Uniform Value in Recursive Games

Eilon Solan[*] and Nicolas Vieille[†]

April 14, 2000

Abstract

We address the problem of existence of the uniform value in recursive games. We give two existence results. (i) The uniform value is shown to exist if the state space is countable, the action sets are finite and if, for some $a > 0$, there are finitely many states in which the limsup value is less than a. (ii) For games with non-negative payoff function, it is sufficient that the action set of player 2 is finite. The finiteness assumption can be further weakened.

[*]Department of Managerial Economics and Decision Sciences, Kellogg Graduate School of Management, Northwestern University, Evanston IL 60208. e-mail: e-solan@nwu.edu

[†]GRAPE, Université Bordeaux-IV Montesquieu, and Laboratoire d'Econométrie de l'Ecole Polytechnique, 1 rue Descartes, 75005 Paris, France. e-mail: vieille@poly.polytechnique.fr

1 Introduction

Two player stochastic games are played in stages. At every stage the game is in some state of the world. Each player, given the whole history, chooses an action independently of the other. The current state together with the pair of actions determine a daily payoff for player 1, as well as a probability distribution according to which a new state of the world is chosen.

The goal of player 1 is to maximize the expected overall payoff, and the goal of player 2 is to minimize the expected overall payoff (note that we did not define yet the "overall payoff". In a moment we will see several possible definitions).

Under very mild assumptions the n-stage game — that is, the game where the overall payoff is the average of the daily payoffs of the first n stages — has a value v_n. When the overall payoff of the players is the λ-*discounted* sum of the infinite sequence of daily payoffs, existence of the value v_λ was proven under some continuity conditions on the transition probability (see, e.g., Nowak (1984a,b, 1985) or Mertens, Sorin and Zamir (1994))

In both cases, the optimal strategies of the players depend crucially on the parameter, the length of the game or the discount factor. A strategy that is optimal for one parameter may yield a low payoff for a different parameter.

A stochastic game has a *uniform value* v_∞ if $\lim_{n\to\infty} v_n$ exists, it is equal to v_∞, and for every ε there exist a positive integer n_0 and a pair of strategies $(\sigma_\varepsilon, \tau_\varepsilon)$ for the two players, each is ε-optimal in every n-stage game, provided $n \geq n_0$.

It can be shown that in this case σ_ε and τ_ε are 2ε-optimal in every discounted game, provided that the discount factor is sufficiently small. That is, if v_∞ exists then $v_\infty = \lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$.

Mertens and Neyman (1981) proved that if the state and action spaces are finite then the game has a uniform value. Their proof uses the fact that the function $\lambda \to v_\lambda$ has bounded variation (see Bewley and Kohlberg (1976) for this result).

If the state space or actions spaces are general, then this function needs not have bounded variation, hence the proof of Mertens and Neyman fails.

Another value that was studied in the literature is the *limsup value.* The limsup value is the value of the game in which the overall payoff to player 1 is the $\limsup$ of the daily payoffs.

Maitra and Sudderth (1993) proved that the limsup value v exists under very mild assumptions. It is easy to see that if the uniform value exists then $v = v_\infty$.

Lehrer and Sorin (1992) gave an example of a Markov decision process (with countable state space) where $\lim_{n\to\infty} v_n$ and $\lim_{\lambda\to 0} v_\lambda$ exist and differ, and they both differ from the limsup value v.

Recursive games are stochastic games where the state space is divided into two sets S and T — non-absorbing states and absorbing states. As long as the game is in S, the payoff is 0, whatever the players play. Once the game reaches a state in T, it remains in it with probability 1, whatever the players play.

Recursive games were introduced by Everett (1957) who proved the existence of the limsup value v and of stationary ε-optimal strategies, when the state space and the action sets are finite.

In the present paper we provide conditions under which the uniform value exists in recursive games. First, we investigate games with countable state space and finite action sets. For such games, Secchi (1997) gave conditions under which one of the players has a stationary ε-optimal strategy (in the limsup sense), but his strategies need not be ε-optimal in a uniform sense. We prove that if the limsup value is positive on S, and bounded away from zero, then the uniform value exists. We use this result to show that if, for some $a > 0$, there are only finitely many states in S where the limsup value is less than a, the game admits a uniform value.

We then show that if the game is positive — that is, if the payoff in absorbing states is always non-negative — then the assumptions on the state space and the limsup value can be dropped, and it is enough to require that the action set of player 2 is finite. This finiteness assumption can be further weakened. It is enough that for every $\varepsilon > 0$ and every state $s \in S$ player 2 has a mixed action that is ε-optimal in the game with continuation payoff $\limsup_{\lambda\to 0} v_\lambda$, and this ε-optimal strategy guarantees he pays (on average) at most $\limsup_{\lambda\to 0} v_\lambda(s) + \varepsilon$ in this one shot game.

The result of Rosenberg and Vieille (1998), who study recursive games with incomplete information, imply that if the values of the discounted games converge uniformly (over the state space) as the discount factor goes to zero then the uniform value exists. Their results are independent of ours.

2 The Model and the Main Results

A recursive game is described by:

- a measurable state space $\Omega = S \cup T$;
- topological action sets A and B for the two players;

- a transition function q from $S \times A \times B$ to Ω;
- a bounded measurable payoff function $g : T \to \mathbf{R}$.

The game is played as follows. An initial state s_1 is given. At any stage $n \geq 1$, the current state s_n is told to the players, the players choose actions a_n and b_n, possibly at random, and the next state s_{n+1} is drawn according to $q(\cdot|s_n, a_n, b_n)$. Once the game reaches a state $s \in T$, player 1 receives from player 2 a stage payoff $g(s)$, and the game remains in s forever.

It is usually important to specify what each player knows in any given stage about the past play of the other player. This is irrelevant for our result: the ε-optimal strategies that we construct have the feature that what a player does depends only on the sequence of states visited so far (including the current one). Therefore, provided the information available to a player enables him to recover this sequence, our results holds. For simplicity, we assume that, in any stage, each player knows the entire past play.

2.1 Strategies

A and B are endowed with the σ-fields of Borel sets. The set of histories of length n is $H_n = \Omega \times (A \times B \times \Omega)^{n-1}$, and the set of finite histories is $H = \cup_{n \in \mathbf{N}} H_n$, where $\mathbf{N}$ is the set of positive integers. The set of plays is $H_\infty = (\Omega \times A \times B)^{\mathbf{N}}$. It is convenient to identify any $h_n \in H_n$ with a cylinder set of H_∞. The σ-algebra induced by H_n over H_∞ is denoted by $\mathcal{H}_n$: it is the information available to the players at stage n. The product σ-field on H_∞ is $\mathcal{H}_\infty = \sigma(\mathcal{H}_n, n \geq 1)$.

We let $\Delta(A)$ and $\Delta(B)$ denote the sets of probability measures over A and B, endowed with the weak-$*$ topology.

A strategy of player 1 is a map $\sigma : H \to \Delta(A)$, (such that the restriction of σ to H_n is measurable), with the interpretation that $\sigma_n(h_n)$ is the lottery used by player 1 to choose his action at stage n, if the history of play up to stage n is h_n. It is called *pure* if $\sigma(h_n)$ is a unit mass, for every $h_n \in H$. A strategy σ can be equivalently viewed as a sequence $(\sigma_n)_{n \geq 1}$, where $\sigma_n : (H_\infty, \mathcal{H}_n) \to \Delta(A)$ is measurable with respect to $\mathcal{H}_n$. Strategies of player 2 are defined analogously.

A strategy σ is stationary if $\sigma(h_n)$ depends only on the current state s_n. Thus, a stationary strategy reduces to a family $(\mathbf{x}(s), s \in S)$, where $\mathbf{x}(s) \in \Delta(A)$ is the mixed move played whenever the current state is s.[1]

[1] This differs from the terminology used in gambling theory, where these strategies are called stationary families.

The letters σ (*resp.* $\mathbf{x}$) will always stand for a strategy (*resp.* stationary strategy) of player 1. τ and $\mathbf{y}$ stand for strategies and stationary strategies of player 2. The sets of strategies of the two players are denoted by $\mathcal{S}$ and $\mathcal{T}$.

We denote by $\mathbf{P}_{s,\sigma,\tau}$ the law of play when the initial state is s, and the players follow the strategies σ and τ: $\mathbf{P}_{s,\sigma,\tau}$ is a probability distribution over $(H_\infty, \mathcal{H}_\infty)$. Expectation w.r.t. $\mathbf{P}_{s,\sigma,\tau}$ is denoted by $\mathbf{E}_{s,\sigma,\tau}$.

Let $t = \inf\{n \geq 1, s_n \in T\}$ be the *termination* stage, and set $g_n = g(s_t)1_{t\leq n}$ (payoff at stage n). Finally, denote by $\overline{g}_n = \frac{1}{n}\sum_{p=1}^n g_p$ the average payoff up to stage n.

We assume w.l.o.g. that $\| g \|_\infty \leq 1$.

2.2 Payoffs and value

Two notions of value have been studied in the literature. The first is based on the payoff function $\gamma : S \times \mathcal{S} \times \mathcal{T} \to \mathbf{R}$, defined as

$$\gamma(s,\sigma,\tau) = \mathbf{E}_{s,\sigma,\tau}[\limsup \overline{g}_n] = \mathbf{E}_{s,\sigma,\tau}[g(s_t)1_{t<+\infty}].$$

Definition 1 *$v : \Omega \to \mathbf{R}$ is the* limsup value *if, for every $s \in S$,*

$$v(s) = \sup_{\sigma\in\mathcal{S}} \inf_{\tau\in\mathcal{T}} \gamma(s,\sigma,\tau) = \inf_{\tau\in\mathcal{T}} \sup_{\sigma\in\mathcal{S}} \gamma(s,\sigma,\tau).$$

A strategy of player 1 which achieves the sup up to ε in the sup inf is called *ε-optimal*. We say that such a strategy *guarantees* $v - \varepsilon$.

We recall a result, which is a particular case of the result of Maitra and Sudderth (1993).

Theorem 2 (Maitra-Sudderth) *Assume that (i) Ω, A and B are Borel subsets of Polish spaces, (ii) B is compact, (iii) g is bounded and upper analytic,*[2] *(iv) $q(E \mid s, a, \cdot)$ is Borel measurable, and continuous over B for every $s \in S$, every $a \in A$ and every $E \subseteq \Omega$. Then v exists, and it is an upper analytic function.*

The second notion of value requires uniformity. Define $\gamma_n(s,\sigma,\tau) = \mathbf{E}_{s,\sigma,\tau}[\overline{g}_n]$, the expected average payoff during the first n stages.

Definition 3 *Let $w : \Omega \to \mathbf{R}$. We say that player 1* uniformly guarantees *w if for every $s \in \Omega$ and every $\varepsilon > 0$ there exists $\sigma_\varepsilon \in \mathcal{S}$ and $N \in \mathbf{N}$, such that*

$$\forall n \geq N, \forall \tau \in \mathcal{T}, \gamma_n(s,\sigma_\varepsilon,\tau) \geq w(s) - \varepsilon$$

[2] That is, the set $\{g > c\}$ is analytic for every $c \in \mathbf{R}$.

We also say that the strategy σ_ε uniformly guarantees $w - \varepsilon$. Similarly, player 2 uniformly guarantees w if for every $s \in \Omega$ and every $\varepsilon > 0$ there exists $\tau_\varepsilon \in \mathcal{T}$ and $N \in \mathbf{N}$ such that

$$\forall n \geq N, \forall \sigma \in \mathcal{S}, \gamma_n(s, \sigma, \tau_\varepsilon) \leq w(s) + \varepsilon$$

Definition 4 $v_\infty : \Omega \to \mathbf{R}$ *is the* uniform value *of the game if both players uniformly guarantee* v_∞.

A strategy that uniformly guarantees $v_\infty - \varepsilon$ is called uniform ε-optimal. We point out that our definition is weaker than the definition in Mertens, Sorin and Zamir (1994), in that we allow N to depend on the initial state s.

By dominated convergence, $\lim_n \gamma_n(s, \sigma, \tau) = \gamma(s, \sigma, \tau)$. Therefore, if the uniform value exists, it coincides with the limsup value.

The value of the n-stage game, that is the game with payoff function $\gamma_n(s, \sigma, \tau)$, is denoted by v_n.

For every $\lambda \in (0, 1)$ and every triplet (s, σ, τ), let

$$\gamma_\lambda(s, \sigma, \tau) = \mathbf{E}_{s,\sigma,\tau} \left[\lambda \sum_{n=1}^{\infty} (1 - \lambda)^{n-1} g_n \right] = \mathbf{E}_{s,\sigma,\tau} \left[(1 - \lambda)^{t-1} g(s_t) 1_{t<+\infty} \right]$$

denote the λ-discounted evaluation of payoffs.

Definition 5 *Let* $\lambda \in (0, 1)$. $v_\lambda : \Omega \to \mathbf{R}$ *is the* λ-discounted value *if*

$$v_\lambda(s) = \inf_{\tau \in \mathcal{T}} \sup_{\sigma \in \mathcal{S}} \gamma_\lambda(s, \sigma, \tau) = \sup_{\sigma \in \mathcal{S}} \inf_{\tau \in \mathcal{T}} \gamma_\lambda(s, \sigma, \tau).$$

Existence of the discounted value and the n-stage value was proved in a general setup (see, e.g., Nowak (1984a,b, 1985) or Mertens, Sorin and Zamir (1994) Proposition VII.1.4).

Theorem 6 *If Ω is Borel, A and B are compact, g is measurable, and for every $S' \subseteq \Omega$, the function $q(S' \mid \omega, a, b)$ is measurable and continuous over $A \times B$ for each fixed ω, then v_n and v_λ exist. Moreover, for every $s \in S$, $v_\lambda(s) = (1 - \lambda)\mathrm{val}G_s(v_\lambda)$ and v_λ is measurable.*

By the definition of the uniform value, whenever it exists we have $v_\infty = \lim_{n\to\infty} v_n$. One can also show that in that case $v_\infty = \lim_{\lambda\to 0} v_\lambda$.

2.3 Known results

In this subsection we review conditions under which the uniform value is known to exist.

The first result, which was proved for general stochastic games, was given by Mertens and Neyman (1981).

Theorem 7 (Mertens and Neyman, 1981) *If the function $\lambda \to v_\lambda$ has bounded variation, then v_∞ exists.*

In Rosenberg and Vieille (1998), recursive games with incomplete information are studied. Their result implies the next theorem.

Theorem 8 (Rosenberg and Vieille, 1998) *If v_λ converge uniformly to a limit, then v_∞ exists.*

Finally, when the transition to states in S is independent of the actions of the players, one can drop the requirement on v_λ. Formally, the next result is a by-product of the last section of Rosenberg et al (1999).

Theorem 9 (Rosenberg, Solan and Vieille, 1999) *If (i) A and B are finite, and (ii) for every $s \in S$ and every $S' \subseteq S$ we have*

$$q(S'|s,a,b)q(S|s,a',b') = q(S' \mid s,a',b')q(S \mid s,a,b) \quad \forall (a,b),(a',b') \in A \times B$$

then v_∞ exists.

2.4 Results and example

Our main result gives a condition on the lim sup value that ensures the existence of v_∞.

Theorem 10 *Assume that Ω is countable and A and B are finite. If the set $\{s \in S, v(s) \leq a\}$ is finite for some $a > 0$, the uniform value exists.*

If the function g happens to be non-negative, then the only condition that is required is that B is finite.

Theorem 11 *If (i) $g \geq 0$, (ii) v_λ exists for every $\lambda \in (0,1)$, and (iii) B is finite, then the uniform value exists.*

One can replace the condition that B is finite by the following weaker condition.

Theorem 12 *If (i) $g \geq 0$, (ii) v_λ exists for every $\lambda \in (0,1)$, and (iii) For every $\varepsilon > 0$ there exists a stationary strategy $\mathbf{y}^\varepsilon = (y_s^\varepsilon)$ for player 2, such that*

$$\int w(s')dq(s' \mid s, a, y_s^\varepsilon) \leq w(s) + \varepsilon \qquad \forall a \in A, s \in S,$$

where $w(s) = \limsup_{\lambda \to 0} v_\lambda(s)$, then the uniform value exists.

As we will see, if $g \geq 0$ then for every $s \in S$, $v_\lambda(s)$ increases when λ decreases. Thus, $w(s) = \sup_{\lambda \in (0,1)} v_\lambda(s)$.

For each $s \in S$, y_s^ε is an ε-optimal strategy for player 2 in the one shot game with payoff $\int w(s')dq(s' \mid s, \cdot, \cdot)$. If B is finite, then any limit of discounted ε-optimal strategies in this game (as the discount factor and ε go to 0), satisfies (iii).

We give now an example that shows that Theorem 10 is in some respect tight. The example is of a game for which the set $\{s \in S, v(s) \leq 0\}$ is empty, but which has no uniform value.

Example 13

The state space is $S = \{1, 2, 3, \dots\} \cup \{t_1, t_{-1}, t_2\} \cup \{1^\star, 2^\star, \dots\}$. Player 1 has a single action, and player 2 has two actions, $\{D, R\}$. Since player 1 is degenerate, we omit him from the notations. States t_1, t_{-1}, t_2 are absorbing, with absorbing payoff 1,-1,2 respectively. The transition function is given by:

$$\begin{array}{ll}
P((k-1)^\star | k^\star, \cdot) = 1 & k > 1 \\
P(t_2 | 1^\star, \cdot) = 1 & \\
P(t_{-1} | k, D) = 1/2 & k \geq 1 \\
P(k^\star | k, D) = 1/2 & k \geq 1 \\
P(t_1 | k, R) = 1/2^{k+4} & k \geq 1 \\
P(k+1 | k, R) = 1 - 1/2^{k+4} & k \geq 1
\end{array}$$

Graphically, the game looks as follows:

If the game reaches a state $k^\star$, then after k stages it reaches state t_2 with probability 1. Hence $v(k^\star) = 2$ for every k. Since $\sum_{k=1}^{\infty} 1/2^{k+4} = 1/16$, it follows that if the initial state is k, then the optimal strategy for player 2 is to play R forever. Hence $v(k) = 1/2^{k+3}$. We shall now see that $\limsup_{n\to\infty} v_n(1) \leq -1/8$. Indeed, for a given $n \in \mathbf{N}$, consider the following strategy of player 2: *play R for the first $n/2$ stages, and then play L once* (afterwards, transitions are independent of the actions played by player 2). It is easy to verify that

$$\gamma_n(1) \leq \frac{1}{16} - \frac{3}{4} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16} - \frac{3}{16} = -\frac{1}{8}.$$

3 Construction of an ε-Optimal Strategy

In this section we consider a recursive game that satisfies the following assumptions:

A.1 The limsup value v exists, and the function $s \mapsto v(s)$ is measurable.

A.2 There exists a stationary strategy $\mathbf{x} = (x_s)$ for player 1 such that for every $s \in S$

$$v(s) \geq \int_S v(s')dq(t' \mid s, x_s, b) \qquad \forall b \in B.$$

A.3 There exists a stationary strategy $\mathbf{y} = (y_s)$ for player 2 such that for every $s \in S$

$$v(s) \leq \int_S v(s')dq(t' \mid s, a, y_s) \qquad \forall a \in A.$$

Thus, for every $s \in S$ the pair of strategies (x_s, y_s) is optimal in the one-shot game with continuation payoff v.

Note that conditions **A**.1-**A**.3 hold under the assumptions of Theorem 10.

We are going to construct a specific ε-optimal strategy for player 1. By symmetry, a similar construction would yield an ε-optimal strategy for player 2. In the next section, we shall argue that under the assumptions of Theorem 10, these strategies are indeed uniformly ε-optimal.

Section 3.1 gives some results on the concatenation of ε-optimal strategies. Section 3.2 deals with recursive games in which the limsup value is bounded away from zero. Section 3.3 deals with general recursive games.

3.1 Preliminary results

We first define terminating strategies

Definition 14 *We say that $\sigma \in \mathcal{S}$ is* terminating *if, for every initial state s and every $\tau \in \mathcal{T}$, $t < +\infty$, $\mathbf{P}_{s,\sigma,\tau}$-a.s.*

For every strategy $\sigma \in \mathcal{S}$ and every finite history $h_n = (s_1, a_1, b_1, ..., s_n) \in H_n$, we denote by σ^{h_n} the strategy induced by σ in the subgame defined by h_n : for every finite history h, $\sigma^{h_n}(h) = \sigma(s_1, ..., s_{n-1}, a_{n-1}, b_{n-1}, h)$.

Let $\sigma_1, \sigma_2 \in \mathcal{S}$, and u be a stopping time, with values in $\mathbf{N} \cup \{+\infty\}$. We define the strategy $\sigma_1 u \sigma_2$ as: play σ_1 up to u, then switch to σ_2 (and forget the history up to u). Formally, for every $n \in \mathbf{N}$ and every $h_n = (s_1, a_1, b_1, \ldots, s_n)$, $(\sigma_1 u \sigma_2)(h_n) = \sigma_1(h_n)$ if $u > n$, and $(\sigma_1 u \sigma_2)(h_n) = \sigma_2(h_n^u)$ if $u \leq n$, where h_n^u stands for the finite history $(s_u, a_u, b_u, \ldots, s_n)$.

Similarly, if $0 < u_1 < u_2 < u_3 < \cdots$ are stopping times and $\sigma_1, \sigma_2, \ldots$ are strategies, we define the strategy $\sigma = \sigma_1 u_1 \sigma_2 u_2 \cdots$ as follows. $\sigma(h_n) = \sigma_1(h_n)$ if $n < u_1$, and $\sigma(h_n) = \sigma_m(h_n^{u_{m-1}})$ if $u_{m-1} \leq n < u_m$.

We start by checking that the concatenation of two ε-optimal strategies is 2ε-optimal (Corollary 16).

Lemma 15 *Let σ_1 be an ε-optimal strategy and $s \in S$. Let u be a stopping time. Assume that for each τ, $u < +\infty$, $\mathbf{P}_{s,\sigma_1,\tau}$-a.s. One has*

$$\forall \tau \in \mathcal{T}, \ \mathbf{E}_{s,\sigma_1,\tau}[v(s_u)] \geq v(s) - \varepsilon.$$

Proof. Otherwise, $\mathbf{E}_{s,\sigma_1,\tau}[v(s_u)] < v(s) - \varepsilon - \eta$ for some $\tau \in \mathcal{T}$ and $\eta > 0$. Let τ_1 be an η-optimal strategy of player 2. One has

$$\gamma_s(\sigma_1, \tau u \tau_1) = \mathbf{E}_{s,\sigma_1,\tau}[\gamma_{s_u}(\sigma^{h_u}, \tau_1)] \leq \mathbf{E}_{s,\sigma_1,\tau}[v(s_u) + \eta] < v(s) - \varepsilon,$$

a contradiction. ■

Corollary 16 *Let σ_1 and σ_2 be respectively ε_1- and ε_2- optimal strategies of player 1. Let u be a stopping time with $\mathbf{P}_{s,\sigma_1,\tau}$-a.s. finite values, for every τ. Then $\sigma_1 u \sigma_2$ is $\varepsilon_1 + \varepsilon_2$-optimal.*

Proof. Observe that

$$\begin{aligned}
\mathbf{E}_{s,\sigma_1 u \sigma_2,\tau}\left[g(s_t)1_{t<+\infty}\right] &= \mathbf{E}_{s,\sigma_1 u \sigma_2,\tau}\left[\mathbf{E}_{s,\sigma_1 u \sigma_2,\tau}\left[g(s_t)1_{t<+\infty}|\mathcal{H}_u\right]\right] \\
&= \mathbf{E}_{s,\sigma_1,\tau}\left[\mathbf{E}_{s,\sigma_2,\tau^{h_u}}\left[g(s_t)1_{t<+\infty}\right]\right] \\
&\geq \mathbf{E}_{s,\sigma_1,\tau}\left[v(s_u) - \varepsilon_2\right] \\
&\geq v(s) - \varepsilon_1 - \varepsilon_2,
\end{aligned}$$

where the first inequality uses the ε_2-optimality of σ_2, and the second one uses Lemma 15. ■

3.2 Recursive games with limsup value bounded away from 0

In this section we prove the following result.

Proposition 17 *Let* Γ *be a recursive game such that, for some* $a > 0$, $v(s) \geq a$, *for each* $s \in S$. *Then player 1 has a terminating* ε*-optimal strategy.*

Observe that a recursive game that satisfies the condition of Proposition 17 need not be positive, and a positive recursive game need not satisfy the condition of Proposition 17.

Fix a recursive game that satisfies the condition of Proposition 17, and an $\varepsilon > 0$.

The proof of Proposition 17 goes as follows. For every m we choose an $\varepsilon/2^{m+1}$-optimal strategy σ_m for player 1. We define a strategy $\bar{\sigma}$ by a suitable concatenation of the σ_m's. We then prove that $\bar{\sigma}$ is terminating and ε-optimal.

Given $\sigma \in \mathcal{S}$, we define the stopping time

$$t_\sigma = \inf\{n \geq 1, \inf_\tau \mathbf{P}_{s,\sigma,\tau}(t < +\infty | \mathcal{H}_n) < \varepsilon\}.$$

Equivalently, $t_\sigma(h_\infty) = \inf\{n \geq 1, \mathbf{P}_{s_n,\sigma^{h_n},\tau}(t < +\infty) < \varepsilon \text{ for some } \tau\}$. It is the first stage after which the residual probability of termination in finite time is very small for some strategy of player 2.

Lemma 18 *For every* σ, τ, $\min(t, t_\sigma)$ *is* $\mathbf{P}_{s,\sigma,\tau}$*-finite.*

Proof. Fix $\tau \in \mathcal{T}$ and set $t_1 = \inf\{n \geq 1, \mathbf{P}_{s,\sigma,\tau}(t < +\infty | \mathcal{H}_n) < \varepsilon\}$. Clearly, $t_1 \geq t_\sigma$, $\mathbf{P}_{s,\sigma,\tau}$-a.s., so it suffices to prove that $\min(t, t_1) < +\infty$, $\mathbf{P}_{s,\sigma,\tau}$-a.s.

Observe that the sequence $(\mathbf{P}_{s,\sigma,\tau}(t < +\infty | \mathcal{H}_n))_n$ is a martingale under $\mathbf{P}_{s,\sigma,\tau}$, which converges $\mathbf{P}_{s,\sigma,\tau}$-a.s. to $1_{t<+\infty}$, hence to 0 on the event $\{t = +\infty\}$. Therefore $t_1 < +\infty$ on the event $\{t = +\infty\}$. ■

We need the following observation.

Lemma 19 *Let $\eta > 0$, and $\sigma \in \mathcal{S}$ be an η-optimal strategy. For every $s \in S$,*

$$\inf_{\tau} \mathbf{P}_{s,\sigma,\tau}(t < +\infty) \geq a - \eta$$

Proof. Fix $\tau \in \mathcal{T}$. Since the payoff function g is bounded by 1, one has

$$a - \eta \leq v(s) - \eta \leq \gamma(s, \sigma, \tau) = \mathbf{E}_{s,\sigma,\tau}(g(s_t) 1_{t<+\infty}) \leq \mathbf{P}_{s,\sigma,\tau}(t < +\infty),$$

as desired. ■

We obtain a terminating ε-optimal strategy $\overline{\sigma}$ of player 1 by concatenation of $\varepsilon/2^n$-optimal strategies. For every $m \geq 1$, choose an $\varepsilon/2^{m+1}$-optimal strategy σ_m. We define inductively a sequence (σ^m) of strategies as follows. Set $\sigma^1 = \sigma_1$. Assume that σ^m is defined. We write t_m instead of t_{σ^m}. Set

$$\sigma^{m+1} = \sigma^m t_m \sigma_{m+1}.$$

In words, $\overline{\sigma}$ plays σ_1 until the residual probability of termination in finite time becomes very small. It then plays σ_2 until the residual probability becomes again very small, and so on up to infinity.

Note that $t_{m+1} > t_m$ on the event $\{t_m < +\infty\}$, and in particular $t_m \geq m$. Moreover, by Lemma 18, $\min\{t, t_m\}$ is $\mathbf{P}_{s,\sigma_m,\tau}$-a.s. finite for every $\tau \in \mathcal{T}$, and therefore $\min\{t, t_m\}$ is $\mathbf{P}_{s,\bar{\sigma},\tau}$-a.s. finite as well. Hence σ^{m+1} is well-defined and coincides with σ^m on H_m. We let $\overline{\sigma}$ be defined by $\bar{\sigma} = \sigma^m$ on H_m.

Lemma 20 *$\overline{\sigma}$ is terminating.*

Proof. Let $\tau \in \mathcal{T}$ be arbitrary. For every $m \in \mathbf{N}$, we have by Lemma 19 and the definition of t_m

$$\mathbf{P}_{s,\sigma_m,\tau}(t \leq t_m) \geq \mathbf{P}_{s,\sigma_m\tau}(t < +\infty) - \mathbf{P}_{s,\sigma_m\tau}(t_m < t < +\infty) \geq a - \varepsilon/2^{m+1} - \varepsilon.$$

As long as $\varepsilon < a$ the result follows by the definition of $\bar{\sigma}$. ■

Lemma 21 *$\overline{\sigma}$ is ε-optimal.*

Proof. By Corollary 16 and since $\min\{t, t_m\}$ is $\mathbf{P}_{s,\bar{\sigma},\tau}$-a.s. finite for every fixed $\tau \in \mathcal{T}$, $\sigma_1 t_1 \sigma_2 \cdots t_{m-1} \sigma_m$ is $\varepsilon/2 + \varepsilon/4 + \cdots + \varepsilon/2^m$-optimal. Since $\overline{\sigma}$ is terminating, $\gamma(s, \overline{\sigma}, \tau) = \lim_{m\to\infty} \gamma(s, \sigma^m, \tau)$ for every τ. In particular, $\overline{\sigma}$ is ε-optimal. ■

We show on an example that the existence of a terminating strategy relies crucially on the fact that v is *uniformly* bounded away from 0.

Example 22 *Consider the following game, with dummy players (a Markov chain).* $T = \{t_1\}$, $S = \mathbf{N}$, *and* $g(t_1) = 1$. *For every* $n \in \mathbf{N}$, $q(t_1|n) = \frac{1}{2^{n+2}}$, *and* $q(n+1|n) = 1 - q(t_1|n)$. *One has* $v(n) = \mathbf{P}_n(t < +\infty) = 1/2^{n+1} > 0$ *for every* $n \in \mathbf{N}$. *However, whatever be the initial state, the probability that the game does not terminate in finite time is strictly positive.*

3.3 The general case

In this section, we let Γ be a general recursive game that satisfies assumptions **A**.1-**A**.3. Our goal is to construct ε-optimal strategies that are not necessarily uniform ε-optimal. For every $\varepsilon > 0$ let $\Gamma(\varepsilon)$ be the game with (i) state space $\Omega_\varepsilon = S_\varepsilon \cup T_\varepsilon$, where $S_\varepsilon = \{s \in S, v(s) \geq 2\varepsilon\}$ and $T_\varepsilon = T \cup \{s \in S, v(s) < 2\varepsilon\}$, (ii) action spaces A and B, and (iii) payoff function that coincides with g on T and defined as $g_\varepsilon(s) = v(s)$ for $s \in T_\varepsilon$. (vi) Transitions on S_ε are unchanged.

Intuitively, states with a value below 2ε are replaced by absorbing states with payoff which is equal to their limsup value.

Denote by $\tilde{v}$ the lim sup value of $\Gamma(\varepsilon)$, and by $\tilde{v}_\lambda$ the λ-discounted value of $\Gamma(\varepsilon)$. As we show below, $\tilde{v} = v$, but v_λ and $\tilde{v}_\lambda$ may differ. In particular, it will follow that for every $s \in S_\varepsilon$, $v(s) \geq 2\varepsilon$. Hence we can apply the results from section 3.2 to $\Gamma(\varepsilon)$. From now on we fix $\varepsilon > 0$, and denote by $\tilde{\gamma}$ the payoff in $\Gamma(\varepsilon)$.

Let (α_n) be a bounded process on $(H_\infty, (\mathcal{H}_n), \mathbf{P})$, and $u \leq \overline{u}$ two stopping times with values in $\mathbf{N} \cup \{+\infty\}$. We say that (α_n) is a submartingale between u and $\overline{u}$ if, for each n, one has $\alpha_n \leq \mathbf{E}(\alpha_{n+1}|\mathcal{H}_n)$ on the event $\{u \leq n < \overline{u}\}$. If (α_n) is a submartingale between u and $\overline{u}$, and $\tilde{u} \leq \overline{u}$ is another stopping time with $\mathbf{P}$-a.s. finite values, $\mathbf{E}[\alpha_{\tilde{u}}|\mathcal{H}_u] \geq \alpha_u$ on the event $\{u \leq \tilde{u}\}$. We say that (α_n) is a submartingale up to $\overline{u}$ if it is a submartingale between 0 and $\overline{u}$.

Lemma 23 $\tilde{v} = v$.

Proof. Let $\sigma \in \mathcal{S}$ be a terminating δ-optimal strategy in $\Gamma(\varepsilon)$. Such a strategy exists by the previous section. Set $\tilde{t} = \inf\{n \geq 1, s_n \in T_\varepsilon\}$. By Lemma 15,

$$\mathbf{E}_{s,\sigma,\tau}\left[v(s_{\tilde{t}})\right] \geq v(s) - \delta \qquad \forall \tau \in \mathcal{T}.$$

Since the left-hand side coincides with $\tilde{\gamma}(s,\sigma,\tau)$, this implies $\tilde{v}(s) \geq v(s)-\delta$. Since δ is arbitrary, this yields $\tilde{v}(s) \geq v(s)$.

Fix $s \in S_\varepsilon$. By assumption **A**.3, for each σ, the sequence $(v(s_n))$ is a (bounded) supermartingale under $\mathbf{P}_{s,\sigma,\mathbf{y}}$. Set $v_\infty = \lim_{n\to\infty} v(s_{\min(n,\tilde{t})})$. By the supermartingale property, $\mathbf{E}_{s,\sigma,\mathbf{y}}[v_\infty] \leq v(s)$. By definition of S_ε, $v_\infty \geq \varepsilon > 0$ on the event $\tilde{t} = +\infty$. Since $v_\infty = v(s_{\tilde{t}})$ on the event $\tilde{t} < +\infty$, one obtains $\tilde{\gamma}(s,\sigma,\mathbf{y}) \leq v(s)$. Hence $\tilde{v}(s) \leq v(s)$. ∎

In particular, $\tilde{v}(s) \geq 2\varepsilon$ for every $s \in S_\varepsilon$. For $s \in S_\varepsilon$, we let $\sigma^*(s)$ denote a terminating ε^2-optimal strategy for the initial state s, in the game $\Gamma(\varepsilon)$. Thus $\sigma^*(s)$ guarantees $v(s) - \varepsilon^2$ in $\Gamma(\varepsilon)$.

The strategy we define now has some features in common with strategies defined in Rosenberg and Vieille (1998). Intuitively, it may be thought of as: play $\mathbf{x}$ whenever the current state belongs to T_ε; whenever the play enters S_ε, say in state s, switch to $\sigma^*(s)$ until the play leaves S_ε. As argued in Rosenberg and Vieille (1998), this might involve too many switches. We refine this idea as follows.

Set $u_1 = 1$, $u_2 = \inf\{n \geq 1, v(s_n) \geq 2\varepsilon\}$. For $p \in \mathbf{N}$, set $u_{2p+1} = \inf\{n \geq u_{2p}, v(s_n) \leq \varepsilon\}$, and $u_{2p+2} = \inf\{n \geq u_{2p+1}, v(s_n) \geq 2\varepsilon\}$. Graphically, one can look at the sequence of real numbers $v(s_n)$. The stopping times u_p (for p even) tell us when this sequence jumps above 2ε, and the stopping times u_p (for p odd) tell us when this sequence jumps below ε.

Define $\overline{\sigma} = \mathbf{x}u_2\sigma^*u_3\mathbf{x}u_4\sigma^*u_5\cdots$ as: play $\mathbf{x}$ from u_1 to u_2, play $\sigma^*(s_{u_2})$ from u_2 to u_3, $\mathbf{x}$ from u_3 to u_4, $\sigma^*(s_{u_4})$ from u_4 to u_5, and so on. We prove below that $\overline{\sigma}$ is ε-optimal. In the next section, we show that it uniformly guarantees $v - \varepsilon$ if $S \setminus S_\varepsilon$ is finite.

In Lemma 24, we prove a submartingale property for the sequence $(v(s_{\min(t,u_p)})_p$. Fix $\tau \in \mathcal{T}$ and set for simplicity $\mathbf{P} = \mathbf{P}_{s,\overline{\sigma},\tau}$, $\mathbf{E} = \mathbf{E}_{s,\overline{\sigma},\tau}$.

Lemma 24 *For every* $p \in \mathbf{N}$*,*

$$\mathbf{E}[v(s_{\min(t,u_{p+1})})|\mathcal{H}_{\min(t,u_p)}] \geq v(s_{\min(t,u_p)}) - \varepsilon^2 1_{t>u_p}.$$

on the event $\min(t,u_p) < +\infty$.

Observe first that by Lemma 18, $\mathbf{P}(u_{2p} < t, \min(t,u_{2p+1}) = +\infty) = 0$. Observe also that since $\mathbf{x}$ is optimal in the local game, $(v(s_n))$ is a submartingale between $\min(t,u_{2p+1})$ and $\min(t,u_{2p+2}))$ for every p. Therefore, on the event $\{u_{2p+1} < +\infty = t = u_{2p+2}\}$, $v(s_\infty) = \lim_n v(s_n)$ exists. Thus, the conditional expectation on the left-hand side is meaningful.

Proof. For even p, on the event $t > u_p$.

$$\mathbf{E}[v(s_{\min(t,u_{p+1})})|\mathcal{H}_{\min(t,u_p)}] = \mathbf{E}_{s_{u_p},\sigma^*(s_{u_p}),\tau^{h_{u_p}}}[v(\tilde{t})] \geq v(s_{u_p}) - \varepsilon^2$$

Let now p be odd. Between $\min(t,u_p)$ and $\min(t,u_{p+1})$, $(v(s_n))_n$ is a submartingale. Therefore $\mathbf{E}[v(s_{\min(t,u_{p+1})})|\mathcal{H}_{\min(t,u_p)}] \geq v(s_{\min(t,u_p)})$ by the sampling theorem. ■

Since $v(s_{u_{2p}}) \geq 2\varepsilon$ and $v(s_{u_{2p+1}}) \leq \varepsilon$, $N = \sup\{p \geq 1, u_{2p} < +\infty\}$ is the number of upcrossings of the interval $[\varepsilon, 2\varepsilon]$ by the sequence $(v(s_n))_n$. An easy adaptation of the standard result on upcrossings (see Rosenberg and Vieille (1998), Proposition 3) gives

$$\mathbf{E}[N] \leq 1/(\varepsilon - \varepsilon^2). \tag{1}$$

Set $\tilde{v}_p = v(s_{\min(t,u_p)}) + \varepsilon^2\tilde{p}$, where $\tilde{p} = p$ if $t > u_p$, and $\tilde{p} = \sup\{k \geq 1, u_k < t\}$ otherwise. By Lemma 24, $(\tilde{v}_p)$ is a submartingale. Moreover, by (1) and since $|v(s_n)| \leq 1$, one has $\sup_p |\tilde{v}_p| \in L^1$. Therefore, $(\tilde{v}_p)$ converges $\mathbf{P}$-a.s. and therefore $(v(s_{\min(t,u_p)}))$ converges as well. Denote $v(s_\infty) = \lim_{p\to\infty} v(s_{\min(t,u_p)})$.

Once again, by definition of σ^*, one has $\mathbf{P}$-a.s. $u_{2p+1} < +\infty$ on the event $u_{2p} < +\infty = t$. Therefore, on the event $t = +\infty$ one has $v(s_\infty) = v(s_{u_{2p+1}}) \leq \varepsilon$ for some $p \in \mathbf{N}$.

Proposition 25 *$\overline{\sigma}$ guarantees v.*

Proof. Recall that $N = \sup\{p \geq 1, u_{2p} < +\infty\}$. One has $\mathbf{E}[\tilde{v}_\infty] \geq \tilde{v}_1$, which reads $\mathbf{E}[v(s_\infty)] \geq v(s) - \varepsilon^2\mathbf{E}[N]$. On the event $\{t < +\infty\}$, $v(s_\infty) = g(s_t)$. On the event $\{t = +\infty\}$, $v(s_\infty) \leq \varepsilon$ $\mathbf{P}$-a.s. Thus,

$$\mathbf{E}[g(s_t)1_{t<+\infty}] \geq \mathbf{E}[v(s_\infty)] - 2\varepsilon \geq v_1 - 2\varepsilon - \varepsilon\frac{1}{1-\varepsilon}.$$

■

4 Uniform Optimality

We prove in this section Theorem 10. Thus, we assume that S is countable, A and B are finite, and that for some $a > 0$, the set $\{s \in S, v(s) \leq a\}$ is finite. Fix $\varepsilon \in (0, a/2)$ such that $2/\varepsilon^2$ is an integer for the rest of the section. The set $S \backslash S_\varepsilon$ defined in the previous section is also finite. We investigate the properties of the strategy $\bar{\sigma}$ that has been defined in the previous section.

We prove (Proposition 26) that under a terminating strategy, termination occurs in fact in *bounded* time.

Proposition 26 *Let $s \in S$ be fixed and $\sigma \in \mathcal{S}$ be a terminating strategy. For every $\eta > 0$, there exists $N \in \mathbf{N}$, such that*

$$\forall \tau \in \mathcal{T}, \mathbf{P}_{s,\sigma,\tau}(t \leq N) > 1 - \eta.$$

Proof. Assume that the result does not hold for some $\eta > 0$. Then, for each $N \in \mathbf{N}$, there exists a *pure* strategy τ_N such that $\mathbf{P}_{s,\sigma,\tau_N}(t \leq N) \leq 1 - \eta$. Obviously,

$$\forall N' \geq N,\ \mathbf{P}_{s,\sigma,\tau_{N'}}(t \leq N) \leq 1 - \eta. \tag{2}$$

Since A and B are finite and S is countable, there exists a *finite* subset Ω_N of Ω such that, for every τ,

$$\mathbf{P}_{s,\sigma_N,\tau}(\forall n \leq N,\ s_n \in \Omega_N) \geq 1 - \eta/2.$$

Clearly, one may choose the sequence $(\Omega_N)_N$ to be non-decreasing. For each N, we partition the set of pure strategies of player 2 as follows: τ_1 and τ_2 in $\mathcal{T}$ are considered equivalent if they coincide on every history of length at most $N-1$ which visits only states in Ω_N: $\tau_1 \simeq_N \tau_2$ if, for every $n \leq N$, and every $h_n = (s_1, a_1, b_1, \dots, s_n) \in H_n$, one has $\tau_1(h_n) = \tau_2(h_n)$ as soon as $s_0, s_1, \dots, s_n \in \Omega_N$. Since Ω_N, A and B are finite, the number of equivalence classes for the relation $\simeq_N$ is finite. Notice that, if $\tau_1 \simeq_N \tau_2$, one has

$$|\mathbf{P}_{s,\sigma,\tau_1}(t \leq N) - \mathbf{P}_{s,\sigma,\tau_2}(t \leq N)| \leq \mathbf{P}_{s,\sigma,\tau_1}(\exists n \leq N, s_n \notin \Omega_N) \leq \eta/2 \tag{3}$$

Since (Ω_N) is non-decreasing, the partition into equivalence classes for $\simeq_{N+1}$ refines the partition obtained for $\simeq_N$. Therefore, one can construct a decreasing sequence $(e_N)_N$ of equivalence classes for $(\simeq_N)$,[3] such that for each N, e_N contains infinitely many of the strategies $(\tau_p)_{p \geq N}$. By this procedure, one gets a pure strategy τ such that for every N there exists $N' \geq N$ with $\tau \simeq_N \tau_{N'}$. From (2) and (3) one obtains $\mathbf{P}_{s,\sigma,\tau}(t \leq N) \leq 1 - \eta/2$ for every N. This contradicts the fact that σ is terminating. ■

[3] That is, each e_N is an equivalence class for $\simeq_N$.

4.1 Player 1 can uniformly guarantee v

We show that, in this case, the strategy $\overline{\sigma}$ that we defined in the previous section is uniformly ε-optimal. We first sketch the idea. Since A, B and $S \setminus S_\varepsilon$ are finite, there exists a finite subset S_1 of S_ε such that $\mathbf{P}_{s,\sigma,\tau}(s_{u_2} \notin S_1, u_2 < +\infty) \leq \varepsilon^3$, for every $s \in S \setminus S_\varepsilon$, every $\sigma \in \mathcal{S}$ and every $\tau \in \mathcal{T}$. For each $s \in S_\varepsilon$ there exists $N_s \in \mathbf{N}$ such that $\mathbf{P}_{s,\sigma^*(s),\tau}(\tilde{t} > N_s) \leq \varepsilon^3$ for every $\tau \in \mathcal{T}$. Since S_1 is finite, $N_1 = \max_{S_1} N_s$ is finite.

This is used in the following way. Define an excursion above 2ε (an *excursion* in short) as the play between u_{2p} and u_{2p+1}, for any p such that $u_{2p} < +\infty$ (recall that these are stages where player 1 follows $\sigma^*(s_{2p})$). One of the arguments of the previous section was that the expected number of excursions is at most $2/\varepsilon$. Therefore, the probability that the total number of excursions during the play exceeds $1/\varepsilon^2$ is small. By definition of S_1, the probability that the number of excursions does not exceed $1/\varepsilon^2$ *and* each of the excursions starts in S_1 is close to 1. Now, given an excursion starts from S_1, the probability that it lasts more than N_1 stages is small. Therefore, the probability that the total number of excursions does not exceed $1/\varepsilon^2$ *and* that no excursion exceeds N_1 stages is close to 1. This implies that, provided n is large, the expected frequency of stages which belong to an excursion is small. This is a crucial observation which allows to compare the average of $\mathbf{E}[v(s_n)]$ over the first n stages to the expected average payoff received up to stage n.

We put this in formal terms. For $n \in \mathbf{N}$, define

$$A_n = \{u_{2p} \leq n < \min(t, u_{2p+1}), u_{2p} < t, \text{ for some } p\} \subseteq H_\infty.$$

These are all infinite plays where stage n is in an excursion.

Lemma 27 *For every $\tau \in \mathcal{T}$ and every $n \geq N_1/\varepsilon^3$,*

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{P}_{s,\overline{\sigma},\tau}(A_k) \leq 5\varepsilon.$$

Proof. Since $\mathbf{E}[N] \leq 2/\varepsilon$, one has $\mathbf{P}(N \geq 1/\varepsilon^2) \leq 2\varepsilon$. By definition of S_1,

$$\mathbf{P}(u_{2/\varepsilon^2} < +\infty, s_{u_{2k}} \notin S_1 \text{ for some } k \leq 1/\varepsilon^2) \leq 1 - (1-\varepsilon^3)^{1/\varepsilon^2} \leq 2\varepsilon.$$

Denote by $d_p = \min(u_{2p+1}, t) - u_{2p}$ if $u_{2p} < t$, $d_p = 0$ otherwise, the length of the p-th excursion. Any excursion which starts in S_1 does not exceed N_1 in length, with high probability:

$$\mathbf{P}_{s,\overline{\sigma},\tau}(d_p > N_1 | s_{u_{2p}} \in S_1, u_{2p} < t) \leq \varepsilon^3.$$

Therefore,

$$\mathbf{P}_{s,\overline{\sigma},\tau}(u_{2/\varepsilon^2} < t, d_p > N_1 \text{ for some } p \leq 1/\varepsilon^2) \leq 4\varepsilon.$$

Denote $E = \{u_{2/\varepsilon^2} < t, d_p > N_1 \text{ for some } p \leq 1/\varepsilon^2\}$, and by $D = \sum_{p=0}^{+\infty} d_p > 0$ the total length of excursions. On the complement E^c of E, $D \leq N_1 \times 1/\varepsilon^2$. Thus,

$$\mathbf{E}_{s,\overline{\sigma},\tau}[1_{E^c} D] = \sum_{k=1}^{+\infty} \mathbf{P}_{s,\overline{\sigma},\tau}(A_k \cap E^c) \leq N_1/\varepsilon^2.$$

One deduces that $\frac{1}{n}\sum_{k=1}^{n} \mathbf{P}_{s,\overline{\sigma},\tau}(A_k) \leq 6\varepsilon + \frac{N_1}{n\varepsilon^2}$, which yields the result. ■

Proposition 28 *The strategy $\overline{\sigma}$ uniformly guarantees $v - 16\varepsilon$.*

Proof. We rewrite the submartingale property of $(v(s_n))$ a bit differently. Our goal is to explicit an estimate of $\mathbf{E}[v(s_n)]$ in terms of $v(s)$ (see inequality (4) below). Fix $n_0 \geq 1$ and for $p \in \mathbf{N}$ set

$$X_p = v(s_{\min(t,u_p,n_0)})$$

Since $v(s_n)$ is a submartingale between $\min(t, u_{2p+1})$ and $\min(t, u_{2p+2})$, one has

$$\mathbf{E}\left[X_{2p+2} | \mathcal{H}_{\min(t,u_{2p+1},n_0)}\right] \geq X_{2p+1}$$

By construction of σ^* one has

$$\mathbf{E}\left[v(s_{\min(t,u_{2p+1})}) | \mathcal{H}_{\min(t,u_{2p})}\right] \geq v(s_{\min(t,u_{2p})}) - \varepsilon^2 \text{ if } u_{2p} < t.$$

If $u_{2p} < \min(n_0, t)$, X_{2p+1} coincides with $v(s_{\min(t,u_{2p+1})})$, except possibly if $u_{2p} < n_0 < \min(t, u_{2p+1})$. Therefore, if $u_{2p} < \min(t, n_0)$

$$\mathbf{E}\left[X_{2p+1} | \mathcal{H}_{\min(t,u_{2p},n_0)}\right] \geq X_{2p} - \varepsilon^2 \mathbf{1}_{u_{2p} < \min(t,n)} - 2\mathbf{P}(u_{2p} < n_0 < \min(t, u_{2p+1}) | \mathcal{H}_{\min(t,u_{2p},n_0)}),$$

and otherwise

$$\mathbf{E}\left[X_{2p+1} | \mathcal{H}_{\min(t,u_{2p},n_0)}\right] \geq X_{2p}.$$

By taking expectations and letting p go to infinity, these inequalities yield

$$\mathbf{E}\left[v(s_{n_0})\right] \geq v(s) - \varepsilon^2 \mathbf{E}[N] - 2\mathbf{P}(A_{n_0}) \tag{4}$$

Observe now that $g_n \geq v(s_n) - \varepsilon$, except on $\cup_k A_k$. One deduces that for every $n \geq N_1/\varepsilon^3$,

$$\gamma_n(s, \overline{\sigma}, \tau) \geq \mathbf{E}\left[\frac{1}{n}\sum_{k=1}^{n} v(s_k)\right] - \varepsilon - \frac{1}{n}\sum_{k=1}^{n}\mathbf{P}(A_k) \geq v(s) - 16\varepsilon,$$

where the second inequality uses (4) and Lemma 27. ■

Remark 29 *Notice that the finiteness of the set* $\{s \in S, v(s) \leq 2\varepsilon\}$ *is needed* only *to ensure the existence of the finite set* $S_1 \subseteq S_{2\varepsilon}$. *In particular, our proof works also if* S_ε *is finite and* $\{s \in S, v(s) \leq 2\varepsilon\}$ *countable.*

4.2 Player 2 can uniformly guarantee v

We prove here that player 2 can uniformly guarantee v. For $\varepsilon > 0$, define $\overline{S}_\varepsilon = \{s \in S, v(s) \leq -\varepsilon\}$, $\overline{T}_\varepsilon = T \cup \{s \in S, v(s) > -\varepsilon\}$, and denote $\overline{\Gamma}_\varepsilon$ the recursive game in which the set of absorbing states is $\overline{T}_\varepsilon$ and the payoff in $s \in \overline{T}_\varepsilon$ is $v(s)$. By Proposition 17, there is a terminating strategy of player 2, which uniformly guarantees $v + \varepsilon^2$ in the game $\overline{\Gamma}_\varepsilon$. By Remark 29 player 2 uniformly guarantees v.

It is not difficult to show that the stationary strategy defined as:

- Play $\mathbf{y}_\lambda$ on $\overline{S}_\varepsilon$ (an optimal strategy in the discounted game).
- Play $\mathbf{y}$ (limit of discounted optimal strategies) on $\overline{T}_\varepsilon$.

uniformly guarantees $v - \varepsilon$, provided λ is close enough to zero. This strategy was used by Thuijsman and Vrieze (1992) for the case $|S| < +\infty$.

5 Positive Recursive Games

In this section we prove Theorem 12.

Proof. First we note that for every fixed state $s \in S$ and every pair of strategies (σ, τ), $\gamma_\lambda(s, \sigma, \tau)$ is increasing in λ. Indeed,

$$\gamma_\lambda(s, \sigma, \tau) = \mathbf{E}_{s,\sigma,\tau}\left[\sum_{t=1}^{\infty}(1-\lambda)^{t-1} g(s_t) 1_{t<+\infty}\right],$$

and all terms are non-negative. We conclude that $v_\lambda(s)$ is increasing in λ. Define $w(s) = \sup_\lambda v_\lambda(s)$. We claim that w is the uniform value. We first check that player 1 can uniformly guarantee w.

Let $\varepsilon \in (0,1)$ and an initial state s be given. Choose $\lambda \in (0,1)$ such that $v_\lambda(s) \geq w(s) - \varepsilon/4$, and choose an $\varepsilon/4$-optimal strategy σ in the λ-discounted game. Let $N_1 = N_1(\lambda, \varepsilon)$ be sufficiently large such that $(1-\lambda)^{N_1} \leq \varepsilon/4$. Since g is bounded by 1,

$$\mathbf{E}_{s,\sigma,\tau}[1_{t<N_1} g_t] \geq \mathbf{E}_{s,\sigma,\tau}[1_{t<N_1}(1-\lambda)^{t-1} g_t] \geq \gamma_\lambda(s,\sigma,\tau) - \varepsilon/4.$$

Let $N_2 \geq 4N_1/\varepsilon$. Then for every $n \geq N_2$

$$\begin{aligned}
\gamma_n(s,\sigma,\tau) &\geq \mathbf{E}_{s,\sigma,\tau}[1_{t<N_1} g_t] - \varepsilon/4 \\
&\geq \gamma_\lambda(s,\sigma,\tau) - \varepsilon/2 \\
&\geq v_\lambda(s,\sigma,\tau) - 3\varepsilon/4 \\
&\geq w(s) - \varepsilon.
\end{aligned}$$

Therefore, player 1 can uniformly guarantee w.

We shall now construct a strategy for player 2 that uniformly guarantees $w(s) + 2\varepsilon$, given any initial state s. Denote for every $n \in \mathbf{N}$, $\varepsilon_n = \varepsilon/2^n$. Define a strategy τ^* for player 2 as follows. At stage n, play the mixed action $y_{s_n}^{\varepsilon_n}$. By (iii),

$$\gamma_n(s,\sigma,\tau^*) \leq \mathbf{E}_{s,\sigma,\tau^*}[g_n] \leq \mathbf{E}_{s,\sigma,\tau^*}[w(s_n)] \leq w(s_1) + 2\varepsilon,$$

as desired. ■

References

[1] Bewley T. and Kohlberg E. (1976) The Asymptotic Theory of Stochastic Games, *Math. Oper. Res.*, **1**, 197-208

[2] Everett H. (1957) Recursive Games, *Contributions to the theory of Games*, **3**, 47-78, Princeton N.J., Annals of Mathematical Studies, **39**, Princeton University Press.

[3] Lehrer E. and Sorin S. (1992) A Uniform Tauberian Theorem in Dynamic Programming, *Math. Oper. Res.*, **17**, 303-307

[4] Maitra A. and Sudderth W. (1993) Borel Stochastic Games with Limsup Payoff, *Ann. Prob.*, **21**, 861-885

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

[6] Mertens J.F., Sorin S. and Zamir S. (1994) Repeated Games, CORE Discussion Papers 9420, 9421, 9422

[7] Nowak A.S. (1984a) On Zero-Sum Stochastic Games with General State Space. I., *Probab. Math. Statis.*, **4**, 13-32

[8] Nowak A.S. (1984b) On Zero-Sum Stochastic Games with General State Space. II., *Probab. Math. Statis.*, **4**, 143-152

[9] Nowak A.S. (1985) Universally Measurable Strategies in Zero-Sum Stochastic Games, *Ann. Probab.*, **13**, 269-287

[10] Rosenberg D., Solan E. and Vieille N. (1999) Stopping Games with Randomized Strategies.

[11] Rosenberg D. and Vieille N. (2000) The Maxmin of Recursive Games with Incomplete Information on One Side, *Math. Oper. Res.,* **25**, 23-35

[12] Secchi P. (1997) Stationary Strategies for Recursive Games, *Math. Oper. Res.*, **22**, 494-512

[13] Sion M. (1958) On General MinMax Theorems, *Pacific J. Math.*, **8**, 171-176

[14] Thuijsman F. and Vrieze K. (1992) A Note on Recursive Games. *Game Theory and Economic Applications,* B. Dutta and al. (eds.) Lecture Notes in Economics and Mathematical Systems 389, Springer Verlag, Berlin, 133-145