U_x} \int_{V_x} K(x, u, v) R(dv) =: J^{n+1}(x)$$

for every $x \in \overline{N(x_0)}$. However, $\hat{K}$ does not feature state-constrained control sets, and therefore the proof of Lemma 2 shows that $J^{n+1}(x)$ is continuous at x_0. Since x_0 was arbitrary this proves the lemma.

Remark 2. Condition (i) of the continuous model can be generalized to:

(i′) For each $x \in X$, U_x and V_x are compact subsets of U and V such that:

(a) There exists a $\varphi: X \times U \times V \to D = \{(x, u, v): x \in X, u \in U_x, v \in V_x\}$ such that φ is continuous, $\varphi(x, u, v) = (x, u'(x, u), v)$ for $v \in V_x$ and $\varphi(x, u, v) = (x, u, v'(x, v))$ for $u \in U_x$.

(b) For every $x_0 \in X$ there exists an open neighborhood of x_0 such that $\cup\{X_x : x \in \overline{N(x_0)}\}$ and $\cup\{V_x : x \in \overline{N(x_0)}\}$ are compact.

The proof is the same as Lemma 4.

4. The value function. We now show that $J^*(\cdot)$ is a lower bound for the lower value of the game.

LEMMA 5. *For every* $x_0 \in X$, $\sup_{F \in M_1} \inf_{G \in D_2} J(x_0; F, G) \geqq J^*(x_0)$.

Proof. Let $F_{n+1} \in M_1$ be defined by

$$F_{n+1} = \{T^n, T^{n-1}, \cdots, T^2, T^1, T^0, T^0, T^0, \cdots\}$$

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where $T^n(du|x)$ is defined as in Lemma 2. Then

$$
\begin{aligned}
&E_{F_{n+1},G}\left[\sum_{k=0}^{n} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right]\\
&\quad = E_{F_{n+1},G}\left[\sum_{k=0}^{n-1} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right] + E_{F_{n+1},G}[\alpha^n c(x_n, u_n, v_n, x_{n+1})|x_0]\\
&\quad = E_{F_{n+1},G}\left[\sum_{k=0}^{n-1} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right]\\
&\qquad + \alpha^n E_{F_{n+1},G}\Big\{E_{F_{n+1},G}\Big[\Big[\int_{V_{x_n}}\int_{U_{x_n}}\int_X c(x_n, u_n, v_n, y) q(dy|x_n, u_n, v_n) T^0(du_n|x_n)\\
&\qquad\qquad \cdot G^n(dv_n|x_0, v_0, x_1, v_1, \cdots, x_n)|x_0, v_0, x_1, v_1, \cdots, x_n\Big]\Big|x_0\Big\}\\
&\quad \geqq E_{F_{n+1},G}\left[\sum_{k=0}^{n-1} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right] + \alpha^n E_{F_{n+1},G}[J^1(x_n)|x_0]\\
&\quad = E_{F_{n+1},G}\left[\sum_{k=0}^{n-2} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right]\\
&\qquad + \alpha^{n-1} E_{F_{n+1},G}[\alpha J^1(x_n) + c(x_{n-1}, u_{n-1}, v_{n-1}, x_n)|x_0]\\
&\quad = E_{F_{n+1},G}\left[\sum_{k=0}^{n-2} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right]\\
&\qquad + \alpha^{n-1} E_{F_{n+1},G}\Big\{E_{F_{n+1},G}\Big[\Big[\int_{V_{x_{n-1}}}\int_{U_{x_{n-1}}}\int_X [\alpha J^1(y) + c(x_{n-1}, u_{n-1}, v_{n-1}, y)]\\
&\qquad\qquad \cdot q(dy|x_{n-1}, u_{n-1}, v_{n-1}) T^1(du_{n-1}|x_{n-1})\\
&\qquad\qquad \cdot G^{n-1}(dv_{n-1}|x_0, v_0, x_1, v_1, \cdots, x_{n-1})|\\
&\qquad\qquad \cdot x_0, v_0, \cdots, x_{n-1}\Big]\Big|x_0\Big\}\\
&\quad \geqq E_{F_{n+1},G}\left[\sum_{k=0}^{n-2} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right] + \alpha^{n-1} E_{F_{n+1},G}[J^2(x_{n-1})|x_0]\\
&\qquad\qquad \vdots\\
&\quad \geqq \alpha^0 E_{F_{n+1},G}[J^{n+1}(x_0)|x_0]\\
&\quad = J^{n+1}(x_0).
\end{aligned}
$$

Hence $J(x_0; F_{n+1}, G) \geqq J^{n+1}(x_0)$ for any $G \in D_2$. Hence $\inf_{G\in D_2} J(x_0; F_{n+1}, G) \geqq J^{n+1}(x_0)$ and therefore, $\sup_{F\in M_1} \inf_{G\in D_2} J(x_0; F, G) \geqq J^{n+1}(x_0)$ for arbitrary n, proving that $\sup_{F\in M_1} \inf_{G\in D_2} J(x_0; F, G) \geqq \sup_n J^{n+1}(x_0) = J^*(x_0)$.

We now provide the following characterization of $J^*(\cdot)$. This characterization is exceedingly useful in many applications.

THEOREM 1.

(D) *$J^*(\cdot)$ is the* unique *solution of* (7).

(P) *$J^*(\cdot)$ satisfies* (8). *If any nonnegative $J(\cdot)$ satisfies* (9) *then $J^*(x) \leqq J(x)$ for every $x \in X$.*

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. We first show that $J^*(\cdot)$ satisfies (8). Since $J^n(y) \leqq J^*(y)$ for every $y \in X$, from (10) we obtain

$$J^{n+1}(x) \leqq \inf_{R \in P(V_x)} \sup_{u \in U_x} \int_{V_x} \int_X [\alpha J^*(y) + c(x, u, v, y)] q(dy|x, u, v) R(dv).$$

Since this is true for every n, we obtain

$$J^*(x) \leqq \inf_{R \in P(V_x)} \sup_{u \in U_x} \int_{V_x} \int_X [\alpha J^*(y) + c(x, u, v, y)] q(dy|x, u, v) R(dv).$$

To prove the reverse inequality consider $R^n(\cdot|x) \in P(V_x)$, which achieves the outer minimum in (10). Then

$$J^*(x) \geqq J^{n+1}(x) = \max_{u \in U_x} \int_{V_x} \int_X [\alpha J^n(y) + c(x, u, v, y)] q(dy|x, u, v) R^n(dv|x). \tag{14}$$

Since V_x is compact, $P(V_x)$ is compact [12, II.6.4] for fixed x, and hence the sequence $\{R^n(\cdot|x)\} \subset P(V_x)$ has a subsequence $\{R^{n_k}(\cdot|x)\}$ which converges to $\tilde{R} \in P(V_x)$. Fix x, relabel the sequence $\{R^{n_k}(\cdot|x)\}$ as $\{R^k\}$ and define

$$L^k(x, u, v) := \int_X [\alpha J^{n_k}(y) + c(x, u, v, y)] q(dy|x, u, v).$$

By (14)

$$J^*(x) \geqq \max_{u \in U_x} \int_{V_x} L^k(x, u, v) R^k(dv) \quad \text{for all } k.$$

Hence

$$J^*(x) \geqq \int_{V_x} L^k(x, u, v) R^k(dv) \quad \text{for all } k \text{ and } u \in U_x,$$

and

$$J^*(x) \geqq \sup_k \int_{V_x} L^k(x, u, v) R^k(dv) \quad \text{for all } u \in U_x. \tag{15}$$

Clearly, $0 \leqq L^1(x, u, v) \leqq L^2(x, u, v) \leqq \cdots \leqq +\infty$, and

$$\lim_{k \to \infty} L^k(x, u, v) = \lim_{k \to \infty} \int_X [\alpha J^{n_k}(y) + c(x, u, v, y)] q(dy|x, u, v)$$
$$= \int_X [\alpha J^*(y) + c(x, u, v, y)] q(dy|x, u, v), \tag{16}$$

by the monotone convergence theorem.

Denote

$$L(x, u, v) := \int_X [\alpha J^*(y) + c(x, u, v, y)] q(dy|x, u, v) \leqq +\infty.$$

We now proceed to show that

$$\sup_k \int_{V_x} L^k(x, u, v) R^k(dv) \geqq \int_{V_x} L(x, u, v) \tilde{R}(dv). \tag{17}$$

By the monotone convergence theorem again, (16) implies

$$\int_{V_x} L(x, u, v)\tilde{R}(dv) = \lim_{k\to\infty} \int_{V_x} L^k(x, u, v)\tilde{R}(dv|x).$$

Hence given $\varepsilon > 0$, there exists N so large that

$$(18)\quad \int_{V_x} L^N(x, u, v)\tilde{R}(dv) \geqq \begin{cases} \int_{V_x} L(x, u, v)\tilde{R}(dv) - \varepsilon & \text{if } \int_{V_x} L(x, u, v)\tilde{R}(dv) < \infty, \\ \dfrac{1}{\varepsilon} & \text{if } \int_{V_x} L(x, u, v)\tilde{R}(dv) = \infty. \end{cases}$$

Now fix N and $u \in U_x$, $L^N(x, u, \cdot)$ is bounded and continuous on V_x and $R^k \to \tilde{R}$ as $k \to \infty$. From [6, Prop. 7.21] we obtain

$$\int_{V_x} L^N(x, u, v)R^k(dv|x) \to \int_{V_x} L^N(x, u, v)\tilde{R}(dv|x) \quad \text{as } k \to \infty.$$

Hence there exists a K such that for all $k \geqq K$

$$(19)\qquad \int_{V_x} L^N(x, u, v)R^k(dv) \geqq \int_{V_x} L^N(x, u, v)\tilde{R}(dv) - \varepsilon.$$

We now consider two cases.

Case 1. Suppose $N \geqq K$. Replacing k by N in (19), we obtain

$$\int_{V_x} L^N(x, u, v)R^N(dv) \geqq \int_{V_x} L^N(x, u, v)\tilde{R}(dv) - \varepsilon.$$

This together with (18) implies that

$$\int_{V_x} L^N(x, u, v)R^N(dv) \geqq \begin{cases} \int_{V_x} L(x, u, v)\tilde{R}(dv) - 2\varepsilon & \text{if } \int_{V_x} L(x, u, v)\tilde{R}(dv) < \infty, \\ \dfrac{1}{\varepsilon} - \varepsilon & \text{otherwise.} \end{cases}$$

Hence

$$(20)\quad \sup_k \int_{V_x} L^k(x, u, v)R^k(dv) \geqq \begin{cases} \int_{V_x} L(x, u, v)\tilde{R}(dv) - 2\varepsilon & \text{if } \int_{V_x} L(x, u, v)\tilde{R}(dv) < \infty, \\ \dfrac{1}{\varepsilon} - \varepsilon & \text{otherwise.} \end{cases}$$

Since ε was arbitrary, (17) is proved.

Case 2. Suppose $N < K$. Then $L^N(x, u, v) \leqq L^K(x, u, v)$,

$$\int_{V_x} L^K(x, u, v)R^K(dv) \geqq \int_{V_x} L^N(x, u, v)R^K(dv),$$

and the latter in turn is greater than or equal to $\int_{V_x} L^N(x, u, v)\tilde{R}(dv) - \varepsilon$ by (19). We thus obtain

$$\sup_k \int_{V_x} L^k(x, u, v)R^k(dv) \geqq \int_{V_x} L^N(x, u, v)\tilde{R}(dv) - \varepsilon.$$

Again, this together with (18) implies (20) and then (17) follows.

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Now (15) and (17) together imply

$$J^*(x) \geqq \int_{V_x} L(x, u, v)\tilde{R}(dv) \quad \text{for every } u \in U_x.$$

Hence

$$J^*(x) \geqq \sup_{u \in U_x} \int_{V_x} L(x, u, v)\tilde{R}(dv),$$

i.e.,

$$J^*(x) \geqq \sup_{u \in U_x} \int_{V_x} \int_X [\alpha J^*(y) + c(x, u, v, y)] q(dy|x, u, v)\tilde{R}(dv).$$

Hence we obtain the reverse inequality and $\tilde{R}$ achieves the minimum. (Note that $\tilde{R}$ depends on x).

(P) Hence $J^*(\cdot)$ satisfies (8). Suppose $J(\cdot)$ is nonnegative and satisfies (9). Then $J(x) \geqq J^0(x) \equiv 0$ for all $x \in X$. Hence, by induction, using (9) and (10), $J(x) \geqq J^n(x)$ for all $x \in X$ and all n, therefore $J(x) \geqq J^*(x)$.

(D) As earlier, $J^*(\cdot)$ satisfies (8). From (13), $J^*(\cdot)$ is bounded by $\theta/(1-\alpha)$. In case B, U_x and V_x are finite for each x and by von Neumann's Fundamental Theorem of Matrix Games [7], (8) and (7) are equivalent. In Case C, by Lemma 3, $J^*(\cdot)$ is continuous and since U_x and V_x are compact for each x, it follows from Lemma 1 that (8) and (7) are equivalent. Hence, in summary, $J^*(\cdot)$ satisfies (7) in case D. To show it is the unique solution of (7), we proceed as follows:

Let $J(\cdot)$ be a nonnegative solution of (7). Given x, let

$$K(T, R) = \int_{V_x} \int_{U_x} \int_X [\alpha J(y) + c(x, u, v, y)] q(dy|x, u, v) T(du) R(dv),$$

$$K^*(T, R) = \int_{V_x} \int_{U_x} \int_X [\alpha J^*(y) + c(x, u, v, y)] q(dy|x, u, v) T(du) R(dv).$$

Clearly,

$$\begin{aligned} K(T, R) - K^*(T, R) &= \int_{V_x} \int_{U_x} \int_X \alpha (J(y) - J^*(y)) q(dy|x, u, v) T(du) R(dv) \\ &\leqq \alpha \|J - J^*\| \quad \text{for all } T, R, \text{ where } \|\cdot\| \text{ is the sup norm.} \end{aligned}$$

Hence $K(T, R) \leqq K^*(T, R) + \alpha\|J - J^*\|$, and therefore

$$\max_{T \in P(U_x)} K(T, R) \leqq \max_{T \in P(U_x)} K^*(T, R) + \alpha \|J - J^*\|,$$

and therefore

$$\min_{R \in P(V_x)} \max_{T \in P(U_x)} K(T, R) \leqq \min_{R \in P(V_x)} \max_{T \in P(U_x)} K^*(T, R) + \alpha\|J - J^*\|;$$

i.e., $J(x) \leqq J^*(x) + \alpha\|J - J^*\|$. We obtain $J(x) - J^*(x) \leqq \alpha\|J - J^*\|$. Similarly, we can show $J^*(x) - J(x) \leqq \alpha\|J - J^*\|$, therefore $|J(x) - J^*(x)| \leqq \alpha\|J - J^*\|$. Since x was arbitrary, we have $\|J - J^*\| \leqq \alpha\|J - J^*\|$. Because $\alpha < 1$, we must have $\|J - J^*\| = 0$, i.e., $J(x) \equiv J^*(x)$, proving that $J^*(x)$ is unique.

5. Optimal stationary strategy for minimizer. In this section we show that $J^*(\cdot)$ is in fact the value of the game. Additionally we show that player II has an optimal stationary strategy and show how such a strategy can be obtained from knowledge of the value function $J^*(\cdot)$.

THEOREM 2.

(i) *$J^*(\cdot)$ is the value function of the game, i.e.,*

$$\inf_{G\in D_2}\sup_{F\in D_1} J(x_0;F,G)=\sup_{F\in D_1}\inf_{G\in D_2} J(x_0;F,G)=J^*(x_0) \quad \textit{for every } x_0\in X.$$

(ii) *Any $G^*=\{G^0,G^0,G^0,\cdots\}\in S_2$ such that $G^0(dv|x)$ achieves the outer minimum in* (8) *for every x is an optimal stationary strategy for player* II, *i.e.,*

$$J(x_0;F,G^*)\leqq J^*(x_0) \quad \textit{for every } F\in D_1 \textit{ and } x_0\in X.$$

There always exists such a G^*.

(iii) *Denoting by $J^*_\alpha(x)$ the value function of the game with discount factor α, we have* $\lim_{\alpha\uparrow 1} J^*_\alpha(x)=J^*_1(x)$.

Proof. (i and ii). We begin by showing the existence of a $G^*=\{G^0,G^0,\cdots\}\in S_2$ such that $G^0(dv|x)$ achieves the outer minimum in (8) for each $x\in X$.

(B) Since $J^*(\cdot)$ is measurable, the proof of existence of a *measurable* $R^n(dv|x)$ achieving the outer minimum in (10) given in Lemma 2 is applicable.

(C) $J^*(\cdot)$ is lower-semicontinuous (l.s.c.), and hence the map

$$k(x,u,v):=\int_X [\alpha J^*(y)+c(x,u,v,y)]q(dy|x,u,v)$$

is l.s.c. [6, Prop. 7.31]. The map

$$l(x,u,R):=\int_V k(x,u,v)R(dv)$$

is therefore l.s.c. (see proof of Lemma 2). The map $m(x,R):=\sup_{u\in U} l(x,u,R)$ is therefore l.s.c. since $\{(x,r)|m(x,r)>\gamma\}=U_{u\in U}\{(x,R)|l(x,u,R)>\gamma\}$ is open for all $\gamma\in\mathbb{R}$. Now [8, Corollary 1] implies the existence of a Borel measurable stochastic kernel, say G^0, which achieves the minimum in $\min_{R\in P(V_x)}(m(x,R))$.

To show $G^*=\{G^0,G^0,\cdots\}$ is optimal, let $F\in D$ be arbitrary. Then

$$\begin{aligned}
&E_{F,G^*}\left[\sum_{k=0}^{n}\alpha^k c(x_k,u_k,v_k,x_{k+1})|x_0\right]+\alpha^{n+1}E_{F,G^*}[J^*(x_{n+1})|x_0]\\
&\quad=E_{F,G^*}\left[\sum_{k=0}^{n-1}\alpha^k c(x_k,u_k,v_k,x_{k+1})|x_0\right]+\alpha^n E_{F,G^*}[\alpha J^*(x_{n+1})+c(x_n,u_n,v_n,x_{n+1})|x_0]\\
&\quad=E_{F,G^*}\left[\sum_{k=0}^{n-1}\alpha^k c(x_k,u_k,v_k,x_{k+1})|x_0\right]\\
&\qquad+\alpha^n E_{F,G^*}\Big\{E_{F,G^*}\Big[\int_{U_{x_n}}\int_{V_{x_n}}\int_X[\alpha J^*(y)+c(x_n,u_n,v_n,y)]\\
&\qquad\qquad\cdot q(dy|x_n,u_n,v_n)G^0(dv_n|x_n)\\
&\qquad\qquad\cdot F^n(du_n|x_0,u_0,x_1,u_1,\cdots,x_n)|x_0,u_0,x_1,u_1,\cdots,x_n\Big]\Big|x_0\Big\}\\
&\quad\leqq E_{F,G^*}\left[\sum_{k=0}^{n-1}\alpha^k c(x_k,u_k,v_k,x_{k+1})|x_0\right]
\end{aligned}$$

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$$(21)\qquad +\alpha^n E_{F,G^*}\left\{E_{F,G^*}\left[\max_{u_n\in U_{x_n}}\int_{V_{x_n}}\int_X [\alpha J^*(y)+c(x_n,u_n,v_n,y)]\right.\right.$$

$$\left.\left.\cdot\, q(dy|x_n,u_n,v_n)G^0(dv_n|x_n)|x_0,u_0,x_1,u_1,\cdots,x_n\right]\Big|x_0\right\}$$

$$=E_{F,G^*}\left[\sum_{k=0}^{n-1}\alpha^k c(x_k,u_k,v_k,x_{k+1})|x_0\right]+\alpha^n E_{F,G^*}[J^*(x_n)|x_0]$$

$$\leqq E_{F,G^*}\left[\sum_{k=0}^{n-2}\alpha^k c(x_k,u_k,v_k,x_{k+1})|x_0\right]$$

$$+\alpha^{n-1}E_{F,G^*}[\alpha J^*(x_n)+c(x_{n-1},u_{n-1},v_{n-1},x_n)|x_0]$$

$$\leqq E_{F,G^*}\left[\sum_{k=0}^{n-3}\alpha^k c(x_k,u_k,v_k,x_{k+1})|x_0\right]$$

$$+\alpha^{n-2}E_{F,G^*}[\alpha J^*(x_{n-1})+c(x_{n-2},u_{n-2},v_{n-2},x_{n-1})|x_0]$$

$$\vdots$$

$$\leqq \alpha^0 E_{F,G^*}[J^*(x_0)|x_0]=J^*(x_0).$$

Hence $E_{F,G^*}[\sum_{k=0}^{n}\alpha^k c(x_k,u_k,v_k,x_{k+1})|x_0]\leqq J^*(x_0)$. Since this holds true for every n, by the monotone convergence theorem as $n\to\infty$ we obtain $J(x_0;F,G^*)\leqq J^*(x_0)$. Since $F\in D_1$ was arbitrary, we obtain $\sup_{F\in D_1}J(x_0;F,G^*)\leqq J^*(x_0)$. This combined with Lemma 5 shows that $J^*(x_0)$ is the value and also that G^* is optimal for every $x_0\in X$, and completes the proof of (i) and (ii).

(iii) Let $\mathrm{Val}\,(x)=\lim_{\alpha\uparrow 1}J^*(x)$; then $\mathrm{Val}\,(x)\leqq J_1^*(x)$. To show the reverse inequality, we first show

$$(8')\qquad \mathrm{Val}\,(x)=\min_{R\in P(V_x)}\sup_{u\in U_x}\int_{V_x}\int_X[\mathrm{Val}\,(y)+c(x,u,v,y)]q(dy|x,u,v)R(dv).$$

This inequality is similar to (8) and the proof is the same; hence we use the notation (8′), (9′), (10′), $\cdots$, etc.

From (8),

$$(22)\qquad J_\alpha^*(x)=\min_{R\in P(V_x)}\sup_{u\in U_x}\int_{V_x}\int_X[\alpha J_\alpha^*(y)+c(x,u,v,y)]q(dy|x,u,v)R(dv),$$

and since $\alpha J_\alpha^*(y)\leqq \mathrm{Val}\,(y)$ for all $\alpha\in(0,1)$, we obtain that the RHS of (22) is less than or equal to the RHS of (8′). Hence

$$J_\alpha^*(x)\leqq\min_{R\in P(V_x)}\sup_{u\in U_x}\int_{V_x}\int_X[\mathrm{Val}\,(y)+c(x,u,v,y)]q(dy|x,u,v)R(dv).$$

Since this is true for all $\alpha\in(0,1)$, as $\alpha\uparrow 1$ we obtain

$$\mathrm{Val}\,(x)\leqq\min_{R\in P(V_x)}\sup_{u\in U_x}\int_{V_x}\int_X[\mathrm{Val}\,(y)+c(x,u,v,y)]q(dy|x,u,v)R(dv).$$

To prove the reverse inequality, let $\{\alpha_n\}$ be a monotonically increasing sequence with $\alpha_n<1$ and $\lim_{n\to\infty}\alpha_n=1$. Fix x, let $\tilde R^n\in P(V_x)$ be such that $\tilde R^n$ achieves the outer minimum in (22) when $\alpha=\alpha_n$. Then

$$(14')\quad \mathrm{Val}\,(x)\geqq J_{\alpha_n}^*(x)=\max_{u\in U_x}\int_{V_x}\int_X[\alpha_n J_{\alpha_n}^*(y)+c(x,u,v,y)]q(dy|x,u,v)\tilde R^n(dv).$$

Now the same argument as in the proof of Theorem 1 shows that (15′)–(20′) are true. Hence (8′) is also true. From Theorem 1 we know that $J_1^*(\cdot)$ is the minimum (at each x) among all the nonnegative functions which satisfy (8). Hence $\mathrm{Val}\,(x) \geqq J_1^*(x)$ for each $x \in X$, proving (iii).

6. Optimal and near optimal strategies for the maximizer. Player I does not always have optimal strategies, stationary or nonstationary, in the positive case. We give an example in § 7. In the positive case we therefore prove the existence of near optimal strategies, appropriately defined in Theorem 3 below. In the discounted case, however, player I always has an optimal stationary strategy.

THEOREM 3.

(D) *Let* $F^* = \{F^0, F^0, F^0, \cdots\} \in S_1$ *be such that* $F^0(du|x)$ *achieves the outer maximum in* (7). *Then* F^* *is an optimal stationary strategy, i.e.,*

$$J(x_0; F^*, G) \geqq J^*(x_0) \quad \textit{for every } G \in D_2 \textit{ and } x_0 \in X.$$

There always exists such an F^*.

(P)(i) *For any* $\lambda \in P(X)$ *and any* $\varepsilon > 0$, *there exists a stationary strategy* $F \in S_1$ *such that*

$$\lambda\left(\left\{x: \inf_{G \in D_2} J(x; F, G) \begin{array}{ll} \geqq J^*(x) - \varepsilon & \text{if } J^*(x) < \infty \\ \geqq \dfrac{1}{\varepsilon} & \text{if } J^*(x) = \infty \end{array}\right\}\right) \geqq 1 - \varepsilon.$$

In particular, for every finite subset $S \subset X$ *and* $\varepsilon > 0$, *there exists an* $F \in S_1$ *such that*

$$J(x; F, G) \geqq \begin{cases} J^*(x) - \varepsilon & \text{if } J^*(x) < \infty \text{ and } x \in S, \\ \dfrac{1}{\varepsilon} & \text{if } J^*(x) = \infty \text{ and } x \in S \end{cases}$$

for every $G \in D_2$.

(ii) *If* X *is compact and* $J^*(\cdot)$ *is continuous, then for every* $\varepsilon > 0$, *player* I *has an* ε*-optimal stationary stationary strategy*; *i.e., there exists an* $F \in S_1$ *such that* $J(x; F, G) \geqq J^*(x) - \varepsilon$ *for every* $x \in X$.

Remark 3. To see that in (P)(i) the λ measure of the set is well defined, we show that $\inf_{G \in D_2} J(\cdot\,; F, G)$ is universally measurable. This is true by Strauch [9], since for fixed F the problem reduces to a dynamic programming problem with maximization of a "negative" cost as the objective.

Proof. (D) From Lemma 3, $J^*(\cdot)$ is bounded; hence, as in the proof of Lemma 2, there exists an F^* satisfying the conditions of the theorem. In Theorem 2, we have shown that G^* is optimal both for $\alpha < 1$ and $\alpha = 1$ by (21). Replacing (F, G^*) by (F^*, G), $\max_{u \in U_x}$ by $\min_{v \in V_x}$ and reversing all the inequalities in (21), we obtain

$$E_{F^*,G}\left[\sum_{k=0}^{n} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right] + \alpha^{n+1} E_{F^*,G}[J^*(x_{n+1})|x_0]$$
$$\geqq J^*(x_0) \quad \text{for all } x_0 \in X, \quad \text{all } G \in D_2.$$

In our case since $\alpha < 1$, $\|J^*\|_\infty < \infty$, letting $n \to \infty$, we obtain

$$E_{F^*,G}\left[\sum_{k=0}^{\infty} \alpha^k c(x_k, u_k, v_k, x_{k+1})|x_0\right] \geqq J^*(x_0) \quad \text{for all } x_0 \in X, \quad \text{all } G \in D_2.$$

Hence F^* is optimal, proving (D).

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(P)(i) From Theorem 2, we have $J_\alpha^*(x)\uparrow J_1^*(x)$ as $\alpha\uparrow 1$. Let $\alpha_n := 1-(1/n)$, $I_n(x) := J_{\alpha_n}^*(x)$; then as $n\to\infty$, $\alpha_n\uparrow 1$ and $I_n(x)\uparrow J_1^*(x)$. Given $\varepsilon>0$, let

$$E_n=\{x\in X: I_n(x)<J_1^*(x)-\varepsilon, \quad \text{if } J_1^*(x)<\infty, \quad \text{or } I_n(x)<\frac{1}{\varepsilon} \quad \text{if } J_1^*(x)=\infty\}.$$

Then $E_1\supset E_2\supset E_3\supset\cdots$ and also $\bigcap_{n=1}^\infty E_n=\varnothing$. Hence for any $\lambda\in P(X)$, $\lim_{n\to\infty}\lambda(E_n)=0$. Choose N so large that $\lambda(E_N)<\varepsilon$. We have for $x\notin E_N$,

$$I_N(x)\geqq\begin{cases} J^*(x)-\varepsilon & \text{if } J^*(x)<\infty,\\ \dfrac{1}{\varepsilon} & \text{if } J^*(x)=\infty.\end{cases}$$

Let F_N be the optimal stationary strategy in the game with discount factor α_N given above in the proof of (D). Then

$$\min_G J(x;F_N,G)\geqq I_N(x)\geqq\begin{cases} J^*(x)-\varepsilon & \text{if } J^*(x)<\infty,\\ \dfrac{1}{\varepsilon} & \text{if } J^*(x)=\infty.\end{cases}$$

This proves the first assertion. For the second, if $S\subset X$ is a finite set with m elements, let λ be the uniform distribution on S. Now let $\bar\varepsilon<\min(\varepsilon,1/m)$. Then the stationary strategy F such that

$$\lambda\left(\left\{\begin{matrix} x: \inf_G J(x;F,G)\geqq J_{\text{opt}}(x)-\bar\varepsilon & \text{if } J^*(x)<\infty\\ \geqq\dfrac{1}{\bar\varepsilon} & \text{if } J^*(x)=\infty\end{matrix}\right\}\right)\geqq 1-\bar\varepsilon$$

is ε-optimal for $x\in S$.

(ii) Let J_α^*, J_α, F_α^*, G_α^* be J^*, J, F^*, G^* respectively in the game with discount factor $\alpha\leqq 1$. Since X is compact and $J_1^*(x)$ is a continuous real function, Dini's theorem implies $J_1^*(x)-J_\alpha^*(x)\downarrow 0$ uniformly on X as $\alpha\uparrow 1$, i.e., $J_\alpha^*(x)\uparrow J_1^*(x)$ uniformly on X as $\alpha\uparrow 1$. Given $\varepsilon>0$, choose α close to 1 so that $J_\alpha^*(x)\geqq J_1^*(x)-\varepsilon$ for all $x\in X$. Then F_α^* is ε-optimal in the Positive ($\alpha=1$) Case, since $J(x;F_\alpha^*,G)\geqq J_\alpha(x;F_\alpha^*,G)\geqq J_\alpha^*(x)\geqq J_1^*(x)-\varepsilon$ for all $x\in X$ and for all $G\in D_2$.

Remark 4. In the proof of Theorem 3, it is clear that if $J_\alpha^*(\cdot)\uparrow J_1^*(\cdot)$ uniformly as $\alpha\uparrow 1$, then player I has an ε-optimal stationary strategy, even if X is not compact. On the other hand, an example in § 7 shows that even if the convergence of $J_\alpha^*(\cdot)$ to $J_1^*(\cdot)$ is not uniform, player I can have an ε-optimal stationary strategy for every $\varepsilon>0$. Whether player I always has an ε-optimal stationary strategy is an open problem. Some related papers are Ornstein [10] and Bertsekas and Shreve [11].

Remark 5. (C) If $J_1^*(\cdot)$ is a continuous real function, then for any compact subset $Y\subset X$, $J_\alpha^*\uparrow J_1^*$ uniformly on Y (Dini's theorem). Hence for any $\varepsilon>0$, player I has a stationary strategy which is ε-optimal on Y. The proof is similar.

7. Some examples. Our first example, adapted from [2], shows that even in particularly simple problems, player I need not have an optimal strategy—stationary or nonstationary.

Example 1. Let

$$
\begin{aligned}
&X=\{x_0, x_1, x_2\}, && \\
&U_x=\{1,2\} && \text{for all } x\in X,\\
&V_x=\{1,2\} && \text{for all } x\in X,\\
&f(x_0,u,v)=x_0 && \text{for all } (u,v),\\
&f(x_1,u,v)=x_1 && \text{for all } (u,v),\\
&f(x_2,u,v)=x_0 && \text{if } u=v,\\
&f(x_2,1,2)=x_2, && \\
&f(x_2,2,1)=x_1, && \\
&c(x,u,v,y)=\begin{cases}1 \\ 0\end{cases} && \begin{matrix}\text{if } x\neq x_0,\ y=x_0,\\ \text{otherwise.}\end{matrix}
\end{aligned}
$$

x_0 and x_1 are absorption points of the system. Hence the only nonzero cost-transition is the transition from x_2 to x_0 which yields a cost $=1$. Clearly $J^*(x_0)=0$, $J^*(x_1)=0$. Hence we need consider only x_2. By Theorem 1, $J^*(x_2)$ is the smallest nonnegative number satisfying

$$J^*(x_2)=\text{value}\begin{bmatrix}1 & J^*(x_2)\\ 0 & 1\end{bmatrix}.$$

Since

$$\text{value}\begin{bmatrix}1 & J^*(x_2)\\ 0 & 1\end{bmatrix}=\frac{1}{2-J^*(x_2)},$$

the equation $J^*(x_2)=1/(2-J^*(x_2))$ has a unique solution $J^*(x_2)=1$, which must be the value at x_2. However there is no optimal strategy for player I. To show this consider two cases.

Case 1. Player I always plays $u=1$. Then if player II always plays $v=2$, the system stays indefinitely in x_2 and the total cost is only 0.

Case 2. At some time n, player I plays $u=1$ with probability $1-\varepsilon$, and $u=2$ with probability $\varepsilon>0$. But then if player II at time n chooses $v=1$, the system ends in state x_1 with probability at least $\varepsilon>0$, hence the total cost is less than or equal to $1-\varepsilon$.

Hence player I has no optimal strategy. Note however that if player I chooses $u=1$ with probability $1-\varepsilon$ and $u=2$ with probability ε, then this stationary strategy is ε-optimal.

Our second example shows that in the continuous model, $J^*(\cdot)$ need not be continuous even when X is compact. This example therefore cannot be solved by the results of [4], but can be by our results.

Example 2. Let

$$X=\left\{\frac{1}{n}\,\middle|\, n \text{ positive integer}\right\}\cup\{0\},$$

$$U_x\equiv V_x\equiv\{1,2\}\quad\text{for all } x\in X,$$

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$$f\left(\frac{1}{n}, u, v\right)=\begin{cases}\dfrac{1}{n-1}, & \text{if } u=v,\\[2ex] \dfrac{1}{n+1}, & \text{if } u\neq v,\end{cases} \qquad \text{for } n\geqq 2,$$

$$f(1, u, v)=\begin{cases}0, & \text{if } u=v,\\ \frac{1}{2}, & \text{if } u\neq v,\end{cases}$$

$$f(0, u, v)=0, \qquad \text{for all } u, v,$$

$$c(x, u, v, y)=\begin{cases}1, & \text{if } x=1, y<\frac{1}{2},\\ 0, & \text{otherwise.}\end{cases}$$

Briefly, the game can be interpreted as follows. Both players play a matrix game at $1/n$, with player I moving one step either towards or away from his goal $x=0$ depending on whether $u=v$ or $u\neq v$. When player I reaches his goal $x=0$ from $x=1$, he gets a reward of 1 unit and the game ends.

To solve the game, we first observe that $J^*(0)=0$. By Theorem 1, $J^*(\cdot)$ is the smallest nonnegative solution of

$$J^*\left(\frac{1}{n}\right)=\text{value of matrix game}\left[J^*\left(f\left(\frac{1}{n}, i, j\right)\right)+c\left(\frac{1}{n}, i, j, f\left(\frac{1}{n}, i, j\right)\right)\right]. \tag{23}$$

Since the value of the symmetric matrix game $\begin{bmatrix} a & b\\ b & a\end{bmatrix}$ is $\frac{1}{2}(a+b)$, (23) is equivalent to

$$J^*(1)=\tfrac{1}{2}[1+J^*(\tfrac{1}{2})],$$

$$J^*\left(\frac{1}{n}\right)=\frac{1}{2}\left[J^*\left(\frac{1}{n-1}\right)+J^*\left(\frac{1}{n+1}\right)\right] \qquad \text{for } n\geqq 2,$$

i.e.,

$$\begin{aligned} J^*(\tfrac{1}{2})&=2J^*(1)-1,\\ J^*\left(\frac{1}{n+1}\right)&=2J^*\left(\frac{1}{n}\right)-J^*\left(\frac{1}{n-1}\right) \qquad \text{for } n\geqq 2.\end{aligned} \tag{24}$$

By induction, it follows that

$$J^*\left(\frac{1}{n+1}\right)=(n+1)J^*(1)-n.$$

Hence

$$J^*(1)=\frac{n}{n+1}+\frac{1}{n+1}J^*\left(\frac{1}{n+1}\right)\geqq\frac{n}{n+1} \qquad \text{for all } n,$$

therefore $J^*(1)\geqq 1$. Also,

$$J^*\left(\frac{1}{n+1}\right)=(n+1)J^*(1)-n\geqq 1.$$

On the other hand $J^*(1/n)\equiv 1$ is a solution of (24), and hence from Theorem 1 we obtain $J^*(1/n)\equiv 1$.

Note now that X is compact with the relative Euclidean topology, $q(\cdot\,|x, u, v)$ and $c(x, u, v, y)$ are continuous, (we only need to check at $x=0$), yet [4] does not work

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

because J^* is not continuous at $x=0$ (hence J^*_α does not converge uniformly to J^* as $\alpha \uparrow 1$). However, we note that player I actually has an optimal stationary strategy which consists of choosing the probability vector $(\frac{1}{2}, \frac{1}{2})$ always.

REFERENCES

[1] L. S. SHAPLEY, *Stochastic games*, Proc. Nat. Acad. Sci., 39 (1953), pp. 1095–1100.

[2] H. EVERETT, *Recursive games*, Ann. Math. Stud., 39 (1957), pp. 47–78.

[3] A. MAITRA AND T. PARTHASARATHY, *On stochastic games*, J. Opt. Theor. Appl., 5 (1970), pp. 289–300.

[4] A. MAITRA AND T. PARTHASARATHY, *On stochastic games* II, J. Opt. Theor. Appl., 8 (1971), pp. 154–160.

[5] T. PARTHASARATHY AND T. E. S. RAGHAVAN, *Some Topics in Two Person Games*, American Elsevier, New York, 1971.

[6] D. P. BERTSEKAS AND S. E. SHREVE, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York, 1978.

[7] J. VON NEUMANN AND O. MORGENSTERN, *Theory of Games and Economic Behavior*, University Press, Princeton, NJ, 1944.

[8] L. D. BROWN AND R. PURVES, *Measurable selections of extrema*, Ann. Stat., 1 (1973), pp. 902–912.

[9] R. E. STRAUCH, *Negative dynamic programming*, Ann. Math. Stat., 3 (1966), pp. 871–890.

[10] D. ORNSTEIN, *On the existence of stationary optimal strategies*, Proc. Am. Math. Soc., 20 (1969), pp. 563–569.

[11] D. P. BERTSEKAS AND S. E. SHREVE, *Existence of optimal stationary policies in deterministic optimal control*, J. Math. An. Appl., 69 (1979), pp. 607–620.

[12] K. R. PARTHASARATHY, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.

Downloaded 01/03/13 to 132.206.27.25. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php