STOCHASTIC CONTROL OF LARGE MARKOV CHAINS

by

Jean-Pierre Forestier and Pravin Varaiya

Memorandum No. UCB/ERL M77/33

17 May 1977

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

To be presented at IFAC Workshop on Control and Management of Integrated Industrial Complexes, Toulouse, France, September 1977.

STOCHASTIC CONTROL OF LARGE MARKOV CHAINS

Jean-Pierre Forestier and Pravin Varaiya
C.E.R.T./D.E.R.A., 2 Av. E. Belin, 31400 Toulouse, France and
Department of Electrical Engineering and Computer Sciences,
University of California, Berkeley, California 94720 U.S.A.

ABSTRACT

We study two problems of control of large Markov chains, and present a different procedure for each of them. The computational burden associated with the large number of states is addressed in the first procedure by a two-level controller, and in the second by a two-layer controller.

INTRODUCTION

While considerable effort has been recently directed at inventing computationally attractive procedures for large deterministic control problems, the comparative effort devoted to large stochastic problems is minute, and most of this has been concentrated on the LQG problem [Ref. 1]. Moreover the experience gained with the LQG problem is unlikely to be "transferable" to other formulations. This paper is concerned with one of these other formulations viz., the control of large, finite state, Markov chains. "Large" means simply that the standard procedures for finding optimal control laws or strategies are too cumbersome and so a practical procedure involving some kind of approximation or decomposition is needed.

We present two such procedures: one uses a two-level decomposition method; the other uses a somewhat novel idea which we have named boundary control or control by exception. As will be seen later the second method can also be regarded as a two-layer controller.[1]

The only paper we know of which also deals with large Markov chains is by Kushner and Chen [Ref. 2]. Starting with the fact [Ref. 3, p. 152] that if the number of controls, as well as the number of states, is finite, then the problem of finding the optimal strategy can be formulated as a linear program (LP), they observe that if the transition probability matrix has a certain form then the LP can be solved by the Dantzig-Wolfe decomposition algorithm. The LP formulation is, generally speaking, not computationally attractive since the number of variables is sxN where s is the number of states and N is the number of controls.

TWO-LEVEL CONTROL

We have in mind the following situation. The state of the Markov chain represents the amount of available resources at any given time. At each time demands are made randomly and some of the resources must be diverted to satisfy them. If too many resources are diverted the current demand can be easily met but future demands cannot. The reverse happens if too few resources are devoted to the current demand. The resources are "renewable" in the sense that once the demand is fulfilled they are once again available. Thus the resources can be considered to be installed plant and equipment capacity or a constant labor pool.

Problem Formulation

Consider a request or demand process r_t, $t = 0,1,\ldots$ with values in a set R different points of which designate different types of request for service. Also consider a state of service process s_t with values in $S = \{1,\ldots,s\}$. Different points in S signify different amounts of resources available for servicing a request. The request process affects the state of service as follows. Suppose the latter is in state $s_t = i$ when the request $r_t = r$ is received. Then a decision must be taken assigning some of the resources indicated by s_t to service r_t. Denote the decision taken by $u_t = u(s_t,r_t)$ and let $U(i,r)$ be the set of all possible decision which can be taken in the condition i,r. u_t affects the next state of service according to

[1] To avoid any misunderstanding we emphasize that we are not concerned here with decentralized control as the term is used in [Ref. 1]. We have in mind a single controller who has access to all the available information.

$$\text{Prob}\{s_{t+1} = j \mid s_t = i, r_t = r, u_t\} = f(i,j,u_t,r) \tag{1}$$

where f is known. The cost associated with this decision is specified by the function $m(i,u_t,r)$. Associated with a strategy $u = \{u(i,r)\} \in U = \chi\{U(i,r) \mid (i,r) \in S \times R\}$ is the long-run average cost

$$J(u) = \lim_{T\to\infty} \frac{1}{T+1} \sum_{0}^{T} Em(s_t,u(s_t,r_t),r_t). \tag{2}$$

To make this expression meaningful two assumptions are made.

A_1. r_t is a stationary, independent process. Denote the distribution of r_t by P(dr). This guarantees that if u is the strategy adopted then s_t becomes a Markov chain with the stationary probability transition matrix $P(u) = \{P_{ij}(u)\}$ where

$$P_{ij}(u) = \text{Prob}\{s_{t+1} = j \mid s_t = i\} = \int_R f(i,j,u(i,r_t),r_t)P(dr_t) \triangleq E_r f(i,j,u(i,r),r). \tag{3}$$

A_2. For each u in U the Markov chain has a single ergodic class in the sense of Doob [4, p. 181].

It is then equivalent to assume that there is a unique (row) vector $\pi(u) = (\pi_1(u),\ldots,\pi_s(u))$ such that

$$\pi(u)P(u) = \pi(u),\ \pi(u)\underline{1} = 1. \tag{4}$$

where $\underline{1} \triangleq (1,\ldots,1)'$. Furthermore the Cesàro limit (2) can then be evaluated as

$$J(u) = \sum_i \pi_i(u)E_r m(i,u(i,r),r). \tag{5}$$

A strategy u is optimal if $J(u) = J^* \triangleq \inf\{J(u) \mid u \in U\}$.

Problem Manipulation

Note from (3) that different u may give rise to the same P(u). Hence the optimal strategy may be sought in two stages. In the first or inner stage P is fixed and we find the u with the least cost subject to P(u) = P. In the second or outer stage we find the best P[2]. So let $P_i(u)$ denote the i^{th} row of P(u). Note that it depends only on $u(i,\cdot)$. Let

$$\mathcal{P}_i = \{P_i \mid P_i(u) = P_i \text{ for some } u\},\ \mathcal{P} = \mathcal{P}_1 x \ldots x\, \mathcal{P}_s.$$

The inner problem can now be defined:

$$k_i(P_i) \triangleq \min_{u(i,\cdot)} E_r m(i,u(i,r),r) \triangleq \int_R m(i,u(i,r),r)P(dr)$$

$$\text{s.t. } P_{ij}(u) \triangleq \int_R f(i,j,u(i,r),r)P(dr) = P_{ij},\ j \in S \tag{6}$$

$$u(i,r) \in U(i,r),\ r \in R.$$

Let $k(P) = (k_1(P_1),\ldots,k_s(P_s))'$. The outer problem then is:

$$J^* \triangleq \min_{P} J(P) = \pi k(P)$$

$$\text{s.t. } \pi P = \pi,\ \pi\underline{1} = 1, \tag{7}$$

$$P \in \mathcal{P}$$

It is assumed that $\mathcal{P}$ is compact and k(·) is continuous. Hence an optimum strategy exists.

Duality conditions

In the problem (7) $\mathcal{P}$ need not be convex. Even if it were convex this is not a convex programming problem since π depends on P in a rather complicated way (π is an eigenvector of P corresponding to the largest eigenvalue.) Nevertheless a duality theorem exists as we will see.

It is convenient to introduce the matrix Q(P) = P-I with $Q_i(P) = Q_i(P_i)$ denoting its i^{th} row. Any s-dimensional vector c is called a dual variable. Define the Hamiltonian

$$H_i(P_i,c) = Q_i(P_i)c + k_i(P_i),$$

[2] Geoffrion (5) calls such problem manipulation "projection". It will be seen later that we combine this with "dualization".

and its minimum

$$h_i(c) = \min\{H_i(P_i,c) \mid P_i \in \mathcal{P}_i\}.$$

The results stated below are proved by Varaiya [6]. The first result shows how J(P) can be computed from Q(P), k(P).

TYPE TITLE OF ARTICLE HERE ON PAGE 1

Lemma 1 Consider the s linear equations in the l+s variables γ, c

$$\gamma \underline{1} = H(P,c). \tag{8}$$

(i) A solution to (8) always exists; (ii) if (γ,c) is a solution then $\gamma = J(P)$ and $(\gamma, c+\delta\underline{1})$ is also a solution for any number δ.

For any c let Authors' Address Here

$$\underline{h}(c) = \min_i h_i(c), \quad \bar{h}(c) = \max_i h_i(c).$$

The next result provides a useful duality bound.

COMMENCE text of article

Lemma 2 Let P,γ,c satisfy (8). Then

$$\underline{h}(c) \le \gamma = J(P) \le J^* \le \bar{h}(c) \tag{9}$$

A "minimum principle" is also known.

Theorem 1 P is an optimal solution to (7) if and only if there exist γ,c such that

$$\gamma \underline{1} = h(c),$$

$$\gamma = H_i(p,c) \text{ whenever } \pi_i(P) > 0.$$

Remark 1. c is said to be an optimal dual variable if these conditions hold for some γ,P. In this case it must be that $\gamma = J^*$. 2. Often A2 may be strengthened to requiring that the unique solution of (4) be strictly positive in which case the minimum principle is more elegant: $\gamma \underline{1} = h(c) = H(P,c)$. This result was known earlier.

The Two-level Controller

With these preliminaries over we can propose a "dual" algorithm for finding the optimal P for the outer problem. Define $\theta(c)$ by

$$\theta_i(c) = h_i(c) - \frac{1}{s}\sum_j h_j(c) \tag{10}$$

and consider the differential equation

$$\dot{c} = \theta(c). \tag{11}$$

Theorem 2 (i) For every initial condition c^0 there is a unique solution c(t) of (11) with $c(0) = c^0$.

(ii) c(t) converges to the set of all optimal dual variables c for which $c'\underline{1} = c^{0\prime}\underline{1}$.

(iii) $\underline{h}(c(t))$, $\bar{h}(c(t))$ converge monotonically to J^* and $\bar{h}(c(t)) - \underline{h}(c(t))$ decreases strictly monotonically to zero.

The following algorithm is suggested by Lemma 2 and Theorem 2.

Step 0. Let P^0 be any initial guess obtained, say, from the currently operating strategy. Find γ^0, c^0 so that $\gamma^0\underline{1} = H(P^0,c^0)$. If P^0 is unavailable choose c^0 arbitrarily.

Step 1. Suppose c^n is known.

(i) Find $P^{n+1} \in \mathcal{P}$ so that $h(c^n) = H(P^{n+1},c^n)$.

(ii) Determine $\underline{h}(c^n)$, $\bar{h}(c^n)$. If $\bar{h}(c^n) - \underline{h}(c^n) < \varepsilon$ stop because $J(P^n) - J^* < \varepsilon$.

(iii) Otherwise, calculate $\theta(c^n)$ according to (10).

Step 2. Set $c^{n+1} = c^n + \Delta\theta(c^n)$ and return to Step 1. Δ is a "small" positive constant.

The only non-trivial procedure in the algorithm is finding P^{n+1} since an explicit formula for H is not availabe. (Recall that H depends on k which is defined by the inner problem (6).) We turn to the calculation of P^{n+1} by a "lower" level. Suppose $c^n = (c_1,\ldots,c_s)'$ is known and we want to find P^{n+1} by solving

$$h_i(c) = \min_{P_i \in \mathcal{P}_i} H_i(P_i,c). \tag{12}$$

Now,

$$H_i(P_i,c) = Q_i(P_i)c + k_i(P_i) = \sum_j P_{ij}c_j - c_i + k_i(P_i). \tag{13}$$

Substituting for $k_i(P_i)$ from the inner problem (6) into (13) gives:

$$H_i(P_i,c) = \min_{u(i,\cdot)} \sum_j P_{ij}c_j - c_i + \int_R m(i,u(i,r),r)P(dr)$$

$$\text{s.t.} \int_R f(i,j,u(i,r),r)P(dr) = P_{ij}, j \in S$$

$$u(i,r) \in U(i,r), \; r \in R.$$

Substituting this version of H_i into (12), and noting the definition of Φ_i, it follows that

$$h_i(c) = \min_{u(i,\cdot)} \int_R [\sum_j f(i,j,u(i,r),r)c_j + m(i,u(i,r),r)]P(dr) - c_i$$

$$\text{s.t. } u(i,r) \in U(i,r), \; r \in R$$

whose solution can be written down by "inspection" as:

$$h_i(c) = \int_R \phi_i(c,r)P(dr) \triangleq E_r\phi_i(c,r), \tag{14}$$

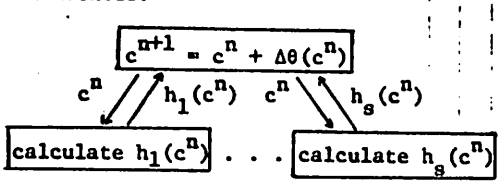
where

$$\phi_i(c,r) \triangleq \min\{\sum_j f(i,j,v,r)c_j + m(i,v,r) \mid v \in U(i,r)\} - c_i. \tag{15}$$

It also follows that the P_i^* which minimizes (12) is given by

$$P_{ij}^* = \int_R f(i,j,u^*(i,r),r)P(ds),$$

where $v = u^*(i,r)$ is the minimizer in (15). The two-level scheme is sketched below. Observe that the lower level calculations can be conducted in "parallel" as is to be expected in dual decomposition methods.



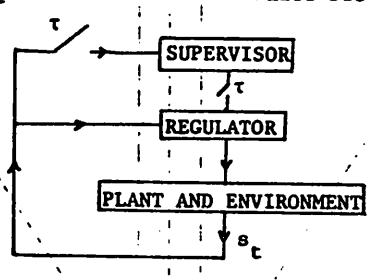
Level 1 updates c^n

Level 2 solves (14), (15).

Remark The scheme presented above was suggested by a practical problem considered in [Ref. 7,8].

TWO-LAYER BOUNDARY CONTROL

We have in mind the situation of a plant being continuously controlled by a local regulator. Every once in a while the "parameters" of the regulator are "reset" by a "supervisor". This can happen in at least two different ways: perhaps some components internal to the plant are malfunctioning and so the supervisor has to carry out some repairs, or there is a change in the external environment and the supervisor intervenes to reset the regulator so as to change the plant's operating point. We represent the state of the plant as well as of the relevant environment by s_t, $t = 0,1,\ldots$. s_t takes values in a finite set $S = \{1,\ldots,s\}$. A



control structure like this one is called a two-layer structure [Ref. 9] in contrast with a two-level structure discussed earlier. In the former the determination of control is split into algorithms which operate at different time scales whereas in the latter there is a "spatial" division into algorithms operating at the same time scale.

There has been little effort devoted to the study of multi-layer structures although they are widely adopted in the control of large processes. The reason for this seems to be the difficulty in combining in a single formulation these essential features of a multilayer structure: (i) the supervisor must itervene less frequently than the regulator (in terms of the figure above $\tau >> 1$), (ii) the supervisor must use considerably less information than the regulator and (iii) the supervisor must solve a "higher" level problem.

Chong and Athans[10] consider an LQG problem in which at the lower layer there are several decentralized regulators, coordinated by a higher layer supervisor. The supervisor intervenes by sending signals which predict the interactions between the regulators. The period between successive supervisor interventions is fixed in advance, and the supervisor has all of the available information. Thus the second feature listed above is absent.[3]

Earlier Donoghue and Lefkowitz [12] had considered a static optimization problem with the same structure. Again the supervisor intervened periodically and had full information. One of the variables to be optimized was the frequency with which the intervention is carried out.

Periodic intervention is appropriate if the lower layer represents a production cycle of fixed duration and the supervisor intervenes at a fixed stage of each cycle. It is less appropriate if the lower layer represents a "continuous" production process. In the model presented here the supervisor intervenes only when the state (of the system and environment) reaches some "extreme" or "boundary" value. That is there is a fixed subset $B \subset S$ such that the supervisor intervenes only at those instances τ for which $s_\tau \in B$. Furthermore the supervisor observes the state only at these instances. Thus the intervention times occur randomly and are determined "intrinsically" by the plant and environment process rather than being arbitrarily preselected. (Of course periodic intervention is a special case.) Finally we pay relatively little attention to the way in which the lower layer regulation is carried out, and concentrate mainly on the supervisor's actions. This permits very different design procedures to be employed for the two layers.

Problem Formulation

The lower layer process is denoted s_t, $t = 0,1,\ldots$ and takes values in $S = \{1,\ldots,s\}$. For each i in S the regulator can select a control $u(i) \in U(i)$ so that a strategy of the regulator is a vector $u \in U(1)x\ldots xU(s)$. For each u the process s_t is Markov with stationary probability transition matrix $P(u) = \{P_{ij}(u(i))\}$ as in (3) above. The cost associated with u is

$$J(u) = \lim_{T\to\infty} \frac{1}{T+1} \sum_0^T E\, k(s_t,u(s_t)). \tag{16}$$

To make the limit meaningful we impose assumption A_2 stated above. Then, using the notation introduced in (4), we have

$$J(u) = \sum_i \pi_i(u)k(i,u(i)) = \pi(u)k(u), \tag{17}$$

where $k(u) = (k(1,u(1)),\ldots,k(s,u(s)))'$.[4]

We now formulate the supervisor's observation and decision processes. A distinguised subset $B \subset S$ is chosen. B is called the set of boundary states. Suppose $B = \{1,\ldots,b\}$. Assume that $s_0 = b_0 \in B$ and let $0 \equiv T_0 < T_1 <\ldots< T_n <\ldots$ be the random times at which s_t enters B i.e.

$$T_{n+1}(\omega) = \min\{t > T_n(\omega) \,|\, s_t(\omega) \in B\} \tag{18}$$

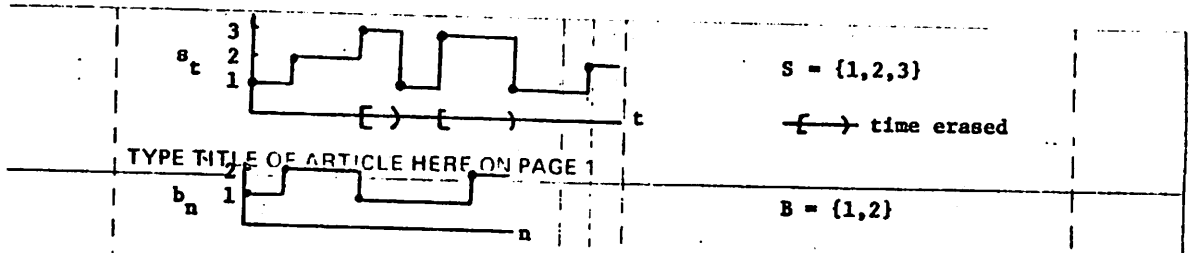
Here ω denotes the sample path. The supervisor's observation process is $b_0,b_1,\ldots,b_n$, where

$$b_n(\omega) = s_{T_n(\omega)}(\omega). \tag{19}$$

Note that the process $\{b_n\}$ operates at a different "time scale" than $\{s_t\}$ as seen in the figure below adapted from [Ref. 13]. $\{b_n\}$ is obtained from $\{s_t\}$ by "erasing" the time that $\{s_t\}$ does not spend in B. The next result if proved in Revuz [14].

[3] The Chong-Athans model is better discussed in [Ref. 11].

[4] Suppose the choice of control u_t depends on the entire past $u_t = u(s_0,\ldots,s_t)$. Then s_t is no longer Markov, but in the case to be examined later (16) will continue to be meaningful.



Type Authors' Address Here

Theorem 3 For each u the supervisor's observation process b_n, $n = 0,1,\ldots$ is Markovian and has a single ergodic class.

Thus $\{b_n\}$ inherits the most important properties of $\{s_t\}$. The next task is to propose a reasonable class of strategies for the supervisor. These strategies must of course be based on the b_n process, and they must reflect the idea that the supervisor resets the regulator each time a boundary state is reached. The reset idea is formulated as follows. Each time a boundary state say β is reached the supervisor selects a regulator strategy u^β from a fixed set $U^\beta \subset U$. Thus a supervisor strategy is a b-tuple of regulator strategies $v = (u^1,\ldots,u^b) \in U^B = U^1 x \ldots x U^b$.

Now suppose $v = (u^1,\ldots,u^b) \in U^B$ is chosen. Then during the random time interval $[T_0,T_1)$, the evolution of s_t is regulated by the transition probability matrix $P(u^{b0})$, during $[T_1,T_2)$ by the matrix $P(u^{b1}),\ldots$, during $[T_n,T_{n+1})$ by $P(u^{bn})$ etc., where $b_0,b_1,\ldots$ is the process observed by the supervisor and $T_0,T_1,\ldots$ are the reset times given by (18).

Several comments are in order before we proceed with the analysis. Firstly, the process s_t will not generally be Markov any more, although it is Markovian within each interval $[T_n,T_{n+1})$. Secondly, if we take $U^\beta = U$ for all β in B then the supervisor basically takes over the task of the regulator and the two layers "collapse". The subset U^β is selected presumably on the basis of simplicity of implementation of the regulator. One way to think about this is to identify a different regulator with each strategy u in U. U^β is then the set of regulators which are available to the supervisor when the boundary state β is reached. Thus the supervisor's task is of a higher level: it consists of selecting a regulator. Alternatively we may imagine a more versatile regulator with some variable parameters which are adjusted by the supervisor.

The supervisor's task is to select an optimal strategy. Before we can do this however we must somehow "lift" the cost function defined at the lower layer to the supervisory layer. This task is accomplished by the next result. All proofs are given in [Ref. 15].

Theorem 4 For each supervisor's strategy $v = (u^1,\ldots,u^b)$ in U^B the process b_n, $n = 0,1,\ldots$ is Markovian with a single ergodic class. Furthermore if $\{s_t\}$ is the (non-Markovian) state process the limit

$$\mathcal{J}(v) \triangleq \lim_{T\to\infty} \frac{1}{T+1} \sum_0^T E\, k(s_t, u^{\beta_t}(s_t)) \tag{19}$$

exists where $\beta_t = b_n$ for t in $[T_n,T_{n+1})$. Moreover there exist functions $K(\beta,u^\beta)$ and $T(\beta,u^\beta)$ defined for $\beta \in B$, $u^\beta \in U^\beta$ such that

$$\mathcal{J}(v) = \frac{\sum_{\beta=1}^b p_\beta(v) K(\beta,u^\beta)}{\sum_{\beta=1}^b p_\beta(v) T(\beta,u^\beta)} \tag{20}$$

where $p(v) = (p_1(v),\ldots,p_b(v))$ is the steady-state probability vector of the Markov chain $b_n, n = 0,1,\ldots$.

The functions K and T can be related to the lower layer state process $\{s_t\}$. Suppose this process reaches the boundary state β at τ, i.e. $s_\tau = \beta$. Between this time and the next time say σ that it reaches a boundary state the $\{s_t\}$ process is governed by the probability transition matrix $P(u^\beta)$. $K(\beta,u^\beta)$ is the expected value of the total cost in the interval $[\tau,\sigma)$,

$$k(s_\tau,u^\beta(s_\tau))+\ldots+k(s_{\sigma-1},u^\beta(s_{\sigma-1})).$$

Note that τ and σ are random times. $T(\beta,u^\beta)$ is the expected value of $\sigma-\tau$. Write $K(v) = (K(1,u^1),\ldots,K(b,u^b))'$ and $T(v) = (T(1,u^1)\ldots,T(b,u^b))'$. Then

$$\mathcal{J}(v) = \frac{p(v)K(v)}{p(v)T(v)} \tag{21}$$

which can be compared with (17). A supervisor's strategy v is optimal if

$$\mathcal{J}(v) = \mathcal{J}^* \triangleq \inf\{\mathcal{J}(v) | v \in U^B\}. \quad (22)$$

It is possible in principle to calculate the terms p(v), K(v), T(v) from the overall data namely $\{P(u),k(u)|u \in U\}$. However it is much more interesting and useful to note that the terms can be estimated from the supervisor's observation process $\{b_n\}$ along. We have already indicated this for K(v), T(v). We turn to p(v).

Theorem 5 Fix $v = (u^1,\ldots,u^b)$ in U^B. p(v) is the unique solution to

$$p(v) = p(v)P^B(v),\ p(v)\underline{1} = 1 \quad (23)$$

where the probability transition matrix $\{P^B_{\alpha\beta}(v)\}$, $\alpha,\beta = 1,\ldots,b$ is such that its αth row depends only on u^α, i.e. $P^B_{\alpha\beta}(v) = P^B_{\alpha\beta}(u^\alpha)$. (Here again $\underline{1} \triangleq (1,\ldots,1)'$).

Type Authors' Address Here

Evidently these transition probabilities can be estimated by the supervisor since $\{b_n\}$ is an ergodic process. The supervisor's decision problem can now be formulated as one of constrained optimization:

COMMENCE text of article

$$\min[p(v)T(v)]^{-1}\,[p(v)K(v)]$$
$$\text{s.t. } p(v) = p(v)P^B(v),\ p(v)\underline{1} = 1 \quad (24)$$
$$v = (u^1,\ldots,u^b) \text{ with } u^\beta \in U^\beta,\ \beta \in B.$$

It is interesting to compare this with the "outer" problem (7). The only difference is in the form of the cost function. We can now state the optimality conditions for (24). Let $Q^B(v) = P^B(v) - I$. Let $Q^B_\alpha(u^\alpha)$, $P^B_\alpha(u^\alpha)$ denote the αth rows. Let c be any b-dimensional vector. Define the Hamiltonian

$$H_\alpha(P^B_\alpha,c) = Q_\alpha(P^B_\alpha)c + K(\alpha,u^\alpha)$$

and its minimum

$$h_\alpha(c) = \min\{H_\alpha(P^B_\alpha,c)|u^\alpha \in U^\alpha\}.$$

Theorem 6 v is an optimal supervisor's strategy if and only if there exist γ,c such that

$$\gamma T(v) = h(c)$$
$$\gamma T_\alpha(v) = H_\alpha(P^B_\alpha,c) \text{ whenever } p_\alpha(v) > 0.$$

Moreover $\gamma = \mathcal{J}^*$.

It is possible to give an algorithm for finding the optimal strategy. This and other results are given in [Ref. 15]. One interesting result is this. Suppose the supervision of the regulator is increased by enlarging the boundary states to $\hat{B} \supset B$. Let $\hat{\mathcal{J}}^*$ be the corresponding minimum cost. Then $\hat{\mathcal{J}}^* \leq \mathcal{J}$.

CONCLUSIONS

Two hierarchical schemes for control of large Markov chains have been presented. The two-level scheme gives a novel application of decomposition techniques used in mathematical programming. The two-layer scheme attempts to incorporate in a formal way many intuitive features of multi-layer control.

ACKNOWLEDGEMENT

Research sponsored in part by the National Science Foundation Grants ENG76-16816, 01P75-04371, and a grant from IRIA (Institut de Recherche en Informatique et Automatique).

REFERENCES

[1] P. Varaiya and J. Walrand, Decentralized stochastic control, preprints, IFAC Workshop on Control and Management of Integrated Industrial Complexes, Toulouse, September 6-8, 1977.

[2] H. J. Kushner and C.-H. Chen, Decomposition of systems governed by Markov chains, IEEE Trans. Auto. Contr. AC-19, 501-507 (1974).

[3] S. M. Ross (1970) Applied Probability Models with Optimization Applications, Holden-Day, San Francisco.

[4] J. M. Doob (1953) Stochastic Processes, Wiley, New York.

[5] A. M. Geoffrion, Elements of large-scale mathematical programming Part I Concepts, Part II Synthesis of algorithms and bibliography, Management Science 16, 652-691 (1970).

[6] P. Varaiya (1977) Optimal and Suboptimal Stationary Controls of Markov Chains, Memo., No. UCB/ERL 77/17, Electronics Res. Lab., University of California, Berkeley, Ca.

[7] P. Varaiya, Dispatching emergency vehicles, Proc. 1977 JACC, San Francisco (June 22-24,1977).

[8] P. Varaiya, U. Schweizer and J. Hartwick, A class of Markovian problems related to the districting problem for urban emergency servies, to appear in Ricerche di Automatica (1977).

[9] W. Findeisen, A survey of problems in hierarchical control, Proc. Workshop on Multilevel Control, Institute of Auto. Contro. Tech. University of Warsaw (1975).

[10] C. Y. Chong and M. Athans, On the periodic coordination of linear stochastic systems, Proc. 6th IFAC World Congress, Boston, Ma. (1975).

[11] A. Benveniste, P. Bernhard and A. Cohen (1976) On the Decomposition of Stochastic Control Problems Rapport de Recherche No. 187, IRIA, Rocquencourt, France.

[12] J. F. Donoghue and I. Lefkowitz, Economic tradeoffs associated with a multilayer control strategy for a class of static systems, IEEE Trans. Auto. Contr. AC-17, 7-15 (1972).

[13] D. Freedman (1971) Approximating Countable Markov Chains, Holden-Day, San Francisco.

[14] D. Revuz (1975) Markov Chains, North Holland, New York.

[15] J. P. Forestier and P. Varaiya, Hierarchical control of large Markov chains, in preparation (1977).