Journal of the Korean Institute of Industrial Engineers
Vol. 16, No. 2, Dec., 1990

System Replacement Policy for A Partially Observable Markov Decision Process Model[+]

Chang-Eun Kim[*]

Abstract

The control of deterioration processes for which only incomplete state information is available is examined in this study. When the deterioration is governed by a Markov process, such processes are known as Partially Observable Markov Decision Processes (POMDP) which eliminate the assumption that the state or level of deterioration of the system is known exactly. This research investigates a two state partially observable Markov chain in which only deterioration can occur and for which the only actions possible are to replace or to leave alone. The goal of this research is to develop a new jump algorithm which has the potential for solving system problems dealing with continuous state space Markov chains.

1. Introduction

It is a common practice to periodically inspect a system that deteriorates over time as a part of a program to keep it operating efficiently and to reduce operating costs. Often, such inspections do not provide perfect information regarding the system status due to the inaccessibility of important system component or due to the expense of a detailed inspection. Therefore it is worthwhile to examine the control of deterioration processes for which only incomplete state information is available. When the deterioration is governed by a Markov process, such a process is known as a Partially Observable Markov Decision Process (POMDP) which is a generalization of a Markov Decision Process (MDP). POMDP eliminates the assumption that the state or level of deterioration of the system is known exactly.

A wide variety of controlled systems are often quantitatively modeled because of elements of uncertainty in their dynamic behavior. For many of these systems, only imperfect observations of the process subject to control are permitted. For example, a physician determines a treatment plan for a patient on the basis of symptoms and laboratory

+ Research Supported by the Korean Research Foundation in 1988
* Department of Industrial Engineering, Myong Ji University, Korea

test results that are only probabilistically related to the state of the patient. With respect to machine maintenance and quality control, an operator decides on the machine.

The study of a POMDP is difficult because the process that is observed is not Markovian. In an effort to develop an efficient approach for obtaining optimal policies for partially observable Markov deterioration processes, a simple structure is assumed. In this research, we investigate a two state partially observable Markov chain in which only deterioration can occur and for which the only actions possible are either to replace it or to leeave it alone. The problem discussed in this paper has grown out of an attempt to introduce a new approach that has potential for solving other problems dealing with continuous state space Markov chains. The objectives for this research is to develop the system replacement policies under a new approach which is a called the jump algorithm.

The problem considered in this research is defined mathematically in section 2, along with some of the basic descriptive measures of the process. A new optimization procedure is given in section 3. In section 4, conclusions are presented.

2. Statement of Problem

The partially observable Markov chain consists of a core process and an observation process. The core process, $\{X_n, n=0, 1, \cdots\}$, is a Markov chain with state space $E=\{1, 2, \cdots, L\}$. The Markov matrix $P=[p_{ij}]$ gives the transition probabilities for the core process, that is

$$P_{ij}=P\{X_{n+1}=j \mid X_n=i\} \text{ for } i, j\in E. \quad \cdots\cdots (2.1)$$

The observation process, $\{Z_n, n=0, 1, \cdots\}$, with state space $\Theta=\{1, 2, \cdots, M\}$ is obtained from the core process through the probabilities given in the matrix $R=[r_{i\theta}]$. If $X_n=i$, then $Z_n=\theta$ with probability $r_{i\theta}$, that is

$$r_{i\theta}=P\{Z_n=\theta \mid X_n=i\} \text{ for } i\in E, \theta\in\Theta. \quad \cdots\cdots (2.2)$$

The matrix R completely describes the output process. If the core process is completely observable, the matrix R is an identity matrix, and the POMDP will be the same as the MDP.

The conceptual idea of these processes is that the core process cannot be directly observed. Instead, the observation process is seen, and the relation between the observation process and the core process is probabilistic. Although not directly observed, control is desired for the core process. The most that can be known about the core process is the probability that it is in a given state based on knowledge of the observation process. The probabilities for the core process at time n will be given by the vector w_n, that is

$$w_n(i)=P\{X_n=i \mid Z_0, \cdots, Z_n\}. \quad \cdots\cdots (2.3)$$

Sondik[1971, 1978] shows that these vectors, $\{w_0, w_1, \cdots\}$, form a Markov chain called the core probability chain, and thus the control problem will be based on these probabilities. We shall refer to $w_n\in W$ as the state of knowledge of the partially observable Markov process, where the state space W is the Cartisian product of the intervals [0, 1]. The control problem is defined through actions and their associated probability matrices and costs. The action space considered in this research contains only two elements and is denoted by $A=\{1, 2\}$, where the action $a=1$ represents the action "not to replace" and $a=2$ represents the action "to replace."

The action is always taken after the observation process has been observed. If action a is taken, the next transition for the core process is according to the matrix P^a, and then the observation process will be determined according to the matrix R^a. For the replacement problem considered for this research, P^1 is an upper triangular Markov matrix, and P^2 is a matrix whose first column is all ones and all other elements are zero. The matrix R^1 is a matrix defined in (2.2) such that the diagonal element is the largest element in the row, and R^2 is the identity matrix. The implicatioins of these conditions are that under the "not to replace" action, the core process cannot improve and under the "to replace" action, the core process is returned to State 1, and its state is known with certainty.

Because of the complexity of the optimal replacement POMDP, we only consider the simplest of problems in an effort to obtain an efficient algorithm. The state space is restricted to contain two elements. State 1 represents a new system and State 2 represents a failed system. The cost structure is defined on the process by a function C, where c(i) is the cost incurred for each time period that the core process is in state i with $c(1)<c(2)$. Additionally, there is a cost of replacement denoted by c_r.

A policy is a decision function, d, that maps the state space of the core probability chain, W into the action space, A. The set of all decision function is denoted by D. For a fixed decision function, a cost is incurred based on the movement of the core process. Therefor, the control problem can be stated as

$$\inf_{d\in D} \Psi_d=\inf_{d\in D}\{\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m-1}[c(X_n)+c_r I_{\{A_n=2\}}]\}, \quad \cdots\cdots (2.4)$$

where A_n is the action taken at time n and I is an idicator function.

Since decisions are based on the core probability chain, some helpful quantities regarding the probability chain are now presented. It is helpful to define the matrix

$$R^a_\theta=\text{diag}[r^a_{1\theta},\cdots,r^a_{L\theta}] \text{ for } a\in A,\ \theta\in\Theta. \quad \cdots\cdots (2.5)$$

For the control problem, the sequence of events is as follows : The decision maker knows the current probability vector, w_n, that gives the core process state probabilities. The decision maker then chooses an action, $A_n=a$, and based on that action, the core process changes according to P^a. He then observes the next state of the observation process, $Z_{n+1}=\theta$, according to R^a. The probability vector is now updated(based on Bayes' rule) according to the transformation given by Sondik[1978]

$$T(w_n\mid\theta,\ a)\equiv w_{n+1}=w_nP^aR^a_\theta/P\{\theta\mid w_n,\ a\}, \quad \cdots\cdots (2.6)$$

where $P\{\theta\mid w_n,\ a\}$ is a shorthand notation used by Sondik to denote the conditional probability that the next observation state will be θ, and it is defined by

$$\{\theta\mid w_n,\ a\}=w_nP^aR^a_\theta 1, \quad \cdots\cdots (2.7)$$

where 1 is a vector of all ones.

For a fixed decision function, d, the process $\{w_n,\ n=0,\ 1,\ \cdots\}$ is a Markov chain and its probability transition function is given by

$$P\{w_{n+1}\in B \mid w_n=w\} = \sum_{\theta : T(w \mid \theta,\ d(w))\in B} P\{\theta \mid w,\ d(w)\}, \quad \cdots\cdots (2.8)$$

where B is a (measurable) set in the state space of the core probability chain. For the basic descriptive measure of the process, the following quantities are specifically defined :

$$P^1=\begin{bmatrix} p & 1-p \\ 0 & 1 \end{bmatrix} \qquad P^2=\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$R^1=\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \qquad R^2=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$C^1=\begin{bmatrix} c(1) \\ c(2) \end{bmatrix} \qquad C^2=\begin{bmatrix} c_r+c(1) \\ c_r+c(2) \end{bmatrix}$$

where $r_{11}>r_{12}$ and $r_{22}>r_{21}$.

The decision process is based on the probabilities $\{w_0, w_1, \cdots\}$ where the components of the vector w are $(1-x, x)$. Since only two states are being considered, the probability vectors can be reduced to scalars by only looking at the second component of the vector w. Thus, a new Markov chain, $\{Y_0, Y_1, \cdots\}$, with continuous state space $[0, 1]$ is defined by

$$Y_n=w_n(2) \text{ for } n=0,\ 1,\ \cdots. \quad \cdots\cdots (2.9)$$

Sondik proved[1978, Theorem 1] that a certain level of deterioration δ exists to minimize the control problem (2.4). Let $B=[0, \delta]$ be a region in the state space where $\delta \in [0, 1]$ denotes the policy to use $A_n=1$ when $Y_n\in B$, and to use $A_n=2$ when $Y_n\notin B$. The value δ is referred to as the control limit.

The Markov chain $Y=\{Y_0, Y_1, \cdots\}$ is a continuous state Markov chain with probability transition function under the control limit policy δ denoted by

$$F_\delta(x,\ y)\equiv P_\delta\{Y_{n+1}\leq y \mid Y_n=x\} \quad \text{for } x, y\in[0,\ 1]. \quad \cdots\cdots (2.10)$$

Before giving an expression for (2.10), it is helpful to rewrite the transformation of (2.6) for the process Y. This is accomplished by substituting $(1-Y_n, Y_n)$ for w_n and then only using the second component of the resulting vector. In particular, we now let $T(Y_n \mid \theta, \delta)$ denote the value of Y_{n+1} after θ is observed and the control limit δ is used ; thus

$$T(Y_n \mid \theta,\ \delta)=\begin{cases} \dfrac{[(1-p)+pY_n]r_{2\theta}}{(1-Y_n)pr_{1\theta}+[(1-p)+pY_n]r_{2\theta}} & \text{for } Y_n<\delta, \\ 0 & \text{for } Y_n\geq\delta. \end{cases} \quad \cdots\cdots (2.11)$$

An expression for the probability transition function (2.10) is similarly obtained from (2.8) and is given as

$$F_\delta(x,\ y)=\begin{cases} \Sigma_{\theta : T(x \mid \theta, \delta)\leq y}[(1-x)pr_{1\theta}+\{(1-p)+px\}r_{2\theta}] & \text{for } x<\delta, \\ 1 & \text{for } x>\delta. \end{cases} \quad \cdots\cdots (2.12)$$

We let $\Phi_\delta(\cdot)$ be the invariant probability function of F ; that is

$$\Phi_\delta(y)=\int_{x\in[0,\ 1]} \Phi_\delta(dx)F_\delta(x,\ y) \text{ and } \Phi_\delta(1)=1. \quad \cdots\cdots (2.13)$$

We derive the expression for $F(x,\ y)$ and $\Phi(\cdot)$

to find the optimal control policy in section 3. Under the existence of these functions, we can derive the optimal long-run average cost for the optimal control policy δ based on the process Y.

Theorem (2.1). Under the policy δ, the long-run average cost is given by

$$\Psi_\delta = c(1)(1-\mu) + c(2)\mu + (1-\Phi_\delta(\delta))c_r \quad \cdots\cdots (2.14)$$

where $\mu = \int_{x\in[0,1]} x\Phi_\delta(dx)$.

Proof. For a fixed policy δ, the expected average cost per unit time can be divided into two categories : (1) to use $A_n=1$ with a cost vector C^1 when $x\in B$, and (2) to use $A_n=2$ with a cost vector C^2 when $x\notin B$. Therefore, by the ergodic property of Markov chains, the long-run average cost can be given by

$$\begin{aligned}\Psi_\delta &= \int_{x\in B} \Phi_\delta(dx)[(1-x)\ c(1) + x\ c(2)] \\ &\quad + \int_{x\in B} \Phi_\delta(dx)\{(1-x)[c(1)+c_r] + x[c(2)+c_r]\} \\ &= \int_{x\in[0,1]} \Phi_\delta(dx)[(1-x)\ c(1) + x\ c(2)] \\ &\quad + \int_{x\in B\delta} \Phi_\delta(dx)C_r \\ &= c(1)(1-\mu) + c(2)\mu + c_r(1-\Phi_\delta(\delta)).\end{aligned}$$

The optimization problem (2.4) restated in terms of the control limit is to find the control limit, δ^*, such that

$$\Psi_{\delta^*} = \inf\{\Psi_\delta :\ \delta \varepsilon [0,\ 1]\}, \quad \cdots\cdots (2.15)$$

where Ψ_δ is defined by Equation (2.14).

The invariant function Φ_δ will be fully exploited to find the long-run average cost defined by Theorem (2.1). This method is a new approach that has potential for solving other continuous state space Markov chains. It is not based on Markov decision theory. The problem in this research is to develop an algorithm to compute δ^* based on an entirely new concept and compare its efficiency to Sondik's algorithm.

3. Optimization Procedure

In this section an algorithm is developed which can be utilized to find an optimal replacement policy defined by Equation (2.15). The approach used by Sondik and others has been based on an iterative procedure derived from Markov decision theory. For a given policy, the cost is not directly determined because the core probability Markov chain has a continuous state space. Although the invariant distribution for continuous state processes is difficult to determine, it makes the new algorithm simple for finding the optimal replacement policy. In order to compute the long-run average cost defined by Theorem (2.1), we need an expected value of invariant function and a value of $\Phi(\delta-)$. In the process of algorithm development, the invariant function Φ_δ will be fully exploited to find a long-run average cost for an optimal replacement policy.

3-1. Analysis of Invariant Function

In this section we return to the function $F_\delta(x, y)$ defined by Equation (2.10) and derive some of the invariant function's characteristics.

Theorem (3.1). The function $F_\delta(x, y)$ is defined, for $x<\delta$, as

$$F_\delta(x,\ y) = \begin{cases} 0 & \text{if } y<\alpha(x) \\ (1-x)pr_{11} + (q+px)r_{21} & \text{if } \alpha(x)\le y<\beta(x) \\ 1 & \text{if } y\ge\beta(x). \end{cases}$$

Proof. The probability transition function $F_\delta(x, y)$ exists in only three cases according to the indicator function $I_0(x, y)$ such that :

Case 1 : $I_1(x, y)=1$, $I_2(x, y)=1$

$$F_\delta(x, y)=\sum_0 \{(1-x)pr_{10}+(q+px)r_{20}\} I_0(x, y)$$
$$=\{(1-x)pr_{11}+(q+px)r_{21}\}\cdot 1+\{(1-x)pr_{12}+(q+px)r_{22}\}\cdot 1=1.$$

Case 2 : $I_1(x, y)=1$, $I_2(x, y)=0$

$$F_\delta(x, y)=(1-x)pr_{11}+(q+px)r_{21}.$$

Case 3 : $I_1(x, y)=0$, $I_2(x, y)=0$

$$F_\delta(x, y)=0.$$

Next, the finvariant function $\Phi_\delta(\cdot)$ defined by Equation (2.13) is examined. It will play a key role in the development of the algorith.

We will construct the invariant function $\Phi_\delta(\cdot)$ as a step function to find a long-run average cost given by Thoerem (2.1). We need to find the expected value μ of the invariant function and the value of $\Phi_\delta(\delta-)$ to find the optimal control limit. Let y_{ij} be a jump point of Φ_δ between 0 and 1. Also, let ϕ_{ij} be the jump size at y_{ij}, and ϕ^*_{ij} be the jump size at $\beta(y_{ij})$, $\phi_{ij}=\phi(y_{ij})-\phi(y_{ij}-)$ and $\phi^*_{ij}=\phi(\beta(y_{ij}))-\phi(\beta(y_{ij})-)$. It is also helpful to define the following notations for simplicity from Fig. 1.

Definition (3.2). The following quantities, for i, j=1, 2, ⋯, are defined :

$$y_{1,0}=0,\ y_{2,0}=\beta(0),$$
$$y_{i,j}=\alpha(y_{i,j-1}),\ y_{i+2,0}=\beta(y_{1,i}),$$
$$F_\delta(y_{ij}, y)=f_{ij}=(1-y_{ij})pr_{11}+(q+p\ y_{ij})r_{21}$$
$$\text{for } \alpha(y_{ij})\leq y<\beta(y_{ij}),$$
$$\phi_{ij}=\Phi_\delta(y_{ij})-\Phi_\delta(y_{ij}-),$$
$$\phi^*_{ij}=\Phi_\delta(\beta(y_{ij}))-\Phi_\delta(\beta(y_{ij})-),$$
$$\lim_{n\to\infty} y_{i,n}=y^*=y_{i,1i},$$

where 1_i is the smallest integer such that $|y^*-y_{i,1i}|<\varepsilon$,

$$\lim_{m\to\infty} y_{m,0}=1\equiv y_{k,0},$$

where k is the smallest integer such that $|1-y_{k,0}|<\varepsilon$.

4-2. The Structure of the Invariant Function

First, we need to find a value of the first jump which occurs at zero. If $y<\alpha(0)$, we can derive the following equation :

$$\Phi_\delta(y)=\int_{x\in[0,1]}\Phi_\delta(dx)F_\delta(x, y)=$$
$$\int_{x\in[0,\delta]}\Phi_\delta(dx)\cdot 0+\int_{x\in[\delta,1]}\Phi_\delta(dx)\cdot 1$$
$$=1-\Phi_\delta(\delta-)\ \text{ for } y\in[0, \alpha(0))$$
$$\Phi_\delta(0)=1-\Phi_\delta(\delta-).$$

Now we will construct the next jumps as follows.

Lemma (3.3). The jump size at $y_{1,n+1}$ for n=0, 1, ⋯, 1_1-1 is given by

$$\phi_{1,n+1}=\Phi_\delta(0)\pi^n_{i=1} f_{1,i},\ \delta\leq\beta(0).$$

Proof. If $\alpha(0)\leq y<\alpha(y_{1,1})$, then

$$\Phi_\delta(y)=\int_{x\in[0,1]}\Phi_\delta F_\delta(x, y)=\int_{x\in[0,\alpha^{-1}(y)]}\Phi_\delta(dx)F_\delta(x, y)+\int_{x\in[\delta,1]}\Phi(dx)$$
$$=\Phi_\delta(0)F_\delta(0, y)+\Phi_\delta(0)=\Phi_\delta(0)f_{1,0}+\Phi_\delta(0)$$
$$\phi_{1,1}=\Phi_\delta(y_{1,1})-\Phi_\delta(y_{1,1}-)=\Phi_\delta(0)f_{1,0}$$

where $\phi_{m,0}=(1-f_{1,m-2})\phi_{1,m-2}$. If $\alpha(y_{1,1})\leq y<\alpha(y_{1,2})$, then

$$\Phi_\delta(y)=\int_{x\in[0,1]}\Phi_\delta(dx)F_\delta(x, y)=$$
$$\int_{x\in[0,\alpha^{-1}(y))}\Phi_\delta(dx)F_\delta(x, y)+\int_{x\in[\delta,1]}\Phi_\delta(dx)$$

$=\Phi_\delta(0)f_{1,0}+\phi_{1,1}f_{1,1}+\phi_\delta(0)$

$\phi_{1,2}=\Phi_\delta(y_{1,2})-\Phi_\delta(y_{1,2}-)=\phi_{1,1}f_{1,1}$

$=\Phi_\delta(0)f_{1,0}f_{1,1}.$

Assuming that $\phi_{1,n}=\Phi_\delta(0)\pi^{n-1}_{i=0} f_{1,i}$ is true, this implies that $\phi_{1,n}=\phi_\delta(y_{1,n})-\phi_\delta(y_{1,n}-)=\Phi_\delta(0)\pi^{n-1}_{i=0}\phi_{1,i}$ when $\alpha(y_{1,n-1})\leq y<\alpha(y_{1,n})$. Then, if $\alpha(y_{1,n})\leq y<\alpha(y_{1,n+1})$,

$\Phi_\delta(y)=\int_{x\in[0,1]}\Phi_\delta(dx)F_\delta(x,y)=\int_{x\in[0,\alpha^{-1}(y)]}\Phi_\delta(dx)$

$F_\delta(x, y)+\int_{x\in[\delta,1]}\Phi_\delta(dx)=\Phi_\delta(0)f_{1,0}$

$+\phi_{1,1}f_{1,1}+\cdots+\phi_{1,n-1}f_{1,n-1}+\phi_{1,n}f_{1,n}$

$+\Phi_\delta(0)$

$\phi_{1,n+1}=\Phi_\delta(y_{1,n+1})-\Phi_\delta(y_{1,n+1}-)=\phi_{1,n}f_{1,n}$

$=\Phi_\delta(0)\ \Pi^{n-1}_{i=0} f_{1,i}f_{1,n}\ (=\Phi_{1,n}=\Phi_\delta(0)$

$\Pi^{n-1}_{i=0} f_{1,i})$

$=\Phi_\delta(0)\prod_{i=0}^{n} f_{1,i}.$

Thus the lemma is proved by induction.

Lemma (3.4). The jump size at $\beta(y_{1,n})$ for $n=1, 2, \cdots, l_1$ is given by

$\phi^*_{1,n}=\Phi_\delta(0)(1-f_{1,n})\ \Pi^{n-1}_{i=0} f_{1,i},\ \delta\leq\beta(0).$

Proof. The jump size, $\phi^*_{1,n}$, can be calculated in the following manner: the jump size of $\phi_{1,n}$ is first calculated by Lemma (3.3) and then multiplied by $(1-f_{1,n})$ to obtain the jump size at $\beta(y_{1,n})$ from $y_{1,n}$.

Theorem (3.5). The jump sizes at $y_{m,n}$ and $\beta(y_{m,n})$ for $m\geq1$ are:

$\phi_{m,n}=\phi_{m,0}\prod_{i=0}^{n-1} f_{m,i},\ y_{m,0}\leq\delta<y_{m+1,0}$

$\phi^*_{m,n}=\phi_{m,0}(1-f_{m,n})\prod_{i=0}^{n-1} f_{m,i},\ y_{m,0}\leq\delta<y_{m+1,0}$

Proof. It is easily proved by combining Lemma (3.3) and Lemma (3.4).

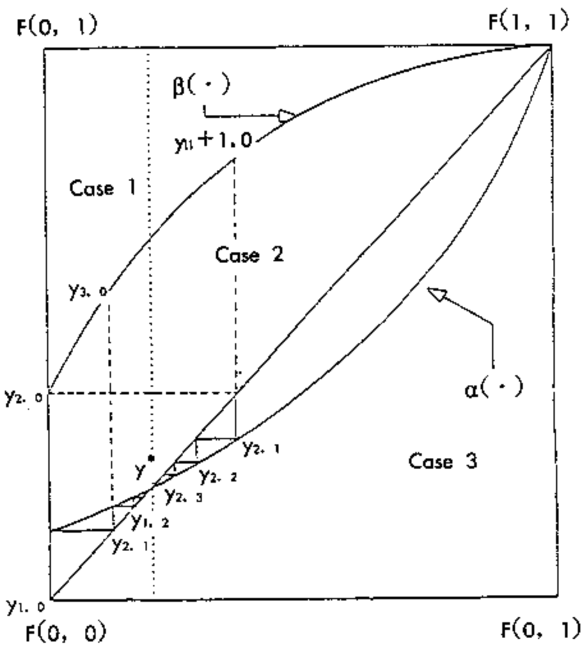

Fig. 1. Analysis of Invariant Function.

Lemma (3.6). The sum of all jumps below δ, if $\delta \leq \beta(0)$, is

$$\Phi_\delta(\delta-)=\frac{\lim_{n\to\infty}\ \Sigma^n_{i=0}\phi_{1,i}}{1+\lim_{n\to\infty}\ \Sigma^n_{i=0}\phi_{1,i}}\ .$$

Proof. To find $\Phi_\delta(\delta-)$ such that $\lim_{n\to\infty}\ \Sigma^n_{i=0}\phi_{1,i}=\Phi_\delta(\delta-)$ when $\delta\leq\beta(0)$, let $\phi_{1,0}=1$. Then, it follows that $(1-\Phi_\delta(\delta-))\lim_{n\to\infty}\ \Sigma^n_{i=0}\phi_{1,i}=\Phi_\delta(\delta-)$. It is proved by solving the equation for $\Phi_\delta(\delta-)$.

Lemma (3.7). The sum of all jumps below δ, if $\beta(y_{1,k-1})\leq\delta<\beta(y_{1,k})$, is

$$\Phi_\delta(\delta-)=\frac{\lim_{n\to\infty}\ \Sigma^{k+1}_{m=1}\Sigma^n_{i=0}\phi_{m,i}}{1+\lim_{n\to\infty}\Sigma^{k+1}_{m=1}\Sigma^n_{i=0}\ \phi_{m,i}}\ .$$

Proof. If $\beta(y_{1,0})\leq\delta<\beta(y_{1,1})$, to find $\Phi_\delta(\delta-)$ such that $\lim_{n\to\infty}\ \Sigma^n_{i=0}(\phi_{1,i}+\phi_{2,i})=\Phi_\delta(\delta-)$, let $\phi_{1,0}=1$. Then,

$$\Phi_\delta(\delta-)=(1-\Phi_\delta(\delta-))\lim_{n\to\infty}\sum_{i=0}^{n}(\phi_{1,i}+\phi_{2,i})$$

$$=\frac{\lim_{n\to\infty}\ \Sigma^2_{m=1}\Sigma^n_{i=0}\ \phi_{m,i}}{1+\lim_{n\to\infty}\ \Sigma^2_{m=1}\Sigma^n_{i=0}\ \phi_{m,i}}\ .$$

If $\beta(y_{1,1})\leq\delta<\beta(y_{1,2})$, the same procedure as given above is followed. To find $\Phi_\delta(\delta-)$ such that $\lim_{n\to\infty}\Sigma^n_{i=0}(\phi_{1,i}+\phi_{2,i}+\phi_{3,i})=\Phi_\delta(\delta-)$, let $\phi_{1,0}=1$. Then,

$$\Phi_\delta(\phi-)=(1-\Phi_\delta(\delta-))\ \lim_{n\to\infty}\sum_{i=0}^{n}(\delta_{1,i},\ \phi_{2,i}+\phi_{3,i})$$

$$=\frac{\lim_{n\to\infty}\ \Sigma^3_{m=1}\Sigma^n_{i=0}\phi_{m,i}}{1+\lim_{n\to\infty}\Sigma^3_{m=1}\Sigma^n_{i=0}\ \phi_{m,i}}$$

In general, if $\beta(y_{1,k-1})\leq\delta<\beta(y_{1,k})$, the same procedure as induction is used to obtain this Lemma.

Lemma (3.8). The expected value of μ defined by Theorem (3.1) is:

$$\mu=\begin{cases}\Sigma^n_{i=0}\{\phi_{1,i}y_{1,i}+\phi^*_{1,i}\ \beta(y_{1,i})\}, & y_{1,n}\leq\delta<y_{1,n+1}\\ \lim_{n\to\infty}\ \Sigma^k_{m=1}\Sigma^n_{i=0}\{\phi_{m,i}y_{m,i}+\phi^*_{m,i}\ \beta(y_{m,i})\}, & y_{k,0}\leq\delta<y_{k+1,0},\end{cases}$$

where $\phi^*_{m,i}=\phi_{m,i}(1-f_{m,i})$.

Proof. The expected value of the step function is the summation of the product of a jump point and a jump size since $E[x]=\int_{x\in[0,1]}x\Phi(dx)=\Sigma_i x_i\phi(x_i)$, where x_i is a jump point and $\phi(x_i)=\Phi(x_i)-\Phi(x_i-)$.

Now, we are ready to compute the long-run average cost defined by Theorem (2.1). To search for the minimum of Ψ_δ defined by Equation (2.15), we use a procedure based on the Golden Section Search method. This procedure will be called the jump algorithm.

Algorithm (3.9). The jump algorithm to find the optimal control limit defined by Equation (2.15) is as follows:

(1) Set $\varepsilon=0.0001$.

(2) Generate the jump points, $y_{1,n}$ and $y_{m,0}$, as defined by Definition (3.2) for $\eta=1, 2, \cdots, 1_1$ and $m=1, 2, \cdots, k$, where k and 1_1 are also defined by Definition (3.2).

(3) Set the control limit value, δ, equal to 0.3819 for which the long-run average cost, Ψ_δ, defined by Theorem (2.1) is to be evaluated.

(4) Temporarily set $\phi_{1,0}=1$ and $\phi_\delta(0)=1$.

(4.1) If $\delta<\beta(0)$, then calculate $\phi_{1,n}$ and $\phi^*_{1,n}$ for $n=1, 2, \cdots, 1_1$ using Lemma (3.3) and Lemma (3.4), and the jump points generated in Step (2).

(4.2) If $\delta\geq\beta(0)$, then calculate $\phi_{m,n}$ and $\phi^*_{m,n}$ for $m=2, 3, \cdots, k$ and $n=1, 2, \cdots, 1_m$ using Theorem (3,5), and the jump points generated in Step (2).

(5) Evaluate the value of $\phi_\delta(\delta-)$ using Lemma (3.6) and Lemma (3.7).

(6) Set $\phi_{1,0}=1-\Phi_\delta(\delta-)$ from Step (5), then recalculate $\phi_{m,n}$ and $\phi^*_{m,n}$ as in Step (4).

(7) Calculate the expected value, μ, using Lemma (3.8), then evaluate the long-run average cost, Ψ_δ, using Theorem (2.1).

(8) Determine δ_1 acoording to the Golden Section Search method, where $\delta \in [0, 1]$. If $|\delta-\delta_1| < \varepsilon$, then STOP; otherwise, let $\delta=\delta_1$ and return to Step (4).

4. Conclusions

The importance of this research lies in the applicability of replacement models and the practical difficulty of their optimization algorithms. In order to obtain data for replacement models and the other controlled systems, statistical data and imperfect instrumentation must frequently be used to establish the state levels. Such information always leads to a partially observable Markov decision process.

The jump algorithm appears fruitful because of the current limited knowledge in optimizing partially observable processes. It is entirely a new approach to obtain the invariant distribution for a continuous state Markov chain and use it for optimizing such processes. We observe that the structure of the invariant function derived in Section 3 enabled the development of a very simple algorithm.

The properties in Theorem (3.1), Lemma (3.3) and Lemma (3.4) played a significant role in the new algorithm.

This research has posed some recommendations that may be fruitful areas for future research. First of all, broadening the Markov matrix greater than two by two would be extremely useful in many applications. It would give the jump algorithm more flexibility in its application to real problems. Second, the computational time of the jump algorithm may also be reduced with a new search method using the structure of the invariant function. In addition, the analysis of the optimizing problem should be extended to include uncertainty about the process parameters.

References

[1] Sondik, E. 1971. The Optimal Control of Partially Observable Markov Processes. Ph. D. Dissertation, Stanford University.

[2] Sondik, E. 1978. The Optimal Control of Partially Observable Markov Processes Over the Infinite Horizon: Discunted Costs. Opns. Res. 26, 282-304.