Discussion Paper No. 14

OPTIMAL DELAYS IN DECISION AND CONTROL †

by

Sudhakar D.Deshmukh (*) and Shirish D.Chikte (**)

September 26, 1972

Revised June 10, 1973

(†) The referees' comments and suggestions on an earlier version of this paper are gratefully acknowledged.

(*) Assistant Professor, Managerial Economics and Decision Sciences, Graduate School of Management, Northwestern University, Evanston, Illinois 60201

(**) Graduate Student, Department of Electrical Engineering, Brooklyn Polytechnic Institute, Brooklyn, New York

ABSTRACT

Computations involved in controlling a system or a decision process are time-consuming in practice. The problem of optimally choosing the estimation and control delays is formulated in the dynamic programming framework and illustrated by examples. Selection of optimal estimation and control algorithms is outlined conceptually.

OPTIMAL DELAYS IN DECISION AND CONTROL

S. D. Deshmukh and S. D. Chikte

INTRODUCTION: Controlling a feedback system or a realistic economic decision process involves observation, measurement, estimation and communication of information about its state, processing it to generate a control and implementation of that control, collectively these operations being called estimation and control. Delays and costs involved in these operations are important practical considerations which affect the system performance. The conceptual stochastic model presented here explicitly takes them into account, the objective being to determine optimal strategies for performing the estimation and control operations.

The estimation and control delays are significant if these operations are slow in comparison with the state changes and if control laws are applied on line. In stochastic systems these delays prevent us from implementing a perfect response to the system state instantaneously and force us to adopt only obsolete estimates and controls. Typically, for a given state, the estimate (control) quality improves with the time spent in generating it but only with respect to that older state and hence the longer delay also results in greater obsolescence, thereby reducing the effectiveness of the computed estimate (control).

In section 2, given the estimation and control algorithms, the objective is to optimally utilize them by determining the stopping policies to dynamically yield optimal estimation and control delays in such a way as to replace an obsolete estimate (control) by a less obsolete one by balancing improvement against obsolescence with delay using the Optimality Principle. Section 3 illustrates applicability of the model by three examples of algorithms operating in stochastic environments and the corresponding optimal stopping policies. Section 4 considers the overall problem of finding the best estimation and control algorithms, taking into account the time delays and operating and purchasing costs.

2. OPTIMAL DELAYS: Consider a discrete time stochastic control process with separation. Let x_n be an m-vector of the system state at time n, $1 \leq n \leq N$, where N is the final time, and $X \subseteq R^m$ the state space. Let y_n be a q-vector of the observation of x_n at time n and $Y \subseteq R^q$ the observation space. Denote the observation process (assumed to be instantaneous) by

$$(2.1) \qquad y_n = G(x_n, \eta_n) \qquad 1 \leq n \leq N$$

where η_n is the white measurement noise. Let $\hat{x}_n$ be the estimate of x_n available at n and $\hat{X}$ the estimate space. (In general, $\hat{x}_n$ may include conditional density of x_n given the history up to n, e.g. see Application C in Section 3). Due to time delay required in the estimation-control process $\hat{x}_n$ is not an estimate of x_n obtained from y_n but it was an estimate generated previously from an observation y_i of an older state x_i, $i < n$, and then predicted to time n. An observation is made only when a new estimate is to be computed, so that $y_{i+1}, \ldots y_{n-1}$ are not observed, because, for example, of measurement costs involved (see [2],[4], where a related problem of finding optimal number and spacing of observations is considered). Let u_n be a p-vector of the latest control available at n and $U \subseteq R^p$ the admissible control space. Again, due to time delay involved in the estimation-control process, u_n was not generated as a response to the current state x_n but to an estimate of an older state x_j, $j < n$. At time n the system earns a return of $W(x_n, u_n)$ and moves according to a Markovian law

$$(2.2) \qquad x_{n+1} = f(x_n, u_n, \omega_n) \qquad 1 \leq n < N$$

where ω_n is the white system noise.

The computations of estimates and controls are carried out consecutively by two algorithms α' and α, which are assumed to be given, in this section the objective being to optimally utilize them. Suppose the k^{th} application of the estimation algorithm α' ("k^{th} estimation substage") starts at time n. Estimation delay $d'_n \in D'$ is then chosen as a decision variable, where $D' \subseteq \{0,1,\ldots,N\}$ is the set of all possible estimation durations. (For notational convenience the subscripts on d' and d below will be suppressed.) The estimation algorithm is a (possibly stochastic) transformation $\alpha': Y \times \hat{X} \times U \times D' \to \hat{X}$. Starting at n the algorithm computes for d' time instants using the fresh observation y_n of x_n to improve upon the latest available estimate $\hat{x}_n$ of x_n. This improved estimate of x_n, the control u_n (which stays in force throughout the estimation substage) and the plant equations (2.2) are then used by a built in predictor to give a prediction $\hat{x}_{n+d'}$ of $x_{n+d'}$. Thus

$$(2.3) \qquad \hat{x}_{n+d'} = \alpha'(y_n, \hat{x}_n, u_n, d')$$

is a predicted estimate of $x_{n+d'}$ which will be available to the successive k^{th} application of the control algorithm α ("k^{th} control substage") starting at $(n+d')$. For example,

$$\alpha'(y_n, \hat{x}_n, u_n, d') = E\{x_{n+d'} | x_n = \hat{x}_n + K'(d')[y_n - E(G(\hat{x}_n, \eta_n))], u_n\}$$

where $K'(d')$ is a correction matrix which approaches the "correct" value as d' increases and $K'(0) = 0$. Usually, the longer the estimation interval d' the better is the estimate of x_n generated but the longer delay also increases the mean squared error of the prediction so that the worse is $\hat{x}_{n+d'}$ as an estimate of $x_{n+d'}$. The objective is to find optimal d' or, more generally, to find an optimal stopping rule for determining the estimation delay. If the k^{th} estimation substage starts at time n upon observing $(y_n, \hat{x}_n, u_n)$ then a stopping rule $\delta'_k: Y \times \hat{X} \times U \to D'$ specifies a delay

$$(2.4) \qquad d' = \delta'_k(y_n, \hat{x}_n, u_n)$$

For example, it is conceivable that optimal d' will be small if, roughly, y_n is close to $E(G(\hat{x}_n, \eta_n))$ (so that the initial estimate appears to be a good one) or if u_n is an inferior control in face of $\hat{x}_n$ (so that it may be optimal to spend little time estimating and introduce a better control early). During these d'

instants the expected estimation cost (including observing y_n) denoted by $C'(y_n,\hat{x}_n,u_n,d')$ will be incurred (for example it is $c_1' + c_2'd'$), control u_n will stay in force and the system will move according to (2.2) with $n + j$ replacing n for $j = 1,2,\ldots,d'-1$, yielding the net expected return denoted by

$$(2.5) \qquad r'(y_n,\hat{x}_n,u_n,d') = \sum_{j=0}^{d'-1} E[W(x_{n+j},u_n)|y_n,\hat{x}_n,u_n] - C'(y_n,\hat{x}_n,d')$$

The expectation is computed using the conditional distribution of x_{n+j} given $(\hat{x}_n,u_n)$ following the law of motion (2.2) for j steps.

Upon the termination of the k^{th} estimation substage the k^{th} control substage begins at time $n + d'$ with the latest estimate $\hat{x}_{n+d'}$ and the (same) control $u_{n+d'} = u_n$. The control delay $d_{n+d'} \in D$ is chosen as another decision variable from the set $D \subseteq \{0,1,\ldots,N\}$ by means of the k^{th} control stopping rule $\delta_k: \hat{X} \times U \to D$, so that

$$(2.6) \qquad d = \delta_k(\hat{x}_{n+d'},u_{n+d'})$$

(Note that d' is in fact d_n' and d is in fact $d_{n+d_n'}$). The control algorithm then starts at $(n+d')$ and computes for d time instants given the latest available control $u_{n+d'} = u_n$ (which serves as an initial solution and which stays in force throughout the control computation) and the latest available estimate $\hat{x}_{n+d'}$, which is further predicted to time $n+d'+d$ by another predictor built in the controller to give $\hat{x}_{n+d'+d}$ (which serves as the parameter). The control algorithm is a (possibly stochastic) transformation $\alpha:\hat{X} \times U \times D \to U$.

$$(2.7) \qquad u_{n+d'+d} = \alpha(\hat{x}_{n+d'},u_{n+d'},d)$$

is thus the control generated by the algorithm at time $(n+d'+d)$ and will stay in force until a new control is computed next time around. For example, $\alpha(\hat{x}_{n+d'},u_{n+d'},d) = u_{n+d'} + K(d)[u^*_{n+d'+d} - u_{n+d'}]$ where $K(d)$ is a correction factor with $K(0) = 0$ and approaching I as d increases and $u^*_{n+d'+d}$ is the optimal control at $(n+d'+d)$ with respect to the prediction $\hat{x}_{n+d'+d} = E[x_{n+d'+d} \mid \hat{x}_{n+d'},u_{n+d'},d]$. Usually the longer the computation interval d the better is the control generated but also more obsolete it gets due to stochastic changes in the system state. It is conceivable that optimal d will be small if, roughly, with information $\hat{x}_{n+d'}$ the control $u_{n+d'}$ is a good one so that any further improvement in the solution is smaller than obsolescence due to further delay. The objective is to find an optimal stopping rule δ_k to give d. During these d instants the expected control computation cost $C(\hat{x}_{n+d'}, u_{n+d'}, d)$ is incurred, the control $u_{n+d'} = u_n$ stays in force and the system moves according to (2.2) with $n + d' + j$ replacing n for $j = 1,2,\ldots,d-1$, yielding the net expected return

$$(2.8) \qquad r(\hat{x}_{n+d'}, u_n, d) = \sum_{j=0}^{d-1} E[W(x_{n+d'+j}, u_n) | \hat{x}_{n+d'}, u_n] - C(\hat{x}_{n+d'}, u_n, d)$$

where the conditional expectation is computed in the same way as in (2.5). The two outputs of the controller $\hat{x}_{n+d'+d}$ and $u_{n+d'+d}$ together with a new observation $y_{n+d'+d}$ serve as the input to the $(k+1)^{st}$ estimation substage, thus completing the "k^{th} stage" and starting the $(k+1)^{st}$.

Setting $n = n + d' + d$ and $k = k + 1$ the process repeats as above until the final time N is reached. The objective is to determine optimal estimation and control delays d''s and d's successively by finding optimal sequences of stopping rules ("policies") $\Pi' = \{\delta_k' : k = 1,2,\ldots\}$ and $\Pi = \{\delta_k : k = 1,2,\ldots\}$ so as to maximize the net contribution of the given algorithms α' and α to the system performance, given the initial conditions $(y_1, \hat{x}_1, u_1)$.

To formulate this problem in the dynamic programming framework let $V_n'(y_n, \hat{x}_n, u_n)$ be the optimal net return from n onwards <u>provided</u> an estimation substage starts at time n with the "state of the model" $(y_n, \hat{x}_n, u_n)$. Similarly, <u>if</u> a control substage starts at time i with the "state of the model" $(\hat{x}_i, u_i)$, let $V_i(\hat{x}_i, u_i)$ be the optimal net return from then on. (State of the model chosen at each substage gives information that is both necessary and sufficient to make an optimal delay decision at that substage and to predict the state of the model at the next substage). In computing (V_n', Π') and (V_i, Π) there are two sequential decision problems, choosing optimal d''s and d's alternately, meshed together. The Principle of Optimality requires that V'_n and V_i satisfy the following functional equations, where $D'_k = D' \cap \{j: j \leq N-k\}$ and $D_k = D \cap \{j: j \leq N-k\}$

$$(2.9) \qquad V'_n(y_n,\hat{x}_n,u_n) = \underset{d'\in D'_n}{\text{Max}} \left\{ r'(y_n,\hat{x}_n,u_n,d') + E\left[V_{n+d'}(\alpha'(y_n,\hat{x}_n,d'),u_n) \,\middle|\, y_n,\hat{x}_n,u_n,d' \right] \right\}$$

$$(2.10) \qquad V_i(\hat{x}_i,u_i) = \underset{d\in D_i}{\text{Max}} \left\{ r(\hat{x}_i,u_i,d) + E\left[V'_{i+d}(y_{i+d},\hat{x}_{i+d},\alpha(\hat{x}_i,u_i,d)) \,\middle|\, \hat{x}_i u_i,d \right] \right\}$$

$$V'_n \equiv 0 \text{ if } n > N \text{ and } V_i \equiv 0 \text{ if } i > N$$

Note that both V' amd V appear in each equation, (2.9) being for the estimation substage and (2.10) for the control substage. To ontain one optimality equation for the complete stage, <u>if</u> it does start at time n, consisting of estimation and the successive control computation, let $i = n + d'$, $\hat{x}_{n+d'} = \alpha'(y_n,\hat{x}_n,d')$ and $u_{n+d'} = u_n$ in (2.10) and substitute it in (2.9) to give the optimal return from time n onwards

$$(2.11) \qquad V'_n(y_n,\hat{x}_n,u_n) = \underset{d'\in D'_n}{\text{Max}} \Bigg\{ r'(y_n,\hat{x}_n,u_n,d') + E\Bigg\{ \underset{d\in D_{n+d'}}{\text{Max}} \Bigg\{ r(\alpha'(y_n,\hat{x}_n,d'),u_n,d) + E\Big[V'_{n+d'+d}(y_{n+d'+d},\hat{x}_{n+d'+d},\alpha(\alpha'(y_n,\hat{x}_n,d'),u_n,d)) \,\Big|\, \alpha'(y_n,\hat{x}_n,d'),u_n,d \Big] \Bigg\} \,\Bigg|\, y_n,\hat{x}_n,u_n,d' \Bigg\} \Bigg\}$$

The dynamic optimization process represented by (2.11) has a direct counterpart with a static decision process as in, for example, [7]. In the latter the "information structure" α': $X \to Y$ partitions the state space X and yields a "message" $y \in Y$, the "decision rule" α: $Y \to U$ then operates on the message y to give an "action" $\alpha(y) \in U$ and as a result the "payoff" is $W(x,\alpha(\alpha'(x)))$, whose expected value is to be maximized by selecting optimal α and α'. This problem is solved by first fixing the information structure α' and finding the best decision rule $\alpha^*(\alpha')$ maximizing $E[W(x,\alpha(\alpha'(x)))]$ to give $W(\alpha', \alpha^*(\alpha'))$, say, and then choosing the best information structure α'^* maximizing the latter;

the solution is then α'^* and $\alpha^*(\alpha'^*)$. In (2.12) the estimation algorithm α' may be interpreted as the information structure which is chosen at each estimation substage paremetrically by choosing d'; similarly the control algorithm may be interpreted as a decision rule, which is selected parametrically at each control substage by choosing d. Thus choosing d''s and d's dynamically corresponds to parametric selection of the information structure and the decision rule and as above, in each control stage for given value of d' the optimal $d^*(d')$ is found by maximizing the quantity in the inner braces of (2.11) and then optimal d'^* is found by maximizing the quantitity in the outer braces. Complications in the dynamic model arise due to interrelations between successive (sub)stages and the delays involved in using the chosen information structures and decision rules.

To obtain optimal stopping policies Π' and Π instead of solving difficult equations (2.11) directly, (2.9) and (2.10) may be solved separately in the usual recursive fashion. For each n, whether a computation will start at n and, if it does, whether it will be an estimation or control computation is unknown a priori. Thus one must solve (2.9) for each n and (2.10) for each i, requiring considerable solution effort. Due to appearance of V' and V in both types of equations (2.9) and (2.10) must be solved simultaneously for each value of $i = n$ starting with $i = n = N$ and working backwards. Then for each n and each state of the model $(y_n, \hat{x}_n, u_n)$ optimal delay will be $d_n'(d_n)$ <u>if</u> an estimation (control) begins at n; the stopping policies are implemented only at beginning of substages as the system moves. The expectation in (2.9) is computed using the distribution of the output $\hat{x}_{n+d'}$, given $(y_n, \hat{x}_n, u_n)$ and the known stochastic algorithm α'. Similarly in (2.10) the stochastic algorithm α specifies conditional distribution of u_{i+d} given $(\hat{x}_i, u_i, d)$ and, in addition, the prediction $\hat{x}_{i+d}$ is obtained from $(\hat{x}_i, u_i)$ and the law of motion (2.2) for d-steps, y_{i+d} being its predicted observation from (2.1). Finally, it should be noted that the maps α' and α must be known completely in determining policies Π' and Π (but not in implementing them).

In spite of these general remarks on the solution procedure it seems clear that the total problem and its functional equations may be computationally very difficult to solve. However, the aim has been to formulate a general model of many realistic decision and control systems where delays may be appreciable and their optimal control may be essential. For special cases of this model with instantaneous and perfect estimation or control operations see [5],[3]. The next section describes some applications of these special cases, which are economically interesting and solvable.

3. APPLICATIONS: (A) Optimal Control of a Queue: The server looks at the length of the queue and decides on how long to service the first customer, longer service means better quality (up to a point) and greater customer satisfaction but also a higher cost of waiting by customers in a queue and the new arrivals. The problem is to determine an optimal quality of service.

Let $X = \hat{X} = Y = \{0,1,2,\ldots\}$, x being the number of people in the system, assumed to be observed perfectly and instantly. $U = D = [0,\infty)$, control u being the time spent on the previous customer and delay d being that spent on the present customer, $\alpha(x,u,d) = d$. With Poisson arrivals, the system moves as $x_{t+d} = x_t - 1 + Z(d)$ if $x_t \geq 1$ and $x_{t+d} = Z(d)$ if $x_t = 0$, where $Z(d)$ is a Poisson random variable with mean λd. The improvement in the generated control with delay is represented by the concave customer satisfaction function $b \times d - ad^2; a,b > 0$. The obsolescence due to delay is represented by the cost of waiting $c[xd + \int_0^d (d-t)\lambda dt]$, where $c < b$ is the cost per customer per unit of his waiting time.

Defining $V_n(x)$ as the optimal return from serving n customers when there are x in the system, the functional equation becomes, for $x \geq 1$,

$$(3.1)\quad V_n(x) = \max_{0\leq d<\infty}\left[(b-c)xd - (a + \frac{\lambda c}{2})d^2 + \sum_{m=0}^{\infty} V_{n-1}(x-1+m)\,\frac{e^{-\lambda d}(\lambda d)^m}{m!}\right]$$

$$V_0(x) = 0$$

If $\delta_n(x)$ is the optimal stopping rule, then $\delta_1(x) = \frac{(b-c)x}{2a+\lambda c}$ linear in x and $V_1(x) = [(b-c)x]^2/2(2a+\lambda c)$ quadratic in x. In general, by induction it can be shown that $\delta_n(x)$ is linear in x and $V_n(x)$ is quadratic in x.

(B) Optimal R and D Under Rivalry: The problem is to determine an optimal development period $d \in D = U = [0,T]$ in order to improve a given product before introducing it into the market, while a competitor is also trying to introduce a similar product. If we introduce it before he does (i.e. $x = 1$) after delay d, our return in the market will be (with $\alpha(x,u,d) = d$)$W(x,d) = R_0 + \frac{(R_1-R_0)d}{T}$ while if he introduces first (i.e. $x = 2$) our return will be $R_0 + \frac{(R_2-R_0)d}{T}$, where $R_1 > R_2 > R_0 > 0$. The state $x \in X = \hat{X} = Y = \{1,2\}$ can be observed instantly and correctly, while the time of the opponent's introduction is assumed to be uniformly distributed on $[0,T]$.

In this single stage version of the problem, optimal d is to be found so as to maximize the expected total return of

$$(3.2)\qquad E[W(x_d,\alpha(x_0,u,d))] = \frac{(R_1-R_0)d}{T}\left(1-\frac{d}{t}\right) + \frac{(R_2-R_0)d}{T}\cdot\frac{d}{T}$$

by striking a balance between the advantage of an improved product and the disadvantage of not being the leader. Then optimal $d^* = \frac{(R_1-R_0)T}{2(R_1-R_2)}$, which is linearly decreasing in R_0 (i.e. the better the initial product the lower is the optimum effort level).

(c) <u>Optimal Observation Strategy</u>: In this example of estimation delay, at each stage the probability distribution of the state is known and its maximum likelihood estimate is used. Suppose a binary-valued state $x_n \in \{0,1\}$ changes with probability $p < \frac{1}{2}$ and, at each n, depending on the prior probability q_n that the state is 0 we decide to observe or not; observation is perfect but delayed by unit time. Then a control is chosen instantaneously to match the estimate; a correct match yields a return $W(x_n,u_n) = \omega$, a wrong one yields 0 and there are no observation costs. The probabilities q_n are changed in the Bayesian fashion upon adjusting for the system motion and the process repeats. The objective is to determine an observation strategy to maximize expected total return.

Thus $X = \hat{X} = Y = U = D' = \{0,1\}$, the state of the model is $(q_n,\hat{x}_n) \in [0,1] \times \hat{X}$. If $d = 0$, $\alpha'(\hat{x}_n,q_n,0) = (\hat{x}_n,q_n)$; if $d = 1$,

$$\alpha'(\hat{x}_n,q_n,1) = \begin{cases} (0,(1-p)) & \text{with probability } q_n \\ (1,p) & \text{with probability } (1-q_n) \end{cases}$$

$\alpha(\hat{x}_n,u_n,0) = \hat{x}_n$ and another output of α is $q_{n+1} = p(1-q_n) + (1-p)q_n$ (new prior probability at next stage) and $\hat{x}_{n+1} = 0$ iff $q_{n+1} \geq \frac{1}{2}$. Since q_n contains all the information needed at n define $V_n(q_n)$ as the optimal value function from n on.

$$(3.3) \qquad V_n(q_n) = \text{Max} \begin{cases} \text{Max}\,\{q_n\omega,(1-q_n)\omega\} + V_{n+1}(p + q_n - 2pq_n) & \text{if } d' = 0 \\ q_nV_{n+1}(1-p) + (1-q_n)\,V_{n+1}(p) & \text{if } d' = 1 \end{cases}$$

$$V_{N+1}(q) = 0 \text{ and } V_N(q) = \text{Max}\,\{q\,\omega,(1-q)\omega\} \text{ with optimal } d' = 0.$$

By induction it can be shown that $V_n(q)$ is convex in q and $V_n(q) = V_n(1-q)$ for all n. From this it can be shown that the observation region is convex so that there exist two numbers q_n' and q_n'' with the property that if $q_n \in [q_n',q_n'']$ it is optimal to observe while if $q_n \leq q_n'$ or $q_n \geq q_n''$ it is not. Also the observation region decreases as p increases. Thus greater the uncertainty about the state or more stable the system is better it is to make an observation, even though a delayed one.

4. <u>OPTIMAL DESIGN</u>: A fast sophisticated estimator or control computer is expensive to purchase and operate but it generates superior responses to the system state changes. The objective is to choose optimal $\alpha'^* \in A'$ and $\alpha^* \in A$, where A' and A are the sets of available algorithms.

Any pair (α',α) is assumed to be operated according to optimal stopping rules as in Section 2 to yield the net total optimal return $V_1'(y,\hat{x}_1,u_1)$ to be denoted by $\overline{V}(\alpha',\alpha)$. Also associated with each pair (α',α) is the purchasing cost $K(\alpha',\alpha)$, which should not exceed a given budget M. Finally, not all members of A' and A may be compatible, since output of α' serves as an input to α. Define, for each $\alpha' \in A'$, $\Gamma(\alpha') \subset A$ as the set of compatible control algorithms.

For each $\alpha' \in A'$ we may choose $\alpha^*(\alpha') \in A$ by solving

$$\text{(4.1)} \qquad \text{Maximize } \overline{V}(\alpha',\alpha)$$
$$\text{subject to } K(\alpha',\alpha) \leq M$$
$$\alpha \in \Gamma(\alpha')$$

to give the optimal worth $\overline{V}(\alpha',\alpha^*(\alpha'))$, which is then maximized by choosing $\alpha'^* \in A'$. This sequential approach then yields the optimal pair $(\alpha'^*,\alpha^*(\alpha'^*))$. The corresponding multiplier associated with the budget constraint in (4.1) has then the usual marginal worth interpretation.

This has been only a conceptual formulation of the problem of optimal design. Its complete solution requires a complete knowledge of sets A',A and functions V(from (2.11)), K and Γ; unfortunately, in practice, this information is very difficult to obtain. Our modest objective has been simply to conceptualize the important and difficult problem of optimal choice and operation of decision and control procedures, taking into account primarily the time delays and then costs involved that are significant in practice.

REFERENCES

1. Aoki, M., Optimization of Stochastic Systems, Academic Press 1967.
2. Aoki, M. and Li, M. T., "Optimal Discrete-Time Systems With Cost for Observation". IEEE Trans. Automat. Contr. Vol. AC-14, April 1969.
3. Chikte, S. D., "Delay Considerations in Decision and Control", unpublished, M. S. Electrical Engineering thesis, University of Utah (June 1972).
4. Cooper, C. A. and Nahi, N. E., "An Optimal Stochastic Control Problem With Observation Cost", IEEE Trans. Automat. Contr. vol. AC-16, April 1971.
5. Deshmukh, S. D., "Optimal Computations in Stochastic Environments" Technical Report No. 11, Collaborative Research on Economic Systems and Organization, Center for Research in Management Science, University of California, Berkeley (August 1971).
6. Kramer, L. C. and Athans, M., "On Simultaneous Choice of Dynamic Control and Measurement Strategies for Stochastic Systems", Joint Automat. Contr. Conf. 1972 pp. 176-181.
7. Marschak, J., "Economics of Information Systems", J. of Am. Stat. Assoc., Vol. GG (1971), No. 333.
8. McGuire, C. B. and Radner, R., (eds.), Decision and Organization, North Holland (1971).