The Annals of Probability
1975, Vol. 3, No. 1, 32–40

CONTROLLED MARKOV CHAINS[1]

BY HARRY KESTEN AND FRANK SPITZER

Cornell University

We propose a control problem in which we minimize the expected hitting time of a fixed state in an arbitrary Markov chains with countable state space. A Markovian optimal strategy exists in all cases, and the value of this strategy is the unique solution of a nonlinear equation involving the transition function of the Markov chain.

We are given a Markov chain with denumerable state space S. Let 0 denote a fixed element of S, and $S' = S\backslash\{0\}$. We assume the transition probabilities $P(x, y)$ are such that 0 can be reached from every $x \in S$, i.e., for each $x \in S$, $P^n(x, 0) > 0$ for some $n \geqq 0$. The control consists in being free to decide whether a transition shall take place or not; the object is to try to reach the point 0 in the shortest possible expected time. Thus a strategy is a rule describing to which states $y \in S$ we are willing to go, from each $x \in S$.

DEFINITION 1. A strategy σ is a family of "go-sets" $G(x)$, $x \neq 0$. Each $G(x)$ is a subset of $S\backslash\{x\}$, containing 0. A transition from x to y will take place, provided $y \in G(x)$. All other transitions are excluded, and the process will then wait at x one unit time, until the next transition is attempted. Let $T \leqq \infty$ be the time of the first visit to 0, under such a strategy. The value of the strategy σ is $h_\sigma(x) = E_\sigma^x[T]$, $x \in S$, with $h_\sigma(0) = 0$.

With each strategy σ, there is associated a new Markov chain, with 0 as absorbing state. Let Q_σ denote the transition function of this Markov chain, restricted to S'. Then

$$(1)\qquad \begin{aligned} Q_\sigma(x, y) &= P(x, y) && \text{if } y \in G(x)\,, && y \neq x \\ &= 0 && \text{if } y \notin G(x)\,, && y \neq x \\ &= \textstyle\sum_{z \in S\backslash G(x)} P(x, z) && \text{if } y = x\,, && x, y \in S'\,. \end{aligned}$$

Let 1 denote the constant (unity) function on S'. It follows from the definition of h_σ, that

$$(2)\text{t}\qquad \begin{aligned} (I - Q_\sigma)^{-1}1(x) &= \textstyle\sum_{k=0}^\infty Q_\sigma^{\,k}1(x) = \sum_{k=0}^\infty E_\sigma^x[I_{[k<T\leqq\infty]}] \\ &= E_\sigma^x[T] = h_\sigma(x)\,, \qquad x \in S'\,. \end{aligned}$$

REMARK 1. There exist strategies σ for which $T < \infty$ with probability one and $h_\sigma(x) < \infty$ for all $x \in S$. To exhibit such a strategy, choose $G(x)$ as follows: for

Received December 6, 1973; revised March 20, 1974.

[1] Research supported by the NSF at Cornell University. The authors are grateful to Anders Martin-Löf for valuable conversations and advice.

AMS 1970 *subject classifications.* Primary 60J05; Secondary 60J20, 93E99.

Key words and phrases. Markov chains, negative dynamic programming, hitting times.

each $x \in S'$ we may pick a minimal positive integer n, and a sequence $x_1, x_2, \cdots, x_n$, depending on x, such that

$$P(x, x_1) P(x_1, x_2) \cdots P(x_n, 0) > 0 .$$

Now let $G(x) = \{0, x_1\}$ if $n \geqq 1$, and $G(x) = \{0\}$ if $P(x, 0) > 0$. This strategy has the desired property.

DEFINITION 2. A strategy $\hat{\sigma}$ is optimal if its value

$$h_{\hat{\sigma}}(x) \leqq h_\sigma(x) , \qquad \text{for all} \quad \sigma \quad \text{and all} \quad x \in S . \tag{3}$$

REMARK 2. In view of Remark 1 the value of an optimal strategy, if one exists, must be everywhere finite. When there exist several optimal strategies, it follows from (3) that they have the same value.

In fact the existence of an optimal strategy was proved by R. Strauch ([1] Theorem 9.1) for a large class of dynamic programming problems, containing the present one. If h is the value of an optimal strategy, then it seems plausible that it should also be the value of the strategy with $G(x) = \{y : h(y) < h(x)\}$, which implies

$$h(x) = 1 + \textstyle\sum_{y \in S'} P(x, y)[h(x) \wedge h(y)] , \qquad x \in S' . \tag{4}$$

This motivates

DEFINITION 3. A real-valued (finite) nonnegative function on S' is called regular, if it satisfies (4).

LEMMA 1. *Let h be regular. Let σ be the strategy whose* $G(x) = \{y : h(y) < h(x)\} \cup \{0\}$, $x \in S'$. *Then h is the value of σ*, i.e., $h_\sigma(x) = h(x)$, $x \in S'$.

PROOF. For σ the strategy in the Lemma, let $Q = Q_\sigma$ denote the matrix associated with σ in (1). Then (1) and (4) imply that $h(x) = 1 + Q_\sigma h(x)$, $x \in S'$, and by iteration

$$h = (I + Q + Q^2 + \cdots + Q^n)1 + Q^{n+1}h .$$

Since h is finite on S', the sequence $(I + Q + \cdots + Q^n)1$ increases to a finite limit, as $n \to \infty$. Hence $Q^n 1(x)$ tends to zero as $n \to \infty$, for each $x \in S'$. But

$$Q^{n+1}h(x) = \textstyle\sum_{y \in S'} Q^{n+1}(x, y)h(y) \leqq h(x)Q^{n+1}1(x) ,$$

since $Q^{n+1}(x, y) > 0$ implies $h(y) < h(x)$. Hence $Q^{n+1}h \to 0$, so that

$$h(x) = \textstyle\sum_{k=0}^{\infty} Q^k 1(x) , \qquad x \in S' . \tag{5}$$

It follows from (2) and (5) that h is the value of σ. □

THEOREM 1 [1]. *There exists an optimal strategy. Its value (the common value of all optimal strategies) is regular.*

PROOF. Let $\mathscr{F}$ be the class of all functions $f: S' \to [0, +\infty]$. Define the operator $U: \mathscr{F} \to \mathscr{F}$ by

$$Uf(x) = 1 + \textstyle\sum_{y \in S'} P(x, y)[f(x) \wedge f(y)] , \qquad x \in S' . \tag{6}$$

Let $f_0 \equiv 0, f_1 = Uf_0, \cdots, f_{n+1} = Uf_n = U^{n+1}f_0$. Since $u \leqq v \in \mathscr{F}$ implies that $Uu \leqq Uv$, we have $f_n \nearrow f$ and also $f = Uf$ by the monotone convergence theorem. Let σ be any strategy and h_σ its value. Then

$$\begin{aligned} h_\sigma(x) &= 1 + \textstyle\sum_{y \in G(x)} P(x, y)h_\sigma(y) + \sum_{y \in S' \setminus G(x)} P(x, y)h_\sigma(x) \\ &\geqq 1 + \textstyle\sum_{y \in S'} P(x, y)[h_\sigma(x) \wedge h_\sigma(y)] = Uh_\sigma(x), \qquad x \in S'. \end{aligned}$$

It follows that

$$h_\sigma \geqq Uh_\sigma \geqq U^2h_\sigma \geqq \cdots \geqq U^nh_\sigma \geqq U^nf_0 = U^n0 = f_n, \qquad n \geqq 0.$$

Letting $n \to \infty$ gives $h_\sigma \geqq f$. Choosing for σ a strategy with everywhere finite value (which exists by Remark 1) we see that f is everywhere finite. To summarize: We have shown that f is regular and that $h_\sigma \geqq f$ for all strategies σ. The proof of Theorem 1 will therefore be complete if we exhibit a strategy whose value is f. (By Remark 2 this will then be the common value of all optimal strategies). It suffices to define a strategy by $G(x) = \{y : f(y) < f(x)\} \cup \{0\}$, $x \in S'$. By Lemma 1 the value of this strategy is f. □

THEOREM 2. *Suppose h is a regular function, and let σ be the strategy with $G(x) = \{y : h(y) < h(x)\} \cup \{0\}$, $x \in S'$. Then σ is an optimal strategy, and h is its value.*

This result will enable us to verify in practice that intuitively appealing candidates for optimal strategies are indeed optimal. It also has the theoretically interesting

COROLLARY 1. *There is one and only one regular function.*

The proof of Theorem 2 will follow from the estimate

LEMMA 2. *Suppose f and g are regular functions, and $f \leqq g$. Then*

$$g(x) \leqq 2[f(x)]^2, \qquad x \in S'. \tag{7}$$

To complete the proof of Theorem 2, suppose that h is regular, and that σ is the strategy in Theorem 2, defined in terms of h. By Lemma 1, h is the value of σ. Let h^* be the value of an optimal strategy σ^*. Then $h^* \leqq h$. To show that σ is optimal it only remains to show that $h \leqq h^*$. Let Q be the matrix associated with σ^* according to (1). Then

$$\begin{aligned} h(x) &= 1 + \textstyle\sum_{y \in S'} P(x, y)[h(x) \wedge h(y)] \\ &\leqq 1 + \textstyle\sum_{y \in A} P(x, y)h(y) + \sum_{y \in S' \setminus A} P(x, y)h(x), \end{aligned}$$

for any subset $A \subset S'$. Choosing $A = G^*(x) \setminus \{0\}$, where $G^*(x)$ is the go-set of σ^*, we have

$$h(x) \leqq 1 + Qh(x), \qquad x \in S'.$$

Iterating this inequality gives

$$h(x) \leqq (I + Q + Q^2 + \cdots + Q^n)1(x) + Q^{n+1}h(x), \quad n \geqq 1, \qquad x \in S'. \tag{8}$$

Lemma 2 implies

$$Q^{n+1}h(x) = \sum_{y\in S'} Q^{n+1}(x, y)h(y) \leqq 2 \sum_{y\in S'} Q^{n+1}(x, y)[h^*(y)]^2$$
$$\leqq 2[h^*(x)]^2 Q^{n+1}1(x) .$$

But $Q^{n+1}1 \to 0$ as $n \to \infty$ (as in the proof of Lemma 1). Therefore $Q^{n+1}h \to 0$ as $n \to \infty$. Hence, using (5) and (8),

$$h(x) \leqq \sum_{k=0}^{\infty} Q^k 1(x) = h^*(x) , \qquad x \in S' . \ \square$$

PROOF OF LEMMA 2. Let us define $f(0) = g(0) = 0$. Let $S_n = \{x : x \in S, f(x) \leqq n/2\}$, $n \geqq 1$. For all $x \in S_n$,

$$g(x) \leqq 1 + \sum_{y\in S'} P(x, y)[g(x) \wedge g(y)]$$
$$\leqq 1 + \sum_{y\in S_n} P(x, y)g(y) + g(x) \sum_{y\in S\setminus S_n} P(x, y)$$
$$\leqq 1 + g(x)[1 - P(x, S_n)] + \sup_{y\in S_n} g(y)P(x, S_n) .$$

For each n let $\varphi(n) = \sup_{y\in S_n} g(y)$. Then we have

$$g(x) \leqq \frac{1}{P(x, S_n)} + \varphi(n) , \qquad x \in S', \ n \geqq 1 . \tag{9}$$

Now we estimate the right-hand side in (9), which might be $+\infty$, if $P(x, S_n) = 0$. For $x \in S_{n+1}\setminus S_n$,

$$\frac{n+1}{2} \geqq f(x) = 1 + \sum_{y\in S'} P(x, y)[f(x) \wedge f(y)]$$

$$\geqq 1 + \sum_{y\in S\setminus S_n} P(x, y)[f(x) \wedge f(y)] \geqq 1 + \frac{n}{2}[1 - P(x, S_n)] ,$$

which gives

$$\frac{1}{P(x, S_n)} \leqq n , \qquad x \in S_{n+1}\setminus S_n, \ n \geqq 1 . \tag{10}$$

Combining (9) and (10) gives $\varphi(n + 1) \leqq n + \varphi(n)$, and since $S_1 = \{0\}$, we have $\varphi(1) = 0$. Hence $\varphi(n + 1) \leqq \sum_1^n k = n(n + 1)/2$. Thus we have shown that $g(x) \leqq n(n + 1)/2$ on the set where $f(x) \leqq (n + 1)/2$. Since $f(x) \geqq 1$ when $x \neq 0$, this implies

$$g(x) \leqq f(x)[f(x) + \tfrac{1}{2}] \leqq 2[f(x)]^2 , \qquad x \in S' . \ \square$$

Theorems 1 and 2 may be rephrased to produce a necessary and sufficient condition for optimality of a given strategy.

THEOREM 3. *A strategy σ with go-sets $G(x)$, $x \neq 0$, and finite value function h is optimal if and only if*

$$P(x, y) > 0 \quad \textit{and} \quad y \in G(x) \Rightarrow h(y) \leqq h(x) \tag{11}$$
$$P(x, y) > 0 \quad \textit{and} \quad y \notin G(x) \Rightarrow h(y) \geqq h(x) .$$

PROOF. Since

$$h(x) = 1 + \sum_{y\in G(x)} P(x, y)h(y) + \sum_{y\notin G(x)} P(x, y)h(x) , \qquad x \in S' ,$$

it is clear that h is regular (satisfies (4)) if and only if (11) holds. If (11) holds, then by Theorem 2, h is the value of an optimal strategy, so that σ is optimal. Conversely, if (11) fails, then h, is not regular, and so by Theorem 1, σ cannot be optimal. □

Now we describe an algorithm which may be used to construct an optimal strategy in certain cases, in particular whenever S is finite.

ALGORITHM. Let $x_0 = 0$, $h(0) = 0$.

Step 1. Choose $x_1 \in S'$, if possible, so as to minimize

$$\frac{1}{P(x, 0)}, \qquad \text{and define} \quad h(x_1) = \frac{1}{P(x_1, 0)}. \tag{12}$$

Step n. Assume that $x_1, x_2, \cdots, x_{n-1} \in S'$ have been chosen in steps $1, 2, \cdots, n-1$, and that $h(x_1), h(x_2), \cdots, h(x_{n-1})$ have been defined. Now choose $x_n \in S' \backslash \{x_1, x_2, \cdots, x_{n-1}\}$, if possible so as to minimize

$$\frac{1 + \sum_{i=0}^{n-1} P(x, x_i) h(x_i)}{\sum_{i=0}^{n-1} P(x, x_i)},$$

and define

$$h(x_n) = \frac{1 + \sum_{i=0}^{n-1} P(x_n, x_i) h(x_i)}{\sum_{i=0}^{n-1} P(x_n, x_i)}. \tag{13}$$

We shall say that the algorithm works, if at each step the desired minimum is assumed, and if every element $x \in S$ is eventually chosen. Then the algorithm will produce an ordering $S = \{0, x_1, x_2, \cdots\}$ of all of S and define a function $h: S \to \mathbb{R}$. The basic property of h is given by

LEMMA 3. *If the algorithm works, then h is non-decreasing in the ordering produced, i.e.,* $0 = h(0) < h(x_1) \leqq h(x_2) \leqq \cdots$.

PROOF. We proceed by induction, assuming $0 = h(x_0) < h(x_1) \leqq h(x_2) \leqq \cdots \leqq h(x_n)$. (We denote $0 = x_0$.) It will follow that $h(x_n) \leqq h(x_{n+1})$ if

$$\varphi_n(x) = \frac{1 + \sum_{i=0}^{n} P(x, x_i) h(x_i)}{\sum_{i=0}^{n} P(x, x_1)} - h(x_n) \geqq 0, \qquad x \notin \{x_0, x_1, \cdots, x_n\}.$$

Now we know from the nth step of the algorithm that

$$h(x_n) \leqq \frac{1 + \sum_{i=0}^{n-1} P(x, x_i) h(x_i)}{\sum_{i=0}^{n-1} P(x, x_i)}, \qquad x \notin \{x_0, x_1, \cdots, x_{n-1}\}.$$

Hence

$$\begin{aligned}
\varphi_n(x) &= \frac{1 + \sum_{i=0}^{n-1} P(x, x_i) h(x_i)}{\sum_{i=0}^{n} P(x, x_i)} - h(x_n)\left[1 - \frac{P(x, x_n)}{\sum_{i=0}^{n} P(x, x_i)}\right] \\
&\geqq \frac{1 + \sum_{i=0}^{n-1} P(x, x_i) h(x_i)}{\sum_{i=0}^{n} P(x, x_i)} - \frac{1 + \sum_{i=0}^{n-1} P(x, x_i) h(x_i)}{\sum_{i=0}^{n-1} P(x, x_i)}\left[1 - \frac{P(x, x_n)}{\sum_{i=0}^{n} P(x, x_i)}\right] \\
&= \left[1 + \sum_{i=0}^{n-1} P(x, x_i) h(x_i)\right] \\
&\quad \times \left[\frac{1}{\sum_{i=0}^{n} P(x, x_i)} - \frac{1}{\sum_{i=0}^{n-1} P(x, x_i)} + \frac{P(x, x_n)}{\sum_{i=0}^{n} P(x, x_i) \sum_{i=0}^{n-1} P(x, x_i)}\right] \\
&= 0.
\end{aligned}$$

□

THEOREM 4. *If the algorithm works, then the function h defined by it in* (12) *and* (13) *is the value of the optimal strategy whose go-sets are* $G(x_n) = \{0, x_1, x_2, \cdots, x_{n-1}\}$, $n \geqq 1$. *Conversely, suppose that the Markov chain has an optimal value function* h^* *with the property that the set* $\{x : h^*(x) < M\}$ *is finite for each* $M > 0$. *Then the algorithm works. This is clearly the case when S is finite.*

PROOF. Suppose that the algorithm works and produces the ordering $S = \{0, x_1, x_2, \cdots\}$. Take for σ the strategy with $G(x_n) = \{0, x_1, \cdots, x_{n-1}\}$, $n \geqq 1$. Let h_σ be its value. Then, denoting T_A the first hitting time of a set $A \subset S$,

$$h_\sigma(x_n) = E_\sigma^{x_n} T_{\{0\}} = E_\sigma^{x_n}\{T_{\{0,x_1,\cdots,x_{n-1}\}} + h_\sigma[X_{T_{\{0,x_1,\cdots,x_{n-1}\}}}]\} \\ = \frac{1 + \sum_{i=0}^{n-1} P(x_n, x_i) h_\sigma(x_i)}{\sum_{i=0}^{n-1} P(x_n, x_i)}, \qquad n \geqq 1 . \tag{14}$$

It follows from (12) and (13) that h_σ is the function h produced by the algorithm. Now Lemma 3 implies that $h = h_\sigma$ satisfies condition (11) in Theorem 3. By Theorem 3, therefore, σ is optimal.

Conversely, suppose the Markov chain has the property of the second part of the Theorem. Suppose the algorithm does not work. Then suppose first it terminates after a finite number of steps, in the sense that the desired minimum is not assumed. So suppose $x_1, x_2, \cdots, x_{n-1}$ have been chosen. Consider the strategy σ_n with $G(x) = \{0, \cdots, x_{n-1}\}$ when $x \notin \{0, x_1, \cdots, x_{n-1}\}$, and $G(x_k) = \{0, x_1, \cdots, x_{k-1}\}$ for $k \leqq n - 1$. Its value is

$$h_{\sigma_n}(x) = \frac{1 + \sum_{i=0}^{n-1} P(x, x_i) h(x_i)}{\sum_{i=0}^{n-1} P(x, x_i)}, \qquad x \notin \{0, x_1, \cdots, x_{n-1}\} .$$

If the algorithm were to terminate at the nth step, then we would have, for some $0 < M < \infty$, $h_{\sigma_n}(x) \leqq M$ for infinitely many $x \in S$. But h_{σ_n} exceeds the value function of the optimal strategy, which contradicts the hypothesis concerning h^*. Thus the selection of the sequence $\{x_n\}$ does not terminate.

The only other way the strategy could fail to work is that $A = \{0, x_1, x_2, \cdots\} \neq S$. If so, consider the strategy σ with go-sets

$$G(x_n) = \{0, x_1, \cdots, x_{n-1}\}, \qquad n \geqq 1, \quad \text{and} \quad G(x) = A \qquad \text{when } x \notin A .$$

Then (14) again holds and

$$h^*(x_n) \leqq h_\sigma(x_n) = \frac{1 + \sum_{i=0}^{n-1} P(x_n, x_i) h_\sigma(x_i)}{\sum_{i=0}^{n-1} P(x_n, x_i)} \\ \leqq \min_{x \notin A} \frac{1 + \sum_{i=0}^{n-1} P(x, x_i) h_\sigma(x_i)}{\sum_{i=0}^{n-1} P(x, x_i)} .$$

Consequently, with σ_n as above, and $x \notin A$

$$\lim_{n\to\infty} h_{\sigma_n}(x) = \lim_{n\to\infty} \frac{1 + \sum_{i=0}^{n-1} P(x, x_i) h_\sigma(x_i)}{\sum_{i=0}^{n-1} P(x, x_i)} \geqq \lim_{n\to\infty} h^*(x_n) = +\infty .$$

This however is impossible, since trivially $h_{\sigma_n}(x) \searrow$ as $n \nearrow$ for $x \notin A$, and $h_{\sigma_n}(x) < \infty$

whenever $\sum_{i=0}^{n-1} P(x, x_i) > 0$. (If $A \neq S$, then there exists some $x \in S\backslash A$ and $n < \infty$ such that $\sum_{i=0}^{n-1} P(x, x_i) > 0$ because 0 can be reached from any $x \in S$.) Thus A is all of S, and the algorithm works. □

Finally, a practical sufficient condition for optimality.

THEOREM 5. *Suppose* $S = \{0, x_1, x_2, \cdots\}$ *and* σ *is the strategy defined by* $G(x_n) = \{0, x_1, \cdots, x_{n-1}\}$, $n \geqq 1$. *Suppose that*

$$(15) \qquad \sum_{i=0}^{k-1} P(x_n, x_i) \geqq \sum_{i=0}^{k-1} P(x_{n+1}, x_i) \geqq 0 , \qquad \text{for all } n \geqq k \geqq 1 , \text{ and } \quad \sum_{i=0}^{n-1} P(x_n, x_i) > 0 .$$

Then σ *is an optimal strategy.*

PROOF. The value function h of σ is given by (14). In view of Theorem 3, σ will be optimal provided h is monotone in the given ordering of S. To simplify the notation, let

$$d_1 = h(x_1) , \qquad d_n = h(x_n) - h(x_{n-1}) , \qquad n \geqq 2 ,$$
$$M_{n,k} = \sum_{i=0}^{k-1} P(x_n, x_i) , \quad 1 \leqq k \leqq n , \quad \text{and} \quad M_{n,0} = 0 .$$

Then (14) reads

$$h(x_n) = \frac{1 + \sum_{i=0}^{n-1} [M_{n,i+1} - M_{n,i}]h(x_i)}{M_{n,n}} , \qquad n \geqq 1 .$$

Summation by parts yields the simple equation

$$(16) \qquad 1 = \sum_{i=1}^{n} M_{n,i} d_i , \qquad n \geqq 1 .$$

We have to show that $d_n \geqq 0$ for $n \geqq 1$. We know $d_1 \geqq 1$. Suppose now that we have $d_1 \geqq 0$, $d_2 \geqq 0$, $\cdots$, $d_n \geqq 0$. Then (15) and (16) imply

$$M_{n+1,n+1} d_{n+1} = 1 - \sum_{i=1}^{n} M_{n+1,i} d_i \geqq 1 - \sum_{i=1}^{n} M_{n,i} d_i = 0 .$$

Since $M_{n+1,n+1} > 0$ we have $d_{n+1} \geqq 0$, which completes the proof by induction. □

EXAMPLE 1. Let $S = \mathbb{Z}$, the integers. Suppose μ is a symmetric probability density on $\mathbb{Z}$, such that $\mu(k) \geqq \mu(k+1)$ for all $k \geqq 1$, and $\mu(0) < 1$. Let $P(i, j) = \mu(i - j) = \mu(j - i)$, $i, j \in \mathbb{Z}$. Then an optimal strategy for control of this random walk toward 0 is defined by

$$G(n) = \{k : |k| < |n|\} , \qquad n \in \mathbb{Z}\backslash\{0\} .$$

PROOF. By symmetry the problem is the same as control toward 0 of the Markov chain on $\mathbb{N} = \{0, 1, 2, \cdots\}$ whose transition function is

$$\tilde{P}(i, j) = \mu(i) \qquad \text{for } j = 0 ,$$
$$= \mu(i - j) + \mu(i + j) \qquad \text{for } j \neq 0 .$$

We have simply identified the states i and $-i$ of the original random walk. The proposed strategy is then to go from k only to the set $\{0, 1, \cdots, k - 1\}$. Now

$$M_{n,k} = \sum_{i=0}^{k-1} \tilde{P}(n, i) = \sum_{i=n-k+1}^{n+k-1} \mu(i) , \qquad n \geqq k \geqq 1 .$$

Hence

$$M_{n+1,k} - M_{n,k} = \mu(n+k) - \mu(n-k+1) \leqq 0\,, \qquad M_{n,n} > 0\,, \qquad n \geqq k \geqq 1\,.$$

It follows from Theorem 5 that the strategy is optimal.

EXAMPLE 2. Let $S = \mathbb{N} = \{0, 1, 2, \cdots\}$, $0 < p_n < 1$ for $n \geqq 0$, and

$$P(n, n+1) = p_n\,, \qquad P(n, 0) = 1 - p_n\,, \qquad n \geqq 0\,.$$

If $p_n \searrow$ as $n \nearrow$ then the uncontrolled strategy is optimal, i.e., $G(n) = \{0, n+1\}$, $n \geqq 1$. But if $p_n \nearrow$ as $n \nearrow$, then the optimal strategy has all $G(n) = \{0\}$, i.e., the process waits for the first transition to 0.

PROOF. Suppose $p_n \searrow$ as $n \nearrow$. Then the process without control has

$$h(n) = E^n T_{\{0\}} = 1 + p_n + p_n p_{n+1} + \cdots < \infty\,,$$

as a simple calculation shows. Hence $h(n) \geqq h(n+1)$ for all $n \geqq 1$. By Theorem 3, h is regular and the uncontrolled strategy is optimal. In this case it may happen that $\lim p_n = p_\infty > 0$. Then

$$\lim_{n\to\infty} h(n) = \frac{1}{1 - p_\infty} < \infty\,,$$

and we have an example of an optimal strategy which cannot be obtained by the algorithm of Theorem 4.

If $p_n \nearrow$ as $n \nearrow$, then the strategy with $G(n) = \{0\}$, $n \geqq 1$, has the value function $h(n) = (1 - p_n)^{-1}$, $n \geqq 1$. Now $h \nearrow$ as $n \nearrow \infty$, so that again h is regular and, by Theorem 3, the strategy is optimal.

EXAMPLE 3. Let $S = \mathbb{Z}_N$, the N-dimensional integers, and let

$$P(x, y) = \frac{1}{2N} \qquad \text{if} \quad |x - y| = 1\,.$$

Then the optimal strategy σ for control toward the origin is that of only allowing transitions which diminish the distance to 0. Thus $G(x) = \{y : 0 \leqq |y| < |x|\}$, $x \neq 0$. When $N = 1$ this follows from Example 1. For $N \geqq 2$, let h be the value function for this strategy. By Theorem 3 it suffices to prove that

$$h(x + e_k) - h(x) = E_\sigma^{x+e_k} T_{\{0\}} - E_\sigma^{x} T_{\{0\}} \geqq 0 \tag{17}$$

for arbitrary $x \geqq 0$ and arbitrary unit vectors e_k in the positive direction of the kth coordinate. To prove (17) consider two particles, one starting at $x + e_k$, the other at x, both using the same strategy. We couple their motions, by making them use the same random transition mechanism. Thus they both undergo the same displacements, unless a displacement is permitted to the first particle and not to the second. This can happen if the first particle is at unit distance from the coordinate plane $x_k = 0$, while the second is already in it. The result will be that the first particle "catches up" with the first and they hit 0 together. In any case the first particle cannot hit 0 before the second, and therefore (17) holds.

REFERENCE

[1] STRAUCH, R. E. (1966). Negative dynamic programming. *Ann. Math. Statist.* **27** 871–890.

DEPARTMENT OF MATHEMATICS
WHITE HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14950