ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Solan, Eilon

Working Paper

Continuity of the Value in Stochastic Games

Discussion Paper, No. 1310

Provided in Cooperation with:
Kellogg School of Management - Center for Mathematical Studies in Economics and Management Science, Northwestern University

Suggested Citation: Solan, Eilon (2000) : Continuity of the Value in Stochastic Games, Discussion Paper, No. 1310, Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science, Evanston, IL

This Version is available at:
https://hdl.handle.net/10419/221664

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Continuity of the Value in Stochastic Games

Eilon Solan *†

November 14, 2000

Abstract

We prove that the undiscounted value of a stochastic game, as a function of the transition, is continuous in the relative interior of the set of transition functions.

*I thank Ehud Lehrer for raising the question and commenting on an earlier version of the paper, and Ron, who had an inactive yet an important role in the solution.

†Department of Managerial Economics and Decision Sciences, Kellogg Graduate School of Management, Northwestern University, *and* The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: eilons@post.tau.ac.il.

1 Introduction

A two-player zero-sum stochastic game is a tuple $G = (S, A, B, r, q)$, where (i) S is a finite set of states, (ii) A and B are finite sets of actions of the two players, (iii) $r : S \times A \times B \to \mathbf{R}$ is a daily payoff function and (iv) $q : S \times A \times B \to \Delta(S)$ is a transition function, where $\Delta(S)$ is the space of probability distributions over S.

The game is played as follows. At every stage $k \in \mathbf{N}$, given the past play (including the current state), each player chooses an action. The pair of chosen actions (a_k, b_k), together with the current state s_k, determine the daily payoff $r(s_k, a_k, b_k)$ player 2 pays player 1, as well as the probability distribution $q(\cdot \mid s_k, a_k, b_k)$ according to which a new state s_{k+1} is chosen.

Shapley (1953) proved that every λ-discounted stochastic game G admits a discounted value $v_\lambda(G)$, and, moreover, both players have optimal stationary strategies.

A fundamental question is whether, fixing the sets of states and actions, the value of the game is continuous in the payoff and transition functions.

Since the discounted payoff is continuous in the daily payoff and in the transition function, one can verify that the discounted value of a game is indeed continuous in these two parameters.

Mertens and Neyman (1981) showed that every stochastic game G admits an undiscounted value $v(G)$, and, moreover, $v(G) = \lim_{\lambda \to 0} v_\lambda(G)$.

One can easily verify that, fixing the transition function, the undiscounted value is continuous in the payoff function. Indeed, if r and r' are two payoff functions that satisfy $\|r - r'\| < \delta$, then any ϵ-optimal strategy under r is $\epsilon + 2\delta$-optimal under r'.

Simple examples show that the undiscounted value is *not* continuous in the transition function. For example, define $S = \{s, t\}$, $|A| = |B| = 1$, and $r(s) = 0$, $r(t) = 1$. Define transition functions q and q_n, $n \geq 1$ by: $q(s \mid s) = q(t \mid t) = 1$, and $q_n(s \mid s) = 1 - 1/n$, $q_n(t \mid t) = 1$. If the initial state is s, the undiscounted value under q is 0, while the undiscounted value under q_n is 1 for every $n \in \mathbf{N}$.

In the present paper we show that, fixing the payoff function, the undiscounted value is continuous in the *relative interior* of the set of transition functions. Formally, define $\mathrm{supp}(q) = \prod_{(s,a,b) \in S \times A \times B} \mathrm{supp}(q(\cdot \mid s, a, b)) \subseteq S \times A \times B \times S$. Our result is:

THEOREM 1.1 *Let $G = (S, A, B, r, q)$ be a stochastic game. For every $n \in \mathbf{N}$*

let $q_n : S \times A \times B \to \Delta(S)$ *be a transition function that satisfies* $\mathrm{supp}(q_n) = \mathrm{supp}(q)$. *Assume that* $q_n \to q$*; that is, for every* $(s, a, b, t) \in S \times A \times B \times S$, $q_n(t \mid s, a, b) \to q(t \mid s, a, b)$, *and define* $G_n = (S, A, B, r, q_n)$. *Then* $\lim_{n\to\infty} v(G_n) = v(G)$.

2 The Proof

To prove the result, we relate the discounted payoff to the transition function. Such a relation was developed by Vieille (2000) for undiscounted recursive games, and it can be used here, since any discounted game is equivalent to an undiscounted recursive game.

From now on we fix the set of states S, the sets of actions of the two players A and B, and the daily payoff function r.

For every discount factor $\lambda \in (0, 1)$, every initial state $s \in S$, every pair of stationary strategies $(x, y) \in (\Delta(A))^S \times (\Delta(B))^S$ and every transition function q, denote by $\gamma_\lambda(s, x, y; q)$ the expected λ-discounted payoff under (x, y) if the initial state is s, provided the transition function is q:

$$\gamma_\lambda(s, x, y; q) = \lambda \sum_{k=1}^{\infty} (1 - \lambda)^{k-1} \mathbf{E}_{s,x,y,q} r(s_k, a_k, b_k).$$

Lemma 2.1 *For every initial state* s *the function* $\gamma_\lambda(s, x, y; q)$ *is the ratio of two polynomials in* λ, x, y *and* q. *Moreover, the coefficients in the denominator are non-negative, and, if the daily payoffs are strictly positive, the coefficients in the numerator are non-negative as well.*

Proof: This result is an easy corollary of the study of Vieille (2000) on recursive games, or of a result due to Freidlin and Wentzell (1984) on Markov chains.

The λ-discounted game is equivalent to an undiscounted recursive game, where the set of non absorbing states is S, at stage k the game is absorbed with probability λ regardless of the current state or chosen actions, and the absorbing payoff is $r(s_k, a_k, b_k)$, while with probability $1 - \lambda$ transitions are as in the original discounted game. This representation is actually Shapley's (1953) original description of a discounted stochastic game.

Note that in this representation, the play eventually reaches an absorbing state with probability 1.

The lemma now follows by Vieille (2000, Eq. (3) and first equation in section 7.2)). Alternatively, one can consider the function that assigns to every $s, t \in S$ the probability that $t \in S$ is the last non absorbing state in the equivalent recursive game that the play visits before absorption, provided the initial state is s. It follows from Freidlin and Wentzell (1984, Lemma 6.3.3) that this function is the ratio of two polynomials in λ, x, y and q, and the coefficients in the denominator are non-negative. ■

Lemma 2.1 implies the following.

COROLLARY 2.2 *Assume payoffs are strictly positive. For every $\epsilon \in (0,1)$, every transition function q, and every transition rule q' that satisfies (i) $\mathrm{supp}(q) = \mathrm{supp}(q')$, and (ii) $1-\epsilon < q(t \mid s,a,b)/q'(t \mid s,a,b) < 1+\epsilon$ for every $(s,a,b,t) \in \mathrm{supp}(q)$, we have*

$$1-\epsilon < \gamma_\lambda(s,x,y;q)/\gamma_\lambda(s,x,y;q') < 1+\epsilon,$$

for every discount factor $\lambda \in (0,1)$, every initial state s, and every pair of stationary strategies (x,y).

Proof: By Lemma 2.1 it is sufficient to prove the following. Let $f(u,v) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=0}^{K}\sum_{l=0}^{L} a_{ijkl}(u_i)^k(v_j)^l$ be a polynomial in $u \in [0,1]^I$ and in $v \in [0,1]^J$ with non negative coefficients. Then for every $\epsilon \in (0,1)$ we have $1-\epsilon < f(u,v)/f(u,v') < 1+\epsilon$ for every $u \in [0,1]^I$, every $v \in [0,1]^J$ and every $v' \in [0,1]^J$ that satisfies for every $j = 1,\ldots,J$ (i) $v_j > 0$ if and only if $v'_j > 0$, and (ii) $1-\epsilon < v_j/v'_j < 1+\epsilon$ whenever the denominator does not vanish. (By convention, $0/0 = 1$.)

This latter claim follows easily from the following fact. If $(a_i)_{i=1}^N$ and $(b_i)_{i=1}^N$ are two sequences of positive real numbers such that $1-\epsilon < a_i/b_i < 1+\epsilon$ for every i, then $1-\epsilon < \sum_{i=1}^N c_i a_i / \sum_{i=1}^N c_i b_i < 1+\epsilon$, for every sequence $(c_i)_{i=1}^N$ of positive real numbers. ■

Proof of Theorem 1.1: Assume w.l.o.g. that daily payoffs are strictly positive and bounded by 1; that is, $0 < r(s,a,b) \leq 1$ for every $(s,a,b) \in S \times A \times B$.

Assume to the contrary that the theorem is not true, and w.l.o.g. $\liminf_{n\to\infty} v(G_n) < v(G)$. Fix $\epsilon > 0$ sufficiently small that satisfies $v(G) - \liminf v(G_n) > 3\epsilon$.

Since $\mathrm{supp}(q_n) = \mathrm{supp}(q)$ for every $n \in \mathbf{N}$, and by Corollary 2.2, there is n_0 sufficiently large such that for every $n > n_0$,

$$1 - \epsilon < \frac{\gamma_\lambda(s, x, y; q_n)}{\gamma_\lambda(s, x, y; q)} < 1 + \epsilon, \tag{1}$$

for every discount factor λ, every initial state s, and every pair of stationary strategies (x, y).

Fix $n > n_0$ such that $v(G_n) < v(G) - 3\epsilon$. Fix now $\lambda \in (0, 1)$ sufficiently small such that $\|v_\lambda(G_n) - v(G_n)\| < \epsilon$ and $\|v_\lambda(G) - v(G)\| < \epsilon$.

Let x_λ be an optimal strategy of player 1 in the λ-discounted version of G, and let y_λ be a best reply of player 2 against x_λ in the λ-discounted version of G_n. In particular,

$$\begin{aligned} \gamma_\lambda(s, x_\lambda, y_\lambda; q_n) &\leq v_\lambda(G_n) < v(G_n) + \epsilon < v(G) - 2\epsilon \\ &< v_\lambda(G) - \epsilon \leq \gamma_\lambda(s, x_\lambda, y_\lambda; q) - \epsilon. \end{aligned}$$

Since payoffs are strictly positive and bounded by 1,

$$\frac{\gamma_\lambda(s, x_\lambda, y_\lambda; q_n)}{\gamma_\lambda(s, x_\lambda, y_\lambda; q)} < 1 - \epsilon,$$

contradicting (1). ■

References

[1] Freidlin M.I. and Wentzell A.D. (1984) Random Perturbations of Dynamical Systems, Springer

[2] Mertens J.F. and Neyman A. (1981) Stochastic Games, *Int. J. Game theory*, **10**, 53-66

[3] Shapley L.S. (1953) Stochastic Games, *Proc. Nat. Acad. Sci. U.S.A.*, **39**, 1095-1100

[4] Vieille N. (2000) Equilibrium in 2-Person Stochastic Games II: The Case of Recursive Games, *Israel J. Math.*, *forthcoming*