The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 765–782

LEARNING WITH FINITE MEMORY[1]

MARTIN E. HELLMAN[2] AND THOMAS M. COVER

Massachusetts Institute of Technology and Stanford University

0. Abstract and summary. This paper develops the theory of the design and performance of optimal finite-memory systems for the two-hypothesis testing problem. Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed random variables drawn according to a probability measure $\mathscr{P}$. Consider the standard two-hypothesis testing problem with probability of error loss criterion in which $\mathscr{P} = \mathscr{P}_0$ with probability π_0; and $\mathscr{P} = \mathscr{P}_1$ with probability π_1. Let the data be summarized after each new observation by an m-valued statistic $T \in \{1, 2, \cdots, m\}$ which is updated according to the rule $T_n = f(T_{n-1}, X_n)$, where f is a (perhaps randomized) time-invariant function. Let $d: \{1, 2, \cdots, m\} \to \{H_0, H_1\}$ be a fixed decision function taking action $d(T_n)$ at time n, and let $P_e(f, d)$ be the long-run probability of error of the algorithm (f, d) as the number of trials $n \to \infty$. Define $P^* = \inf_{(f,d)} P_e(f, d)$. Let the a.e. maximum and minimum likelihood ratios be defined by $\bar{l} = \sup(\mathscr{P}_0(A)/\mathscr{P}_1(A))$ and $\underline{l} = \inf(\mathscr{P}_0(A)/\mathscr{P}_1(A))$ where the supremum and infimum are taken over all measurable sets A for which $\mathscr{P}_0(A) + \mathscr{P}_1(A) > 0$. Define $\gamma = \bar{l}/\underline{l}$. It will be shown that $P^* = [2(\pi_0\pi_1\gamma^{m-1})^{\frac{1}{2}} - 1]/(\gamma^{m-1} - 1)$, under the nondegeneracy condition $\gamma^{m-1} \geqq \max\{\pi_0/\pi_1, \pi_1/\pi_0\}$; and a simple family of ε-optimal (f, d)'s will be exhibited. In general, an optimal (f, d) does not exist; and ε-optimal algorithms involve randomization in f.

1. Introduction. Let $X_1, X_2, \cdots$ be a sequence of independent, identically distributed random observations drawn according to a probability measure $\mathscr{P}$ defined on an arbitrary probability space $(\mathfrak{X}, \mathscr{B}, \mathscr{P})$. Consider the simple hypothesis testing problem

$$H_0: \mathscr{P} = \mathscr{P}_0 \quad \text{vs.} \quad H_1: \mathscr{P} = \mathscr{P}_1. \tag{1}$$

Let the prior probabilities of the null hypothesis H_0 and the alternative hypothesis H_1 be denoted by π_0 and π_1 respectively.

Let $d_n \in \{H_0, H_1\}$ denote the decision made at time n. If d_n is allowed to depend on $X_1, X_2, \cdots, X_n$, it is well known that a standard likelihood ratio test yields a probability of error tending exponentially to zero as the sample size n tends to infinity. If the statistician has but a limited memory, several models come to mind. For example, d_n may be restricted to depend only on the last k observations [12]. This model fails to suit the author's notion of limited memory because the cardinality of the space in which $(X_1, X_2, \cdots, X_k)$ takes its values is unrestricted (and usually infinite). For the same reason, models in which the statistician is allowed to re-

Received January 9, 1969; revised January 5, 1970.

[1] This work was supported under contract AF 49(638)1517, and under the National Science Foundation Graduate Fellowship Program.

[2] Formerly with IBM Research and Stanford University.

member a single real number [15], [5] (e.g., a sufficient statistic such as the current likelihood ratio) fail to be nontrivial or practical constraints on memory. For even in those cases in which X takes on only a finite number of values, the likelihood ratio can take on an infinite number of possible values in the infinite-sample case.

For these reasons, the problem of learning under a *finite* memory constraint will be considered. Specifically, consider the family of all learning algorithms of the type

$$(2) \qquad \begin{aligned} T_n &= f(T_{n-1}, X_n); \qquad & f&:\{1, 2, \cdots, m\} \times \mathfrak{X} \to \{1, 2, \cdots, m\} \\ d_n &= d(T_n); & d&:\{1, 2, \cdots, m\} \to \{H_0, H_1\} \end{aligned}$$

where X_n is the nth observation, T_n is the state of the memory at time n, d_n is the nth decision, and f is a function (perhaps randomized), independent of n and the data. The algorithm is said to have *finite memory* of size m if T is m-valued (i.e., $T_n \in \{1, 2, \cdots, m\}$ for $n = 1, 2, \cdots$). The goal is to minimize the expected asymptotic proportion of errors

$$(3) \qquad P(e) = E\{\lim_{n\to\infty} n^{-1} \textstyle\sum_{i=1}^{n} e_i\},$$

where $e_i = 1$ or 0 accordingly as $d_i \neq H_t$ or $d_i = H_t$, where H_t denotes the true hypothesis. (In the case in which f describes an aperiodic ergodic process on $\{1, 2, \cdots, m\}$, it may be seen that $P(e) = \lim_{n\to\infty} \Pr\{d_n \neq H_t\}$.)

From elementary decision theoretic considerations it is clear that no randomization of d is required for optimal procedures. However, it will be shown to be generally true that an ε-optimal f requires artificial randomization.

The pair (f, d) describes a finite-state machine (automaton) with inputs X_n, outputs $d_n = d(T_n)$, and state space $S = \{1, 2, \cdots, m\}$. The state of the machine at time n is T_n. Under hypothesis H_t, $t = 0$ or 1, the sequence T_n, together with some specified initial state, forms a Markov chain over the state space S. The action of f may be prescribed by a (perhaps infinite) family of stochastic transition matrices indexed by x,

$$(4) \qquad \mathbf{P}(x) = [\Pr\{f(i, x) = j\}] = [p_{ij}(x)], \qquad i, j = 1, 2, \cdots, m,$$

where $\sum_{j=1}^{m} p_{ij}(x) = 1$, and $p_{ij}(x) \geqq 0$, $\forall i, j, x$. Here $p_{ij}(x)$ is the probability that $T_n = j$ given that $T_{n-1} = i$ and that $X_n = x$ is observed. Taking the expectation over x, it is found that the state transition probability matrices under H_0 and H_1 are given by

$$(5) \qquad \mathbf{P}^t = \int \mathbf{P}(x)\, d\mathscr{P}_t(x), \qquad t = 0, 1.$$

The stationary probability distributions on the states, denoted by

$$(6) \qquad \boldsymbol{\mu}^t = (\mu_1^{\ t}, \mu_2^{\ t}, \cdots, \mu_m^{\ t}), \qquad t = 0, 1$$

are then solutions to the matrix equations

$$(7) \qquad \boldsymbol{\mu}^t = \boldsymbol{\mu}^t \mathbf{P}^t, \qquad t = 0, 1.$$

The resulting long-run probability of error $P(e)$ is now simply given by

$$P(e) = \pi_0 \sum_{i \in S_1} \mu_i^{\,0} + \pi_1 \sum_{i \in S_0} \mu_i^{\,1}, \tag{8}$$

where $S_j = \{i : d(i) = H_j\}, j = 0, 1$, are the decision regions induced by the decision rule d. Note that the (f, d) description and the $(\mathbf{P}(x), d)$ description are equivalent. Define P^* to be the greatest lower bound on $P(e)$:

$$P^* = \inf_{(f,d)} P(e). \tag{9}$$

In this paper P^* will be found for every memory size m, and an ε-optimal class of (f, d)'s will be demonstrated. (That is, it will be shown that for any $\varepsilon > 0$ there exists an (f, d) in this class for which $P(e) \leqq P^* + \varepsilon$.) It will also be shown that, in general, no optimal (f, d) exists.

Section 2 establishes a lower bound P^* on $P(e)$ in terms of the statistics of the problem and the memory size m. Lemmas 1 and 2 and Theorem 3 are of primary importance, although Lemma 3 and Theorem 1, which treat the mathematically annoying case of reducible automata, are necessary for completeness. The section concludes with Theorem 4 which establishes P^* to be unachievable.

In Section 3 an ε-optimal class of automata is derived. The special case $P^* = 0$ is treated in Section 4, and examples are given in Section 5.

The uniqueness of the ε-optimal class is discussed in Section 6. Necessary and sufficient conditions are given for a class of automata to be ε-optimal, completing the solution of the problem.

Using different methods, the time-varying learning with finite memory algorithm

$$\begin{aligned} T_n &= f(T_{n-1}, X_n, n), \qquad T_n \in \{1, 2, \cdots, m\} \\ d_n &= d(T_n), \qquad d : \{1, 2, \cdots, m\} \to \{H_0, H_1\} \end{aligned} \tag{10}$$

has been shown by Cover [2] to yield $P^* = 0$ for a memory of size $m = 4$. Thus there exist learning rules for a time-varying finite memory which yield asymptotically zero probability of error. No such hope exists (except in special cases) in the time-invariant problem treated here. References [2] and [3] contain a further discussion of finite memory constraints and sufficient statistics.

Several authors in the Russian literature [16], [10], [17] have investigated the behavior of automata in random media. However, their work is primarily devoted to the analysis of the behavior of various *ad hoc* machine designs. Moreover, the problem formulations are more properly in the area of the sequential design of experiments (the so-called two-armed bandit problem) than in the area of hypothesis testing. This work is nonetheless interesting because of the similarity of the formalism to that of the problem considered here (see [8]). Under an alternative definition of finite memory (in which the memory consists solely of the last k observations) the two-armed bandit problem has been attacked by Robbins [12], Isbell [9], Smith and Pyke [14], Samuels [13], and Cover [1]. Cover and Hellman [3] have solved the two-armed bandit problem under a definition of finite memory similar to that used in this paper. The methods used are based on those developed in this paper. The hypothesis testing problem under the constraint that the memory be one dimensional (a single updatable real number) has been discussed by Spragins [15] and Fralick [5].

2. A lower bound for $P(e)$. Let f_0 and f_1 be the Radon–Nikodym derivative density functions of $\mathscr{P}_0$ and $\mathscr{P}_1$ with respect to a dominating measure ν, such as $\nu = \mathscr{P}_0 + \mathscr{P}_1$. Define the likelihood ratio (l.r.) to be $l(x) = f_0(x)/f_1(x)$. Let the ess sup and ess inf of the l.r. be denoted by

(11a) $$\bar{l} = \sup \mathscr{P}_0(A)/\mathscr{P}_1(A) \quad \text{and}$$

(11b) $$\underline{l} = \inf \mathscr{P}_0(A)/\mathscr{P}_1(A)$$

where the supremum and infimum are taken over all measurable sets A such that $\mathscr{P}_0(A) + \mathscr{P}_1(A) > 0$.

LEMMA 1. *The ratio of the probability p_{ij}^0 of transition from state i to state j under H_0, to the probability p_{ij}^1 of the same transition under H_1 satisfies the inequality*

(12) $$\underline{l} \leqq p_{ij}^0/p_{ij}^1 \leqq \bar{l}, \qquad \forall_{i,j}.$$

If both p_{ij}^0 and p_{ij}^1 are zero, their ratio is undefined.

PROOF. p_{ij}^k is equal to $\int_{\mathfrak{X}} p_{ij}(x) f_k(x)\, d\nu(x)$, for $k = 0, 1$. Since $f_0(x) = l(x) f_1(x)$,

(13) $$\frac{p_{ij}^0}{p_{ij}^1} = \frac{\int_{\mathfrak{X}} p_{ij}(x) l(x) f_1(x)\, d\nu(x)}{\int_{\mathfrak{X}} p_{ij}(x) f_1(x)\, d\nu(x)} \leqq \frac{\bar{l} \int_{\mathfrak{X}} p_{ij}(x) f_1(x)\, d\nu(x)}{\int_{\mathfrak{X}} p_{ij}(x) f_1(x)\, d\nu(x)} = \bar{l}.$$

Similarly, replacing $l(x)$ by its a.e. lower bound $\underline{l}$, the other inequality is obtained.

DEFINITION. Let

(14) $$\gamma = \bar{l}/\underline{l}.$$

Note that $\gamma \geqq 1$, with $\gamma = 1$ if and only if $\mathscr{P}_0 = \mathscr{P}_1$, a.e. The parameter γ will be seen to be a natural measure of the resolvability of the two hypotheses in the finite memory case.

Henceforth, it will be assumed that $\gamma < \infty$. From Lemma 1 it follows that for all i and j, p_{ij}^0 and p_{ij}^1 are either both positive or both zero. Hence, the classification into transient and recurrent states in the Markov chain is the same under either hypothesis. The case $\gamma = \infty$ will be disposed of in Section 4, and will be shown to be in agreement with the results for finite γ.

DEFINITION. The state likelihood ratio for state i is defined to be

(15) $$\lambda_i = \mu_i^0/\mu_i^1, \qquad i = 1, 2, \cdots, m,$$

the ratio of the stationary occupancy probabilities under H_0 and H_1.

LEMMA 2. *For an irreducible automaton in which the state likelihood ratios λ_i are arranged in nondecreasing order, the following relation holds:*

(16) $$1 \leqq \lambda_{i+1}/\lambda_i \leqq (\bar{l}/\underline{l}) = \gamma, \qquad i = 1, 2, \cdots, m-1.$$

REMARK. Irreducibility implies that λ_i is defined for all states, since $\mu_i^0 > 0$ and $\mu_i^1 > 0$; $i = 1, 2, \cdots, m$.

PROOF. The lower bound of (16) follows from the assumption that the state likelihood ratios have been arranged in nondecreasing order. To establish the upper bound the following fact [6] will be needed: If the state space S is partitioned into two sets C and C', then in the steady state the probability of transition from C to C' must equal the probability of transition from C' to C. This condition must be satisfied separately under H_0 and H_1. Precisely stated,

(17a) $$\sum_{j \in C} \sum_{k \in C'} \mu_j^0 p_{jk}^0 = \sum_{j \in C'} \sum_{k \in C} \mu_j^0 p_{jk}^0,$$

(17b) $$\sum_{j \in C} \sum_{k \in C'} \mu_j^1 p_{jk}^1 = \sum_{j \in C'} \sum_{k \in C} \mu_j^1 p_{jk}^1.$$

To establish the upper bound in (16), suppose the lemma were false. Then for some $i \in S$

(18a) $$\mu_j^0/\mu_j^1 \leqq c \qquad \text{for all } \ j \in C \equiv \{1, 2, \cdots, i\}$$

(18b) $$\mu_j^0/\mu_j^1 > c\bar{l}/\underline{l} \qquad \text{for all } \ j \in C' \equiv \{i+1, i+2, \cdots, m\}$$

where $c = \lambda_i$.

But using the inequalities (12) and (17a)

(19) $$\sum_{j \in C} \sum_{k \in C'} \mu_j^0 p_{jk}^0 \leqq \sum_{j \in C} \sum_{k \in C'} (c\mu_j^1)(\bar{l} p_{jk}^1),$$

so that

(20) $$\sum_{j \in C} \sum_{k \in C'} \mu_j^0 p_{jk}^0 \leqq c\bar{l} \sum_{j \in C} \sum_{k \in C'} \mu_j^1 p_{jk}^1.$$

Similarly,

(21) $$\sum_{j \in C'} \sum_{k \in C} \mu_j^0 p_{jk}^0 > (c\bar{l}/\underline{l})\underline{l} \sum_{j \in C'} \sum_{k \in C} \mu_j^1 p_{jk}^1.$$

But using (17a) the left sides of (20) and (21) are equal, and using (17b) the right sides are equal, a contradiction.

Note that the irreducibility of the machine is used in obtaining (21), for if the machine is reducible there exists a partition of S into two sets such that there is no flow from either one to the other. Then none of the ratios $\mu_j^0 p_{jk}^0/\mu_j^1 p_{jk}^1$ used to obtain (21) is defined.

This is not to say that reducible automata have no restrictions on the rate of increase of their state likelihood ratios. But a different condition and proof are needed (see Lemma 3 and Theorem 1).

COROLLARY 1. *The second inequality in* (16) *becomes an equality*

$$\lambda_{i+1}/\lambda_i = \gamma \qquad i = 1, 2, \cdots, m-1 \tag{22}$$

if and only if

$$p_{ij}^0 = \bar{l} p_{ij}^1 \qquad \textit{for} \quad 2 \leqq j = i+1 \leqq m \tag{23a}$$

$$= \underline{l} p_{ij}^1 \qquad \textit{for} \quad 1 \leqq j = i-1 \leqq m-1 \tag{23b}$$

$$= p_{ij}^1 = 0 \qquad \textit{for} \quad |i-j| \geqq 2. \tag{23c}$$

PROOF. Follows from the proof of Lemma 2.

It is intuitively clear that an irreducible automaton is better than a reducible automaton because of the lack of use of the transient states in the reducible case. However, there seems to be no simple way to dispose of the reducible case short of proving the following Lemma and Theorem.

LEMMA 3. *For a reducible automaton with several recurrent communicating classes $\mathscr{R}_1, \mathscr{R}_2, \cdots, \mathscr{R}_k$ and set $\mathscr{T}$ of transient states, with initial state $i_0 \in \mathscr{T}$,*

$$\frac{P^0(\mathscr{R}_i)P^1(\mathscr{R}_j)}{P^1(\mathscr{R}_i)P^0(\mathscr{R}_j)} \leqq \gamma^{m_T} \tag{24}$$

where m_T is the number of states in $\mathscr{T}$ and $P^t(\mathscr{R}_l)$ is the probability of absorption by $\mathscr{R}_l$ under hypothesis H_t, for $t = 0, 1$ and $l = 1, 2, \cdots, k$.

PROOF. Consider a new m-state automaton $\mathscr{A}'$ having the same state transitions as $\mathscr{A}$ between states $i \in \mathscr{T}$ and $j \in \mathscr{T}$. However, $\mathscr{A}'$ differs from $\mathscr{A}$ in that all transitions of $\mathscr{A}$ from $i \in \mathscr{T}$ to $j \notin \mathscr{T}$ are changed to transitions from i to i_0. Let $S' \subseteq \mathscr{T}$ denote the set of states which are accessible to $\mathscr{A}'$ from i_0. Let τ_i^0 and τ_i^1 denote the stationary distributions of $\mathscr{A}'$. Since $\mathscr{A}'$ restricted to S' is irreducible, Lemma 2 can be applied to yield

$$c \leqq \tau_i^0/\tau_i^1 \leqq c\gamma^{(m_T-1)}. \tag{25}$$

But, as is known in the theory of Markov Chains [7],

$$P^t(\mathscr{R}_l) = \sum_{i \in S'} \tau_i^t \sum_{n \in \mathscr{R}_l} p_{in}^t \tag{26}$$

where p_{in}^t is the state transition probability for the original automaton $\mathscr{A}$ conditioned on H_t, $t = 0, 1$.

Then applying (12), (25) and (26)

$$\begin{aligned} \frac{P^0(\mathscr{R}_i)P^1(\mathscr{R}_j)}{P^1(\mathscr{R}_i)P^0(\mathscr{R}_j)} &= \frac{[\sum_{l \in S'} \tau_l^0 \sum_{n \in \mathscr{R}_i} p_{ln}^0][\sum_{l \in S'} \tau_l^1 \sum_{n \in \mathscr{R}_j} p_{ln}^1]}{[\sum_{l \in S'} \tau_l^1 \sum_{n \in \mathscr{R}_i} p_{ln}^1][\sum_{l \in S'} \tau_l^0 \sum_{n \in \mathscr{R}_j} p_{ln}^0]} \\ &\leqq \frac{[c\gamma^{m_T-1}\bar{l} \sum_{l \in S'} \tau_l^1 \sum_{n \in \mathscr{R}_i} p_{ln}^1][\sum_{l \in S'} \tau_l^1 \sum_{n \in \mathscr{R}_j} p_{ln}^1]}{[\sum_{l \in S'} \tau_l^1 \sum_{n \in \mathscr{R}_i} p_{ln}^1][c\underline{l} \sum_{l \in S'} \tau_l^1 \sum_{n \in \mathscr{R}_j} p_{ln}^1]} \\ &= \gamma^{m_T-1}\bar{l}/\underline{l} = \gamma^{m_T}. \end{aligned} \tag{27}$$

□

DEFINITION. The *spread* σ of an automaton is the ratio of its maximum state likelihood ratio to its minimum state likelihood ratio. If the states are ordered as usual, so that $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_m$, then $\sigma = \lambda_m/\lambda_1$.

THEOREM 1. *The spread of a reducible automaton is less than or equal to* γ^{m-2}.

PROOF. *Case* I. If the automaton has only one recurrent communicating class, $\mathscr{R}$ (with $m_1 < m$ states) and a set of transient states $\mathscr{T}$ (with $m_T = m - m_1$ states), the automaton must eventually reach $\mathscr{R}$ independently of the initial state. Thus the machine effectively reduces to an m_1 state irreducible automaton. Then by repeated application of Lemma 2, the spread is less than or equal to γ^{m_1-1}, which is less than or equal to γ^{m-2}.

Case II. The automaton has no transient states, but several recurrent communicating classes $\mathscr{R}_1, \mathscr{R}_2, \cdots, \mathscr{R}_k$ having $m_1, m_2, \cdots, m_k$ states respectively. If the automaton starts in state $i \in \mathscr{R}_j$, it never leaves $\mathscr{R}_j$. Thus the machine is effectively irreducible with m_j states. Again by Lemma 2 the spread is less than or equal to $\gamma^{m_j-1} \leqq \gamma^{m-2}$.

Case III. There are several recurrent communicating classes, $\mathscr{R}_1, \mathscr{R}_2, \cdots, \mathscr{R}_k$, having $m_1, m_2, \cdots, m_k$ states respectively. In addition there is a set of transient states $\mathscr{T}$ having m_T states, where $m_T \geqq 1$. If the automaton starts in a recurrent state, Case II applies. If, however, the automaton starts in a state $i_0 \in \mathscr{T}$, several of the $\mathscr{R}_i$'s may be accessible from i_0.

Let v_i^t be the stationary distribution for the m_i-state recurrent class $\mathscr{R}_i$. Then for $r \in \mathscr{R}_i$,

$$\mu_r^t = P^t(\mathscr{R}_i)v_{ir}^t, \quad \text{and} \tag{28}$$

$$1/\gamma^{m_i-1} \leqq v_{ir}^0/v_{ir}^1 \leqq \gamma^{m_i-1}. \tag{29}$$

Thus, for $r \in \mathscr{R}_i$, $s \in \mathscr{R}_j$,

$$\lambda_r/\lambda_s = \left[\frac{P^0(\mathscr{R}_i)P^1(\mathscr{R}_j)}{P^1(\mathscr{R}_i)P^0(\mathscr{R}_j)}\right]\frac{[v_{ir}^0/v_{ir}^1]}{[v_{js}^0/v_{js}^1]} \tag{30}$$

$$\leqq \gamma^{m_i-1} < \gamma^{m-2} \qquad \text{if } i = j$$

$$\leqq \gamma^{m_T}\gamma^{m_i-1}\gamma^{m_j-1} \leqq \gamma^{m-2} \qquad \text{if } i \neq j. \tag{31}$$

This is the desired result.

THEOREM 2. *The spread of an m-state automaton is less than or equal to* γ^{m-1}.

PROOF. If the automaton is irreducible then $m-1$ applications of Lemma 2 yield the desired result. If the automaton is reducible, Theorem 1 gives an even stronger bound.

THEOREM 3. *For an m-state automaton, $P(e) \geqq P^*$ where*

$$P^* = \frac{2(\pi_0 \pi_1 \gamma^{m-1})^{\frac{1}{2}} - 1}{(\gamma^{m-1} - 1)}, \quad \text{if} \quad \gamma^{m-1} \geqq \max\{\pi_0/\pi_1, \pi_1/\pi_0\}; \tag{32}$$

$$= \min\{\pi_0, \pi_1\}, \qquad \text{otherwise.}$$

REMARK. This theorem shows P^* to be a lower bound on $P(e)$. In the next section P^* will be shown to be the greatest lower bound.

REMARK. If $\pi_0 = \pi_1 = \frac{1}{2}$, (32) reduces to

$$P^* = (\gamma^{\frac{1}{2}(m-1)} + 1)^{-1}. \tag{33}$$

REMARK. It has been assumed that $\gamma < \infty$. However, as shown in Section 4, if $\gamma = \infty$ then $P^* = 0$, in agreement with this theorem. Thus the theorem holds without restriction on γ.

REMARK. The case $\gamma^{m-1} < \max\{\pi_0/\pi_1, \pi_1/\pi_0\}$ is trivial, as it can be seen from Theorem 2 that then no state likelihood ratio will cause a reversal of the a priori decision. No machine is needed since the trivial rule of deciding upon the hypothesis with the larger prior achieves the lower bound $\min\{\pi_0, \pi_1\}$. Note that $\gamma^{m-1} = \max\{\pi_0/\pi_1, \pi_1/\pi_0\}$ implies that $P^* = \min\{\pi_0, \pi_1\}$, in agreement with this heuristic discussion.

PROOF OF THEOREM 3. If k is the minimum state likelihood ratio, then by Theorem 2

$$k \leqq \mu_i^0/\mu_i^1 \leqq k\gamma^{m-1}, \qquad \forall\, i \in S. \tag{34}$$

Using this equation and letting $\alpha = P(e \mid H_0)$ and $\beta = P(e \mid H_1)$, it follows that

$$\alpha = \textstyle\sum_{i \in S_1} \mu_i^0 \geqq k \sum_{i \in S_1} \mu_i^1 = k(1-\beta), \qquad \text{or}$$

$$\alpha \geqq k(1-\beta); \qquad \text{and} \tag{35}$$

$$\beta = \textstyle\sum_{i \in S_0} \mu_i^1 \geqq (1/k\gamma^{m-1}) \sum_{i \in S_0} \mu_i^0, \qquad \text{or}$$

$$\beta \geqq (k\gamma^{m-1})^{-1}(1-\alpha). \tag{36}$$

Multiplying (35) and (36) one obtains

$$\alpha\beta \geqq \gamma^{-(m-1)}(1-\alpha)(1-\beta). \tag{37}$$

Equation 37 gives a lower boundary for the operating characteristic of a decision rule (automaton) with memory of size m. Thus the results of this analysis apply to a Neyman–Pearson formulation of the problem, as well as to the Bayesian approach that is being followed. The algorithms that will be demonstrated in the next section can approach any point on the operating characteristic.

Straightforward Lagrange minimization [7] of $P(e) = \pi_0\alpha+\pi_1\beta$, subject to the inequality constraint (37), yields minimizing values

$$\alpha^* = [((\pi_1/\pi_0)\gamma^{m-1})^{\frac{1}{2}}-1]/(\gamma^{m-1}-1), \tag{38}$$
$$\beta^* = [((\pi_0/\pi_1)\gamma^{m-1})^{\frac{1}{2}}-1]/(\gamma^{m-1}-1)$$

and the resulting minimum value of $P(e)$

$$P^* = [2(\pi_0\pi_1\gamma^{m-1})^{\frac{1}{2}}-1]/(\gamma^{m-1}-1) \tag{39}$$

for $\gamma^{m-1} \geqq \max\{\pi_0/\pi_1, \pi_1/\pi_0\}$. Thus $P(e) \geqq P^*$ as was to be shown.

COROLLARY 2. *A reducible* $(m+1)$*-state automaton obeys the same bound* (32) *on $P(e)$ as an m-state irreducible automaton.*

PROOF. If the bound of Theorem 1 is used in the proof of Theorem 3, it is seen that a lower bound on $P(e)$ for an $(m+1)$-state reducible automaton is the same as the lower bound derived for a general m-state automaton. Thus a reducible automaton "wastes" at least one state.

The following theorem shows the unachievability of the lower bound P^* in all but degenerate cases.

THEOREM 4. *If $m > 2$ and $\gamma^{m-1} > \max\{\pi_0/\pi_1, \pi_1/\pi_0\}$, then $P(e) > P^*$.*

PROOF. From the proof of Theorem 3 it is seen that $P(e) = P^*$ implies that equality must hold in (37). This in turn implies that

$$\mu_i^0 = \mu_i^1 = 0, \qquad i \neq 1 \quad \text{or} \quad m. \tag{40}$$

Since irreducible automata have all $\mu_i > 0$, any automaton achieving $P(e) = P^*$ with $m > 2$ must be reducible. Corollary 2 rules out this possibility.

3. A class of ε-optimal automata. In this section an ε-optimal class of automata will be demonstrated, i.e., for every $\varepsilon > 0$, it will be shown that there exists an automaton in this class with $P(e) \leqq P^*+\varepsilon$. Thus, the lower bound P^*, although shown to be unachievable by Theorem 4, is established as the greatest lower bound, completing the solution of the problem.

The structure of an ε-optimal automaton is almost given away by the following theorem. However, an understanding of this theorem is not necessary for what follows.

THEOREM 5. *Let $\{\mathscr{A}(n)\}$ be a sequence of m-state automata. Further let $\{\boldsymbol{\mu}^t(n)\}$, $t = 0, 1$, and $\{P_e(n)\}$ be the associated sequences of $\boldsymbol{\mu}^t$ and $P(e)$. If*

$$\gamma^{m-1} > \max\{\pi_0/\pi_1, \pi_1/\pi_0\},$$

then $P_e(n) \to P^*$ if and only if

1. $\{[\mu_m^0(n)][\mu_1^1(n)]\}/\{[\mu_m^1(n)][\mu_1^0(n)]\} \to \gamma^{m-1}$. *(It is assumed that for each n the states of $\mathscr{A}(n)$ are numbered in order of increasing state likelihood ratio.)*
2. $\mu_i^t(n) \to 0$, $2 \leqq i \leqq m-1$, $t = 0, 1$ *and*
3. $\mu_1^0(n) \to [((\pi_1/\pi_0)\gamma^{m-1})^{\frac{1}{2}}-1]/[\gamma^{m-1}-1]$.

REMARK. It is assumed for each n, that the decision rule d is optimal with respect to the state transition function f; i.e., $d(i) = H_0$ if and only if $\pi_0\mu_i^0(n) > \pi_1\mu_i^1(n)$.

PROOF. The sufficiency of the conditions will be proved first. Since the expression (8) for $P(e)$ is a finite sum,

$$P_e \equiv \lim_{n\to\infty} P_e(n) = \pi_0 \textstyle\sum_{i\in S_1} \mu_i^0 + \pi_1 \sum_{i\in S_0} \mu_i^1 \tag{41}$$

where

$$\mu_i^t \equiv \lim_{n\to\infty} \mu_i^t(n), \qquad t = 0, 1; \quad i \in S. \tag{42}$$

Thus condition 2 implies

$$P_e = \pi_0 \mu_1^0 + \pi_1 \mu_m^1 \tag{43}$$

and

$$\mu_1^t + \mu_m^t = 1, \qquad t = 0, 1. \tag{44}$$

Conditions 1 and 3, together with (44) result in

$$\mu_m^1 = \frac{[(\pi_0/\pi_1)\gamma^{m-1}]^{\frac{1}{2}} - 1}{\gamma^{m-1} - 1}. \tag{45}$$

Hence

$$P_e = \frac{2(\pi_0\pi_1\gamma^{m-1})^{\frac{1}{2}} - 1}{\gamma^{m-1} - 1} = P^*, \tag{46}$$

proving the sufficiency of the conditions.

The necessity of the conditions follows from the proof of Theorem 3. For $P(e)$ to approach P^* it is necessary that (37) approach equality and that $\alpha \to \alpha^*$, $\beta \to \beta^*$ as given in (38). □

We shall now propose an automaton which achieves the maximum spread of Condition 1 of Theorem 5. We will subsequently modify this automaton in order to meet Conditions 2 and 3. Consider an automaton which moves up on high-likelihood-ratio events and down on low-likelihood-ratio events. To this end define the sets

$$\mathscr{H}_\varepsilon = \{x \in \mathfrak{X} : l(x) \geqq [(1/\bar{l}) + \varepsilon]^{-1}\} \tag{47a}$$

$$\mathscr{T}_\varepsilon = \{x \in \mathfrak{X} : l(x) \leqq (\underline{l} + \varepsilon)\} \tag{47b}$$

$$\mathscr{S}_\varepsilon = \{x \in \mathfrak{X} : x \notin (\mathscr{H}_\varepsilon \cup \mathscr{T}_\varepsilon)\}. \tag{47c}$$

Further define, for $t = 0, 1$

$$h_\varepsilon^t = \mathscr{P}_t(\mathscr{H}_\varepsilon) \tag{48a}$$

$$t_\varepsilon^t = \mathscr{P}_t(\mathscr{T}_\varepsilon). \tag{48b}$$

By definition of $\bar{l}$ and $\underline{l}$, for any $\varepsilon > 0$, h_ε^t and t_ε^t are strictly greater than zero. Note also that

$$\bar{l} h_\varepsilon^1 \geqq h_\varepsilon^0 = \textstyle\int_{\mathscr{H}_\varepsilon} f_0 \, dv \geqq [(1/\bar{l}) + \varepsilon]^{-1} \int_{\mathscr{H}_\varepsilon} f_1 \, dv \tag{49}$$

or equivalently

$$\bar{l}h_\varepsilon^{\,1} \geqq h_\varepsilon^{\,0} \geqq [(1/\bar{l})+\varepsilon]^{-1}h_\varepsilon^{\,1}. \tag{50}$$

Similarly

$$\underline{l}t_\varepsilon^{\,1} \leqq t_\varepsilon^{\,0} \leqq (\underline{l}+\varepsilon)t_\varepsilon^{\,1}. \tag{51}$$

Consider the automaton with state transition function

$$\begin{aligned} f(i,x) &= i+1 && \text{if } \; i < m, && x \in \mathscr{H}_\varepsilon; \\ &= i-1 && \text{if } \; i > 1, && x \in \mathscr{T}_\varepsilon; \\ &= i && \text{otherwise.} \end{aligned} \tag{52}$$

Using (17) and letting $C = \{1, 2, \cdots, i\}$ and $C' = \{i+1, i+2, \cdots, m\}$ it is seen that

$$\mu_i^t h_\varepsilon^{\,t} = \mu_{i+1}^t t_\varepsilon^{\,t}, \qquad t = 0, 1; \quad i = 1, 2, \cdots, m-1. \tag{53}$$

From (50), (51) and (53)

$$\gamma = (\bar{l}/\underline{l}) \geqq \frac{\mu_{i+1}^0/\mu_{i+1}^1}{\mu_i^{\,0}/\mu_i^{\,1}} = \frac{h_\varepsilon^{\,0}t_\varepsilon^{\,1}}{t_\varepsilon^{\,0}h_\varepsilon^{\,1}} \geqq \frac{1}{[(1/\bar{l})+\varepsilon](\underline{l}+\varepsilon)}. \tag{54}$$

Therefore, for a sequence of automata $\{\mathscr{A}(n)\}$ with $\varepsilon(n) \to 0$, $\lambda_{i+1}(n)/\lambda_i(n) \to \gamma$ and the spread $\sigma(n) \to \gamma^{m-1}$. Condition 1 of Theorem 5 is thus satisfied.

In order to satisfy conditions 2 and 3 a slight modification of the state transition function will suffice. Let $f(i, x)$ be specified as follows (see Figure 1):

For $2 \leqq i \leqq m-1$ let

$$\begin{aligned} f(i,x) &= i+1, && x \in \mathscr{H}_\varepsilon \\ &= i, && x \in \mathscr{S}_\varepsilon \\ &= i-1, && x \in \mathscr{T}_\varepsilon. \end{aligned} \tag{55a}$$

In the end states let

$$\begin{aligned} f(1,x) &= 2, \quad \text{with probability} \quad \delta > 0 \quad \text{if} \quad x \in \mathscr{H}_\varepsilon; \\ &= 1, \quad \text{otherwise.} \end{aligned} \tag{55b}$$

$$\begin{aligned} f(m,x) &= m-1, \quad \text{with probability} \quad k\delta > 0 \quad \text{if} \quad x \in \mathscr{T}_\varepsilon; \\ &= m, \quad \text{otherwise.} \end{aligned} \tag{55c}$$

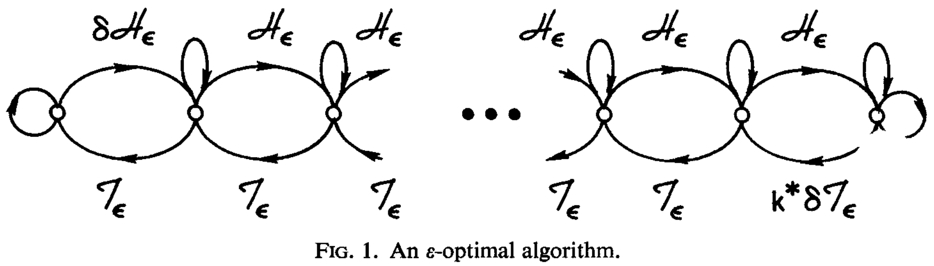

FIG. 1. An ε-optimal algorithm.

If in this modified sequence of automata $\{\mathscr{A}'(n)\}$, $\delta(n) \to 0$ with $\delta(n) > 0$ for all n, then Condition 2 of Theorem 5 is satisfied. Furthermore, by varying k it is possible to force $\mu_1{}^0$ to any desired value. Thus for some value $k = k^*$, Condition 3 of Theorem 5 is satisfied. It is easily verified that these modifications do not affect λ_{i+1}/λ_i. Hence, the spread $\sigma(n)$ is unaffected and $P_\varepsilon(n) \to P^*$. This demonstrates the ε-optimality of the class of automata depicted in Figure 1.

The value of k^*, obtained by straightforward minimization of $P(e)$, depends on the relative probabilities of $\mathscr{H}_\varepsilon$ and $\mathscr{T}_\varepsilon$ (and hence is a function of ε) and on the a priori probabilities.

Defining

$$\gamma_0 = \mathscr{P}_0(\mathscr{H}_\varepsilon)/\mathscr{P}_0(\mathscr{T}_\varepsilon) \tag{56a}$$

$$\gamma_1 = \mathscr{P}_1(\mathscr{H}_\varepsilon)/\mathscr{P}_1(\mathscr{T}_\varepsilon), \tag{56b}$$

$k^*(\varepsilon)$ is given by

$$\begin{aligned} k^*(\varepsilon) &= \frac{\pi_1 \gamma_0^{m-1} - \pi_0 \gamma_1^{m-1}}{(\pi_0 - \pi_1) + (\gamma_0^{m-1} - \gamma_1^{m-1})\sqrt{\dfrac{\pi_0 \pi_1}{(\gamma_0 \gamma_1)^{m-1}}}}, \\ &\qquad\qquad \text{if } \gamma^{m-1} > \max\{\pi_0/\pi_1, \pi_1/\pi_0\}; \\ &= 0, \qquad \text{if } \gamma^{m-1} \leqq \max\{\pi_0/\pi_1, \pi_1/\pi_0\} \text{ and } \pi_0 > \pi_1; \\ &= \infty, \qquad \text{if } \gamma^{m-1} \leqq \max\{\pi_0/\pi_1, \pi_1/\pi_1\} \text{ and } \pi_0 < \pi_1. \end{aligned} \tag{57}$$

It is interesting to note that although the state transition matrices $\mathbf{P}(x)$ were allowed to differ with x, and therefore could have been infinite in number, it is found that only three are needed, one each for the regions $\mathscr{H}_\varepsilon$, $\mathscr{T}_\varepsilon$ and $\mathscr{S}_\varepsilon$.

It is also of interest to note that if X has a continuous probability distribution, it is possible to satisfy the conditions of Theorem 5 without recourse to artificial randomization. Define sets $\mathscr{H}_\varepsilon' \subset \mathscr{H}_\varepsilon$, $\mathscr{T}_\varepsilon' \subset \mathscr{T}_\varepsilon$ such that, for $t = 0, 1$,

$$\begin{aligned} \mathscr{P}_t(\mathscr{H}_\varepsilon')/\mathscr{P}_t(\mathscr{H}_\varepsilon) &\approx \delta \\ \mathscr{P}_t(\mathscr{T}_\varepsilon')/\mathscr{P}_t(\mathscr{T}_\varepsilon) &\approx k^*\delta. \end{aligned} \tag{58}$$

For ε sufficiently small, the approximations (58) can be made as accurate as desired. Then if the automaton exits from state 1 with probability one when $x \in \mathscr{H}_\varepsilon'$ is observed, and exits from state m with probability one when $x \in \mathscr{T}_\varepsilon'$ is observed, the desired behavior is achieved, and no artificial randomization is required. Thus in the continuous case, but not in the discrete case, deterministic rules can be optimal.

A more detailed derivation of this ε-optimal class is given in [7].

4. The case of infinite γ. Although it was assumed that $\gamma < \infty$ for the derivation of P^*, it is easy to show that if $\gamma = \infty$ then $P^* = 0$ for any memory at all ($m \geqq 2$), in agreement with (32). If $\gamma = \infty$, either $\bar{l} = \infty$ or $\underline{l} = 0$. We shall treat the case $\bar{l} = \infty$ since interchanging the role of H_0 and H_1 yields the other case.

Let $\mathscr{H} = \{x \in \mathscr{X} : l(x) = \infty\}$. Then $\mathscr{P}_1(\mathscr{H}) \equiv 0$. If $\mathscr{P}_0(\mathscr{H}) > 0$, then the following two-state machine achieves $P(e) = 0$. Start in state 1; transit to state 2 only if $x \in \mathscr{H}$ is observed; and once in state 2 never leave. Decide H_0 in state 2 and H_1 in state 1. Then, with probability one, this automaton makes only a finite number of errors.

If $\mathscr{P}_0(\mathscr{H}) = 0$, then no machine achieves $P(e) = P^* = 0$. However, the ε-optimal class of Section 3 will ε-achieve $P^* = 0$. All the formulae of Section 3 may be applied. For example, $\mathscr{H}_\varepsilon = \{x \in \mathscr{X} : l(x) \geqq 1/\varepsilon\}$ which is the extension of (47a) to the case $\bar{l} = \infty$.

5. Examples.

EXAMPLE 1. Let X be a Bernoulli random variable with distribution

$$(59) \qquad \begin{aligned} X &= \text{Heads} \quad , \quad p \\ &= \text{Tails} \quad , \quad 1-p. \end{aligned}$$

Consider the two-hypothesis testing problem $H_0 : p = p_0$ vs. $H_1 : p = p_1$, under equal priors $\pi_0 = \pi_1 = \frac{1}{2}$. Recall in the case $\pi_0 = \pi_1$ that the ε-achievable lower bound on the probability of error reduces to $P^* = 1/(1 + \gamma^{\frac{1}{2}(m-1)})$.

(a) Let $p_0 = .99 \cdots 99$ and $p_1 = .99 \cdots 90$ (with the same number of 9's in between). This problem appears difficult because of the large number of trials necessary to obtain a significant test of the small difference between p_0 and p_1. Since an m-state automaton can only "count to m," it seems that memory will be exhausted before the test reaches an interesting level of significance. However, in this problem $\bar{l} = p_0/p_1 \cong 1$, $\underline{l} = q_0/q_1 = .1$, and $\gamma = \bar{l}/\underline{l} = p_0 q_1/p_1 q_0 \cong 10$. Thus for $m = 5$ (a five-state memory), $P^* = 1/101 \cong .01$.

(b) Now let $p_0 = \frac{3}{4}$, $p_1 = \frac{1}{4}$. Here $\gamma = p_0 q_1/p_1 q_0 = 9$, and $P^* = 1/82$ (for a 5-state automaton). This probability of error is actually higher than that of the previous example in which $p_0 = .99 \cdots 99$ and $p_1 = .99 \cdots 90$.

(c) Of peculiar interest is the case $p_0 = .501$, $p_1 = .499$. Here $\gamma \cong 1.008$, which yields $P^* \cong .496$ for a 5-state automaton–little better than using no memory at all. In fact, approximately 500 states are required to obtain $P^* = .01$. Clearly the difference $|p_0 - p_1|$ is a poor measure of the resolvability of H_0 vs. H_1 in the finite-memory case.

In Examples 1(a), 1(b) and 1(c) the optimal algorithm moves up one state on Heads and down one state on Tails, with appropriate randomization in the end states. Also in Examples 1(b) and 1(c) it is seen from the symmetry $\pi_1 = \pi_0 = \frac{1}{2}$ and $p_0 = 1 - p_1$ that $k^* = 1$. In Example 1(a), $k^* \gg 1$ to offset the drift to the right caused by $p_0, p_1 \approx 1$.

The difference between examples 1(a) and 1(c) is that in Example 1(a) there is an event (Tails) which occurs much more frequently under one hypothesis (H_1) than under the other. By essentially disregarding the other events, the high information content of the extreme event is well utilized. Neither Heads nor Tails is a high information event in Example 1(c).

EXAMPLE 2. Let X be a univariate normal random variable with mean $\mu = +1$ (under H_0) and $\mu = -1$ (under H_1) and fixed variance $\sigma^2 = 1$. Let $\pi_0 = \pi_1 = \frac{1}{2}$. In this case the likelihood ratio is given by $l(x) = \exp(2x)$. Therefore $\bar{l} = \infty$, $\underline{l} = 0$, and $\gamma = \infty$—resulting in $P^* = 0$ for any memory whatsoever ($m \geqq 2$). To achieve this, let $\mathscr{H}_\varepsilon = \{x : x \geqq T\}$ and $\mathscr{T}_\varepsilon = \{x : x \leqq -T\}$. Move to state 1 for $x \in \mathscr{T}_\varepsilon$, to state 2 for $x \in \mathscr{H}_\varepsilon$, and remain in the current state otherwise. Then $P(e)$ tends to zero as $T \to \infty$.

EXAMPLE 3. X has a Cauchy distribution with pdf $f(x) = 1/\pi(1+(x-\mu)^2)$. Test $H_0 : \mu = 1$ vs. $H_1 ; \mu = -1$ with $\pi_0 = \pi_1 = \frac{1}{2}$. This example is of interest because the Cauchy and the normal distributions have similar shapes and have comparable convergence rates for the probabilities of error in the infinite-memory case. However, calculation shows that $\bar{l} = \underline{l}^{-1} \cong 5.8$ and $\gamma \cong 33.6$. Thus a 2-state memory yields $P^* \cong .15$ for the Cauchy distribution, in marked contrast to the $P^* = 0$ obtainable in the normal case. The optimal algorithm for the Cauchy distribution also differs markedly from that for the normal distribution. Here $\mathscr{H}_\varepsilon$ and $\mathscr{T}_\varepsilon$ are small intervals centered at $x = \pm 2^{\frac{1}{2}}$, in contrast with the semi-infinite intervals in Example 2.

6. Characterization of all ε-optimal algorithms. The real purpose of this somewhat technical section is to show which learning algorithms are bad. P^* has now been established as a lower bound on $P(e)$, and the ε-achievability of P^* has been demonstrated by the family of automata depicted in Figure 1. Theorem 5 gives necessary and sufficient conditions for an algorithm to be ε-optimal. However, these conditions only relate to gross properties of the algorithm and do not adequately characterize the structure of all ε-optimal algorithms. In this section more explicit conditions will be given. From Theorem 5 and Corollary 1 it appears that any ε-optimal algorithm must resemble that of Figure 1. However, this is not entirely true, as is demonstrated by the class of six-state machines depicted in Figure 2. Using (17) to solve for μ^0 and μ^1, it can be verified that $P_e \to P^*$ as $\delta \to 0$, $\varepsilon \to 0$ for $\delta > 0$, $\varepsilon > 0$.

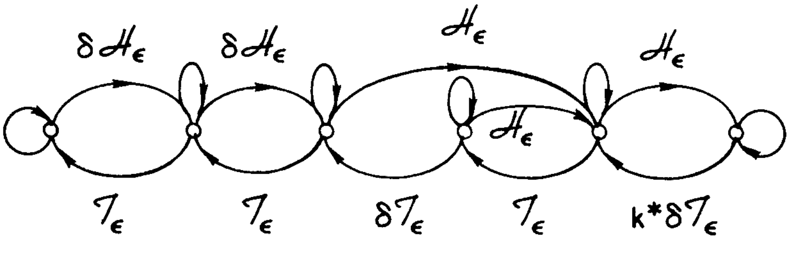

FIG. 2. An alternative ε-optimal algorithm.

In particular, it is seen that the tempting conditions of Corollary 1 are not necessary, even in the limit, for a sequence of automata $\{\mathscr{A}(n)\}$ to have $P_e(n) \to P^*$. Of the three conditions given in Theorem 5, only Condition 1 is nontrivial, since given condition 1, conditions 2 and 3 can always be achieved by multiplying the matrices

$[p_{ij}(x)]$ in the first and last rows by δ and $k\delta$ respectively. It remains only to specify the conditions under which the spread approaches γ^{m-1}.

THEOREM 6. *Consider a sequence of irreducible m-state automata* $\{\mathscr{A}(n)\}$ *with stationary occupation probabilities* $\mu^t(n)$, *state transition probabilities* $p_{ij}^t(n)$, *state likelihood ratios* $\lambda_1(n) \leqq \lambda_2(n) \leqq \cdots \leqq \lambda_m(n)$ *and spreads* $\sigma(n) = \lambda_m(n)/\lambda_1(n)$; *with respective limiting values (as* $n \to \infty$*)* μ^t, p_{ij}^t, λ_i *and* σ. *Then* $\sigma = \gamma^{m-1}$ *if and only if the following conditions are satisfied*:

1.

$$\frac{\sum_{j=k+1}^{m} \mu_k^t(n) p_{kj}^t(n)}{\sum_{i=1}^{k} \sum_{j=k+1}^{m} \mu_i^t(n) p_{ij}^t(n)} \to 1 \tag{60a}$$

$$\frac{\sum_{j=1}^{k} \mu_{k+1}^t(n) p_{k+1,j}^t(n)}{\sum_{i=k+1}^{m} \sum_{j=1}^{k} \mu_i^t(n) p_{ij}^t(n)} \to 1 \tag{60b}$$

for $t = 0, 1$ *and* $k = 1, 2, \cdots, m-1$; *and*

2.

$$[\textstyle\sum_{j=k+1}^{m} p_{kj}^0(n)/[\sum_{j=k+1}^{m} p_{kj}^1(n)] \to \bar{l} \tag{61a}$$

$$[\textstyle\sum_{j=1}^{k} p_{k+1,j}^0(n)]/[\sum_{j=1}^{k} p_{k+1,j}^1(n)] \to \underline{l} \tag{61b}$$

for $k = 1, 2, \cdots, m-1$.

REMARK. Let $C = \{1, 2, \cdots, k\}$ and $C' = \{k+1, \cdots, m\}$. Then Condition 1 says that the fraction of probability flow from C to C' which is from state k must tend to one; similarly, the fraction of flow from C' to C, from state $k+1$, must tend to one. Condition 2 states that the probability ratio of upward (or downward) transitions under H_0 and H_1 must approach $\bar{l}$ (or $\underline{l}$).

PROOF OF THEOREM 6. The sufficiency of the Conditions will be proved first. Considering $k \in \{1, 2, \cdots, m-1\}$ as fixed, using (17),

$$\textstyle\sum_{i=1}^{k} \sum_{j=k+1}^{m} \mu_i^t(n) p_{ij}^t(n) = \sum_{i=k+1}^{m} \sum_{j=1}^{k} \mu_i^t(n) p_{ij}^t(n) \equiv F^t(n) \tag{62}$$

for $t = 0, 1$. Let

$$\varepsilon_1(n) = [\textstyle\sum_{i=k+2}^{m} \sum_{j=1}^{k} \mu_i^0(n) p_{ij}^0(n)]/F^0(n) \tag{63}$$

denote the fraction of the probability flow from C' to C which is not from state $k+1$, under H_0. From (62) and (63)

$$(1-\varepsilon_1(n))^{-1} \mu_{k+1}^0(n) \textstyle\sum_{j=1}^{k} p_{k+1,j}^0(n) = F^0(n) \geqq \mu_k^0(n) \sum_{j=k+1}^{m} p_{kj}^0 \tag{64}$$

where the inequality follows from (62). Similarly, defining $\varepsilon_2(n)$ to be the fraction of flow from C to C' which is not from state k under H_1 results in

$$(1-\varepsilon_2(n))^{-1} \mu_k^1(n) \textstyle\sum_{j=k+1}^{m} p_{kj}^1(n) \geqq \mu_{k+1}^1(n) \sum_{j=1}^{k} p_{k+1,j}^1(n). \tag{65}$$

Combining (64) and (65) and taking the limit as $n \to \infty$ yields

$$\frac{\mu_{k+1}^0/\mu_{k+1}^1}{\mu_k^0/\mu_k^1} \geqq (1-\varepsilon_1)(1-\varepsilon_2)\frac{\sum_{j=k+1}^m p_{k,j}^0/\sum_{j=k+1}^m p_{k,j}^1}{\sum_{j=1}^k p_{k+1,j}^0/\sum_{j=1}^k p_{k+1,j}^1} = l/\underline{l} \tag{66}$$

where the limiting behavior follows from Conditions 1 and 2. Thus $\lambda_{k+1}(n)/\lambda_k(n) \to \gamma$ and therefore $\sigma(n) \to \gamma^{m-1}$, completing the proof of the sufficiency.

The necessity of the first part of Condition 1 will be proved by contradiction. Let $\varepsilon_2(n)$ be defined as before. Assume that $\varepsilon_2(n) \to \varepsilon_2 > 0$ in contradiction to Condition 1. Now $\sigma(n) \to \gamma^{m-1}$ implies that for any $\varepsilon > 0$ there exists a number $N(\varepsilon)$ such that, for $n \geqq N(\varepsilon)$,

$$\mu_i^0(n)/\mu_i^1(n) > \lambda_1(n)(\gamma^{i-1}-\varepsilon). \tag{67}$$

Using Lemmas 1 and 2 and (62), for $n > N(\varepsilon)$,

$$\lambda_1(n)\gamma^{k-1}\bar{l}F^1(n) \geqq F^0(n) > \lambda_1(n)\underline{l}\sum_{i=k+1}^m \sum_{j=1}^k (\gamma^{i-1}-\varepsilon)\mu_i^1(n)p_{ij}^1(n). \tag{68}$$

Using the expression (62) for $F^1(n)$ results in

$$\sum_{i=k+1}^m \sum_{j=1}^k (1-\gamma^{i-(k+1)}+\varepsilon\gamma^{-k})\mu_i^1(n)p_{ij}^1(n) > 0. \tag{69}$$

Since $1-\gamma^{i-(k+1)} \leqq 1-\gamma$ for $i \geqq k+2$, one obtains

$$\varepsilon\gamma^{-k}[1-\varepsilon_2(n)]F^1(n)+(1-\gamma+\varepsilon\gamma^{-k})\varepsilon_2(n)F^1(n) > 0. \tag{70}$$

Since $F^1(n) > 0$, $\gamma > 1$, and $\varepsilon_2(n) \to \varepsilon_2 > 0$, it is seen that for sufficiently small ε the left-hand side is negative in the limit; a contradiction. A similar argument proves that the second part of Condition 1 is also necessary for $\sigma(n) \to \gamma^{m-1}$.

Proceeding to the necessity of Condition 2, recall, from Lemma 2 that

$$\mu_k^0(n) \leqq \lambda_1(n)\gamma^{k-1}\mu_k^1(n) \tag{71}$$

for all n. Thus

$$F^0(n) = \sum_{i=1}^k \sum_{j=k+1}^m \mu_i^0(n)p_{ij}^0(n) \leqq \lambda_1(n)\gamma^{k-1}\bar{l}F^1(n). \tag{72}$$

Then from (64) and (67)

$$\begin{aligned} F^0(n) &\geqq \frac{1}{1-\varepsilon_1(n)}\lambda_1(n)(\gamma^k-\varepsilon)\left[\frac{\sum_{j=1}^k p_{k+1,j}^0(n)}{\sum_{j=1}^k p_{k+1,j}^1(n)}\right]\mu_{k+1}^1(n)\sum_{j=1}^k p_{k+1,j}^1(n) \\ &= \frac{1-\varepsilon_3(n)}{1-\varepsilon_1(n)}\lambda_1(n)(\gamma^k-\varepsilon)\left[\frac{\sum_{j=1}^k p_{k+1,j}^0(n)}{\sum_{j=1}^k p_{k+1,j}^1(n)}\right]F^1(n), \end{aligned} \tag{73}$$

where $\varepsilon_3(n)$ is the fraction of flow from C' to C not from state $k+1$, under H_1. (Note that $\gamma^{-(m-1)}\underline{l}\varepsilon_1(n) \leqq \varepsilon_3(n) \leqq \gamma^{m-1}\bar{l}\varepsilon_1(n)$, so that $\varepsilon_1(n) \to 0$ is equivalent to $\varepsilon_3(n) \to 0$.) From Condition 1, which has already been established, it is seen that $\sigma(n) \to \gamma^{m-1}$ implies $\varepsilon_1(n) \to 0$. Thus combining (72) and (73) one obtains

$$[\sum_{j=1}^k p_{k+1,j}^0(n)]/\sum_{j=1}^k p_{k+1,j}^1(n)] \leqq \frac{\gamma^{k-1}\bar{l}[1-\varepsilon_1(n)]}{[1-\varepsilon_3(n)](\gamma^k-\varepsilon)} \to \underline{l}, \tag{74}$$

where the limiting behavior (as $n \to \infty$ followed by $\varepsilon \to 0$) follows from $\varepsilon_1(n) \to 0$, $\varepsilon_3(n) \to 0$. Thus (61b) has been shown to be a necessary consequence of $\sigma(n) \to \gamma^{m-1}$. The necessity of (61a) has a similar proof.

7. Conclusions. The principal results of this discussion may be stated in a way which emphasizes the independence of the results from a Bayesian formulation. Let $\alpha = \Pr\{\text{Decide } H_1 \mid H_0\}$ and $\beta = \Pr\{\text{Decide } H_0 \mid H_1\}$. Then every algorithm with time-invariant m-state memory must satisfy (by (37)) the inequality.

$$(\alpha+t)(\beta+t) \geqq t(1+t) \tag{75}$$

where $t = (\gamma^{m-1}-1)^{-1}$ and $\gamma = \bar{l}/\underline{l}$. Also, for any α', β' such that $(\alpha'+t)(\beta'+t) > t(1+t)$ there exists an algorithm of the form shown in Figure 1 for which $\alpha < \alpha'$ and $\beta < \beta'$. It has also been shown that an optimal algorithm, i.e., one for which equality is achieved in (75), does not in general exist. Finally, ε-optimal algorithms require artificial randomization, or, what amounts to the same thing, a randomization induced by the observation itself through the action of $\mathscr{H}_\varepsilon'$ and $\mathscr{T}_\varepsilon'$.

Although stochastic transition rules are unnecessary in the continuous case, their introduction allows the unification of the discrete and continuous cases in much the same way that stochastic decision procedures unify the continuous and discrete cases in the Neyman–Pearson formulation of the hypothesis testing problem. It is somewhat surprising that randomization is needed at all, since randomization usually decreases information.

The form of the ε-optimal class provides insight into the optimal decision making process. Essentially the automaton changes state only on maximal or minimal likelihood ratio events. Furthermore, once in an extreme state the automaton leaves only with small probability. In the case of discrete distributions, this requires artificial randomization.

It is noticed that the automaton waits for extreme events before changing state. This shows that in many cases roundoff schemes are far from optimal, since they put emphasis on small changes. Thus taking a sufficient statistic and rounding it off to a finite number of decimal places will in general not be close to an optimal finite memory strategy.

Transitions may be restricted to extreme events, even if they occur infrequently, since the number of trials is infinite. If the number of trials N is finite, the automaton will not be able to neglect events of moderate information. The problem of finding an optimal machine when N is finite is an interesting one, for, except in certain degenerate cases, if $P(e)$ approaches P^*, δ must approach zero. The resulting convergence time increases without bound. Thus a machine which is nearly optimal for an infinite number of samples is far from optimal in the small sample case. However, we conjecture that for finite N the optimal machine will still resemble Figure 1 in certain respects. We believe that high likelihood ratio events will cause upward transitions, and low likelihood ratio events downward transitions although the events need not be as extreme as before and transitions need not be between adjacent states. Furthermore, we also believe that artificial randomization will still be needed in the discrete case; although the values of δ will not be near zero.

Let $\mathscr{A}^*(N)$ denote the optimal m-state algorithm for N observations, having associated probability of error $P^*(N)$. It is easily seen that $P^*(N) \to P^*$, and we believe that $\mathscr{A}^*(N)$ will in some sense approach the structure of Figure 1 as opposed to a more complicated structure such as that of Figure 2.

It would also be of interest to see whether human beings, in problems to which they have allotted finite memory (such as "like," "indifference" and "dislike") demonstrate an optimal randomized learning procedure similar to that suggested by this paper.

Acknowledgment. The authors wish to express their gratitude to D. Sagalowicz for his substantial contributions to this work. The authors also wish to thank S. Samuels and the referee for their detailed comments and suggestions.

REFERENCES

[1] Cover, T. M. (1968). A note on the two-armed bandit problem with finite memory. *Information and Control* **12** 371–377.

[2] Cover, T. M. (1969). Hypothesis testing with finite statistics. *Ann. Math. Statist.* **40** 828–835.

[3] Cover, T. and Hellman, M. (1970). The two-armed bandit problem with time invariant finite memory. *IEEE Trans. Information Theory* **IT-16** 185–195.

[4] Feller, W. (1957). *An Introduction to Probability Theory and its Applications* **1** 2nd ed. Wiley, New York, 318.

[5] Fralick, S. C. (1967). Learning to recognize patterns without a teacher. *IEEE Trans. Information Theory* **IT-13** 57–65.

[6] Hellman, M. (1969a). A faster solution for random walk problems. IBM Research Report, RC 2389 No. 11615.

[7] Hellman, M. (1969b). Learning with finite memory. Ph.D. Dissertation, Stanford Univ.

[8] Hellman, M. and Cover, T. M. (1970). Comments on automata in random media. Submitted to *Problems of Information Transmission*.

[9] Isbell, J. R. (1959). On a problem of Robbins. *Ann. Math Statist.* **30** 606–610.

[10] Krylov, V. Y. (1963). On one automaton that is asymptotically optimal in a random medium. *Avtomat. Telemeh.* **24** 1226–1228. Translated in *Automat. Remote Control.*

[11] Parzen, E. (1962). *Stochastic Processes*. Holden-Day, San Francisco.

[12] Robbins, H. (1956). A sequential decision problem with a finite memory. *Proc. Nat. Acad. Sci.* **42** 920–923.

[13] Samuels, S. M. (1968). Randomized rules for the two-armed bandit with finite memory. *Ann. Math. Statist.* **39** 2103–2107.

[14] Smith, C. V. and Pyke, R. (1965). The Robbins–Isbell two-armed bandit problem with finite memory. *Ann. Math. Statist.* **36** 1375–1386.

[15] Spragins, J. (1966). Learning without a teacher. *IEEE Trans. Information Theory* **IT-12** 223–229.

[16] Tsetlin, M. L. (1961). On the behavior of finite automata in random media. *Avtomat. Telemeh.* **22** 1345–1354.

[17] Varshavskii, V. I. and Vorontsova, I. P. (1963). On the behavior of stochastic automata with a variable structure. *Avtomat. Telemeh.* **24** 353–360.