Discussion Paper No. 837

DISCOUNTING VERSUS UNDISCOUNTING IN DYNAMIC PROGRAMMING

by

Ehud Lehrer*

and

Dov Monderer*

June 1989

*Department of Managerial Economics and Decision Sciences, J.L. Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston, Illinois 60208.

DISCOUNTING VERSUS UNDISCOUNTING IN DYNAMIC PROGRAMMING

by

Ehud Lehrer and Dov Monderer

Abstract

We explore the various relationships between the limits of the values of the discounted, or the finitely truncated dynamic programming problems and the values of the undiscounted problems.

Discounting Versus Undiscounting

In Dynamic Programming

EHUD LEHRER AND DOV MONDERER
Northwestern University
June, 1989

1. Introduction. The following problem was suggested by J. F. Mertens at the Open Problems session in the International Conference of Game Theory held in Columbus, Ohio in June 1987. This is a particular case of more general problems to be found, with context and motivation, in [**2**]. For a general discussion about dynamic programming the reader is refered to [**1**].

The Problem. Let $\emptyset \neq \Gamma(s) \subseteq S$ for all $s \in S$. And let f be a bounded function defined on S. Consider the dynamic programming problem where the decision maker on day t has to choose a new state $s_{t+1} \in \Gamma(s_t)$, and recieves a payoff $f(s_t)$. Let $V_\lambda(s)$ be the value of the discounted problem:

$$V_\lambda(s_0) = \sup_{s_{t+1} \in \Gamma(s_t)} (1-\lambda) \sum_{t=0}^{\infty} \lambda^t f(s_t).$$

Assume $V_\lambda(s)$ converges uniformly (in $s \in S$) when $\lambda \longrightarrow 1$, say to $V(s)$.

Can the decision maker get $V(s)$ in the liminf undiscounted problem ? That is, does there exist for each s_0 and each $\varepsilon > 0$ a sequence $s_{t+1} \in \Gamma(s_t)$ such that

$$\liminf_{T \to \infty} \frac{1}{T+1} \sum_{t=0}^{T} f(s_t) \geq V(s_0) - \varepsilon.$$

In section 3 we construct a counter example , where $V(s_0) = 1$ and the lower Cesaro limit of any sequence $s_{t+1} \in \Gamma(s_t)$ equals 0.

In Section 2 we will show that (under the uniform convergence assumption) an "optimistic" decision maker can get $V(s)$ in the limsup undiscounted problem. However, we show that even an optimistic decision maker cannot get more than $V(s)$ in the undiscounted problem. Pointwise convergence of $V_\lambda(s)$ is proved to be sufficient for the latter result.

Section 4 is devoted to concluding remarks about the various relationships between the limits of the values of the discounted, or the finitely truncated dynamic programming problems and the values of the undiscounted problems. These relationships depend on the type of convergence of V_λ , and on the type of covergence of V_T, where V_T is the value of the Tth truncated problem.

2. The Upper Limit Value. In this section we prove that if $V_\lambda \longrightarrow V$ uniformly then $V = \overline{V}$, where

$$\overline{V}(s_0) = \sup_{s_{t+1}\in\Gamma(s_t)} \limsup_{T\to\infty} \frac{1}{T+1}\sum_{t=0}^{T} f(a_t).$$

In Theorem A we prove that $\overline{V} \geq V$, and in Theorem B we prove (under the assumption of pointwise convergence only) that $V \geq \overline{V}$. We now give an outline of the proof of Theorem A.

Let $s_0 \in S$ and let $\varepsilon > 0$. Our task is to find a sequence $s_0 \prec s_1 \prec s_2 \prec \ldots$ (where $s_t \prec s_{t+1}$ means $s_{t+1} \in \Gamma(s_t)$) such that the upper Cesaro limit of the payoffs' sequence $(f(s_t))_{t=0}^{\infty}$ is at least $V(s_0) - \varepsilon$. We define sequences $\lambda_n \longrightarrow 1$ and $\delta_n \longrightarrow 0$ with properties that will be described later and proceed as follows:

λ_1 is chosen sufficiently close to 1 so that $V_{\lambda_1}(s_0) \geq V(s_0) - \delta_1$. We then find a sequence $s_0 = a_0 \prec a_1 \prec a_2 \prec \ldots$ of states which is δ_1–optimal with respect to the discount factor

λ_1. That is, the λ_1–Abel series of $(f(a_t))_t$ is at least $V_{\lambda_1}(s_0) - \delta_1$. However, for our sequence $(s_t)_t$ we take only the first $t_1 + 1$ states, where $t_1 \geq 0$ is chosen in such a way that both the partial average of the head up to t_1 and $V_{\lambda_1}(a_{t_1+1})$ are at least $V_{\lambda_1}(s_0) - 2\delta_1$. Such t_1 exists because each Abel series can be written as a convex combination of the partial averages (see (2.3)).

λ_2 is chosen so that $V_{\lambda_2}(a_{t_1+1}) \geq V_{\lambda_1}(a_{t_1+1}) - (\delta_1 + \delta_2)$ (here we are using the uniform convergence of V_λ). We then find another sequence $a_{t_1+1} = b_0 \prec b_1 \prec b_2 \prec \ldots$, which is δ_2–optimal with respect to λ_2. We take the first $t_2 + 1$ states of this sequence to be the next states in our sequence $(s_t)_t$, where t_2 is chosen so that both the partial average of the head up to t_2 and $V_{\lambda_2}(b_{t_2+1})$ are at least $V_{\lambda_2}(b_0) - 2\delta_2$. The next $t_3 + 1$ states in $(s_t)_t$ come from a sequence corresponding to λ_3, and so on. At the end of the induction process we have an infinite sequence

$$s_0 = a_0 \prec a_1 \prec \cdots \prec a_{t_1+1} = b_0 \prec b_1 \prec \cdots \prec b_{t_2+1} = c_0 \prec c_1 \prec \cdots \prec c_{t_3+1} = d_0 \prec \ldots \quad .$$

This sequence has the property that the partial averages from the $(T_n + 1)$th state to the T_{n+1}th state are at least $V(s_0) - \varepsilon$, where $T_n = (\sum_{k=1}^{n} t_k) + n - 1$. Hence, $\overline{V}(s_0) \geq V(s_0) - \varepsilon$.

THEOREM A. *Let (S, Γ, f) be a dynamic programming problem as defined in the introduction. Assume $V_\lambda(s)$ converges to $V(s)$ uniformly in $s \in S$. Then, for every $s_0 \in S$ and for every $\varepsilon > 0$ there exists a path $s_{t+1} \in \Gamma(s_t)$ such that*

$$\limsup_{T \to \infty} \frac{1}{T+1} \sum_{t=0}^{T} f(s_t) \geq V(s_0) - \varepsilon.$$

PROOF: Without loss of generality we may assume that $0 \leq f(s) \leq 1$ for all $s \in S$. Let $(\delta_k)_{k=1}^{\infty}$ be a sequence of positive numbers such that $\sum_{k=1}^{\infty} \delta_k < \frac{\varepsilon}{8}$. Since $V_\lambda \longrightarrow V$

uniformly, we can find an increasing sequence $(\lambda_k)_{k=1}^{\infty}$ converging to 1 and satisfying the following two properties:

$$1-\lambda_k < \delta_k \qquad \text{for all } k \geq 1 \; ; \tag{2.1}$$

and

$$\|V_{\lambda_k} - V\| < \delta_k \qquad \text{for all } k \geq 1, \tag{2.2}$$

where for $x \in R^S$ $\|x\| = \sup_{s \in S} |x(s)|$.

Before proceeding to the construction of the sequence $s_0 \prec s_1 \prec s_2 \prec \ldots$ (that is, $s_{t+1} \in \Gamma(s_t)$) we need the following lemmas.

LEMMA 2.1. *For every sequence $b = (b_t)$ of real numbers, for every $0 \leq \lambda < 1$, and for every $T \geq 0$*

$$(1-\lambda)\sum_{t=0}^{T} \lambda^t b_t = (1-\lambda)^2 \sum_{t=0}^{T-1} \lambda^t (t+1) S_t(b) + (1-\lambda)\lambda^T (T+1) S_T(b),$$

where $S_t(b) = \frac{1}{t+1}\sum_{i=0}^{t} b_i$, and $\sum_{t=0}^{-1} = 0$.

PROOF OF LEMMA 2.1: The proof is based on a simple direct computation, and therefore will be omitted. ∎

Lemma 2.1 implies that for every bounded sequence $b = (b_t)$

$$(1-\lambda)\sum_{t=0}^{\infty} \lambda^t b_t = (1-\lambda)^2 \sum_{t=0}^{\infty} \lambda^t (t+1) S_t(b). \tag{2.3}$$

LEMMA 2.2. *Let $0 \leq \lambda < 1$ and let $b = (b_t)$ be a bounded sequence. Then, there exists $T \geq 0$ such that*

$$S_T(b) \geq (1-\lambda)\sum_{t=0}^{\infty} \lambda^t b_t .$$

PROOF OF LEMMA 2.2: To get the result Combine $(1-\lambda)^2 \sum_{t=0}^{\infty} \lambda^t (t+1) = 1$ with (2.3). ▮

We now construct the path $s_0 \prec s_1 \prec s_2 \prec \ldots$.

Step 1. Denote $v_0 = s_0$. Take a path $v_0 = a_0 \prec a_1 \prec a_2 \prec \ldots$ such that

$$V_{\lambda_1}(a_0) - \delta_1 \leq (1-\lambda_1) \sum_{t=0}^{\infty} \lambda_1^t f(a_t).$$

By Lemma 2.2 there exists $T \geq 0$ with

$$S_T(f(a)) \geq (1-\lambda_1) \sum_{t=0}^{\infty} \lambda_1^t f(a_t) \geq V_{\lambda_1}(v_0) - \delta_1,$$

where $f(a) = (f(a_t))_{t=0}^{\infty}$.

Let t_1 be the smallest t for which $S_t(f(a)) \geq V_{\lambda_1}(v_0) - \delta_1$.

Define $s_1 = a_1$, $s_2 = a_2$, $\ldots$ $s_{t_1+1} = a_{t_1+1}$. And for use in the next step define $v_1 = a_{t_1+1}$.

Step 2. Take a path $v_1 = a_0 \prec a_1 \prec \ldots$ with

$$V_{\lambda_2}(v_1) - \delta_2 \leq (1-\lambda_2) \sum_{t=0}^{\infty} \lambda_2^t f(a_t).$$

Let $t_2 \geq 0$ be the smallest integer satisfying

$$s_{t_2}(f(a)) \geq V_{\lambda_2}(v_1) - \delta_2.$$

Define, $s_{t_1+i} = a_{i-1}$ for $2 \leq i \leq t_2 + 2$, and $v_2 = a_{t_2+1}$.

Step n. Denote $q_k = \sum_{i \leq k} t_i$.

v_{n-1} has been already defined in the $(n-1)$th step (as $s_{q_{n-1}+n-1}$).

Take a sequence $v_{n-1} = a_0 \prec a_1 \prec a_2 \prec \ldots$ with

$$V_{\lambda_n}(v_{n-1}) - \delta_n \leq (1-\lambda_n) \sum_{t=0}^{\infty} \lambda_n^t f(a_t). \tag{2.4}$$

Let $t_n \geq 0$ be the smallest integer satisfying

$$S_{t_n}(f(a)) \geq V_{\lambda_n}(v_{n-1}) - \delta_n. \tag{2.5}$$

Define $s_{q_{n-1}+i} = a_{i-n+1}$ for $n \leq i \leq t_n + n$, and $v_n = a_{t_n+1}$.

At the end of the induction process we have a path

$$s_0 \prec s_1 \prec s_2 \prec \ldots,$$

a sequence $(t_n)_{n \geq 1}$ of nonnegative integers, and a sequence $(v_n)_{n \geq 0}$ of states such that $v_n = s_{q_n+n}$, where $q_0 = 0$ and for $n \geq 1$ $q_n = \sum_{i \leq n} t_i$.

We now need the following two lemmas. The first one asserts that by moving from v_{n-1} to v_n and from λ_n to λ_{n+1} we may lose only a small payoff. In the second lemma we show that $V_{\lambda_n}(v_{n-1})$ is close to $V(s_0)$.

LEMMA 2.3. *For every $n \geq 1$,*

$$V_{\lambda_n}(v_{n-1}) - V_{\lambda_{n+1}}(v_n) \leq 3\delta_n + \delta_{n+1}.$$

PROOF OF LEMMA 2.3: By the construction, and by Lemma 2.1 we have at the nth step:

$$V_{\lambda_n}(v_{n-1}) - \delta_n \leq (1-\lambda_n) \sum_{t=0}^{\infty} \lambda_n^t f(a_t) = \tag{2.6}$$

$$(1-\lambda_n) \sum_{t=0}^{t_n-1} \lambda_n^t f(a_t) + \lambda_n^{t_n}(1-\lambda_n) \sum_{t=t_n}^{\infty} \lambda_n^{t-t_n} f(a_t) =$$

$$(1-\lambda_n)^2 \sum_{t=0}^{t_n-2} \lambda_n^t (t+1) S_t + (1-\lambda_n)\lambda_n^{t_n-1} t_n S_{t_n-1} + \lambda_n^{t_n}(1-\lambda_n) \sum_{t=t_n}^{\infty} \lambda_n^{t-t_n} f(a_t).$$

The right-hand-side of (2.6) is a convex combination of all S_t, $0 \le t \le t_n - 1$, and $\sum_{t=t_n}^{\infty}(1-\lambda_n)\lambda_n^{t-t_n} f(a_t)$. Since $S_t < V_{\lambda_n}(v_{n-1}) - \delta_n$, for all $0 \le t \le t_n - 1$, we have:

$$(1-\lambda_n) \sum_{t=t_n}^{\infty} \lambda_n^{t-t_n} f(a_t) \ge V_{\lambda_n}(v_{n-1}) - \delta_n.$$

Hence,

$$V_{\lambda_n}(v_n) = V_{\lambda_n}(a_{t_n+1}) \ge (1-\lambda_n) \sum_{t=t_n+1}^{\infty} \lambda_n^{t-(t_n+1)} f(a_t) \ge$$

$$(1-\lambda_n) \sum_{t=t_n}^{\infty} \lambda_n^{t-t_n} f(a_t) - (1-\lambda_n) f(a_{t_n}) \ge V_{\lambda_n}(v_{n-1}) - 2\delta_n.$$

Therefore,

$$V_{\lambda_n}(v_{n-1}) - V_{\lambda_{n+1}}(v_n) \le$$

$$V_{\lambda_n}(v_{n-1}) - V_{\lambda_n}(v_n) + V_{\lambda_n}(v_n) - V(v_n) + V(v_n) - V_{\lambda_{n+1}}(v_n) \le$$

$$3\delta_n + \delta_{n+1}.$$

This complete the proof of Lemma 2.3. ∎

LEMMA 2.4. *For all $n \ge 1$*

$$V(s_0) - V_{\lambda_n}(v_{n-1}) \le 4 \sum_{k \le n} \delta_k.$$

PROOF OF LEMMA 2.4:

$$V(s_0) - V_{\lambda_n}(v_{n-1}) \le V(s_0) - V_{\lambda_1}(s_0) + \sum_{k=1}^{n-1} (V_{\lambda_k}(v_{k-1}) - V_{\lambda_{k+1}}(v_k)).$$

Therefore by Lemma 2.3

$$V(s_0) - V_{\lambda_n}(v_{n-1}) \le \delta_1 + \sum_{k=1}^{n-1} (3\delta_k + \delta_{k+1}) \le 4 \sum_{k \le n} \delta_k. \quad \blacksquare$$

End of Proof of Theorem A. Denote $T_n = q_n + n - 1$, where $q_n = \sum_{k=1}^{n} t_k$ $(q_0 = 0)$.

Then

$$\frac{1}{T_n+1}\sum_{t=0}^{T_n} f(s_t) = \sum_{j=1}^{n} \frac{t_j+1}{T_n+1}\left[\frac{1}{t_j+1}\sum_{t=T_{j-1}+1}^{T_j} f(s_t)\right] \geq$$

$$\sum_{j=1}^{n} \frac{t_j+1}{T_n+1}(V_{\lambda_j}(v_{j-1}) - 2\delta_j) \geq V(s_0) - \varepsilon$$

by Lemma 2.4. Therefore

$$\limsup_{T\to\infty} \frac{1}{T+1}\sum_{t=0}^{T} f(s_t) \geq V(s_0) - \varepsilon. \qquad \blacksquare \; \blacksquare$$

THEOREM B. *Let (S,Γ,f) be a dynamic programming problem as described at the introduction. Assume $V_\lambda(s) \longrightarrow V(s)$ for all $s \in S$. Then for every sequence $(s_t)_{t=0}^{\infty}$ with $s_{t+1} \in \Gamma(s_t)$*

$$V(s_0) \geq \limsup_{T\to\infty} \frac{1}{T+1}\sum_{t=0}^{T} f(s_t).$$

PROOF: Without loss of generality assume $0 \leq f(s) \leq 1$. Denote $u = \limsup \frac{1}{T+1}\sum_{t=0}^{T} f(s_t)$. Let $\varepsilon > 0$. We can find a sequence $0 = T_0 < T_1 < T_2 < \ldots$ such that for all $k \geq 1$

$$A_k = \frac{1}{T_k}\sum_{t=q_{k-1}}^{q_k - 1} f(s_t) \geq u - \frac{\varepsilon}{2}, \tag{2.7}$$

where $q_k = \sum_{i=0}^{k} T_i$.

Note that for all $k \geq 1$

$$\lim_{\lambda\to 1}\left(\frac{1-\lambda}{1-\lambda^{T_k}}\right)\sum_{t=0}^{T_k-1} \lambda^t b_t = \frac{1}{T_k}\sum_{t=0}^{T_k-1} b_t \tag{2.8}$$

uniformly in all sequences (b_i) of numbers in the interval $[0,1]$.

Let $(\delta_k)_{k=0}^{\infty}$ be a sequence of positive numbers with $\sum_{k=0}^{\infty} \delta_k < \frac{\varepsilon}{6}$.

Combine (2.8) with $V_\lambda \longrightarrow V$ to form an increasing sequence $(\lambda_k)_{k=0}^\infty$ with $\lambda_k \longrightarrow 1$ such that for all $k \geq 1$ and for all sequences (b_i),

$$\left(\frac{1-\lambda_k}{1-\lambda_k^{T_k}}\right)\sum_{t=0}^{T_k-1}\lambda^t b_t \geq \frac{1}{T_k}\sum_{t=0}^{T_k-1} b_t - \delta_k; \tag{2.9}$$

and such that for all $k \geq 1$

$$|V_\lambda(v_k) - V(v_k)| < \delta_k \qquad \text{for all } \lambda \geq \lambda_k, \tag{2.10}$$

where $v_k = s_{q_k}$ for all $k \geq 0$.

Obviously, for all $k \geq 1$,

$$V_{\lambda_k}(v_{k-1}) \geq (1-\lambda_k^{T_k})B_k + \lambda_k^{T_k} V_{\lambda_{k+1}}(v_{k-1}),$$

where

$$B_k = \left(\frac{1-\lambda}{1-\lambda^{T_k}}\right)\sum_{t=0}^{T_k-1}\lambda^t b_t.$$

Therefore, by (2.9) and (2.10),

$$V_{\lambda_k}(v_{k-1}) \geq (1-\lambda_k^{T_k})A_k + \lambda_k^{T_k} V_{\lambda_{k+1}}(v_{k+1}) - 3\delta_k. \tag{2.11}$$

Combine now $V(s_0) \geq V_{\lambda_1}(v_0) - \delta_0$ with the inequalities in (2.11) to get:

$$V(s_0) \geq \sum_{k=1}^{\infty}\left(\prod_{t=0}^{k-1}\lambda_t^{T_t}\right)(1-\lambda_k^{T_k})A_k - 3\sum_{k=0}^{\infty}\delta_k \geq u - \varepsilon. \tag{2.12}$$

(Note: (2.7) and $\sum_{k=1}^{\infty}(\prod_{t=0}^{k-1}\lambda_t^{T_t})(1-\lambda_k^{T_k}) = 1$). As (2.12) is true for all $\varepsilon > 0$ the result follows. ∎ ∎

COROLLARY C. *Let (S, Γ, f) be a dynamic programming problem as defined in the introduction. Assume $V_\lambda(s)$ converges to $V(s)$ for all $s \in S$, and that the function $\lambda \longrightarrow V_\lambda(s)$ is non-decreasing. Then $V = \overline{V}$.*

PROOF: The proof goes exactly along the lines of the two previous theorems' proofs and therefore will be omitted. ∎

To conclude this section we remark that Theorem A, Theorem B, and Corollary C, as well as their proofs, remain valid if the assumption $V_\lambda \longrightarrow V$ is replaced by the weaker assumption:

There exists a sequence $\lambda_n \nearrow 1$ such that $V_{\lambda_n} \longrightarrow V$ (uniformly in Theorem A, pointwise in Theorem B, and monotonicly in Corollary C).

3. The Lower Limit Value. In this section we construct a dynamic programming problem in which $V_\lambda \longrightarrow V$ uniformly, but $V > \underline{V}$, where

$$\underline{V}(s_0) = \sup_{s_{t+1} \in \Gamma(s_t)} \liminf_{T \to \infty} \frac{1}{T+1} \sum_{t=0}^{T} f(a_t).$$

The following terminology and notations will be used in the construction of the example.

A *numerical tree* is a directed tree with a real valued function (called, the payoff function) defined on the set of vertices. Every numerical tree can be naturally identified with a dynamic programming problem:

The set of states is identified with the set of vertices, and $\Gamma(s)$ is the set of all vertices t for which (s, t) is an edge.

We will denote by V_T, E_T, s_T, and f_T the set of vertices of the numerical tree T, set of edges, root, and its payoff function respectively. We will say that $T_1 \leq T_2$, if $V_{T_1} \subseteq V_{T_2}$,

$E_{T_1} \subseteq E_{T_2}$, $s_{T_1} = s_{T_2}$, and f_{T_1} is the restriction of f_{T_2} to V_{T_1}. The union numerical tree $T = \vee_{i \geq 1} T_i$ of the sequence $T_1 \leq T_2 \leq \ldots$ is naturally defined.

Let T_1, T_2 be numerical trees with disjoint set of vertices, and let $v \in V_{T_1}$ satisfy $f_{T_1}(v) = f_{T_2}(s_{T_2})$; The numerical tree T, obtained from attaching T_2 to T_1 at v, is defined as follows: Identify v with s_{T_2}, and then define $V_T = V_{T_1} \cup V_{T_2}$, $E_T = E_{T_1} \cup E_{T_2}$, and $f_T(u) = f_{T_i}(u)$ if $u \in V_{T_i}$.

The tree T in our example will be the union of an inductively defined increasing sequence $T_1 \leq T_2 \leq T_3 \leq \ldots$ of numerical trees.

A few more definitions will help to facilitate the description of the induction step.

Let $(A^n)_{n=3}^{\infty}$ be a sequence of mutually disjoint copies of the set of non-negative integers. The kth element in A^n will be denoted by a_k^n. We identify all a_0^n with $a_0^1 = a_0$. We now define a numerical tree $T(c, b)$ for every $c, b \geq 0$. The vertices' set is $\cup_{n=3}^{\infty} A^n$. The edges are all pairs (a_k^n, a_{k+1}^n), and the payoff function f is defined as follows: $f(a_0) = b$, and $f(a_k^n) = \max(c - \frac{1}{\sqrt{\log n}}, 0)$, if $k \in I_n$, and $f(a_k^n) = 0$ otherwise, where I_n is the interval $(n, n[\log n])$ of integers.

Stage 1. T_1 is $T(1, 1)$.

Stage 2. Let D_2 be the set of all vertices v of T_1 whose distance from a_0 is 1 (i.e., there exists a directed path with one edge from a_0 to v). For every $v \in D_2$ let

$$c(v) = \max\{f(u) : u \text{ is a vertex of } T_1 \text{ and } u \gg v\},$$

where $u \gg v$ means that there is a directed path from v to u. The tree T_2 is obtained from T_1 by attaching the tree $T(c(v), f(v))$ to T_1 at each vertex $v \in D_2$.

Stage n. T_n is obtained from T_{n-1} exactly as T_2 is obtained from T_1. That is, we define D_n to be the set of all vertices in T_{n-1} whose distance from a_0 is n. For each $v \in D_n$ we define

$$c(v) = \max\{f(u) : u \in T_{n-1} \text{and } u \gg v\}.$$

To get T_n we attach the tree $T(c(v), f(v))$ to T_{n-1} at each $v \in D_n$.

Finally, define $T = \bigvee_{n=1}^{\infty} T_n$. Note that $c(v)$ is defined now for every $v \in V_T$.

Few Observations.

(a). If $u \gg v$ then $f(u) \leq c(v)$ (where $c(a_0) = 1$). Therefore $V_\lambda(v) \leq c(v)$ for all $0 < \lambda < 1$.

(b). Denote $\lambda_m = 1 - \frac{1}{m\sqrt{\log m}}$. Then,

$$\lambda_m(I_m) = (1 - \lambda_m) \sum_{i \in I_m} \lambda_m^i = \lambda_m^m (1 - \lambda_m^{m[\log m] - m + 1}) = \gamma_m \longrightarrow 1,$$

when $m \longrightarrow \infty$.

(c). Because of (b)

$$V_{\lambda_m}(v) \geq \gamma_m \left(c(v) - \frac{1}{\sqrt{\log m}} \right) \longrightarrow c(v).$$

(d). For all $\lambda_m < \lambda < \lambda_{m+1}$,

$$V_\lambda(v) \geq \lambda(I_m) \left(c(v) - \frac{1}{\sqrt{\log m}} \right) \geq \lambda_m(I_m) \left(\frac{1-\lambda}{1-\lambda_m} \right) \left(c(v) - \frac{1}{\sqrt{\log m}} \right)$$

$$\geq \lambda_m(I_m) \left(\frac{1-\lambda_{m+1}}{1-\lambda_m} \right) \left(c(v) - \frac{1}{\sqrt{\log m}} \right).$$

(e). Since $\frac{1-\lambda_{m+1}}{1-\lambda_m} \longrightarrow 1$ we get from (a) ,(b) , and the last inequality that $V_\lambda(v) \longrightarrow c(v)$ as $\lambda \longrightarrow 1$.

(f). Observe that if $\lambda_m \leq \lambda \leq \lambda_{m+1}$, then by (d)

$$c(v) - V_\lambda(v) \leq c(v) - \lambda_m(I_m)\left(\frac{1-\lambda_{m+1}}{1-\lambda_m}\right)\left(c(v) - \frac{1}{\sqrt{\log m}}\right).$$

(g). Set

$$\varepsilon_m = 1 - \lambda_m(I_m)\frac{1-\lambda_{m+1}}{1-\lambda_m} + \frac{\lambda_m(I_m)(\frac{1-\lambda_{m+1}}{1-\lambda_m})}{\sqrt{\log m}}.$$

Combining $c(v) \leq 1$ for all v and (f) yields that

$$c(v) - V_\lambda(v) \leq \varepsilon_m \text{ for all } \lambda_m \leq \lambda \leq \lambda_{m+1}.$$

(h). Since $\varepsilon_m \longrightarrow 0$, $V_\lambda(v) \longrightarrow c(v)$ uniformly in all vertices v of T.

Averaging on Infinite Paths.

Every sequence $((r_k, y_k))$ of pairs of integers defines a path in the tree T (starting from a^0) as follows:

y_1 paces on the r_1th branch of T_1, up to a vertex v_1. Then, y_2 paces on the r_2th branch of T_{v_1} (the subtree that was attached to v_1 at stage y_1), up to the vertex v_2. Then y_3 paces on the r_3th branch of T_{v_2} (the subtree that was attached to v_2 at stage $y_1 + y_2$). And so on.

Note that if a path $a_0 \prec a_1 \prec a_2 \prec \ldots$ cannot be described as above, then it actually means that for some k $y_k = \infty$, and therefore $\lim_{T\to\infty} \frac{1}{T+1}\sum_{t=0}^{T} f(a_t) = 0$.

Therefore, if $\liminf \frac{1}{T+1}\sum_{t=0}^{T} f(a_t) > 0$ then the path can be described as above.

Suppose now that $\liminf \frac{1}{T+1}\sum_{t=0}^{T} f(a_t) > 0$. Since

$$1 - \sum_{k<K} \frac{1}{\sqrt{\log r_k}} \geq \liminf \frac{1}{T+1}\sum_{t=0}^{T} f(a_t)$$

for all K, we get that

$$1-\sum_{k=1}^{\infty}\frac{1}{\sqrt{\log r_k}}\geq 0.$$

In particular

$$\sum_{k=1}^{\infty}\frac{1}{\sqrt{\log r_k}}<\infty. \tag{3.1}$$

W.l.o.g. we can assume that for all k $y_k \geq r_k$. Otherwise $y_{k_0} < r_{k_0}$ for some k_0 and we can enlarge the lower Cesaro limit of the path by replacing it with the path $((r_k, y_k))_{k\neq k_0}$.

PROPOSITION. *If (3.1) holds then* $\liminf_{K\to\infty} d_K = 0$, *where*

$$d_K = \frac{\sum_{k<K} r_k \log r_k}{r_K}.$$

PROOF OF THE PROPOSITION: If the assertion does not hold, then there exists $d > 0$ such that

$$\frac{\sum_{k<K} r_k \log r_k}{r_K} > d \qquad \text{for all } K.$$

Define $f(x) = Cx\log x$.

CLAIM 1. *For C big enough* $\sum_{k<K} r_k \log r_k \leq f^{(K-1)}(r_1)$ *for all* $K \geq 2$. *($f^{(j)}$ is the jth iterate of f).*

PROOF OF CLAIM 1: We prove the claim by induction on K.

For $K = 2$

$$r_1 \log r_1 \leq C r_1 \log r_1.$$

Suppose $\sum_{k<K} r_k \log r_k \leq f^{(K-1)}(r_1)$, then

$$\sum_{k<K+1} r_k \log r_k \leq f^{(K-1)}(r_1) + r_{K+1}\log r_{K+1} \leq$$

$$f^{(K-1)}(r_1) + \frac{1}{d}\left(\sum_{k<K} r_k \log r_k\right) \log\left(\frac{1}{d}\sum_{k<K} r_k \log r_k\right) \leq$$

$$f^{(K-1)}(r_1) + \frac{1}{d} f^{(K-1)}(r_1) \log\left(\frac{1}{d} f^{(K-1)}(r_1)\right) \leq f^{(K)}(r_1)$$

if C is big enough and $f^{(K-1)}(r_1) \longrightarrow \infty$. The latter is true because by the induction hypothesis

$$f^{(K-1)}(r_1) \geq \sum_{k<K} r_k \log r_k > dr_K$$

and by (3.1) r_K tends to ∞. ▮

CLAIM 2. *If $f(r_1) \leq A^{A^2}$ for some $A > 0$, then for all $K \geq 0$*

$$f^{(K+1)}(r_1) \leq (A+K)^{(A+K)^2}.$$

PROOF OF CLAIM 2: The proof can be obtained by a simple induction and therefore will be omitted. ▮

Using claims 1 and 2 we deduce:

$$\sum_{K=1}^{\infty} \frac{1}{\sqrt{\log r_K}} \geq \sum_{K=1}^{\infty} \frac{1}{\sqrt{\log\left(\frac{1}{d} \sum_{k<K} r_k \log r_k\right)}} \geq \sum_{K=1}^{\infty} \frac{1}{\sqrt{\log\left(\frac{1}{d} f^{(K-1)}(r_1)\right)}}$$

$$\geq \sum_{K=1}^{\infty} \frac{1}{\sqrt{\log\left(\frac{1}{d}(A+K-2)^{(A+K-2)^2}\right)}}$$

$$\geq \sum_{K=1}^{\infty} \frac{1}{\sqrt{(A+K)^2 \log(A+K) - \log d}} = \infty.$$

This contradicts (3.1) and the proposition is established. ▮ ▮

We now return to our example. Observe that the average of the payoffs after taking $r_K - 1$ zeros in the Kth step (before walking into the nonzero zone of the r_Kth branch) is bounded from above by

$$\frac{0(r_K - 1) + 1 \sum_{k<K} r_k \log r_k}{r_K - 1 + \sum_{k<K} r_k \log r_k} = \frac{d_K}{1 - \frac{1}{r_K} + d_K}.$$

This is because at each former step k the path goes through at most $r_k[\log r_k]$ positive numbers which are smaller than 1.

Now, if $\liminf \frac{1}{T+1} \sum_{t=0}^{T} f(a_t) > 0$, then (3.1) holds, and therefore, by the proposition, $\liminf d_K = 0$. This in its turn implies that $\liminf \frac{1}{T+1} \sum_{t=0}^{T} f(a_t) = 0$. A contradiction. This shows that the lower Cesaro limit of any path equals 0 while $V(a_0) = 1$.

4. An Overview. In this section we describe the various relationships between the limits of the values of the discounted, or the finitely truncated dynamic programming problems and the values of the undiscounted problems. These relationships are described in connection with the type of convergence of V_λ , or with that of V_T, where V_T is the value of the Tth truncated problem.

Let $\underline{V}$ and $\overline{V}$ be the values of the undiscounted problem. That is,

$$\underline{V}(s_0) = \sup_{s_{t+1} \in \Gamma(s_t)} \liminf_{T \to \infty} \frac{1}{T+1} \sum_{t=0}^{T} f(a_t);$$

and

$$\overline{V}(s_0) = \sup_{s_{t+1} \in \Gamma(s_t)} \limsup_{T \to \infty} \frac{1}{T+1} \sum_{t=0}^{T} f(a_t).$$

It is obvious that $\overline{V} \geq \underline{V}$. Combining $\overline{V} \geq \underline{V}$ with theorem B yields:

$$V_\lambda \longrightarrow V \qquad \text{(pointwise)} \Longrightarrow V \geq \overline{V} \geq \underline{V}. \tag{4.1}$$

Any of the inequalities in (4.1) may be strict. The example of Section 3 shows that $V = \overline{V} > \underline{V}$ is possible. The numerical tree $T(1,1)$ (defined in Section 3) is an example to the feasibility of $V > \overline{V} = \underline{V}$, and a slight modification of $T(1,1)$ yields $V > \overline{V} > \underline{V}$.

Theorem A yields:

$$V_\lambda \longrightarrow V \qquad \text{(uniformly)} \Longrightarrow V = \overline{V} \geq \underline{V}. \tag{4.2}$$

As we showed, strong conditions are necessary to ensure $V \leq \underline{V}$. Several such sufficient conditions were given by Mertens and Neyman [**3**, 1981]. E.g., $V \leq \underline{V}$ if the vector valued function $\lambda \longrightarrow V_\lambda$ has a bounded variation with respect to the Supremum norm on R^S. Finally, for every $T \geq 0$ denote

$$V_T(s_0) = \sup_{s_{t+1} \in \Gamma(s_t)} \frac{1}{T+1} \sum_{t=0}^{T} f(a_t).$$

It can be easily seen that Theorem A, Theorem B, Corollary C, the counter example of Section 3, (4.1), and (4.2) remain valid (with the necessary notational changes), if $V_\lambda \longrightarrow V$ is everywhere replaced by $V_T \longrightarrow V$.

REFERENCES

1. Bertsekas,D.P., "Dynamic Programming and Stochastic Control," Academic Press, Inc, New York, 1976.
2. Mertens, J–F., *Repeated Games*, CORE Reprint No. 8624 (1986).
3. Mertens, J–F. and Neyman, A., *Stochastic games*, International Journal of game Theory **10** 2 (1981), 53–66.