ON THE DETERMINATION OF AN ON-DEMAND POLICY FOR A MULTILAYER CONTROL SYSTEM

Kiichiro Tsuji and Irving Lefkowitz

April 1975 WP-75-35

This paper is to be published in the IEEE Transactions on Automatic Control, Vol. 20, No. 4 (August 1975).

On the Determination of an On-Demand Policy for a Multilayer Control System

Abstract

The cost-performance tradeoff problem associated with a multilayer control system for controlling a class of static, nonlinear, multivariable systems is considered. The multilayer control system has a number of layers of control functions each of which updates different subsets of the manipulated variables at different costs.

A favorable cost-performance tradeoff is achieved by determining at each control decision time which subset of the control variables is to be updated. In this paper, we present a mathematical model which describes the operation of the multilayer control system. Also we show that the problem of determining a decision rule (policy) which results in an optimal cost-performance tradeoff can be formulated as a problem in Markovian Decision Processes. Consequently, an optimal policy can be identified by solving a linear program.

In order to reduce the computational effort required for identifying the optimal policy, a class of parameterized policies is introduced based on a measure of deviation of the disturbance. This approach provides a designer with a practical method of determining a control policy which achieves a favorable cost-performance tradeoff.

An example is given for demonstrating a possible application to process control.

I. INTRODUCTION

In many cases, industrial systems are subject to uncontrollable disturbances, and the performance of the system depends on these disturbances as well as on the variables which can be controlled. Due to changes in the disturbances, the implemented values for the control variables† at a particular instant of time (for example, the values of m which maximize a performance function $P(m,u)$ for a particular value of disturbance u) may not be very appropriate at some later time. In order to compensate for this, we consider a control strategy which updates the vector of control variables from time to time, responding to the observed changes in the disturbance. Since the complexity of the system may result in significant costs for computation, measurement and implementation each time an update is performed, the effectiveness of performing an update becomes very important. That is, there exists an economic tradeoff between the averaged performance achieved and the averaged cost of control over a long period of system operation. This tradeoff depends on the relative frequency of carrying out the updates and also on the structure of the control system. For instance, we would expect that more frequent updates would achieve a better averaged performance at the expense of a higher cost. Also, we expect that it would cost more, in general, to perform an optimization and control action with respect to all of the control variables than with respect to only a subset of them.

The general multilayer control system proposed by Donoghue[1] provides one way of incorporating the cost-performance tradeoff

---

† We use the term "control variable" to denote generally the output of a decision making or control process.

associated with the structure of the control system. In his general multilayer control system, he considers various control functions each of which affects only a subset of the control variables; the control variables are ordered and the control system is structured so that those variables to which the performance of the system is most sensitive tend to be updated more frequently and by simpler functions. The structure provides a basis for investigating the heuristic ideas a designer might consider with regard to the cost-performance tradeoffs as mentioned before.

In this paper, we present a mathematical model of both the multilayer control system and the cost-performance tradeoff for the case in which the control functions at the different layers in the system are based on updating different subsets of the control variables at different costs. A favorable cost-performance tradeoff is achieved by deciding at each control decision time which subset of the control variables is to be updated. We refer to this decision rule, in the subsequent sections, as an updating policy, or more simply, as a policy. Thus, the tradeoff problem is reduced to determining an updating policy which achieves a favorable cost-performance tradeoff.

In section IV, we show that the tradeoff problem can be formulated as a Markovian Decision Process and that an optimal policy can be obtained as a solution to a linear program. In section V, a class of parameterized policies are introduced and the design problem is reduced to determining a set of parameters. This approach may not lead us to an optimal solution. Nevertheless, it is considered to be practical for determining a policy which gives a favorable cost-performance tradeoff because of the significant reduction in computational effort.

II. Multilayer Control System

The structure of the multilayer control system whose cost-performance tradeoff is to be examined is shown in Fig. 1. The process performance is assumed to be given by P(m,u), where m is the vector of control variables and u is the vector of disturbances. The block G is the measurement and data processing unit where the set of raw data describing the disturbance input is transformed into an information vector (e.g., current observation u(t), mean and variance values, density function, etc.), which is denoted by θ. It is assumed that there is a pre-determined ordering† among the control

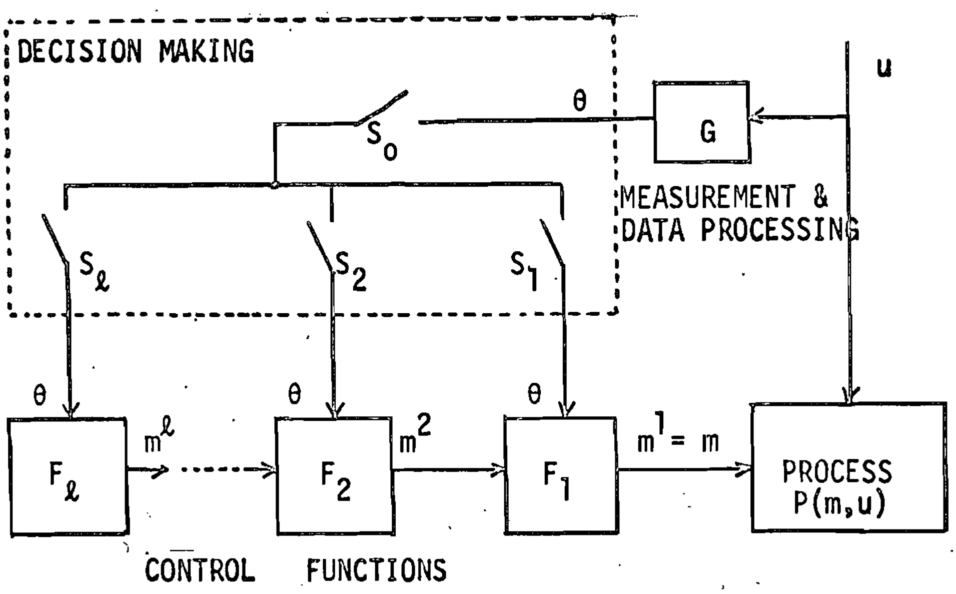

Fig. 1 Multilayer Control System

† The ordering is based on the sensitivity of system performance to each variable. If $i < j$, then system performance is more sensitive to m_i than m_j.

variables so that the following partitioned form of the vector m is given:

$$m = (m_1,\ldots,m_i,m_{i+1},\ldots,m_\ell). \tag{1}$$

Note that m_i may itself be a vector. Let m^i be a subset of m defined as follows:

$$m^i = (m_i,m_{i+1},\ldots,m_\ell) = (m_i,\ m^{i+1}). \tag{2}$$

The control function F_i, in general, determines the subset of m, i.e., m^i, based on the current value of θ and m^{i+1} which is the output of the control function F_{i+1}. However, we assume that F_i, in effect, changes only m_i, i.e., the i-th (partitioned)element of control variables. Therefore, we may express the relationship among the variables m^i, m^{i+1}, and θ as follows:

$$m^i = (m_i,\ m^{i+1}) = (\ f_i(\theta,m^{i+1}),\ m^{i+1}\ ) = F_i(\theta,m^{i+1}),$$

$$i=1,2,\ldots,\ell-1,$$

$$m^\ell = m_\ell = f_\ell(\theta) = F_\ell(\theta), \tag{3}$$

where f_i, $i=1,2,\ldots,\ell$ are given functions. As an example, f_i may be the result of a maximization operation on some performance index with respect to the indicated subset of m (see [1],[2]).

We also assume that there is a pre-determined period on which the operation of the control system is based, which will be referred to as the "basic period". In the decision making block, we assume the following: the information vector, θ, is made available to the control system every time S_0 is closed; F_i is performed every time S_i is closed. The decision maker determines which switches are to be closed every basic period of time. In the most general situation, the way the switches are closed is completely arbitrary, however, without loss of generality, we can restrict our attention to those decisions described below for convenience.

First, it is assumed that the decision as to which control function should be performed will be made only when S_0 is closed.

Second, we define the control actions as follows:

Control action i denotes that all of the control functions $F_i, F_{i-1}, \ldots, F_1$ are performed in this order, where $i=1,\ldots,\ell$. (4)

We identify the control action i simply by i, and define the set of alternative decisions Δ as follows:

$$\Delta = \{ 0,1,2,\ldots,\ell \} \quad (5)$$

where 0 denotes the decision that none of the control functions is to be performed. It should be noted, from the definitions (3) and (4), that control action i results in an update of only a subset of m, i.e., $m_1, m_2, \ldots, m_i$. It is assumed that once the subset of m is updated, then the values are kept constant until the next time of measurement and decision.

The following control costs are considered explicitly:

1) C_{GH} = Cost of measurement, data processing and decision making which is incurred every time S_0 is closed.

2) C_i = Cost of the control action i, which is incurred every time the control action i, $i \in \Delta$ is taken.

The cost C_i includes the cost of computing the new set of values for the control variables and the costs associated with implementing the results. We assume zero cost for no control action, i.e., $C_0=0$. After each decision making, the process produces a performance which is measured by the given function $P(m,u)$. Thus, the cost-performance tradeoff problem is now reduced to the problem of determining an updating policy (a decision rule which produces a sequence of integers from Δ), which gives a favorable balance between the performance achieved and the cost of control. It should be noted that the structure of the multilayer control system described here reflects the various ideas of the cost-performance tradeoff mentioned in section I.

III. Formalization of the Tradeoff Problem

In this section, we derive a mathematical model of the control system and define the cost-performance tradeoff problem explicitly.

Let the basic period be normalized to unity and let $t=0,1,\ldots$ be the discrete time index for the operation of the control system. Let $\mathcal{s}$ be an integer which represents the decision as to whether the switch S_0 is closed or not. That is,

$$\mathcal{s} = \begin{cases} 1, & \text{if } S_0 \text{ is closed,} \\ 0, & \text{otherwise.} \end{cases} \qquad (6)$$

Let δ be an integer which represents the control action δ, $\delta \in \Delta$. Let m^{Ψ} be the value of the vector of control variables after taking a control action. Then, there is a function Ψ such that

$$m^{\Psi} = \Psi(m,\theta,\mathcal{s},\delta),$$

where Ψ is defined as follows:

1) $(\mathcal{s}=0,\ \delta\in\Delta)$ or $(\mathcal{s}=1,\ \delta=0) \Rightarrow m_j^{\Psi} = m_j$, for all $j\in\{1,\ldots,\ell\}$.

2) $(\mathcal{s}=1,\ \delta\in\{1,\ldots,\ell-1\}) \Rightarrow$

$$m_j^{\Psi} = f_j(\theta,\ m_{j+1}^{\Psi},\ldots,m_{\ell}^{\Psi}) \quad \text{for all } j\in\{1,\ldots,\delta\},$$

$$m_j^{\Psi} = m_j, \qquad \text{for all } j\in\{\delta+1,\ldots,\ell\}.$$

3) $(\Delta=1, \delta=\ell) \Rightarrow \; m_j^\Psi = f_j(\theta, m_{j+1}^\Psi, \ldots, m_\ell^\Psi)$, for all $j \in \{1, \ldots, \ell-1\}$.

$$m_\ell^\Psi = f_\ell(\theta). \tag{7}$$

The table is derived by considering that control action δ implies the execution of $F_\delta, F_{\delta-1}, \ldots, F_1$ in this order, resulting in only a subset of m being updated. Supposing now that m, θ, Δ, m^Ψ, δ are functions of time and considering the assumption that the implemented value of m is kept constant over the basic period, the dynamic behaviour of the control system is described by

$$m(t+1) = m^\Psi(t) = \Psi(\, m(t), \theta(t), \Delta(t), \delta(t)\,), \; t=0,1,\ldots \; . \tag{8}$$

Now, we consider an expression for a measure of the cost-performance tradeoff associated with the control system operation. Let $\{u(t)\}$, $t=0,1,\ldots$ be a discrete time stochastic process representing the disturbance. Suppose at time t, the values $m(t)$, $\theta(t)$, $\Delta(t)$ and $\delta(t)$ are given and suppose the actual value of the disturbance is $u(t)$. Then, the performance actually achieved over the next basic period can be expressed as

$$W_t = P(m^\Psi(t), u(t)) - \Delta(t) \cdot (\, C_{GH} + C_{\delta(t)}\,), \; t=0,1,\ldots \; . \tag{9}$$

We will refer to this quantity as the net performance over the basic period. W_t, $t=0,1,\ldots$ is a sequence of random variables whose stochastic nature depends on $\{u(t)\}$, the initial value $m(0)$ and the policy denoted by H. An appropriate measure of the cost-performance tradeoff is represented by taking the expected average net performance over an infinite planning horizon, i.e., we define

Expected Average Net Performance over an Infinite

$$\text{Planning Horizon} = P_{net}(H, m(0)) = \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} E[W_t], \tag{10}$$

where E denotes the mathematical expectation. It is now possible to define our cost-performance tradeoff problem formally as follows.

Cost-Performance Tradeoff Problem:

$$\text{maximize} \quad P_{net}(H, m(0)) \tag{11}$$

where maximization is taken over all feasible policies and over all possible initial values for m. In the subsequent sections, we will show methods for solving this problem.

IV. Markovian Decision Process Approach

The cost-performance tradeoff problem formulated by (11) is a sequential decision process. One of the most powerful tools for sequential decision processes is the theory for Markovian Decision Processes which has been developed extensively over the last ten years.

In this section, we show that the behaviour of the multilayer control system can be described as a Markovian Decision Process by introducing an alternative state expression. Some assumptions are made for this purpose.

1) Measurement and decision making is performed every basic period of time, i.e., $\Delta(t)=1$, for all $t=0,1,\ldots$.
2) The vector of disturbance information $\theta(t)$ may take on only a finite number of possible values denoted by θ^i, $i=1,\ldots,N$, where N is the total number of possible values for $\theta(t)$. Alternatively, we will characterize the disturbance as the i-th disturbance level when $\theta(t) = \theta^i$. We denote the set of possible disturbance levels by S, i.e., $S = \{1,\ldots,N\}$.

3) The stochastic nature of the process $\{\theta(t)\}^{\dagger}$, $t=0,1,\ldots$ is Markovian and ergodic. Let the transition probability for the process $\{\theta(t)\}$, denoted by q_{ij}, $i,j \in S$, be given.

Under the assumption 2), let $x_i(t)$ represent the "level" of disturbance at the most recent time prior to t when control action i was performed, where $x_i(t) \in S$ for $i=1,\ldots,\ell$ and for $t=0,1,\ldots$. Then it can be shown that the vector $X(t) = (x_0(t), x_1(t), \ldots, x_\ell(t))$ may be considered as the "state" of the multilayer control system. Here, $x_0(t)$ represents the present level of the disturbance at time t. From the assumptions 2) and 3), $X(t)$ takes on only a finite number of possible values and, therefore, the state space $\hat{S}$ is finite, i.e., $\hat{S} = S^{\ell+1}$.

In terms of this state expression, the operation of the control system under the assumptions 1) through 3) can be described as follows. Suppose at some time t, the state of the control system is found to be $X = (x_0, x_1, \ldots, x_\ell)$, where the argument t is suppressed for notational convenience. If the decision maker now selects a control action, say δ, then X changes immediately to some new state X^ϕ because, by definition, the first $(\delta+1)$ elements of the state vector X are reset to the present level of the disturbance. That is, there exists a function†† ϕ defined as follows which expresses the results of control actions.

$$X^\phi = (x_0^\phi, x_1^\phi, \ldots, x_\ell^\phi) = \phi(X,\delta), \text{ where}$$

$$x_j^\phi = \begin{cases} x_0, & j=0,1,\ldots,\delta \\ x_j, & j=\delta+1,\ldots,\ell, \quad X \in \hat{S}, \qquad \delta \in \Delta. \end{cases} \tag{12}$$

---

† $\{\theta(t)\}$ is the process induced from $\{u(t)\}$ through G.

†† This function ϕ corresponds to the function Ψ defined by (7) where $\delta = 1$.

The state may then move to some new state due to a change in the current disturbance level.

Note that the last ℓ elements of the state vector $X(t)$ contain sufficient information for determining the values of m uniquely. That is, for each component x_j, $j=1,2,\ldots,\ell$, there is a corresponding value of the information vector which can be denoted as θ^{x_j}, and it can be shown that there exists a function g such that $m(t) = g(\theta^{x_0(t)},\theta^{x_1(t)},\ldots,\theta^{x_\ell(t)})$ for all t, where $m(t)$ is the state of the control system defined by (8). The function g can be easily derived from (7).

We may consider, in general, the transition from a state $X\in\hat{S}$ to a state $Y\in\hat{S}$. Then, the behaviour of the control system can be described by the following diagram.

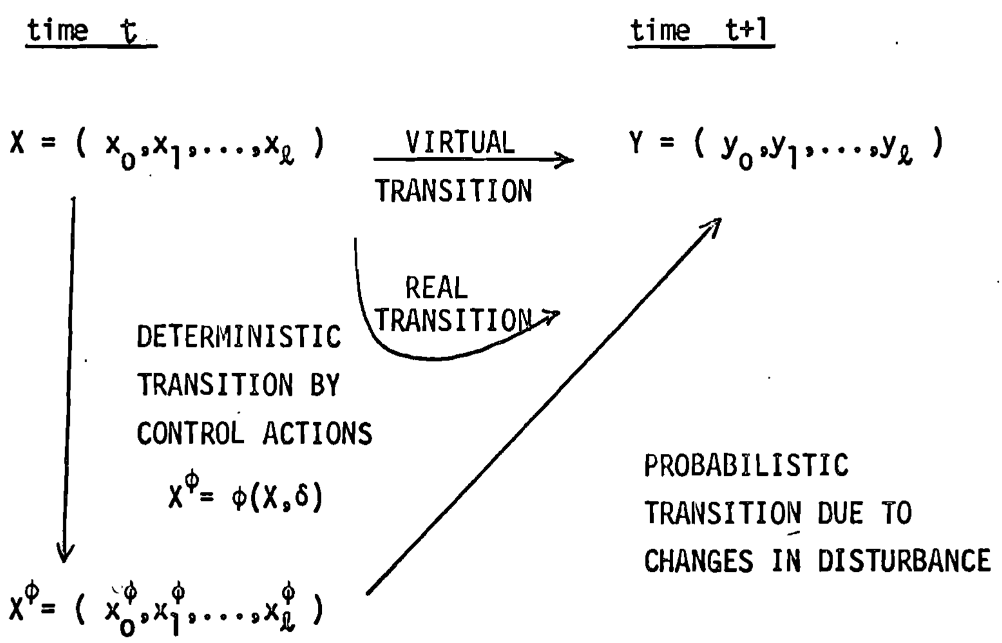

Fig. 2 Transition Diagram

Fig. 2 shows that the operation of the control system is a Markovian Decision Process explicitly. The transition probability to state Y given that the present state is X and the control action is δ, denoted by P^{δ}_{XY}, can be expressed by

$$P^{\delta}_{XY} = \begin{cases} q_{x_o y_o}, & \text{if } y_j = x^{\phi}_j, \text{ for all } j \in \{1, \ldots, \ell\}. \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

where $X^{\phi} = \phi(X, \delta)$, as defined in (12) and $\delta \in \Delta$, $X, Y \in \hat{S}$.

The cost associated with the transition from state X to state Y by taking control action δ, denoted by r^{δ}_{XY}, can be expressed by the following equation:

$$r^{\delta}_{XY} = \sum_{k=1}^{K} P(m^{X^{\phi}}, u^k) \cdot \eta^{x_o}_k - (C_{GH} + C_{\delta}), \tag{14}$$

where X^{ϕ} represents the state right after control action δ, and $m^{X^{\phi}}$ represents the unique numerical value for the control variable vector corresponding to the state X^{ϕ}. $\eta^{x_0}_k$ represents the discrete probability corresponding to the present level of the disturbance x_0, i.e.,

$$\eta^{x_0}_k = \text{Prob}\{ u(t) = u^k \mid \theta(t) = \theta^{x_0} \}. \tag{15}$$

Note that if the information vector θ provides the numerical value of u directly, then (14) reduces simply to

$$r^{\delta}_{XY} = P(m^{X^{\phi}}, u^{x_o}) - (C_{GH} + C_{\delta}). \tag{16}$$

The net performance W_t given by (9) becomes

$$W_t = r^{\delta(t)}_{X(t), X(t+1)}, \tag{17}$$

and the expected average net performance over an infinite planning horizon can be written as

$$P_{net}(H, X(0)) = \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} E[W_t]$$

$$= \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \sum_{Y\in \hat{S}} \sum_{k\in\Delta} r_Y^k \, \text{Prob}(X(t)=Y,\delta(t)=k;H,X(0)), \quad (18)$$

where

$$r_Y^k = \sum_{Y'\in \hat{S}} r_{YY'}^k \, P_{YY'}^k \quad . \qquad (19)$$

Thus, the problem is to determine a policy which maximizes P_{net} over all possible policies and over all initial states.

It is well known that if a Markovian Decision Process is completely ergodic, then the linear programming formulation proposed by Manne[3] determines an optimal policy. When a Markovian Decision Process is not completely ergodic, Manne's linear program only identifies an ergodic chain which gives an optimal expected average criterion[4]. It can easily be shown that the tradeoff problem is not completely ergodic. However, it is enough to identify only an optimal ergodic chain, because for any state $X\in\hat{S}$, we can calculate the value of $\bar{m}$ a priori and therefore we can always set the state of the multilayer control sytem to one of the states in the identified ergodic chain. Thus, Manne's linear program is considered to be appropriate and it is shown in the following:

$$\text{Find } \{\pi_Y^\delta\} \text{ to maximize } \quad \sum_{Y\in\hat{S}} \sum_{\delta\in\Delta} r_Y^\delta \, \pi_Y^\delta \qquad (20)$$

subject to

$$\sum_{\delta\in\Delta} \pi_Y^\delta - \sum_{X\in\hat{S}} \sum_{\delta\in\Delta} P_{XY}^\delta \pi_X^\delta = 0, \quad Y \in \hat{S}, \tag{21}$$

$$\sum_{\delta\in\Delta} \sum_{Y\in\hat{S}} \pi_Y^\delta = 1, \qquad \pi_Y^\delta \geq 0, \ \forall\, \delta\in\Delta, \ Y\in\hat{S}. \tag{22}$$

The objective function represents the quantity given by (18). The constraints are derived from the familiar steady state relationships between the state distribution and the set of transition probabilities in Markov chains where P_{XY}^δ is given by (13). It can be shown that for each $Y\in\hat{S}$, at most one $\delta\in\Delta$ such that $\pi_Y^\delta>0$, will appear in every basic feasible solution of the above program. Therefore, an optimal policy can be identified by taking the pairs (Y,δ) such that $\hat{\pi}_Y^\delta > 0$, where $\hat{\pi}_Y^\delta$ is an optimal solution of the linear program.

V. Parameterized Policy Approach

The linear program (20)$\sim$(22) identifies an optimal ergodic chain whose expected average net performance is a maximum. However, the size of the linear program may become very large as N and ℓ increase, because the number of rows is given by $N^{\ell+1}$ (which is the same as the number of possible states). Therefore, a technique which requires much less computational effort is desired. In this section, we develop a method which determines an approximation to an optimal policy with considerably less computational effort.

Parameterized Policy

Let ρ be a function which represents a measure of difference between two disturbance levels, i.e.,

$$\rho(a,b) \geq 0, \text{ equality holding if } a=b.$$
$$\rho(a,b) = \rho(b,a), \quad a,b \in S. \tag{23}$$

It should be noted that the number of possible values that ρ can take on is finite because S is finite. We refer to this function as a testing function. An example of ρ is $\rho(a,b) = (a-b)^T \beta (a-b)$, where $\beta > 0$ is a weighting matrix. The idea of a parameterized policy is as follows: Let a be the disturbance level at the last time of update and let b be the present disturbance level. Then, using the function ρ, we can consider the following updating rule:

$$\begin{array}{lll} \text{update} & \text{if} & \rho(a,b) \geq \alpha \\ \text{do not update} & \text{if} & \rho(a,b) < \alpha \end{array} \tag{24}$$

where $\alpha > 0$ is some real number which is referred to as the testing criterion. In other words, an update will be carried out only if the value of the testing function equals or exceeds a certain prescribed limit.

We can extend this idea to the general multilayer control system under consideration. Now, let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be the vector of testing criteria where α_i is the criterion for control action i. Then a parameterized policy $H\alpha$ can be defined by

$$H\alpha(X) = \begin{cases} 0, & \text{if } \rho(x_0,x_j) < \alpha_j, \quad j=1,2,\ldots,\ell. \\ k, & \text{if } \rho(x_0,x_k) \geq \alpha_k \text{ for some } k \in \{1,\ldots,\ell-1\} \\ & \text{and if} \\ & \quad \rho(x_0,x_j) < \alpha_j \text{ for all } j \in \{k+1,\ldots,\ell\}. \\ \ell, & \text{if } \rho(x_0,x_\ell) \geq \alpha_\ell. \end{cases}$$

where $X = (x_0, x_1, \ldots, x_\ell) \in \hat{S}$. (25)

The parameterized policy determines the control action as follows: starting from control action ℓ, we calculate the value of the testing function. If the value equals or exceeds α_ℓ, control action ℓ is to be performed. If the value is less than α_ℓ, then we evaluate the testing function again for control action ℓ-1. In this way, the control action to be taken is determined as the first one (counting from control action ℓ) whose value of the testing function is not less than its testing criterion. Thus, this class of parameterized policies assigns a unique control action to each state in $\hat{S}$, and gives priority to the higher layer control actions. Note that the number of possible parameterized policies is finite for given ρ and ℓ.

As an example, consider a 2-layer control system in which the disturbance takes on only three values, i.e., S= {1,2,3}. A state for this case is represented by the vector (x_0,x_1,x_2), where $x_i \in S$, i=0,1,2. Let ρ be chosen as

$$\rho(a,b) = |a - b| , \quad a \in S, \quad b \in S. \tag{26}$$

Then, according to (25), the control action for the state (x_0,x_1,x_2) is determined by

$$H\alpha(x_0,x_1,x_2) = \begin{cases} 0 , & \text{if } |x_1-x_0|<\alpha_1 \text{ and } |x_2-x_0| < \alpha_2 \\ 1 , & \text{if } |x_1-x_0|\geqq\alpha_1 \text{ and } |x_2-x_0| < \alpha_2 \\ 2 , & \text{if } |x_2-x_0|\geqq\alpha_2, \end{cases} \tag{27}$$

where $\alpha = (\alpha_1, \alpha_2)$.

For instance, if the policy vector α =(1,2), then the control actions associated with the states (1,1,1), (1,2,2), (1,2,3) are given by 0,1, 2, respectively.

Having defined the class of parameterized policies, our next concern is to develop a method for evaluating the expected average net performance (P_{net}) for a given parameterized policy. Based on the virtual transition illustrated in Fig. 2, the states of the multilayer control system operating under the class of parameterized policies form a Markov Chain which, in general, may contain a number of ergodic chains depending on the initial starting state X^*. Let us denote the ergodic chain for given α and X^* by $A(\alpha,X^*)$†. Then the expected average net performance over an infinite horizon defined by (18) exists and is given by

$$P_{net}(\alpha,X^*) = \sum_{X \in A(\alpha,X^*)} r_X^{H\alpha(X)} \pi_X(\alpha,X^*), \tag{28}$$

where r_X^{δ} is defined by (19) and $\pi_X(\alpha,X^*)$ represents the steady state probability that the state of the multilayer control system is X, and is given as a solution to the following set of equations.

$$\pi_Y(\alpha,X^*) - \sum_{X \in A(\alpha,X^*)} P_{XY}^{H\alpha(X)} \pi_X(\alpha,X^*) = 0, \; Y \in A(\alpha,X^*) \tag{29}$$

$$\sum_{X \in A(\alpha,X^*)} \pi_X(\alpha,X^*) = 1. \tag{30}$$

Thus, the tradeoff problem is reduced to determining the set of values α and the initial state X^* so that P_{net} is maximized over all possible α and X^*.

For given α and X^*, P_{net} can be calculated by first identifying $A(\alpha,X^*)$ and then solving (29) and (30) for $\pi_X(\alpha,X^*)$. Since the number

† Since an arbitrary initial starting state may be a transient state and may lead to two or more distinct ergodic chains, the range of A must be appropriately restricted.

of states contained in the ergodic chain $A(\alpha, X^*)$ may be very large, solving (29) and (30) directly may still require an excessive computational effort.

The amount of computation can be reduced by considering an alternative "virtual transition". Consider the state transition of the multilayer control system under the class of parameterized policies as shown in Fig. 3, where X'_t, X'_{t+1} denote the states before taking control actions, and where X_t, X_{t+1} denote the states immediately after taking control actions.

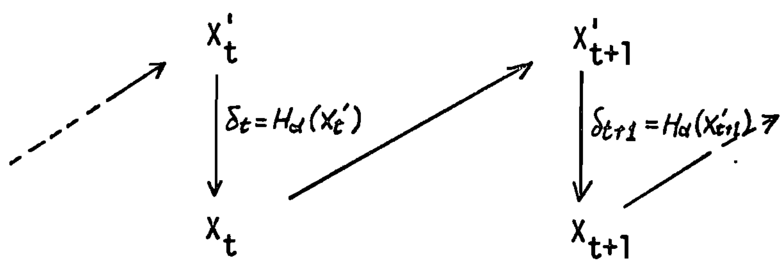

Fig. 3 State Transition under a Parameterized Policy

Since the control actions associated with the states X'_t, X'_{t+1} are determined by the given parameterized policy, we may consider the virtual transition $X_t \to X_{t+1}$ instead of $X'_t \to X'_{t+1}$ †. Then, it is clear that for each ergodic chain $A(\alpha, X^*)$ generated by the transition $X'_t \to X'_{t+1}$, there is an ergodic chain (we denote this chain by $A^s(\alpha, X^*)$) generated by the virtual transition $X_t \to X_{t+1}$. Based on the chain $A^s(\alpha, X^*)$, we can show that (28)~(30) may be modified to the following set of equations from which the averaged performance P_{net} can be determined.

† Note that this new interpretation of virtual transition is not appropriate in the formulation of the Markovian Decision Process developed in section IV.

$$P_{net}(\alpha,X^*) = \sum_{X\in A^S(\alpha,X^*)} r_X(\alpha)\ \mu_X(\alpha,X^*) \tag{31}$$

where

$$r_X(\alpha) = \sum_{y_0\in S} \{ \sum_{k=1}^{K} P(m^X,u^k)\eta_k^{x_o} - (C_{GH} + C_{H\alpha(Y)})\}\ q_{x_o y_o} \tag{32}$$

where x_o, y_o are the first elements of X,Y, respectively, and $\mu_X(\alpha,X^*)$ is given as a solution to the equations

$$\mu_Y(\alpha,X^*) - \sum_{X\in A^S(\alpha,X^*)} P^S_{XY}(\alpha)\ \mu_X(\alpha,X^*) = 0, \quad Y\in A^S(\alpha,X^*) \tag{33}$$

$$\sum_{X\in A^S(\alpha,X^*)} \mu_X(\alpha,X^*) = 1, \tag{34}$$

where $P^S_{XY}(\alpha)$ denotes the transition probability between states X and Y in $A^S(\alpha,X^*)$. $P^S_{XY}(\alpha)$ is given by

$$P^S_{XY}(\alpha) = \begin{cases} q_{x_o y_o}, & \text{if the last } \ell\text{-elements of } Y \text{ and } X'^{\phi}, \\ & \text{respectively, are identical, where} \\ & X'^{\phi} = \phi(X', H_\alpha(X')), \text{ and where} \\ & X' = (y_0, x_1, x_2, \ldots, x_\ell). \\ 0, & \text{otherwise.} \end{cases} \tag{35}$$

In the expression (35), the function ϕ is defined by (12) in section IV.

The advantage of this formulation is the fact that the number of states in A^S is usually much less than the number of states in A (See Appendix). Hence, it requires much less computational effort to use (31) through (35) instead of (28) through (30).

Search Method

The discussion so far is concerned with the evaluation of $P_{net}(\alpha,X^*)$ for given α and X^*. In order to obtain the best parameterized policy, a search procedure must be considered. Since P_{net} is not continuous with respect to its arguments, we have to rely on so-called direct search methods, such as Hooke and Jeeves[5], Rosenbrock[5], Nelder and Mead[5]. For each index state X^*, these methods may be applied to determine the best value for α. The best parameterized policy is then determined by taking the best combination of α and X^*.

It should be noted here that due to a special property of the multi-layer control system operating under parameterized policies, the range of α for which an optimal value is searched can be restricted. We call this the "Collapsing Property" and it is described in the following.

Collapsing Property

Let us consider a two-layer control system as in section V. The parameterized policy is defined by (27), and let us assume that $\alpha_1 \geqq \alpha_2$. Suppose at some instant of time, control action 2 was taken. Immediately follows the control action, the state satisfies the condition $x_0=x_1=x_2$, because control action 2 has reset the values of x_1 and x_2 to the current disturbance level x_0. Therefore, at the next time testing, we observe that $|x_1-x_0|=|x_2-x_0|$. However, since $\alpha_1 \geqq \alpha_2$, $|x_1-x_0| \geqq \alpha_1$ always implies $|x_2-x_0| \geqq \alpha_2$. Thus, the next control action cannot be control action 1. By repeating this argument for each time of testing, we can show that control action 1 will never be carried out under the condition $\alpha_1 \geqq \alpha_2$. As a result, the two-layer control system, in effect, "collapses" to a one-layer control system for which only control action 2 is implemented.

The above property is valid in the general ℓ-layer control system. That is, if $\alpha_i \geqq \alpha_{i+1}$ for some i, then control action i will not be implemented.

Suppose a given parameterized policy $H\alpha$ has a testing criterion α_i ,α_{i+1} such that $\alpha_i \geqq \alpha_{i+1}$. Then, the collapsing property quarantees that we can obtain an equivalent policy by replacing α_i by α_{i+1} . Therefore, we conclude that it is sufficient to consider only the values of α which satisfy

$$\alpha_1 \leqq \alpha_2 \leqq \cdots\cdots \leqq \alpha_\ell, \tag{36}$$

and this reduces the range of α to be searched considerably.

VI. An Example

In order to demonstrate a possible application of the multilayer control approach, the static control of a simple stirred tank reactor process is considered.

The reactor process has been studied in [7] from which the description is taken. The inflow to the reactor contains two components R_X and R_Y with concentrations γ_{X_0} and γ_{Y_0} , respectively. The outflow contains components R_X, R_Y and R_Z with concentrations γ_X, γ_Y, and γ_Z , respectively. The only reactions taking place in the reactor are given by

$$R_X \xrightarrow{K_1} R_Y \xrightarrow{K_2} R_Z \tag{37}$$

where K_i are the reaction rate coefficients

$$K_i = \frac{V}{q} A_i \, e^{-B_i/T} , \quad i=1,2, \tag{38}$$

and T is the temperature, and V is the volume of the reaction mixture, q is the throughput rate and A_1, A_2, B_1 and B_2 are given constants. The steady state relationships are given by

$$\begin{aligned} &\gamma_{X_0} - \gamma_X (1+K_1) = 0 \\ &\gamma_{Y_0} - \gamma_Y (1+K_2) + K_1\gamma_X = 0 \\ &K_2\gamma_Y - \gamma_Z = 0 \\ &\gamma_{X_0} + \gamma_{Y_0} = 1 \\ &\gamma_X + \gamma_Y + \gamma_Z = 1. \end{aligned} \tag{39}$$

The measure of system performance is taken to be

$$\begin{aligned} y(t) = \beta_1 \gamma_Y(t) q(t) - \langle \beta_2 K_2(t)\gamma_Y(t) - \beta_3 \rangle^2 \\ -\beta_4 T^4(t) - \beta_5 T(t) q(t) - (\beta_6 q(t) - \beta_7) \end{aligned} \tag{40}$$

where β_i, i=1,2,...,7 are given constants and

$$\langle \xi \rangle = \begin{cases} \xi, & \xi > 0 \\ 0, & \xi \leq 0. \end{cases} \tag{41}$$

The first term in (40) is the value of the desired product R_Y, the second term is a loss due to the high concentration of the side product R_Z, the third term represents the heat loss due to radiation, the fourth term is the cost of heating the mixture, the fifth term is the cost associated with the input stream.

The volume V and the temperature T are considered as the control variables and the throughput rate q is considered to be the disturbance. Using the relationships (39), the measure of system performance (40) can be expressed in terms of only V, T and q, i.e.,

$$y(t) = P(V(t), T(t), q(t)). \tag{42}$$

If the throughput rate q is a constant, then the values for V and T can be chosen so as to maximize (42), and they can be kept constant throughout the operation of the system. When q varies with respect to time, however, the values for V and T should be updated in order to keep the system in its best operating condition. In this example, we assume that there are costs associated with updating V and/or T and the tradeoff problem considered in the previous sections becomes important.

We consider a two-layer control system where the first layer updates T and the second layer updates V (See Fig. 4), and define control actions 1 and 2 according to (4) as follows:

Control action 1 = Calculate the value of T so that P(V,T,q) is maximized with respect to T and implement the result on the system.

Control action 2 = Calculate the values of both T and V so that P(V,T,q) is maximized with respect to T and V, and implement the result on the system.

We assume that there are costs C_1 and C_2 associate with control actions 1 and 2, respectively. For simplicity, q(t) is assumed to be measured directly with no cost(this implies that G is an identity and C_{GH} in section II is zero).

A sample record for q is assumed to be given as in Fig. 5. Since q is continuous valued, it is necessary to quantize the value of q in order to apply the techniques in the previous sections. The number of disturbance levels N is chosento be 3 and the quantum values for each level are determined by the method suggested in [7] and the transition probabilities are calculated from this sample record. These results are shown in Table 1.

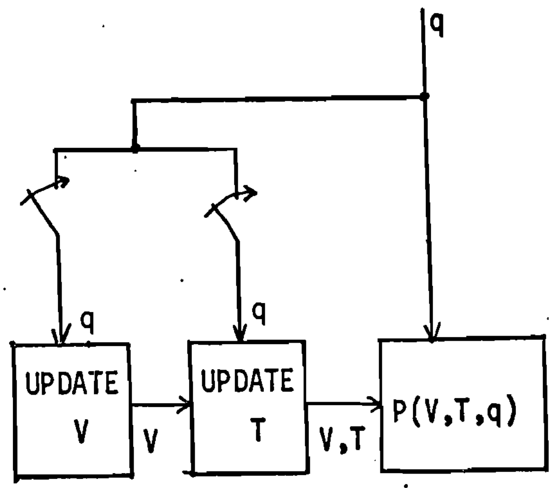

Fig. 4 Two Layer Control System

DISCRETIZATION OF DISTURBANCE

DISTURBANCE LEVEL	RANGE R_i, $i \in \{1,2,3\}$	QUANTUM VALUES u^i, $i \in \{1,2,3\}$
1	$q \leq \mu - 0.432\sigma = 10.78$	$u^1 = \mu - 0.969\sigma = 10.52$
2	$10.78 < q \leq \mu + 0.432\sigma = 11.19$	$u^2 = \mu = 10.98$
3	$q > 11.19$	$u^3 = \mu + 0.969\sigma = 11.45$

TRANSITION PROBABILITY: $q_{ij} = \text{Prob}\{u(t+1) \in R_j \mid u(t) \in R_i\}$

$$\begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} = \begin{bmatrix} 0.904 & 0.096 & 0.000 \\ 0.095 & 0.833 & 0.072 \\ 0.000 & 0.068 & 0.932 \end{bmatrix}$$

Table.1 Discretization of Disturbance Record

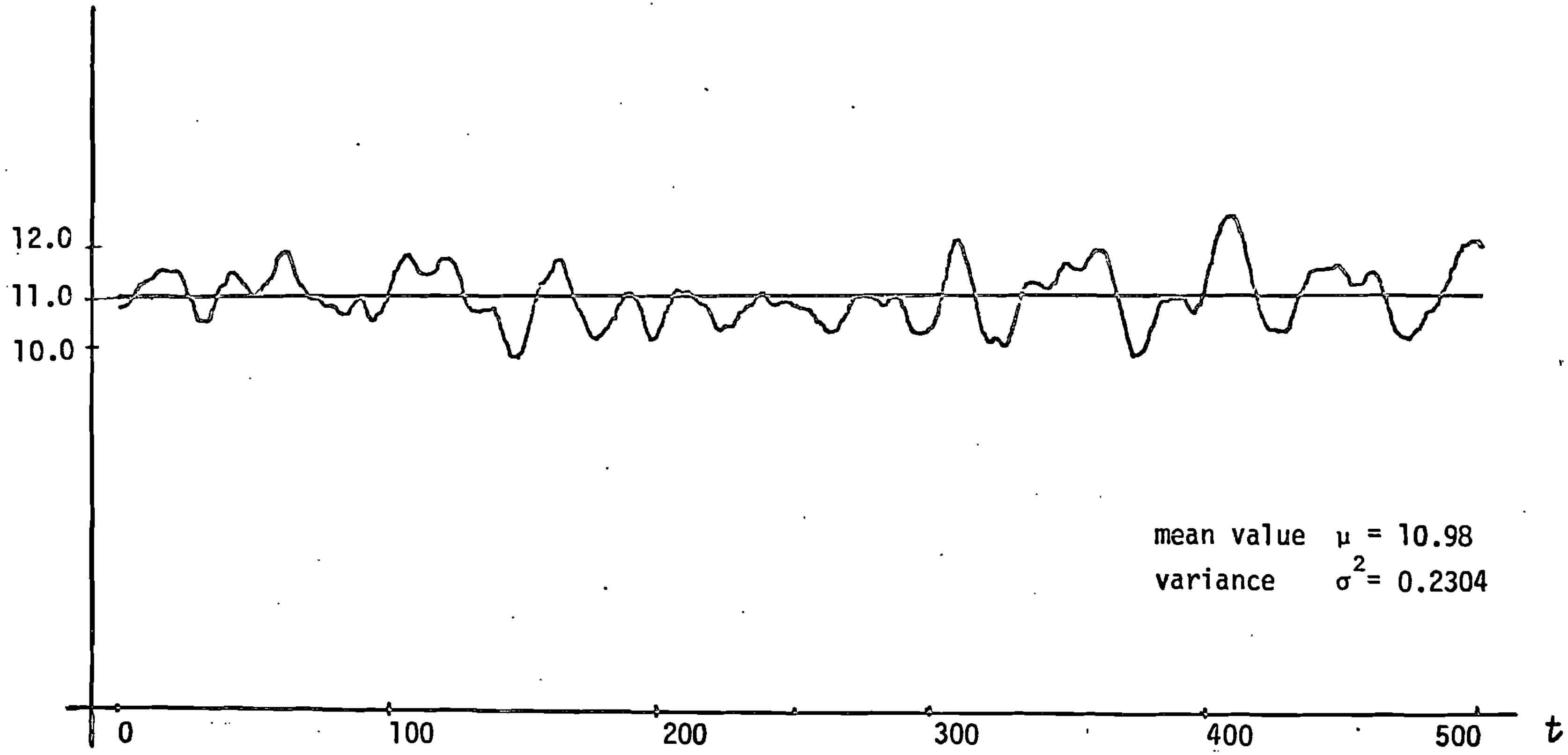

Fig. 5 Disturbance Sample Record

The state of the two-layer control system is expressed by $X = (x_0, x_1, x_2)$, where x_0 is the present disturbance level, x_1 is the disturbance level at most recent time of control action 1, and x_2 is the disturbance level at most recent time of control action 2, where $x_i \in \{1,2,3\}$, $i=0,1,2$. There are 27 states in the state space. The net performance associated with the transition from state X to Y taking control action δ, i.e., r_{XY}^{δ} is approximated by the following expression:

$$r_{XY}^{\delta} = \frac{1}{2}\left[P(V^{X^\phi}, T^{X^\phi}, q^{x_0}) + P(V^{X^\phi}, T^{X^\phi}, q^{y_0})\right] - C_\delta , \qquad (43)$$

where $X^\phi = \phi(X,\delta)$ as in (12) and V^X is the result of performing

$$\underset{V,T}{\text{Maximize}} \quad P(V,T,q^{x_2}), \qquad (44)$$

and T^X is the result of performing

$$\text{Maximize} \quad P(V^X,T,q^{x_1}). \qquad (45)$$

The following numbers are used in the example.

$A_1=14000$, $A_2=80$, $B_1=4000$, $B_2=2500$, $\gamma_{Y_0}=0.1$, $\gamma_{X_0}=0.9$,

$\beta_1=25000$, $\beta_2=5.0\times10^6$, $\beta_3=5.0\times10^{-3}$, $\beta_4=3.0\times10^{-6}$, $\beta_5=5.0$,

$\beta_6=3590$, $\beta_7=26500$;

$C_1=50$, $C_2=200$, $C_0=0$, $C_{GH}=0$.

The linear program (20)~(22) was set up using these numbers. An optimal policy was determined by taking those variables in the optimal basic feasible solution whose values are strictly positive. This procedure resulted in identification of the following policy:

State	Optimal Action	State	Optimal Action
(1,1,1)	0	(1,2,3)	2
(2,1,1)	1	(2,2,3)	0
(1,2,1)	1	(3,2,3)	1
(2,2,1)	0	(2,3,3)	1
(3,2,1)	2	(3,3,3)	0

The parameterized policy approach in section V was also applied to the example. Here, we used the testing function (26) and the policy is determined by (27). The best parameterized policy identified was $\alpha = (1,2)$ with the initial state (1,1,1). This policy happened to be exactly the same as the policy identified by solving the linear program. This may not always be true in general, because an optimal policy may not be expressed in the form of a parameterized policy.

VII. Conclusions

A general multilayer control system is developed which improves the balance between the control cost and the performance achieved for a class of static, nonlinear, multivariable systems. The control system is formalized as a generalization of the multilayer control approach. A convenient mathematical model describing the behaviour of the control system is obtained which admits a simple state expression. The problem of determining an updating policy is then shown to be formulated as a Markovian Decision Process under some assumptions. Consequently, a policy which is optimal over all possible policies can be identified as a solution to a linear program. Since the computations of an optimal policy become quite tedious, a parameterized policy approach is proposed, which results in an identification of a suboptimal control policy with much less computational effort.

It should be noted that the development described in this paper did not really take the effect of the measurement cost on the expected average net performance into account. This follows from the fact that we have made the assumption that measurement and decision making are performed every basic period (refer to assumption 1) in section IV). However, it is rather straightforward to extend both the Markovian Decision Process Approach and the Parameterized Policy Approach to the case where the effect of C_{GH} on P_{net} is significant [2]. The key is to include the interval of two successive measurements into the set of decision alternatives. This results in defining a Markovian Decision Process similar to the one discussed in [8].

The above investigation provides an extension of the multilayer control strategy in Donoghue's development, and also formalizes an important notion of controlling on-demand (i.e., controlling only when it is economically worthwhile to do so) for the class of static systems.

Some of the important questions such as the optimality of the best parameterized policy, the treatment of non-Markovian disturbance and extensions to the class of dynamic systems need further investigation.

Appendix

The superiority of using the ergodic chain $A^S(\alpha,X^*)$ instead of using $A(\alpha,X^*)$ is illustrated in Table 2 in which the numbers of states in $A^S(\alpha,X^*)$ and in $A(\alpha,X^*)$ are compared for some values of α and X^*. Note also that this table shows an indication of the computational reduction in the Parameterized Policy Approach, because, in this example, the number of rows in the linear program (20)~(22) is given by 625.

Table 2 Comparison of the number of states in A and A^S
3-layer example (ℓ=3). Number of disturbance levels N=5.
$\{q_{ij}\}$ is given by

$$\begin{bmatrix} 0.5 & 0.3 & 0.2 & 0.0 & 0.0 \\ 0.2 & 0.6 & 0.2 & 0.0 & 0.0 \\ 0.3 & 0.4 & 0.1 & 0.2 & 0.0 \\ 0.0 & 0.0 & 0.4 & 0.1 & 0.5 \\ 0.0 & 0.0 & 0.0 & 0.6 & 0.4 \end{bmatrix}$$

Initial State X^* / Policy α	The number of states in			
	$A(\alpha,X^*)$		$A^S(\alpha,X^*)$	
	(1,1,1,1)	(3,3,3,3)	(1,1,1,1)	(3,3,3,3)
(1,1,1)	15	15	5	5
(1,2,3)	29	21	9	7
(1,2,4)	31	21	10	7
(1,3,4)	39	15	12	5
(2,3,4)	33	11	16	7

References

1. Donoghue,J.F., Lefkowitz,I., "Economic Tradeoffs Associated with a Multilayer Control Strategy for a Class of Static Systems", IEEE, Trans. on AC, Vol. AC-17, No.1, Feb. 1972.

2. Tsuji,K., "Determination of On-Demand Policy for a Multilayer Control System with Cost-Performance Tradeoff Considerations", Ph.D. Dissertation, Case Western Reserve University, June 1973.

3. Manne,A.,"Linear Programming in a Markov Chain", Mgmt. Sci., Vol.6, 1960.

4. Denardo, E.V.,"On Linear Programming in a Markov Chain", Op. Res., Vol.10, 1962.

5. Kowalik and Osborn,"Methods for Unconstrained Optimization Problems", Elsevier, 1968.

6. Lefkowitz, I.,"Multi-Level Approach Applied to Control System Design", Trans. ASME, 88, 1966.

7. Burghart, J.H., Lefkowitz,I., "A Technique for On-Line Steady-State Optimizing Control", IEEE,Trans. on SCC, Vol. SCC-5, No.2, April 1969.

8. Klein, M.,"Inspection-Maintenance-Replacement Schedules under Markovian Deterioration", Mgmt. Sci., Vol.9, 1962.