האוניברסיטה העברית בירושלים

THE HEBREW UNIVERSITY OF JERUSALEM

ABSORBING GAMES WITH COMPACT ACTION SPACES

by

JEAN-FRANÇOIS MERTENS, ABRAHAM NEYMAN
and DINAH ROSENBERG

Discussion Paper # 456 July 2007

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

ABSORBING GAMES WITH COMPACT ACTION SPACES

JEAN-FRANÇOIS MERTENS†, ABRAHAM NEYMAN‡, AND DINAH ROSENBERG§

ABSTRACT. We prove that games with absorbing states with compact action sets have a value.

INTRODUCTION

Stochastic games are Markov decision processes in which the transitions of the state are controlled by the actions of the decision makers. In a stochastic game the players interact repeatedly. At each stage the players observe the current state, next choose an action independently, and are then informed of the chosen actions. According to these actions and the current state the chain moves to a new state that is observed by all players. The stage payoff is a function of the current state and the actions chosen. We focus on two-player zero-sum stochastic games.

Stochastic games were introduced by Shapley [10]. He proved the existence of the value of λ-discounted two-player zero-sum stochastic games with finitely many states and actions.

In the case where the sets of actions and of states are finite, the existence of the limit, say v, of the values of the λ-discounted games as λ goes to 0 (i.e., as players become more and more patient) was proved in [1], using an algebraic argument. It is proved in [5] that v is a value: for each $\varepsilon > 0$, there exists $N(\varepsilon)$ and $\lambda(\varepsilon)$ and each of the players has a strategy that 1) guarantees him a payoff of v up to an error of ε in any n-stage or λ-discounted game provided $n \geq N(\varepsilon)$ or $\lambda \leq \lambda(\varepsilon)$, and 2) the strategy of player one (respectively, player two) guarantees that the expectation of the liminf (respectively, limsup) of the average payoff in the first n-stages as n goes to infinity is at least $v - \varepsilon$ (respectively, at most $v + \varepsilon$).

In the case of stochastic games with finite state space but with compact action sets, there is no general result ensuring the convergence of the values of the n-stage or λ-discounted games (and a fortiori none ensuring the existence of the value).

An absorbing state of a stochastic game is a state such that for any profile of actions chosen by the players the state remains the same almost surely. An absorbing

Date: June 28, 2007.

2000 *Mathematics Subject Classification.* 91A15

J.E.L. *Classification numbers.* C73.

Key words and phrases. Games, stochastic games, zero-sum, minimax, absorbing states, value.

This research was supported in part by the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Prime Minister's Science Policy Office. Financial support from European TMR Network FMRX CT 960055 is gratefully acknowledged. This research was supported in part by Israel Science Foundation grants 263/03 and 1123/06, and by the Center for rationality of the Hebrew University of Jerusalem.

†CORE, Université Catholique de Louvain, 34 Voie du Roman-Pays, B-1348 Louvain-la-Neuve, Belgique. E-mail: jfm@core.ucl.ac.be.

‡Institute of Mathematics and Center for Rationality, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel. E-mail: aneyman@math.huji.ac.il.

§LAGA, Institut Galilée, Université Paris Nord, Avenue Jean Baptiste Clément, 93430 Villetaneuse, France and Laboratoire d'Econométrie de l'École Polytechnique, Paris, France. E-mail: dinah@math.univ-paris13.fr.

game is a stochastic game in which all states but one are absorbing. [3] proved the existence of the value for (two-person zero-sum) absorbing games with finite action sets (this is now a particular case of [5]).

For absorbing games with compact action sets the algebraic approach of [1] does not apply and convergence of the values of the λ-discounted games was proved using an operator approach by [9].

In this paper we prove that the value of absorbing games with compact action sets exists. The proof relies on the characterization of the limit of the values of the λ-discounted games provided in [9], and on [5].

1. The Model

Here a *stochastic game* is a two-player zero-sum game determined by:

- Three sets: S (the set of states), I (the set of actions of player 1), and J (the set of actions of player 2). We will assume throughout that I and J are compact metric; we will see that when focusing on absorbing games we can assume without loss of generality that S is finite.
- A bounded payoff function $g\colon S \times I \times J \to \mathbb{R}$ that is separately continuous. Note that this implies measurability (see [6] [I.1.Ex.7a]).
- A transition probability q from $S \times I \times J$ to S, where $q(\, z' \mid z, i, j\,)$ denotes the probability of reaching state z' from state z given the pair of actions (i, j), where q is separately continuous on $I \times J$. Note that this implies measurability (see [6] [I.1.Ex.7a]).
- An initial state $z_1 \in S$.

The game is played in stages. At each stage $n \in \mathbb{N}$, player 1 and player 2 choose an action, $\mathtt{i}_n \in I$ and $\mathtt{j}_n \in J$, knowing the whole past history, including current state z_n. Then the current payoff is $g_n = g(z_n, \mathtt{i}_n, \mathtt{j}_n)$, and $q(s \mid z_n, \mathtt{i}_n, \mathtt{j}_n)$ is the conditional probability, given $z_1, i_1, j_1, \ldots, z_{n-1}, i_{n-1}, j_{n-1}$, that the next state $z_{n+1} = z$. We denote by h_n the history up to stage n, more precisely $h_n = (z_1, \mathtt{i}_1, \mathtt{j}_1, ..., z_{n-1}, \mathtt{i}_{n-1}, \mathtt{j}_{n-1}, z_n)$. Let H_n denote the set of histories up to stage n, and H the set of infinite length histories. Let $\mathcal{H}_n$ be the σ-algebra on H induced by histories h_n (of length n) and $\mathcal{H}_\infty$ be the σ-algebra spanned by $\cup_n \mathcal{H}_n$.

Thus, a player's behavioral strategy (σ of player 1, τ of player 2) specifies a probability distribution over his actions at each stage conditional on the current state and the past history. A pair (σ, τ) induces a probability $P_{\sigma,\tau}$ on histories $(H, \mathcal{H}_\infty)$. The corresponding expectation is denoted by $\mathsf{E}_{\sigma,\tau}$.

Notation 1. *We use* $\|\cdot\|$ *for the* sup*-norm, and let* $A = \|g\| := \mathsf{sup}_{z,i,j} |g(z, i, j)|$.

We are interested in the existence of infinite-game strategies that guarantee a given payoff in all sufficiently long games, as well as in the infinite undiscounted game.

Definition 1. Player I can guarantee $v \in \mathbb{R}^S$ if for every $\delta > 0$ and $s_1 \in S$ there is a strategy σ of player 1 and $N > 0$ such that for any strategy τ of player II,

$$\mathsf{E}_{\sigma,\tau}\left[\liminf_{n\to\infty} \frac{1}{n} \sum\nolimits_{i<n} g_i\right] \geq v(z_1) - \delta$$

$$\forall n \geq N,\ \mathsf{E}_{\sigma,\tau}\left[\frac{1}{n} \sum\nolimits_{i<n} g_i\right] \geq v(z_1) - \delta$$

Player II can guarantee $v \in \mathbb{R}^S$ if for every $\delta > 0$ and $s_1 \in S$ there is a strategy

τ of player 2 and $N > 0$ such that for any strategy σ of player I,

$$\mathsf{E}_{\sigma,\tau}\left[\limsup_{n\to\infty}\frac{1}{n}\sum\nolimits_{i=1}^{n} g_i\right] \le v(z_1)+\delta$$
$$\forall n \ge N,\ \mathsf{E}_{\sigma,\tau}\left[\frac{1}{n}\sum\nolimits_{i=1}^{n} g_i\right] \le v(z_1)+\delta$$

A stochastic game has a value $v \in \mathbb{R}^S$ if both players can guarantee v.

All stochastic games with finite state space S and finite action sets I and J have a value. It is unknown whether all stochastic games with a finite state space have a value. This paper proves that all absorbing games have a value.

Absorbing games with finite action sets I and J were studied in [3]. An absorbing state is a state $z \in S$ such that for all $\mathtt{i} \in I$ and all $\mathtt{j} \in J$, $q(z \mid z, \mathtt{i}, \mathtt{j}) = 1$. An absorbing game is a stochastic game in which all states but one are absorbing. We are going to study absorbing games with compact action sets.

Remark 1. As soon as an absorbing state z is hit, both players know it; the rest of the game is therefore a repeated zero-sum game, with value $v(z)$. We add to the game two absorbing states s^+ and s^- with constant payoff A and $-A$. Then one can replace the transitions to any absorbing state z by transitions with probabilities $(1 \pm A^{-1}v(s))/2$ to s^+ and s^-. Thus, in an absorbing game we need without loss of generality at most 3 states: s^0 (the nonabsorbing state), s^+, and s^-.

This note proves the following result:

Theorem 1. *Absorbing games have a value.*

2. The proof

2.1. Reminder of λ-discounted games.

Notation 2. *For a compact set X, $\Delta(X)$ denotes the set of probability distributions over X.*

The λ-discounted payoff function is, for $\lambda \in (0,1)$:

$$\mathsf{E}_{\sigma,\tau}\sum\nolimits_{i=1}^{\infty}\lambda(1-\lambda)^{i-1}g_i$$

The value $v_\lambda(s_1)$ and stationary optimal strategies (x_λ, y_λ) (of the λ-discounted game) exist ([6, Chapter VI proposition 1.4]), and v_λ is characterized by:

$$v_\lambda = T(\lambda, v_\lambda) \tag{1}$$

where for $u \in \mathbb{R}^S$,

$$T(\lambda,u)(s) = \max_{x\in\Delta(I)}\min_{y\in\Delta(J)} \mathsf{E}_{x,y,q}\left[\lambda g(s,i,j) + (1-\lambda)u(\tilde{s})\right] \tag{2}$$

where $\mathsf{E}_{x,y,q}$ is the expectation operator where, independently, i and j are distributed according to x and y, and then $\tilde{s}$ is distributed according to $q(\tilde{s}|s,i,j)$.

2.2. Reminder of the Mertens-Neyman theorem.

The proof in Section 2 of [5] proves the following[1] theorem:

Theorem 2. *If $\lambda \mapsto w_\lambda \in \mathbb{R}^S$ is a function defined on $]0,1[$ with*

$$\|w_\lambda - w_{\bar\lambda}\| \le \int_\lambda^{\bar\lambda}\psi(x)dx \qquad \textit{for all } 0<\lambda,\bar\lambda<1 \tag{3}$$

[1]In [5, Section 2] the function v_λ indeed stands for the value of the λ-discounted game, and thus condition (4) of Theorem 2 (equivalently, [5, Inequality 2.1]) and [5, Inequality 2.2] follow. These two conditions are the only use in [5, Section 2] of the fact that v_λ is the value of the λ-discounted game. Other examples where the proof in [5] is applied to functions v_λ that are not necessarily the values of the discounted games appear in [7] and [8].

where $\psi :]0,1] \to \mathbb{R}_+$ is integrable and for every $\lambda \in]0,1]$ sufficiently small we have

$$T(\lambda, w_\lambda) \geq w_\lambda \tag{4}$$

then player 1 can guarantee $\lim_{\lambda\to 0+} w_\lambda$.

2.3. **The auxiliary function w_λ.** Recall that we can assume w.l.o.g. that the absorbing game has 3 states, s^0, s^+, and s^-. Fixing in (2) $w(s^+)$ and $w(s^-)$ at the values A and $-A$, one gets a map $T(\lambda, \cdot)\colon \mathbb{R} \to \mathbb{R}$; this map is the restriction of $T(\lambda, .)$ to vectors that take values A and $-A$ at their s^+ and s^- coordinates; unless otherwise specified $T(\lambda, .)$ will now denote this restriction. Then, by [9], v_λ converges to v, characterized by $v(s^+) = A$, $v(s^-) = -A$ and:

$$\begin{aligned} T(0, v) &= v && \text{and} \\ \lim_{\lambda\to 0}\big(T(\lambda, w) - w\big)/\lambda &< 0 && \text{for } w > v \\ \lim_{\lambda\to 0}\big(T(\lambda, w) - w\big)/\lambda &> 0 && \text{for } w < v \end{aligned} \tag{5}$$

Take $\varepsilon > 0$. Our goal is to apply the previous theorem with $w_\lambda = v_\varepsilon$ (for any λ) defined by $v_\varepsilon(z) = v(z) - \varepsilon \mathbb{I}_{z=s^0}$.

2.4. **Proof of Theorem 1.** First, we prove that w_λ $(= v_\varepsilon(s))$ satisfies the conditions of Theorem 2. As w_λ is independent of λ the function $\lambda \mapsto w_\lambda$ satisfies condition (3) with the function $\psi(\lambda) = 0$. Condition (4) holds trivially for $s_1 = s^+$ and for $s_1 = s^-$. It remains to prove that the condition holds for $s_1 = s^0$. However, equation (5) implies that, for λ small enough,

$$T(\lambda, v_\varepsilon) \geq v_\varepsilon \tag{6}$$

Therefore, for every $\varepsilon > 0$, player 1 can guarantee v_ε, and therefore player 1 can guarantee v. This completes the proof that continuous absorbing games have a value.

3. An explicit strategy

In this section we construct an explicit strategy, based on the construction in [5, Section 2] and using the auxiliary function $w_\lambda = v_\varepsilon$ that obeys inequality (6).

Fix $\varepsilon > 0$. We define a sequence $(\lambda_i)_{i=1}^\infty$ so that $0 < \lambda_i < 1$ is a function of past history, i.e., measurable w.r.t. the σ-algebra $\mathcal{F}_i$ of all events preceding time i (including the choice of a new state z_i after the play at time $i-1$), and so that all λ_i are sufficiently small so that $T(\lambda_i, v_\varepsilon) \geq v_\varepsilon$. The $(\lambda_i)_{i=1}^\infty$-strategy of player 1 is to play on time i a strategy such that inequality (2.1) of [5],

$$E(w_{\lambda_i}(z_{i+1}) - w_{\lambda_i}(z_i) + \lambda_i(x_i - w_{\lambda_i}(z_{i+1})) \mid \mathcal{F}_i) \geq 0 \tag{7}$$

holds with $w_{\lambda_i}(z) := v(z) - \varepsilon \mathbb{I}_{z=z_1}$, where z_1 is the initial nonabsorbing state.

Let $M > 1/\varepsilon$ be a sufficiently large constant so that (6) holds for $\lambda < 1/M^2$ and $(6A)^2/M < \varepsilon$ and thus for all $s_i, s_{i+1} \geq M$ with $|s_{i+1} - s_i| \leq 6A$ we have

$$|(s_{i+1} - s_i)(\frac{s_i}{s_{i+1}} - 1)| < 2\varepsilon \tag{8}$$

Defining inductively (as in [5, Section 2]), $s_{i+1} = \max[M, s_i + x_i - v_{\lambda_i}(z_{i+1}) + 4\varepsilon]$ and $\lambda_i = \frac{1}{s_i^2}$, starting with $s_1 \geq M$ arbitrary. Set $Y_i = w_{\lambda_i}(z_i) - 1/s_i$ $(= v_\varepsilon - 1/s_i)$. We have

$$\begin{aligned} Y_{i+1} - Y_i \quad &= \quad w_{\lambda_i}(z_{i+1}) - w_{\lambda_i}(z_i) + \lambda_i(x_i - v_{\lambda_i}(z_{i+1})) \\ &\quad\; -1/s_{i+1} \;+\; 1/s_i \quad - \lambda_i(x_i - v_{\lambda_i}(z_{i+1})) \end{aligned}$$

and therefore, using (7), we have

$$\begin{aligned}E(Y_{i+1}-Y_i \mid \mathcal{F}_i) &\geq E(1/s_i - 1/s_{i+1} - \lambda_i(x_i - v_{\lambda_i}(z_{i+1})) \mid \mathcal{F}_i)\\ &\geq E(\frac{s_{i+1}-s_i}{s_i s_{i+1}} - \lambda_i(s_{i+1}-s_i-4\varepsilon) \mid \mathcal{F}_i)\\ &= E(4\varepsilon\lambda_i - \lambda_i(s_{i+1}-s_i)(\frac{s_i}{s_{i+1}}-1) \mid \mathcal{F}_i)\\ &\geq 2\varepsilon\lambda_i\end{aligned}$$

where the second inequality follows from $s_{i+1}-s_i \geq x_i - v_{\lambda_i}(z_{i+1}) + 4\varepsilon$, and the last inequality follows from (8).

Since $\lambda_i \geq 0$, the inequality $E(Y_{i+1}-Y_i \mid \mathcal{F}_i) \geq 2\varepsilon\lambda_i$ implies that Y_i is a submartingale. Obviously, Y_i is bounded and thus it converges a.s., say to Y_∞, with $E(Y_\infty \mid \mathcal{F}_1) > Y_1$. It follows that $4A \geq 2A + 2/M \geq E(Y_k - Y_1) \geq 2\varepsilon E(\sum_{1\leq i<k}\lambda_i)$ so that by the monotone convergence theorem

(9) $$E(\sum_{i<\infty}\lambda_i) < 2A/\varepsilon, \text{ and}$$

(10) $$E(\#\{i \mid \lambda_i \geq \eta\}) \leq \frac{2A}{\varepsilon\eta}.$$

Thus, a.s., $\lambda_i \to 0$, $s_i \to \infty$ (and hence also $\mathbb{I}_{s_i=M} \to 0$), as $i \to \infty$, and therefore

(11) $$w_{\lambda_i}(z_{i+1}) \to Y_\infty \text{ with } E(Y_\infty \mid \mathcal{F}_1) > Y_1 \geq v(z_1) - \varepsilon - 1/M$$

Also

(12) $$E(w_{\lambda_i}(z_{i+1})) \geq v(z_1) - \varepsilon - 1/M$$

From the definition of s_{i+1} it follows that

$$x_i \geq v_{\lambda_i}(z_{i+1}) + s_{i+1} - s_i - 4\varepsilon - 2A\mathbb{I}_{s_{i+1}=M}$$

Summing these inequalities over $1 \leq i < n$, we have

(13) $$\sum_{i<n} x_i \geq \sum_{i<n} v_{\lambda_i}(z_{i+1}) + s_n - s_1 - 4\varepsilon n - \sum_{i<n} 2A\mathbb{I}_{s_{i+1}=M}$$

implying that

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{i<n} x_i \geq Y_\infty - 4\varepsilon$$

and thus

$$E(\liminf_{n\to\infty} \frac{1}{n}\sum_{i<n} x_i) \geq E(Y_\infty \mid \mathcal{F}_1) - 4\varepsilon \geq v(z_1) - 5\varepsilon - 1/M$$

and

$$E(\frac{1}{n}\sum_{i<n} x_i) \geq v(z_1) - 2\varepsilon - 4\varepsilon - s_1/n$$

Altogether, we deduce that the $(\lambda_i)_{i=1}^\infty$-strategy of player 1 guarantees $v(z_1) - 7\varepsilon$. Thus, player 1 can guarantee $v(z_1)$. Similarly, player 2 can guarantee $v(z_1)$ and therefore $v(z_1)$ is the value of the absorbing game.

The $(\lambda_i)_{i=1}^\infty$-strategy of player 1 (is a constant mixed action on the absorbing states, and) has a simplified form on the nonabsorbing state. Indeed, as $w_\lambda(z) = v(z) - \varepsilon\mathbb{I}_{z=s^0}$ we can define $s_{i+1} = \max[M, s_i + x_i - v(s^0) + 5\varepsilon]$.

References

[1] Bewley, T. and E. Kohlberg (1976), The Asymptotic Solution of a Recursion Equation occurring in Stochastic Games, *Mathematics of Operations Research*, 1, 321-336.

[2] Blackwell, D. and T.S. Ferguson (1968), The Big Match, *Annals of Mathematical Statistics*, 39, 159-168.

[3] Kohlberg, E. (1974), Repeated Games with Absorbing States, *The Annals of Statistics*, 2, 724–738.

[4] Mertens, J.F. and A. Neyman (1980), Stochastic Games, Core Discussion Paper 8001, Université Catholique de Louvain, Belgium.

[5] Mertens, J.F. and A. Neyman (1981), Stochastic Games, *International Journal of Game Theory*, 10, 53–66.

[6] Mertens J.F., S. Sorin and S. Zamir (1994), *Repeated Games*, Core Discussion Papers 9420-9421-9422, Université Catholique de Louvain la Neuve, Belgium.

[7] Neyman, A. (2003), Stochastic Games: Existence of the Minmax, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Dordrecht, Kluwer Academic Publishers, Chapter 6, pp. 57–75.

[8] Rosenberg, D., E. Solan and N. Vieille (2003), The Maxmin Value of Stochastic Games with Imperfect Monitoring, *International Journal of Game Theory*, 32, 133–150.

[9] Rosenberg, D. and S. Sorin (2001), An Operator Approach to Stochastic Games, *Israel Journal of Mathematics*, 121, 221–246.

[10] Shapley, L. (1953), Stochastic Games, *Proceedings of the National Academy of Sciences, U.S.A.*, 39, 1095–1100.