DECENTRALIZED STOCHASTIC CONTROL

by

P. Varaiya and J. Walrand

Memorandum No. UCB/ERL M77/32

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.

DECENTRALIZED STOCHASTIC CONTROL

Pravin Varaiya and Jean Walrand
Department of Electrical Engineering and Computer Sciences, and the Electronics Research Laboratory, University of California, Berkeley, CA 94720

ABSTRACT

The literature on decentralized stochastic control is classified in terms of a few paradigmatic problem formulations. The status of research in each problem formulation is presented, and an attempt is made to assess the limit of these formulations.

INTRODUCTION

Over the past thirty years many procedures were invented for analyzing the behavior of and for designing control strategies for systems operating under uncertainty. We call this set of procedures the "classical" theory and divide it into three classes according to the nature of the problem which is addressed. Thus we have procedures for

- modelling systems e.g. state-space models in the form of stochastic difference or differential equations, input-output models such as ARMA;
- describing system behavior e.g. quantitative descriptions such as state or parameter estimators and qualitative descriptions such as stability, ergodicity and identifiability;
- controlling system behavior e.g. stochastic optimal control, adaptive control and stabilizing feedback control.

It is not the intention to suggest that each class of procedures was developed without reference to the others. Nevertheless, it is possible to detect, as our field progresses, an unfortunate division of labor among researchers resulting in each problem formulation gaining a certain degree of antonomy.

The classical theory presupposes the perspective of a centralized controller. That is it assumes that all the information about the system is acquired, and the calculations based upon it are made, at the same location. Two kinds of information can be distinguished.

- Information about the system model and the environment in which it operates. We call this *a priori* or off line information. For example, in the case where the model is a stochastic difference equation the off line information consists of the statistics of the disturbance or "noise" process and of the values of the parameters of the equation; when the latter are not exactly known then a prior probability distribution of the parameter values is specified.
- At each instant that the system is operating, the set of measurements on the system made up to that time. It is customary to call this on line information.

The range of systems and problems to which the classical theory is applicable is limited in a technical and theoretical manner by the centralization assumption. The limitation is technical when the theory is applicable in principle but the system is so large that centralized information acquisition and calculation are infeasible. An example of this arises in the attempt to control the traffic in a highway network. To describe such a system requires hundreds of state and control variables, and to implement a centralized control scheme requires an elaborate communication network to transmit to the controller all of the traffic measurements made at different network locations.

When the system is such that the acquisition and processing of information is decentralized or shared among several locations then the classical theory is inapplicable. A good example of this is the problem of routing of messages over a date communication network such as ARPANET.[1] In this system both the control signals ("protocols") and messages are transmitted over the same links of the network and at the same speed. One state description of such a network consists of the number of different kinds of messages at each node and the states of the buffer storage located there. It follows that at each node the on line information available from nearby nodes is more up-to-date than that available from more remote nodes. This happens because the latter is obtained after a longer transmission delay and this delay is necessarily of such a magnitude that the system state changes significantly during this time.[2]

[1]For a brief description of ARPANET and *ad hoc* decentralized control strategy used there see Kahn [14].

[2]This state description is due to Segall [33]. An imaginative analysis of the effects of transmission delay on the optimal decentralized control strategy is given by Schoute [32] for the simpler case of the ALOHA system.

Our objective here is to review the literature on decentralized stochastic control. This literature is motivated by the desire to overcome the limits of the classical theory. Since these limits are most felt in the problem of control most of the contributions are addressed to it. But there is some material addressed to the problems of system modelling and system behavior description and we shall begin with this.

SYSTEM MODELS

There is an extensive literature concerned with simplification of models for large systems (see (Ref. 28)). Most of it is, however, still within the framework of deterministic centralized control and little attention is paid to the fact that in large systems information and control are often decentralized. Hence the simplified model will be useful only to the extent that it is "compatible" with the way this sharing occurs. It is unfortunately not possible to be more precise about this compatibility consraint. Some illustrative examples and suggestive comments are given by Sandell [27]. For our purpose the only useful model simplification procedures are those using decomposition. But since these are invented for the purpose of addressing the other two problem formulations, we will consider them later.

SYSTEM BEHAVIOR DESCRIPTIONS

We consider research addressed to questions of the form: How can qualitative or quantitative properties of system behavior be deduced when the available information is decentralized? The qualitative property which has been studied is that of stability and the quantitative property is state estimation.

In analyzing stability, Michel [19] starts by supposing a model of interconnected subsystems

$$\dot{x}_i = f_i(x_i) + \sigma_i(x_i)\xi_i + \Sigma_j\, g_{ij}(x_j), \qquad i = 1,\dots,k \tag{1}$$

where x_i is the state of the ith subsystem and ξ_i is white noise. Suppose that f_i, σ_i, g_{ij} all vanish at 0 so that 0 is an equilibrium state of (1), as well as an equilibrium state of each "isolated" subsystem i

$$\dot{x}_i = f_i(x_i) + \sigma_i(x_i)\xi_i. \tag{2}$$

Next it is supposed that the system (2) is asymptotically stable in the large with probability 1 (ASIL), and furthermore that there is available a Lyapunov function $V_i(x_i)$ which guarantees ASIL. Then in a way which is by now standard it is possible to provide a test, involving only the bounds on $V_i(x_i)$, $\mathcal{L}V_i(x_i)$ and the g_{ij}, which guarantees ASIL of (1). (Here $\mathcal{L}V_i$ denotes the differential generator of (2) applied to V_i.) The result has been extended in [20] to permit stochastic variations in the interaction terms g_{ij} as well.

In performing the test the analyst does not need all of the off line information i.e. the functions f_i, σ_i, g_{ij}, but only the various bounds which presumably can be supplied by controllers of each subsystem. (The on line information plays no role here.) Thus a certain degree of decentralization is possible possible. There are two reservations: the equilibrium state of (1) must be known, and it does not seem possible to calculate this on a decentralized basis, and the test is known to be quite conservative since it involves only the magnitudes of the g_{ij}.

To study the literature dealing with decentralized state estimators consider a linear version of (1) namely

$$\begin{aligned}\dot{x}_i &= A_i x_i + \xi_i + B_i x,\\ y_i &= H_i x_i + \eta_i, \qquad i = 1,\dots,k\end{aligned} \tag{3}$$

where $x = (x_1,\dots,x_k)$ is the state of the composite system, ξ_i, η_i are independent white noises and y_i is the on line observation at the ith controller. Syppose the "interaction" matrix B_i is small. Then an obvious state estimator can be proposed:

$$\dot{\hat{x}}_i = A_i\hat{x}_i + K_i(t)[y_i - H_i\hat{x}_i]. \tag{4}$$

The advantage of (4) is that the processing of on line information is decentralized. One may fix the structure (4) and ask for the gain matrices $K_i(t)$ which minimize the mean square error $x_i - \hat{x}_i$. This minimization requires centralized knowledge of the off line information and the (off line) computation necessary to find the best gain is much greater than that required to compute the gain matrix for the centralized case [26]. It should also be noted that the estimate $\hat{x}_i$ is most likely to be biased. Finally, unless B_i is indeed small, the estimator will behave poorly even to the extent that if may be unstable.

A more clever observation was made in [29]. Consider instead of (3) the system

$$\begin{aligned}\dot{x}_i &= A_i x_i + \xi_i + g_i(x)\\ y_i &= H_{i1}x_i + H_{i2}g_i(x) + \eta_i.\end{aligned}$$

for which Sanders, Tacker and Linton propose the decentralized estimator

$$\dot{\hat{x}}_i = A_i\hat{x}_i + K_i(t)[y_i - H_{i1}\hat{x}_i]. \tag{5}$$

If the gain satisfies the condition

$$[K_i(t)H_{i2} - I]g_i(x) \equiv 0, \tag{6}$$

then $\hat{x}_i$ will be an unbiased estimate. Moreover, within the class of G_i which satisfy (6) the one which minimizes the mean square error can be found easily and to do so only the ith subsystem model need be known. Thus both off line and one line information can be decentralized. The proposed estimator, which of course gives a larger error variance than the best centralized estimator, is likely to be most useful when the interaction input $g_i(x)$ is directly measured so that y_i has two components, $y_i = (y_{i1}, y_{i2})$ where

$$y_{i2} = g_i(x) + \eta_{i2} .$$

Conversely the estimator (5) loses all significance when $g_i(x)$ is not measured i.e. when $H_{i2} = 0$. The stability of the estimator (5) is investigated in [30].

Reconsider now the system (3) again. It seems natural to see what improvement in the estimator can be achieved when we allow the dimension of the estimator (4) to be arbitrary, but still insisting that it depend on y_i alone. The optimal estimator in this case will, for each i, have the same dimension as x. In fact each will attempt to estimate x directly using its "own" measurement. Of course if the local controllers can exchange the estimates of their own states then the estimates can be improved without increasing the computational burden, but there is an increase in the cost of communication. For some additional discussion of this trade-off and examples see [31]. Yet another possibility for computing a decentralized estimator is to use the "ε-coupling method" introduced in [14a]. In the version as given by Cline [7] this method does not permit decentralization of either on line or off line information.[3] However if we impose a priori a decentralized estimator such as (4), then the ε-coupling technique would appear to be applicable.

As the final example which may be relevant here consider the estimation problem for the singularly perturbed system

$$\dot{x}_1 = A_{11}x_1 + A_{12}x_2 + A_{13}x_3 + \xi_1$$
$$\dot{x}_2 = A_{21}x_1 + A_{22}x_2 + \xi_2$$
$$\dot{x}_3 = A_{21}x_1 + A_{33}x_3 + \xi_3,$$

with measurements

$$y_2 = C_1x_1 + C_2x_2 \qquad \eta_2$$
$$y_3 = C_1x_1 + C_3x_3 + \eta_3.$$

Suppose A_{22}, A_{33} are stable. Setting $\varepsilon = 0$ gives the "slowly-varying steady states"

$$\bar{x}_2 = -A_{22}^{-1}[A_{21}x_1+\xi_2], \quad \bar{x}_3 = -A_{33}^{-1}[A_{31}x_1+\xi_3]$$

and hence the reduced-order degenerate system

$$\dot{\bar{x}}_1 = [A_{11}-A_{12}A_{22}^{-1}A_{21}-A_{13}A_{33}^{-1}A_{31}]\bar{x}_1 + \xi_1 - A_{12}A_{22}^{-1}\xi_2 - A_{13}A_{33}^{-1}\xi_3$$
$$y_2 = [C_1-C_2A_{22}^{-1}A_{21}]\bar{x}_1 - C_2A_{22}^{-1}\xi_2 + \eta_2$$
$$y_3 = [C_1-C_3A_{33}^{-1}A_{31}]\bar{x}_1 - C_3A_{33}^{-1}\xi_3 + \eta_3.$$

Let $\hat{\bar{x}}_1$ be the estimate for this degenerate system. Then the estimate for the boundary layer system is

$$\hat{\bar{x}}_2 = -A_{22}^{-1}A_{21}\hat{\bar{x}}_1 + \hat{\theta}_2, \quad \hat{\bar{x}}_3 = -A_{33}^{-1}A_{31}\hat{\bar{x}}_1 + \hat{\theta}_3$$

where

$$\varepsilon\dot{\hat{\theta}}_2 = A_{22}\hat{\theta}_2 + K_2[C_2\hat{\theta}_2+C_1\hat{x}_1-y_2]$$
$$\varepsilon\dot{\hat{\theta}}_3 = A_{33}\hat{\theta}_3 + K_3[C_3\hat{\theta}_3+C_1\hat{x}_1-y_3],$$

giving rise to a hierarchical estimator design. Haddad and Kokotovic [10a] who proposed this estimator have shown that $\hat{\bar{x}}_i - \hat{x}_i \to 0$ in quadratic mean as $\varepsilon \to 0$ where $\hat{x}_i$ is the optimal centralized estimator. Their result has been extended by Teneketzis and Sandell [35]. This technique is useful for decentralized estimation provided that the variable x_1, x_2, x_3 describe physically different subsystems, for then the hierarchical estimator above does reduce communication of on line information between subsystems.[4] Note also that if this is not the case then it is a non-trivial matter to determine the degenerate and boundary layer systems.

CONTROL OF SYSTEM BEHAVIOR

As mentioned before the bulk of the work on decentralized control is addressed to the problem of control rather than systems modelling or analysis of system behavior. The literature can be subdivided into three subclasses: that concerned with optimal stochastic control; with exploring decentralized strategies within a restricted structure; and with quasi-static optimization.

Optimal Stochastic Control

Consider the following centralized control problem:

$$x_t = f_t(x_{t-1}, v_t, u_t), \quad t = 1,..,T \qquad (7)$$

$$y_t = g_t(x_{t-1}, w_t), \quad t = 1,..,T \qquad (8)$$

where x_t is the state, u_t is the control input to be selected from a prespecified set U and y_t is the on line measurement; and v_t, w_t are random disturbances. Assume that the primitive random variables $x_0, v_1, .., v_T, w_1, .., w_T$ are independent. A feedback law is a sequence of functions $\gamma = (\gamma_1, ..., \gamma_T)$ where $\gamma_t : (y_1, .., y_t) \to u_t \in U$. Let Γ be the set of all feedback laws. The cost associated with γ is

$$J(\gamma) = E \sum_{1}^{T} h_t(x_t, u_t). \qquad (9)$$

[3] This may not be obvious from a cursory reading of [7]. To see it note that in the notation of the reference the term x_i^k depends on all of the previous x_j^{k-1} via the dependence of A_i^{k-1}, B_i^{k-1}.

[4] Of course even if this is not the case, the hierarchical estimator does reduce (centralized) processing of on line information.

$\gamma^* \in \Gamma$ is said to be optimal if it minimizes $J(\gamma)$. Two sets of results characterize optimal laws. The first set consist of "separation results." These effectively describe a small subset Γ_s of Γ which includes an optimal law. For example suppose we have complete information i.e. $y_t = x_{t-1}$ in (8). Then Γ_s consists of only those $\gamma = (\gamma_1, \ldots, \gamma_T)$ in which γ_t depends only on $y_t = x_{t-1}$. A more general result asserts that in the incomplete information case Γ_s consists of those γ in which γ_t has the form $\gamma_t = \phi_t \circ F_t(y_1, \ldots, y_t, u_1, \ldots, u_{t-1})$ where $F_t(y_1, \ldots, u_{t-1})$ is the conditional distribution of x_{t-1} given $y_1, \ldots u_{t-1}$ [36,37]. In the special case where f_t, g_t are linear and the primitive random variables are Gaussian then γ_t can be restricted even further to be of the form $\phi_t(\hat{x}_{t-1})$ where $\hat{x}_{t-1}$ is the conditional mean of x_{t-1} given $y_1, \ldots, u_{t-1}$. The second set of results consists of necessary and sufficient conditions for the otpimality of a proposed γ. These normally take the form of "minimum principles" and are usually derived via dynamic programming techniques. For example, in the complete information case $y_t = x_{t-1}$, we have this minimum principle:

Let $V_t(x)$ be functions defined recursively by $V_{T+1}(x) \equiv 0$, and

$$V_t(x) = \inf_{u \in U} \int [h_t(f_t(x, v_t, u)) + V_{t+1}(f_t(x, v_t, u))] P(dv_t), \quad t = T, \ldots, 1,$$

where $P(dv_t)$ is the probability distribution of v_t. Then $x = \{x_t(x_{t-1})\}$ is optimal if and only if for every t, x the infimum above is achieved at $u_t = \gamma_t(x)$.

With the appearance of Witsenhausen's example in [45] it became clear that the separation results for the decentralized control problem would be quite different from the centralized case. To discuss this it is more efficient to proceed in a formal setting following [46]. We replace (7), (8) by

$$x_t = f_t(x_{t-1}, v_t, u_t^1, \ldots, u_t^K), \quad t = 1, \ldots, T, \quad (10)$$

$$y_t^m = g_t^m(x_{t-1}, w_t^m), \quad m = 1, \ldots, M \text{ and } t = 1, \ldots, T. \quad (11)$$

The primitive random variables are now the x_0, v_t, w_t^m. y_t^m is the measurement made at t by the mth "observation post," and u_t^k is the control input selected at time t from a set U^k by the kth "control station." The decentralized information structure is described by specifying for each (t,k) a subset $Y_{t,k}$, $U_{t,k}$ from $\{y_\tau^m, u_\sigma^k \mid \tau \leq t, m \leq M, \sigma \leq t-1, k \leq K\}$ which is the set of all available measurements made and actions taken up to t. A feedback law is now of the form $\gamma = \{\gamma_t^k\}$ where $\gamma_t^k : (Y_{t,k}, U_{t,k}) \to u_t^k \in U^k$. Again γ^* is optimal if it minimizes $J(\gamma)$ given by (9) with $u_t = (u_t^1, \ldots, u_t^k)$. Witsenhausen's counterexample shows that even when the f_t, g_t^m are linear and the primitive random variables are all Gaussian there may be no optimal law $\{\gamma_t^k\}$ of the form $\gamma_t^k = \phi_t(\hat{x}_{t-1} | t,k)$ where $\hat{x}_{t-1}|t,k$ is the conditional mean of x_{t-1} given $(Y_{t,k}, U_{t,k})$. This example runs counter to the understanding gained from experience in the study of the centralized control problem.

Since the appearance of this counterexample work has progressed in two directions. On the one hand there have been attempts to demarcate those decentralized problems for which separate results, suggested by the centralized problem, continue to hold. On the other hand there are attempts to generate counterexamples to other plausible separation results. In the first category we must include the work of Ho and Chu [12, 13] and Hartman [11], some results of Witsenhausen in [46], Sandell and Athans [25], Kurtaran and Sivan [15] and Schoute [32].

In the second category we must include some perceptive observations made by Rhodes and Luenberger [22] and developed by Benveniste Bernhard and Cohen [4], a possible "ill-posedness" of some information structures noted in [5,25] and a counterexample by Varaiya and Walrand [41]. We review these results now.

Consider the model (9)-(11) with the information structure $\{Y_{t,k}, U_{t,k}\}$. Let Γ be the set of feedback laws compatible with this information structure. Let ω denote the collection of the primitive random variables, $\mathfrak{F}$ the σ-field generated by ω. From Assertions 1 and 3 of [46] it is easy to prove the following separation result.

Proposition Suppose that for each (t,k) the sub-σ-field of $\mathfrak{F}$ generated by $\{(Y_{t,k}, U_{t,k})\}$ is the same for all γ in Γ. Then it is enough to limit attention to the set Γ_s consisting of $\gamma = \{\gamma_t^k\}$ of the form $\gamma_t^k = \phi_t^k \circ F_t(Y_{t,k}, U_{t,k})$ where $F_t(Y_{t,k}, U_{t,k})$ is the conditional probability distribution of ω given $Y_{t,k}$, $U_{t,k}$.

Witsenhausen calls such a problem feedforward while Ho and Chu call it a static team. The condition of the proposition hold for the centralized control problem (7), (8) if f_t, g_t are linear. Ho and Chu [13] have discovered information structures for the linear versions of (10), (11) for which the condition also holds. Suppose now that in addition to this condition it also happens that $F_t(Y_{t,k}, U_{t,k})$ is Gaussian. Then of course Γ_s can be further restricted to consist of $\gamma = \{\gamma_t^k\}$ of the form $\gamma_t^k = \phi_t^k(\hat{\omega}_{t,k})$ where $\hat{\omega}_{t,k}$ is the conditional mean of ω given

$(Y_{t,k},U_{t,k})$. If moreover the cost functions h_t in (9) are positive semi-definite quadratic forms in x_t, u_t, then, using Radner's result [21], Ho and Chu have shown that the ϕ_t^k can be restricted to be affine functions. (Radner's result has been extended by Hartman [11]).

One interesting case is when f_t, g_t are linear, ω is Gaussian and the information structure corresponds to the "one-step delay sharing pattern." As Ho and Chu have shown, this case meets the condition of the proproposition. Hence for the quadratic cost function the optimal feedback law is an affine function of $\hat{\omega}_{t,k}$. A more easily implementable affine feedback law is also possible and has been worked out by Kurtaran and Sivan [15].

We now turn to the "negative" results. It turns out that for the n-step delay sharing pattern if $n \geq 2$ then the condition of the proposition fails to hold. Moreover, even the weaker forms of a separation result which was conjectured in (Ref. 46, Assertions 8,9) do *not* hold, as the example in [41] shows.

Examples have also been given to show that the specification of an information structure $\{Y_{t,k},U_{t,k}\}$ can easily become ill-posed or ambiguous if one is not careful. To see this suppose $u_{\tau,\sigma} \in U_{t,k}$ but that $(Y_{\tau,\sigma},U_{\tau,\sigma})$ is not contained in $(Y_{t,k},U_{t,k})$, in words: (t,k) knows what (τ,σ) does but not what (τ,σ) knows. It is then possible for (τ,σ) to encode all the information that is known namely $(Y_{\tau,\sigma},U_{\tau,\sigma})$ into the control action $u_{\tau,\sigma}$ and thereby transmit it to (t,k). In this way the intent of the specification of the information structure is violated. This observation has been made by Bismut [5] and Sandell and Athans [25].[5] This ambiguity can be easily removed in a number of ways as suggested in the footnote or by simply deleting certain kinds of specification of information structure.

A problem which seems much more intrinsic to the formulation of optimal decentralized control was apparently first noted by Rhodes and Luenberger [22] and later elaborated by Benveniste, Bernhard and Cohen [4]. Consider a linear version of (10), (11) with $M = K = 2$, with the information structure $(Y_{t,k},U_{t,k}) = \{(y_1^k,\dots,y_t^k),\ (u_1^k,\dots,u_{t-1}^k)\}$. In words: each controller k remembers its own previous observations but there is no "sharing" of on line information between the two. Suppose the cost function is quadratic, and suppose we restrict attention to those $\gamma = \{\gamma_t^k\}$ which are affine functions. Within this class the best γ generally has the following structure: At each time t, controller k will make an estimate of the action of controller j; since the latter's action depends on its own estimates of k's control, we get a recursive situation in which as t increases each controller has to make estimates of random variables of exponentially growing dimension! This example demonstrates in a dramatic manner an observation made earlier by Varaiya [40] and Athans [2]: while decentralization reduces communication of on line information it tend to increase the processing of this information. Because of this researchers have turned to study subsets of Γ in which the processing is restricted *a priori*. This literature is examined later.

As we have seen some effort has been devoted to discovering separation results. To our knowledge, the result of Varaiya and Walrand [42] is the only one which seeks to determine the counterpart of the minimum principle of centralized control. Unfortunately, due to limitations of space we cannot present this result here.

Decentralized Control with Fixed Structure

The search for optimal decentralized control laws has to date led to few practical design procedures. Hence research has been directed towards finding reasonable control laws from within a prespecified (and very small) subset Γ_F of the set of all feedback laws Γ which are compatible with the way the on line information is decentralized.[6] In general, such a subset Γ_F is specified to consist of all $\gamma \in \Gamma = \{\gamma_t^k\}$ where the γ_t^k have a fixed functional form. This form has inveriably been taken to be linear so that each $\gamma \in \Gamma_F$ is described by a finite set of parameters and the search within Γ_F becomes computationally feasible.

The first work in this direction is due to Chong and Athans (6) who considered the problem

$$\begin{aligned} \dot{x} &= Ax + B_1u_1 + B_2u_2 \\ y_1 &= C_1x + \eta_1 \\ y_2 &= C_2x + \eta_2, \end{aligned} \tag{12}$$

where as usual η_1, η_2 are independent white noises which are independent of the initial state $x(0)$. The on line information available to controller k at time t is $Y^k(t) = \{y_k(\tau)\,|\,\tau \leq t\}$. It is assumed that the off line information is available to both controllers. Thus Γ consists of all pairs $\{\gamma_1(t),\gamma_2(t)\}$ where $\gamma_k(t) : Y^k(t) \mapsto u_k(t)$ is

[5] A similar ambiguity arises in specifying the class of computational algorithms with a fixed finite memory. This was first pointed out by Winograd and Wolfe [43], and an example of an encoding scheme was given by Cohen [8]. In this context, it is possible to overcome the ambiguity by insisting on a smoothness condition [23] or a more careful definition [8].

[6] The idea of limiting the search to a strategy with fixed structure was first used in the context of differential games by Rhodes and Luenberger [22].

arbitrary. The fixed structure selected is the subset Γ_F consisting of all $\{\gamma_1, \gamma_2\}$ of the form $u_k(t) = D_k(t)\hat{x}_k(t)$ where $\hat{x}_k(t)$ is an estimate of $x(t)$ based on $Y^k(t)$:

$$\dot{\hat{x}}_k(t) = F_k(t)\hat{x}_k(t) + G_k(t)y_k(t) + H_k(t)u_k(t).$$

The "parameter" matrices D_k, F_k, G_k, H_k are arbitrary. For a quadratic cost index the authors determine the optimal parameter values of D_k, G_k, H_k.[7] It should be observed that the computational burden in finding these optimal values is much greater than finding the optimal centralized feedback law.

In the example above the off line information is available to all players and no communication of the decentralized on line information is permitted. Singh, Hassan and Titli [35] consider the opposite case: they calculate, for a deterministic linear quadratic problem, the optimal centralized feedback law using a hierarchical computational procedure in which the off line information is decentralized. Even though the procedure appear to be extendable to the stochastic case this and similar work is not discussed here because it seems that the critical obstacle is the decentralization of on line information.

If a certain amount of communication between the various controllers is permitted then the performance can certainly be improved. This was seen above in the case of decentralized state estimators. It should be possible, although such work has not yet been reported, to cascade the decentralized state estimators with decentralized feedback controllers for deterministic linear quadratic problems [18, 26].

A more novel idea was developed by Chong and Athans [6a]. They take a model of a linear interconnected system similar to (3) above:

$$\dot{x}_i = A_{ii}x_i + B_{ii}u_i + \sum_{j \neq i} [A_{ij}x_j + B_{ij}u_j],$$
$$y_i = H_i x_i + n_i \qquad , i = 1,..,k. \qquad (13)$$

and assume a quadratic cost function. They propose a two-layer structure. In the lower layer there are k local controllers; the ith local controller assumes a local model of the form

$$\dot{x}_i = A_{ii}x_i + B_{ii}u_i + v_i$$
$$y_i = H_i x_i + n_i . \qquad (14)$$

Here v_i is the "prediction" of the neglected interaction terms received from the higher layer coordinator. At time 0 $v_i(t)$, $0 \leq t \leq T$ is received, at time T the next prediction $v_i(t)$, $T \leq t \leq 2T$ is received and so on. The on line information $y_i(t)$ is also continuously available. At each time nT the coordinator receives all of the on line information, $y_i(t)$ $0 \leq t \leq nT$, $i=1,..,k$ and makes the predictions for the next period. Since the ith local controllers assumes (14) to be the "true" model its optimal control problem is of the classical LQG form. The coordinator recognizes this and then finds reasonable prediction values $v_i(t)$. A variety of "reasonable" choices are possible [4], each of which can be intuitively rationalized. It is however not yet possible to decide whether the proposed structure is of greater theoretical or practical interest than many other fixed structures which can be proposed. A good discussion of the issues is given in [4]. Another structure, similar to that of Chong and Athans and with some more theoretical analysis has been recently proposed by Forestier and Varaiya [9].

It is too early to be able to pass judgement on the research in the study of fixed structures. The only assertion which seems to be valid is that almost all of the proposed structures result in greatly increased computational burden and almost none of them have been rationalized on theoretically secure grounds.

Quasi-static Optimization

Consider the centralized problem

$$\begin{aligned} \text{minimize} \quad & E\, h(u_t, w_t, \theta) \\ & y_t = g(u_t, w_t, \theta) \\ & f(u_t, \theta) \leq 0 \\ & u_t \in U \end{aligned} \qquad (15)$$

where w_t is a sequence of independent, identically distributed random variables, u_t is the control input, y_t is the observed output and θ is an unknown parameter. f, g are known a priori (off line information). If $\theta = \theta^o$ is known this is a straightforward static optimization problem, the on-line information plays no role, and the optimal input is $u_t = u^o$ which is the solution of

$$\begin{aligned} \text{minimize} \quad & E\, h(u, w_t, \theta^o) \\ & f(u, \theta^o) \leq 0, \quad u \in U. \end{aligned}$$

When θ is unknown, repeated observation of y_t can be processed to yield improved estimates $\hat{\theta}_t$ of θ. In many cases these estimates can be used to select u_t in such a way that u_t converges to u^o. We call such a procedure for selecting u_t quasi-static optimal (QSO). The most well-known QSO procedure is the method of stochastic approximation first invented by Robbins and Munro [24].[8]

When θ is unknown, it is evident that the cost

[7] By requiring that $x_k(t)$ be an unbiased estimator of $x(t)$, $F_k(t)$ is no longer a free parameter. The authors however give no justification for this restriction.

[8] An excellent discussion of the literature and new results are given by Ljung [17]. It is useful and important to note that the assumption of independence of w_t can be considerably weakened.

function (15) is meaningless. We can see that a QSO procedure does minimize the cost function

$$\lim_{t} E\, h(u_t, w_t, \theta), \tag{16}$$

and if the convergence of u_t to u^o is rapid enough it may even minimize the average cost per unit time

$$\lim_{T} \frac{1}{T} \sum_{1}^{T} E\, h(u_t, w_t, \theta). \tag{17}$$

The point to observe is that both of these two cost functions "ignore" the choice of u_t during the "transient." Hence the set of QSO procedures is very large and so we may be able to find a QSO $\gamma = \{\gamma_t(y_1, \ldots, y_t)\}$ which is also computationally attractive. To appreciate this suppose we modify the cost (16) or (17) to

$$\sum_{1}^{\infty} \rho^t E\, h(u_t, w_t, \theta) \tag{18}$$

where $0 < \rho < 1$ is a discount factor. This function does weigh the transient behavior. In most cases this choice of cost will result in a unique optimal feedback law with an extremely complicated structure. Moreover to make sense of (18) it is generally necessary to specify a prior distribution on θ, whereas QSO procedures will minimize (16) or (17) for arbitrary θ.

A decentralized version of (15) would be this:

$$\begin{aligned} \text{minimize} \quad & E\, h(u_t^1, \ldots, u_t^K, w_t \theta) \\ & y_t^k = g^k(u_t^1, \ldots, u_t^K, w_t, \,), \; k = 1, \ldots, K. \\ & f(u_t^1, \ldots, u_t^K, \theta) \le 0 \\ & u_t^k \in u^k, \quad k = 1, \ldots, K. \end{aligned} \tag{19}$$

A decentralized control law is now of the form $\gamma = \{\gamma_t^k(y_1^k, \ldots, y_t^k)\}$. Suppose $\theta = \theta^o$ in unknown. A QSO γ would be such that for each k u_t^k converges to $u^{k,o}$ which solves the static problem

$$\begin{aligned} \text{minimize} \quad & E\, h(u^1, \ldots, u^K, w_t, \theta^o) \\ & f(u^1, \ldots, u^K, \theta^o) \le 0, \; u^k \in U^k. \end{aligned}$$

Barta and Varaiya [3] use stochastic approximation to generate QSO procedures for some special cases of (19) which arise in some simple microeconomic models. A very interesting model of "satisficing" behavior [34] is proposed in [44].

In [16] Lau, Persiano and Varaiya propose a QSO scheme to find the maximal flow of a single commodity through a capacitated network. This scheme is based on the labeling algorithm. Gallager [10] has proposed a QSO scheme to find the optimal flow for a multicommodity problem. His scheme is based on a gradient method and assumes a strictly convex cost function. Both schemes require some communication between controllers. These two papers suggest that many of the methods of decentralized deterministic optimization (see e.g. [1]) should be extendable to the stochastic quasi-static problem.

CONCLUSIONS

Research in decentralized stochastic control (DSC) has been largely concerned with decentralized versions of the centralized LQG problem. However the attempt to find optimal DSC laws was, with some exceptions, soon frustrated and this led to the study of feedback control laws with fixed linear structures. The choice of these structures has not been rationalized since we do not know how to compare these with the class of all possible decentralized laws. The choice has been motivated simply by our familiarity with the centralized LQG problem or more generally with the finite-time horizon centralized stochastic control problem. It seems to us that we should reexamine all of the components of this latter paradigmatic problem namely the way in which the system dynamics, the on-line measurements and the cost function are modelled. In the discussion of quasi-static optimization we observed how a modest change in the formulation of the cost function can lead to dramatic simplification in the structure of optimal strategies. From the results on adaptive control [47] it would appear that similar simplification is possible for some appropriately formulated dynamic problems. The point is not that we should abandon the study of optimal DSC laws but that we should not let ourselves be guided too closely by the centralized LQG problem.

ACKNOWLEDGEMENTS

The research reported here was sponsored in part by the National Science Foundation Grant ENG76-16816. Very fruitful discussions with Professors M. Athans, N. Sandell, Jr. and I. Rhodes are gratefully acknowledged. This paper is a greatly expanded version of one section of [28].

REFERENCES

[1] K. J. Arrow and L. Hurwicz: Decentralization and computation in resource allocation. R. W. Pfouts (ed.)(1960) Essays in Economics and Econometrics, U. North Carolina Press.

[2] M. Athans: Survey of decentralized control methods. Ann. Econ. Soc. Meas. 4, 345-355 (1975).

[3] S. Barta and P. Varaiya: Stochastic models of price adjustment. Ann. Econ. Soc. Meas. 5, 267-281 (1976).

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

[5] J. M. Bismut: An example of interaction between information and control: the transparency of the game. IEEE Trans. Auto. Contr. AC-18, 518-522 (1973).

[6] C-Y. Chong and M. Athans: On the stochastic control of linear systems with different information sets. IEEE Trans. Auto. Contr. AC-16, 423-430 (1971).

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

[7] T. Cline: Near-optimal state estimation for interconnected systems. IEEE Trans. Auto. Contr. AC-20, 348-351 (1975).

[8] A. I. Cohen: Rate of convergence and optimality conditions of root finding and optimization algorithms. Ph.D. Dissertation, Univ. of California, Berkeley, Ca. (1970).

[9] J. P. Forestier and P. Varaiya: Control of large Markov chains. Proc. IFAC Workshop on Control and Management of Integrated Industrial Complexes, Toulouse (1977).

[10] R. G. Gallager: Local optimization of routing tables for data networks. Proc. 1976 IEEE Conf. on Decision and Control, 956-961.

[10a] A. H. Haddad and P. V. Kokotovic: On singular perturbations in linear filtering and smoothing. Proc. 5th Symp. on Nonl. Estim. and its Appl. (1974).

[11] W. H. Hartman: An extension to Radner's theorem, in G. Leitmann (ed.)(1976), Multicriteria Decision Making and Differential Games, Plenum Press, New York.

[12] Y-C. Ho and K-C. Chu: Team decision theory and information structures in optimal control problems, Parts I and II, IEEE Trans. Auto. Contr. AC-17, 15-28 (1972).

[13] Y-C. Ho and K-C. Chu: Information structure in dynamic multi-person control problems. Automatica 10, 341-351 (1974).

[14] R. E. Kahn: Resource-sharing computer communication networks. Proc. IEEE 60, 1397-1407 (1972).

[14a] P. V. Kokotovic, W. R. Perkins, J. B. Cruz, Jr. and G. D'Ans: ε-coupling method for near optimal design of large scale linear systems. Proc. IEE 116, 889-892 (1969).

[15] B. Z. Kurtaran and R. Sivan: Linear quadratic gaussian control with one step delay sharing pattern. IEEE Trans. Auto. Contr. AC-19, 571-574 (1974).

[16] R. Lau, R. M. Persiano and P. Varaiya: Decentralized information and control: a network flow example. IEEE Trans. Auto. Contr. AC-17, 466-473 (1972).

[17] L. Ljung: Convergence of Recursive Stochastic Algorithms, Lund Inst. Tech., Div. Auto. Contr. Report 7403, Lund, Sweden (1974).

[18] T. D. Linton, E. C. Tacker and C. W. Sanders: Computational and performance aspects of static decentralized controllers. Proc. 1976 IEEE Conf. on Decision and Control.

[19] A. W. Michel: Stability analysis of stochastic composite systems. IEEE Trans. Auto. Contr. AC-20, 246-250 (1975).

[20] A. N. Michel and R. D. Rasmussen: Stability of stochastic composite systems. IEEE Trans. Auto. Contr. AC-21, 89-94 (1976).

[21] R. Radner: Team decision problems. Ann. Math. Statist. 33, 857-881 (1962).

[22] I. B. Rhodes and D. G. Luenberger: Stochastic differential games with constrained state estimators. IEEE Trans. Auto. Contr. AC-14, 476-481 (1969).

[23] J. Rissanen: On Optimum Root-Finding Algorithms, IBM Report RJ 726 (No. 13830)(1970).

[24] H. Robbins and S. Munro: A stochastic approximation method. Ann. Math. Statist. 22, 400-407 (1951).

[25] N. Sandell Jr. and M. Athans: Solution of some nonclassical LQG stochastic decision problems. IEEE Trans. Auto.

Contr. AC-19, 108-116 (1974).

[26] N. R. Sandell, Jr.: Decomposition vs. decentralization in large-scale system theory. Proc. 1976 IEEE Conf. on Decision and Control.

[27] N. R. Sandell, Jr.: Information flow in decentralized systems, in Y. C. Ho and S. K. Mitter (eds.)(1976) Directions in Large-Scale Systems, Plenum Press, New York.

[28] N. R. Sandell, Jr., P. Varaiya and M. Athans: A survey of decentralization control methods for large scale systems, submitted to IEEE Trans. Auto. Contr. (1977).

[29] C. W. Sanders, E. C. Tacker and T. D. Linton: A new class of decentralized filters for interconnected systems. IEEE Trans. Auto. Contr. AC-19, 259-262 (1974).

[30] C. W. Sanders, E. C. Tacker and T. D. Linton: Stability and performance of a class of decentralized filters. Int. J. Contr. 23 (1976).

[31] C. W. Sanders, E. C. Tacker, T. D. Linton and R. Y-S. Ling: Design and evaluation of large-scale estimation algorithms having information exchange. Proc. 1976 IEEE Conf. on Decision and Control, 306-311 (1976).

[32] F. C. Schoute: Information and control in packet-switched satellite communication. Proc. 1976 IEEE Conf. on Decision and Control, 297-305 (1976).

[33] A. Segall: The modelling of adaptive routing in data-communication networks. IEEE Trans. Comm. C-25 (1977).

[34] H. A. Simon: A behavioral model of rational choice. Quart. J. Econ. 69, 99-118 (1955).

[35] M. G. Singh, M. F. Hassan and A. Titli: Multilevel feedback control for interconnected systems using the prediction principle. IEEE Trans. Syst. Man Cyber. SMC-6, 233-239 (1976).

[36] C. Striebel: Sufficient statistics in the optimal control of stochastic systems. J. Math. Anal. Appl. 12, 576-592 (1965).

[37] C. Striebel: (1975) Optimal Control of Discrete Time Systems, Lecture notes in economics and mathematical systems, v. 110, Springer.

[38] D. Teneketzis: Perturbation Methods in Decentralized Stochastic Control, Report ESL-R-664, Electronics Systems Lab, M.I.T., Cambridge, Ma. (1976).

[39] D. Teneketzis and N. R. Sandell, Jr.: Linear regulator design for stochastic systems by a multiple time scale method, to appear in IEEE Trans. Auto. Contr. (1977).

[40] P. Varaiya: Trends in the theory of decision-making in large systems. Ann. Econ. Soc. Meas. 1, 493-499 (1972).

[41] P. Varaiya and J. Walrand: A counterexample for delay sharing patterns, unpublished (1977).

[42] P. Varaiya and J. Walrand: An optimality condition for decentralized stochastic control, in preparation (1977).

[43] S. Winograd and P. Wolfe: Private communication (1969).

[44] S. G. Winter: Satisficing, selection and the innovating remnant. Quart. J. Econ. 85, 237-261 (1971).

[45] H. S. Witsenhausen: A counterexample in stochastic optimum control. SIAM J. Contr. 6, 131-147 (1968).

[46] H. Witsenhausen: Separation of estimation and control for discrete time systems. Proc. IEEE 59, 1557-1566 (1971).

[47] B. Wittenmark: Stochastic adaptive control methods: a survey. Int. J. Contr. 21, 705-730 (1975).