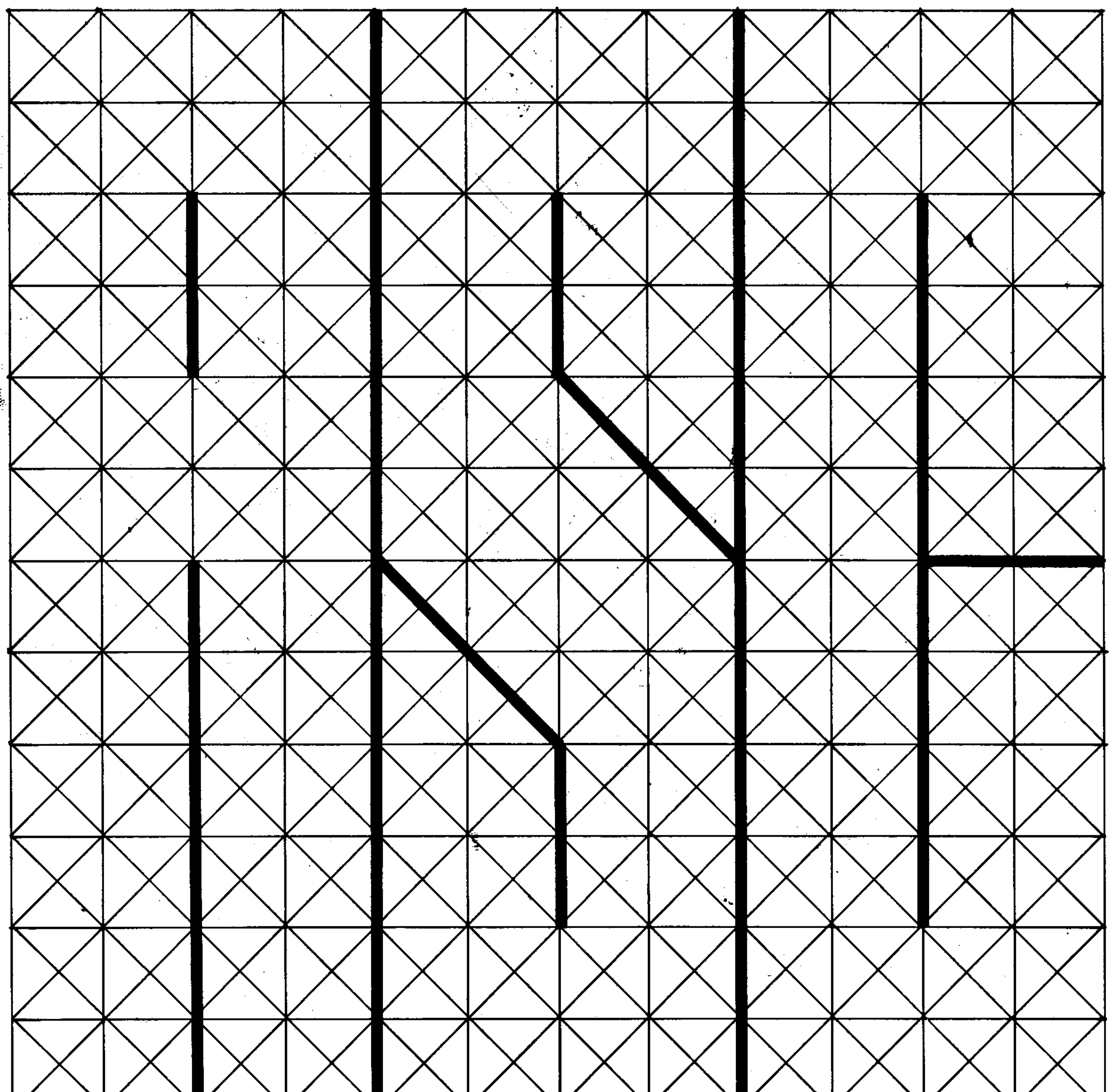
Journal of Cybernetics and Information Science

ASC
American Society for Cybernetics

VOLUME 1 NUMBER 2 THRU 4
SPRING - SUMMER - FALL 1977



SPECIAL ISSUE ON LEARNING AUTOMATA

AN EXPANDING AUTOMATON FOR USE IN STOCHASTIC OPTIMIZATION

Luc P. Devroye*
School of Computer Science
McGill University
Montreal, Canada H3A 2K6

Abstract

A probabilistic automaton with an expanding memory is presented. Its asymptotic properties as a stochastic optimization technique are studied. The procedure is shown to be convergent under very mild conditions on the statistical characteristics of the random environment.

1. PROBLEM FORMULATION

A finite random environment is a finite collection of distribution functions on $\mathbb{R}^1$, say $\{F_1,..,F_N\}$. It is assumed that $F_1,..,F_N$ are unknown. To gather some knowledge about we can apply a strategy x_i ($i\in\{1,..,N\}$) to the environment which responds with a number Y where Y is a random variable with distribution function F_i in $\mathbb{R}^1$ and mean.

$$Q(x_i) = \int y.dF_i(y) = E_{x_i}\{Y\} \quad ,i=1,...,N . \tag{1}$$

$Q(x_i)$ is the expected loss with strategy x_i.

The problem of the sequential selection of the best of the N strategies x_i, i=1,..,N has been extensively dealt with in the literature. One of the most popular methods to tackle this problem was the stochastic automaton with a variable structure (for a survey, see [1]). To describe the strategy selection process, we assume that there exists a discrete probability density on $\{x_1,...,x_N\}$, say $P_n=(P_{1n},...,P_{Nn})$ so that

$$P_{in} = P\{X_n=x_i\} \; ; i=1,...,N \; ; n=1,2,...$$
$$\sum_{i=1}^{N} P_{in} = 1 \; ; n=1,2,... \tag{2}$$

where $X_n\in\{x_1,...,x_N\}$ is the strategy that is picked at epoch n. In general, the selection vector P_n is a random vector. Let Y_n be the loss that is observed after X_n is applied to the environment. The expected loss with P_n is

$$M_n = E\{Y_n|P_n\} = \sum_{i=1}^{N} E_{x_i}\{Y\}.P\{X_n=x_i|P_n\} = \sum_{i=1}^{N} P_{in}.Q(x_i) \; . \tag{3}$$

A probabilistic automaton is a set of rules for computing P_{n+1} given (P_j,X_j,Y_j), j=1,..,n (see [1-5]). A probabilistic automaton is said to be optimal if

* The research was partially supported by Air Force Grant No. 72-2371.

$$\lim_{n} E\{M_n\} = \inf \{Q(x_1),\ldots,Q(x_N)\} . \quad (4)$$

We remark that in general random environments (i.e., no restrictions are imposed upon the F_i except for the existence of the $Q(x_i)$), only the performance directed probabilistic automaton [5] is known to be optimal.

The problem naturally arises of whether it is possible to find a probabilistic automaton that is optimal in countably infinite (c.i.) random environments $\mathcal{E}=\{F_1,F_2,..\}$. We will show in this paper that if the variances associated with the $F_1,F_2,..$ are uniformly bounded and $\inf \{Q(x_1), Q(x_2),..\} > -\infty$, then the answer to this problem is affirmative. We will first classify the random environments according to the characteristics of the F_i in $\mathcal{E}$. Thereafter, we describe two probabilistic automata, P1 and P2, and prove their optimality as well as some other asymptotical properties. The emphasis is on the new techniques employed to prove the convergence of the said procedures.

2. A GENERAL SETTING

The importance of the class of c.i.random environments is the following. One way to tackle the multimodal stochastic optimization problem in $\mathbb{R}^m$ $(m \geq 1)$ is to partition $\mathbb{R}^m$ into small compact sets (for instance, rectangles) and consider each rectangle as one strategy to be applied to a c.i. random environment. We can thus reduce the optimization problem to the problem of finding the best strategy in a c.i. random environment provided that the rectangles are small enough. The c.i. random environments are but a special case of general random environments which are characterized as follows. Let $(\Omega,\mathcal{G},P)$ be a probability space and let B be a closed set from $\mathbb{R}^m$. Let $\mathcal{B}^m_B$ be the σ-algebra of all the Borel sets that are contained in B. Let $\mathcal{B}^1$ be the σ-algebra of all the Borel sets from $\mathbb{R}^1$ and let h be a measurable mapping from $(\Omega\times B,\mathcal{G}\times\mathcal{B}^m_B)$ to $(\mathbb{R}^1,\mathcal{B}^1)$. Notice that for every $x\in\mathbb{R}$: $y=h(\omega,x)$ is a random variable on $(\Omega, ,P)$. We say that a collection

$$\mathcal{E} = \{F_x(.) \mid x\in B\} \quad (5)$$

of distribution functions is <u>a randon environment with search domain B</u> if B is a closed set from $\mathbb{R}^m$ for some m, and if there exists a probability space $(\Omega,\mathcal{G},P)$ and a $(\Omega\times B,\mathcal{G}\times\mathcal{B}^m_B)-(\mathbb{R}^1,\mathcal{B}^1)$ measurable function h such that for all $y\in\mathbb{R}^1$:

$$F_x(y) = P\{\omega \mid \omega\in\Omega, h(\omega,x)\leq y\} ,$$

Notice that if B is countable or finite, then such a probability space and measurable function h can always be found. Thus it makes sense to say that a <u>c.i. random environment</u> is a countable collection of distribution functions on $\mathbb{R}^1$.

The reason for this definition is the following. Let X be any random vector (on some probability space $(\Omega',\mathcal{G}',P')$ that is different from $(\Omega,\mathcal{G},P)$) taking values in B, then $Y=h(\omega,X)$ is a random variable on the product of both probability spaces.

We will assume, for all the sequences $X_1,..,X_n$ of random vectors that are applied to the environment, that the corresponding <u>observed losses</u> are independent random variables given that $X_1=x_1,\ldots,X_n=x_n$ for all $x_i\in\mathbb{R}^m, i=1,\ldots,n$. We will refer to

$$Q(x) = \int y.dF_x(y) \overset{\Delta}{=} E_x\{Y\} \quad (6)$$

as the <u>stochastic performance index</u>. Q is assumedly a Borel measurable function from B to $\mathbb{R}^1$. Of course except for B, no knowledge is available about $\mathcal{E}$ and Q. The problem is to sequentially find a value $x\in B$ for which Q is minimal or nearly minimal. But this is exactly the <u>stochastic optimization problem</u>. The reader is referred to [6] for a survey of the most popular stochastic optimization techniques. The choice of a particular technique depends upon the a priori knowledge about $\mathcal{E}$ and Q (is Q smooth ? unimodal ? differentiable ? etc.). For unknown $\mathcal{E}$ and Q, random search is probably the most frequently used optimization method (see [7],[9] for surveys), the asymptotic behavior of which is studied in [8]. In this paper an expanding automaton is presented which generalizes the finite automaton of [5] for use in general random environments.

We assume that there is a random generator covering B, i.e. a device for generating a sequence $X_1,\ldots,X_n,\ldots$ of iid random vectors taking values in B and distributed as X where X has a (known or unknown) distribution function G. The minimum of Q with respect to G is

$$q_{min} \triangleq \text{ess inf } Q(X) \,. \tag{7}$$

For a definition of the essential infimum, see [10]. Actually, q_{min} is the unique number with the property that for all $\varepsilon > 0$: $P\{Q(X) \le q_{min} - \varepsilon\} = 0$ and $P\{Q(X) \le q_{min} + \varepsilon\} > 0$ provided that $q_{min} > -\infty$. We remark that if B_∞ is a countable set $\{x_1, x_2, \ldots\}$ and G puts mass g_i at x_i such that $\sum_{i=1}^{\infty} g_i = 1$; $0 \le g_i \le 1$; $i=1,2,\ldots$,then

$$q_{min} = \inf_{i:g_i>0} Q(x_i) \,.$$

In this case q_{min} is independent of G as long as every x_i receives positive probability from G.

We distinguish between the following types of random environments:

(i) $\mathcal{E}$ is $\mathcal{D}$ (deterministic, noiseless, etc.) if

$$\sup_{x \in B} E_x\{(Y-Q(x))^2\} = 0 \tag{8}$$

which is equivalent to saying that for all x in B: $Y=Q(x)$ WP1 (with probability one).

(ii) $\mathcal{E}$ is $\mathcal{E}_t$ $(t>0)$ with parameter $L<\infty$ if

$$\sup_{x \in B} E_x\{|Y-Q(x)|^t\} = \sup_{x \in B} \int |y-Q(x)|^t d\Gamma_x(y) \le L < \infty, \tag{9}$$

(iii) $\mathcal{E}$ is $\mathcal{X}$ (exponential) if for every $\varepsilon>0$ there exists a $c(\cdot)>0$ such that

$$\sup_{x \in B} E_x\{e^{\lambda(Y-Q(x))}\} = \sup_{x \in B} \int e^{\lambda(y-Q(x))} dF_x(y) \le e^{|\lambda|\varepsilon} \quad \text{for all } \lambda \in [-c(\varepsilon), +c(\varepsilon)] \,. \tag{10}$$

If an environment is $\mathcal{D}$ then it is $\mathcal{X}$ and if it is $\mathcal{X}$ then it is $\mathcal{E}_t$ for all $t>0$. If $\mathcal{E}$ is $\mathcal{E}_t$ then $\mathcal{E}$ is $\mathcal{E}_s$ for all s with $0<s<t$. If $\mathcal{E}$ is $\mathcal{E}_2$ with parameter $L=0$ then $\mathcal{E}$ is $\mathcal{D}$. It should be pointed out that most environments of any practical interest are $\mathcal{X}$. For instance, if F_x puts mass 1 on $[Q(x)-a, Q(x)+a]$ for some $a<\infty$ and for all $x \in B$, then the environment is $\mathcal{X}$. Also, if all F_x are gaussian with a variance that is not greater than some $a<\infty$, then $\mathcal{E}$ is $\mathcal{X}$.

3. DESCRIPTION OF THE AUTOMATON

In an iterative optimization procedure, one generates a sequence of random vectors $Z_0, Z_1, \ldots$ where for all n, $Z_n \in \mathbb{R}^{m+k}$, i.e. $Z_n=(X_n, Z_n')$ with $X_n \in \mathbb{R}^m$ and $Z_n' \in \mathbb{R}^k$. X_n is the best estimate of the minimum (or: basepoint) at epoch (or: iteration) n.

We describe two very similar iterative optimization procedures, P1 and P2. Let $\{\alpha_n\}, \{\beta_n\}, \{\gamma_n\}, \{\delta_n\}$ and $\{\eta_n\}$ be sequences from [0,1] that are picked by the designer in a special way in order to insure the convergence of the procedures P1 and P2 and to obtain the desired rates of convergence. For procedure P1, let $\{b_n\}$ be a sequence of integers with $1 \le b_n \le n$ for all n. For procedures P2, let $\{k_n\}$ be a sequence of integers with $1 \le k_n \le n$ for all n. Then the expanding probabilistic automaton can be described as follows.

(i) Z_0 (and thus X_0) is either given or selected by the designer. Z_0 is sometimes referred to as the initial state. Notice that one can always randomly generate X_0 with distribution function G in $\mathbb{R}^m$. X_0 is then applied to the environment and a loss Y_0 is observed. Let $H_0=(X_0, Y_0, 1, 0)$ and let $\{H_n\}$ $(n \ge 0)$ be a sequence of random vectors composed of a growing number of quadruples. H_n can be thought of as the memory at epoch n. Let L_n be the number of quadruples in H_n, say

$$H_n = (W_1, \bar{Y}_1^n, \bar{N}_1^n, T_1), \ldots, (W_{L_n}, \bar{Y}_{L_n}^n, \bar{N}_{L_n}^n, T_{L_n})$$

where $W_i \in \mathbb{R}^m$ and where N_i^n is the experience gained with W_i up to epoch n, i.e. N_i^n is the number of times that W_i was applied to the environment up to epoch n. T_i is the iteration at which W_i was first generated and added to H_n. We say that T_i is the birth date of W_i. $\bar{Y}_i^n$ is the average of the N_i^n losses that were observed after W_i was applied to the environment. $\bar{Y}_i^n$ obviously serves as an estimate of $Q(W_i)$. We will see that $1=L_0 \le L_1 \le L_2 \le \ldots$.

(ii) Proceed to the next iteration, say the n-th. H_{n-1} and Z_{n-1} are known. X_{n-1} is the basepoint before the n-th iteration and H_{n-1} , as we know, contains L_{n-1} quadruples. We generate an independent random variable U_n' where

$$U_n' = \begin{cases} 1 & \text{with probability } \alpha_n \\ 0 & \text{otherwise .} \end{cases} \tag{11}$$

If $U_n'=0$, no new quadruple is generated so that $L_n=L_{n-1}$. We proceed to (iv). If $U_n'=1$, a new quadruple is generated and tested. We let $L_n=L_{n-1}+1$ and proceed to (iii) for the generation of this new quadruple.

(iii) Generate an independent random variable U_n where

$$U_n = \begin{cases} 1 & \text{with probability } \eta_n \\ 0 & \text{otherwise ,} \end{cases} \tag{12}$$

If $U_n=1$, we generate W_{L_n} at random in B using the random generator with the distribution function G. If $U_n=0$, then W_{L_n} is a random vector taking valuesin B and having an arbitrary distribution function in $\mathbb{R}^m$. W_{L_n} may depend in an arbitrary fashion upon H_{n-1} and Z_{n-1}. We require that $P\{W_{L_n} \in B\}=1$ for all n.

W_{L_n} is then applied to the environment and Y_n is the observed loss. We can now obtain H_n as follows. Let

$$H_n=H_{n-1},(W_{L_n},\overline{Y}^n_{L_n},\overline{N}^n_{L_n},T_{L_n})$$

where $\overline{Y}^n_{L_n}=Y_n$, $N^n_{L_n}=1$ and $T_{L_n}=n$. Proceed to (v).

(iv) If no new quadruple is generated then the experience with one of the W_i in H_{n-1} has to be increased. We describe first how to pick a member W_i from H_{n-1}. Generate an independent random variable V_n where

$$V_n = \begin{cases} 1 & \text{with probability } \beta_n \\ 0 & \text{with probability } \gamma_n \\ -1 & \text{with probability } \delta_n=1-\beta_n-\gamma_n . \end{cases} \tag{13}$$

If $V_n=1$, we pick the basepoint X_{n-1} from $W_1,\ldots,W_{L_{n-1}}$. If $V_n=0$, a point is picked at random from $W_1,\ldots,W_{L_n}$, i.e. a uniform distribution is used over the L_n points in H_{n-1}. If $V_n=-1$, then a point is picked from $W_1,\ldots,W_{L_{n-1}}$ in some clever way (in order to achieve some goal, accelerate the rate of convergence, etc.) but it is not specified how the selection is to be made.

Let the selected point be W_{I_n} where $1 \le I_n \le L_{n-1}=L_n$. Apply W_{I_n} to the environment, observe a loss Y_n and update H_{n-1} in the obvious way. That is, let $H_n=H_{n-1}$ except that $(W_{I_n},\overline{Y}^{n-1}_{I_n},N^{n-1}_{I_n},T_{I_n})$ is replaced by $(W_{I_n},\overline{Y}^n_{I_n},N^n_{I_n},T_{I_n})$ where

$$\overline{Y}^n_{I_n} = (Y_n + N^{n-1}_{I_n}.\overline{Y}^{n-1}_{I_n})/(1 + N^{n-1}_{I_n}) \tag{14}$$
$$N^n_{I_n} = 1 + N^{n-1}_{I_n} .$$

(v) Now that we have made one observation at the n-th iteration and have obtained H_n and L_n, we have to decide which of the points $W_1,\ldots,W_{L_n}$ is most likely to have the lowest corresponding value $Q(W_i), i=1,\ldots,L_n$. The new basepoint X_n is picked from $W_1,\ldots,W_{L_n}$ in the following way (notice that it is at this point that the procedures P1 and P2 are different from each other). With the procedure P1, we look for all the quadruples in H_n for which

$$N^n_i \ge b_n \quad (i=1,\ldots,L_n) . \tag{15}$$

Among these quadruples, pick the one that corresponds to the lowest value $\overline{Y}^n_i$ and let the corresponding W_i be X_n. Ties are broken randomly. If there are no quadruples with $N^n_i \ge b_n$, let $X_n=X_{n-1}$. With procedure P2, we look for all the quadruples in H_n for which

$$T_i \le k_n \quad (i=1,\ldots,L_n) \tag{16}$$

and proceed in a similar fashion. We remark here that the procedures P1 and P2 can be carried out recursively, i.e. we do not have to check all the quadruples in H_n all over again at every iteration. The methods for reducing the computational burden are standard and are left to the reader. Note that P1 selects the basepoint among the W_i of H_n with a large experience while P2 selects the basepoint among the W_i of H_n with the highest "ages" (i.e., earliest birth dates).

(vi) We remark that it is up to the designer to specify the random vector Z_n'. How Z_n' is updated or computed from Z_{n-1}', H_n, etc. is left in the middle. These updating mechanisms can play an important role in obtaining a high rate of convergence. In fact, it is in this stage that the vast experience of the designer can pay off. For instance, Z'_{n-1} can be used to help generate W_{L_n} in a promising small subset of B. In any case, the nature of Z'_{n-1} and of the updating mechanisms is of no importance whatsoever to establish the convergence of the algorithm.

(vii) (ii-vi) constitute one basic cycle (iteration) of the search process. Go back to (ii).

If the random environment is countably infinite, then the algorithm can be modified because the identification problem for points in B can be solved. That is, if $U'_n=1$, (and thus W_{L_n} is some random vector taking values in B), then it is decidable whether $W_{L_n}=W_i$ for some i with $1\le i\le L_{n-1}$. If this should happen, then of course $L_n=L_{n-1}$ (not : $L_n=L_{n-1}+1$) and we can proceed from step (iii) to step (iv) for updating H_{n-1}. For this slightly modified procedure, all the theorems of this paper remain valid. Another modification for which the theorems of convergence remain valid (the proofs need minor modification), consists of rejecting W_{L_n} if $U'_n=1$ and $W_{L_n}=W_i$, some $1\le i\le L_{n-1}$. In case of a rejection, other points are generated (using new and independent U_n) until one point is found outside $W_1,\ldots,W_{L_{n-1}}$. If the random environment is not finite, if the support of G is infinite and if $\eta_n>0$ then this procedure is bound to stop in finite time.

Both modifications are geared to prevent a loss of information in the sense that, with the modifications, we will have for all n and all i,j,with $i=1,\ldots,L_n$;$j=1,\ldots,L_n$;$i\neq j$, that $W_i\neq W_j$.

Let $W'_0,W'_1,W'_2,\ldots$ be the sequence of inputs to the environment. The reader will have no difficulty, assuming that the random environment is countably infinite, finding the description of the infinite dimensional discrete probability vector according to which the W'_n are to be generated, both with the original procedure and the modified procedures.

4. CONVERGENCE OF THE PROCEDURE

The random variables that are of interest to us all have the form $f(Z_n)$ where f is a Borel measurable mapping from $\mathbb{R}^{m+k}$ to $\mathbb{R}^1$. Let $X_0,X_1,X_2,\ldots$ be the sequence of basepoints in B and let $W'_0,W'_1,W'_2,\ldots$ be the sequence of inputs to the environment. In classical optimization one is mainly interested in

$$Q_n \triangleq Q(X_n) \tag{17}$$

while in automata theory, the expected loss at epoch n is also important :

$$\begin{aligned} M_n &\triangleq E\{Y_n \mid (H_{n-1},Z_{n-1})\} = E\{Q(W'_n) \mid (H_{n-1},Z_{n-1})\} \\ &= \alpha_n . E\{Q(W_{L_n}) \mid (H_{n-1},Z_{n-1})\} + (1-\alpha_n).E\{Q(W_{I_n}) \mid (H_{n-1},Z_{n-1})\} \\ &\qquad (n=1,2,\ldots) . \end{aligned} \tag{18}$$

Further we can also define

$$M'_n = E\{Y_n \mid (H_{n-1},Z_{n-1},U'_n,U_n,V_n)\} = \begin{cases} Q(W_{L_n}) & \text{if } U'_n = 1 \\ Q(W_{I_n}) & \text{if } U'_n = 0 \end{cases} \tag{19}$$

where $M'_1,M'_2,\ldots$ is the sequence of observed losses if $\mathcal{E}$ were a noiseless environment. We will see in theorem 1 that the convergence of Q_n to q_{min} is of crucial importance in the study of the convergence of M_n and M'_n.

Let $a \vee b \triangleq \text{Max}\,(a,b)$ and introduce the condition

$$\sup_{x\in B} |Q(x)| \le K_1 < \infty . \tag{20}$$

We already know that if $Q_n \vee q_{min} \to q_{min}$ in L_r (where $r>0$) as $n\to\infty$ (i.e. $\lim_n E\{|Q_n \vee q_{min} - q_{min}|^r\}=0$) ,then $Q_n \vee q_{min} \to q_{min}$ in probability as $n\to\infty$ [10]. If (20) holds, then the converse is also true. We further have

$$M_n = E\{M'_n \mid (H_{n-1},Z_{n-1})\} \tag{21}$$

with probability one. In all the following theorems we assume that the condition C1 holds :

Condition C1 : Let $Z_0, Z_1, \ldots$ be generated through the procedure (i-vii) where $B \subseteq \mathbb{R}^m$, is a random environment with search domain B, $\{\alpha_n\}, \{\beta_n\}, \{\gamma_n\}$ and $\{\eta_n\}$ are sequences from [0,1], $\{b_n\}$ (for P1) and $\{k_n\}$ (for P2) are integer sequences such that $1 \le b_n \le n$, $1 \le k_n \le n$ for all n and $q_{min} > -\infty$.

Theorem 1 :

(i) Let the conditions C1 and (22) hold.

$$\lim_n \alpha_n = 0 \;\; ; \; \lim_n \beta_n = 1 \, . \tag{22}$$

If

$$Q_n v q_{min} \to q_{min} \text{ in probability as } n \to \infty \tag{23}$$

,then

$$M'_n v q_{min} \to q_{min} \text{ in probability as } n \to \infty \, .$$

If in addition (20) holds ,then $M'_n v q_{min} \to q_{min}$ in L_r as $n \to \infty$ for all $r > 0$ and $M_n v q_{min} \to q_{min}$ in probability and in L_r (for all $r>0$) as $n \to \infty$.

(ii) Let the conditions C1 and (22) hold. If

$$Q_n v q_{min} \to q_{min} \text{ WP1 as } n \to \infty \tag{24}$$

then

$$M_n v q_{min} \to q_{min} \text{ WP1 as } n \to \infty \, .$$

Theorem 1 is proved in the Appendix. Note that in general it will be impossible to insure that $M'_n v q_{min} \to q_{min}$ WP1 as $n \to \infty$. Indeed, for the latter type of convergence we need that $\Sigma \alpha_n < \infty$ but this contradicts the condition $\Sigma \alpha_n = \infty$ that is needed, as we will see, for the convergence of $Q_n v q_{min}$ to q_{min} as $n \to \infty$ in the sense of (23). We will now show under which conditions we can insure that (23) is true.

Theorem 2 (procedure P1) :Let the conditions C1, (25) and (26) hold.

$$\sum_{n=1}^{\infty} \alpha_n \cdot \eta_n = \infty \tag{25}$$

$$\lim_n \left(\sum_{i=1}^{n} \gamma_i \cdot (1-\alpha_i) \right) / \left(b_n \cdot \sum_{i=1}^{n} \alpha_i \right) = \infty \tag{26}$$

If the environment is either $\mathcal{S}$ (in which case (26) is replaceable by the condition that $b_n = 1$ for all n) ,or $\mathcal{E}_t$ (t>1) and (27) holds, or $\mathcal{K}$ and (28) holds,

$$\lim_n b_n / \left(\sum_{i=1}^{n} \alpha_i \right)^{1/(t-1)} = \infty \tag{27}$$

$$\lim_n b_n / \log \left(\sum_{i=1}^{n} \alpha_i \right) = \infty \tag{28}$$

then $Q_n v q_{min} \to q_{min}$ in probability as $n \to \infty$.

Theorem 3 (procedure P2) :Let the conditions C1, (25) and (29) hold.

$$\lim_n k_n = \infty \, , \tag{29}$$

If the environment is either $\mathcal{S}$ (in which case we can, but need not, let $k_n = n$ for all n) or $\mathcal{E}_t$ for some t>1 and (30) holds, or $\mathcal{K}$ and (31) holds,

$$\lim_n \left(\sum_{i=k_n+1}^{n} \gamma_i \cdot (1-\alpha_i) \Big/ \sum_{i=1}^{n} \alpha_i \right) \cdot \left(\sum_{i=1}^{k_n} \alpha_i \right)^{-1/(t-1)} = \infty \tag{30}$$

$$\lim_n \left(\sum_{i=k_n+1}^{n} \gamma_i \cdot (1-\alpha_i) \Big/ \sum_{i=1}^{n} \alpha_i \right) \cdot \left(\log \sum_{i=1}^{k_n} \alpha_i \right)^{-1} = \infty \, , \tag{31}$$

then $Q_n v q_{min} \to q_{min}$ in probability as $n \to \infty$.

The proofs of theorems 2 and 3 are given in the appendix. Let us briefly discuss some of the conditions of convergence. Notice that b_n can be considered as the minimum experience required for any W_i in H_n to be a candidate basepoint with procedure P1 ; on the other hand, $n-k_n$ is the minimum age required for any W_i in H_n to be a candidate basepoint with procedure P2.

Condition (25) not only insures that,WP1, $L_n \to \infty$ but also that with probability one there is an infinite sequence of points W_i that are generated by the "random generator" (thus having distribution function G). Notice that if L_n is large, then all the N_i^n are small, $i=1,\ldots,L_n$ and thus the $\bar{Y}_i^n$ are relatively noisy estimates of the $Q(W_i)$ (how noisy depends of course upon the type of environment). If L_n is small, then the N_i^n are large, but at the same time, the probability that any of the $Q(W_i)$ is close to q_{min} is small because H_n contains so few members. Thus there should be a trade-off between the size of H_n (L_n roughly increases as $\sum_{i=1}^{n} \alpha_i$ as we know) and the minimum experience or age of the candidate basepoints. This is exactly expressed in the conditions (27) and (28) of theorem 2. Condition (26) insures that given b_n, enough points are available in H_n for which $N_i^n \geq b_n$. Conditions (30) and (31) in theorem 3 are the counterparts of (27) and (28) if one remarks that

$$\sum_{i=k_n+1}^{n} \gamma_i . (1-\alpha_i) \Big/ \sum_{i=1}^{n} \alpha_i$$

can be considered as the minimum experience associated with any W_i in H_n with birth date $T_i \leq k_n$ (i.e. points that are candidate basepoints with procedure P2).

It is not hard to see that P1=P2 if $b_n=1$ with P1 and $k_n=n$ with P2. This procedure is easily recognized as the classical random search algorithm for deterministic environments (see [6],[8]).

Let us give an example of sequences $\{\alpha_n\}$,$\{\beta_n\}$,$\{\gamma_n\}$, $\{\eta_n\}$,$\{b_n\}$ and $\{k_n\}$ satisfying the conditions of theorems 2,3. For any nonnegative number sequences $\{a_n\}$ and $\{c_n\}$, we say that $a_n=O(c_n)$ if there exists a K with $0<K<\infty$ such that $a_n \leq K.c_n$ for all n. Let for some K_2,K_4,K_5 from $(0,\infty)$:

$$\alpha_n = K_2/n^a \;;\; \alpha_n \leq 1/2 \text{ for all large } n \;;$$
$$\beta_n \text{ arbitrary} \qquad (32)$$
$$b_n = K_4.n^b \;;\; \gamma_n^{-1} = O(n^g) \;;\; \eta_n^{-1} = O(n^h) \;;\; k_n = K_5.n^k$$

where, obviously, $a \geq 0, b \geq 0, k \geq 0, g \geq 0, h \geq 0$.

It is a straightforward exercise to show that (25) is implied by $a+h \leq 1$, (26) is implied by $a > g+b$, (27) is implied by $b > (1-a)/(t-1)$, (28) is implied by $b > 0$, (29) is implied by $k > 0$, (30) follows from $g > a - (1-a)k/(t-1)$ and $k_n < n/2$ for all n large enough , and (31) follows from $g > a$ and $k_n < n/2$ for all n large enough.

7. CONCLUSION

A probabilistic automaton with an expanding memory is presented. The important feature of the probabilistic automaton is that all the past observations are used which makes the technique information intensive. The automaton is constructed in such a way that the size of the memory is continuously growing (which enables the automaton to act as a search procedure) and that the accuracy of the information that is stored in the memory is continuously improving with time by virtue of an averaging process (which makes the automaton suited for use in stochastic environments).

A detailed study is made of the properties of convergence of the automaton. Among other things we proved the optimality of the automaton in a large class of random environments. To achieve this optimality, it was shown that the rate of increase of the size of the memory and the rate of increase of the accuracy of the information that is stored in the memory have to satisfy a trade-off condition that depends upon the noise characteristics of the random environment. The advantage of the technique is that the class of allowable stochastic performance indices includes the class of all Borel measurable functions on $\mathbb{R}^m$ that are bounded from below. We did not intend to present a procedure that is quickly convergent. For this, it is necessary to use the freedom in the design of the automaton as well as possible. In particular it seems natural to use the information contained in H_{n-1}

(i) to direct the search process (see stage (iii) of the algorithm). If $U_n=0$, use the data in H_{n-1} to determine promising subregions of B, outside $W_1,\ldots,W_{L_{n-1}}$. One can for instance consider a variable distribution for W_{L_n} that puts all of its weight in a small neighborhood of those points $W_i, 1 \leq i \leq L_{n-1}$ with low corresponding estimates $\hat{Y}_i^{n-1}$.

(ii) to control the sampling process (see stage (iv) of the algorithm). If $V_n=-1$, use the data in H_{n-1} to determine which W_i of H_{n-1} need more sampling. To do this we can be guided by the same sampling techniques that are in use for finite probabilistic automata in finite random environments (see,e.g. [5]).

The heuristics used in (i) and (ii) are of the utmost importance to obtain high rates of convergence. One of the reasons we are particularly interested in high rates of convergence is an economical one. Given a certain stopping rule, it is hoped that Q_n is close to q_{min} at the stopping time and that, at the same time, the size of H_n is not excessively large (because of the limitations for the active memory in the computer).

We remark that in general random environments, our technique is competitive with random search which requires only a fixed finite amount of memory from the computer. Thus, the expanding automaton should be used when the effort of storing all the past information can pay off, i.e. when the cost of making observations is relatively high or when Q is very "abnormal" or when the environment is extremely "noisy".

8. APPENDIX

Lemma 1: Let $Y_1,\ldots,Y_n$ be a sequence of independent random variables with

$$Y_i = \begin{cases} 1 & \text{with probability } \alpha_i \\ 0 & \text{otherwise} \end{cases} \qquad (i=1,\ldots,n)$$

where α_i [0,1]. Then,

$$P\{\sum_{i=1}^{n} Y_i \le \sum_{i=1}^{n} \alpha_i\} \le \exp\{-\sum_{i=1}^{n} \alpha_i / 10\}$$

and

$$P\{|\sum_{i=1}^{n}(Y_i-\alpha_i)| \ge (1/2)\sum_{i=1}^{n}\alpha_i\} \le 2\exp\{-\sum_{i=1}^{n}\alpha_i/10\}.$$

Proof: Let

$$\sigma^2 \triangleq n^{-1}\cdot\sum_{i=1}^{n} E\{(Y_i-E\{Y_i\})^2\} = n^{-1}\cdot\sum_{i=1}^{n}\alpha_i(1-\alpha_i) \le n^{-1}\cdot\sum_{i=1}^{n}\alpha_i.$$

From Bennett's inequality (e.g., see equation 43 of [13] and use the inequality $\log(1+u)\ge 2u/(2+u)$ for $u>0$) we know that for every $\varepsilon>0$:

$$P\{n^{-1}\cdot\sum_{i=1}^{n}(Y_i-\alpha_i) \ge \varepsilon\} \le \exp\{-n\varepsilon^2/(2\sigma^2+\varepsilon)\}$$

$$P\{n^{-1}\cdot\sum_{i=1}^{n}(Y_i-\alpha_i) \le -\varepsilon\} \le \exp\{-n\varepsilon^2/(2\sigma^2+\varepsilon)\}.$$

With $\sigma^2 \le n^{-1}\cdot\sum_{i=1}^{n}\alpha_i$ and $\varepsilon=(1/2n)\cdot\sum_{i=1}^{n}\alpha_i$, we obtain the bounds

$$P\{n^{-1}\cdot\sum_{i=1}^{n}Y_i \le (1/2n)\cdot\sum_{i=1}^{n}\alpha_i\} \le \exp\{-\sum_{i=1}^{n}\alpha_i/10\}$$

and

$$P\{n^{-1}\cdot\sum_{i=1}^{n}Y_i \ge (3/2n)\cdot\sum_{i=1}^{n}\alpha_i\} \le \exp\{-\sum_{i=1}^{n}\alpha_i/10\}$$

from which lemma 1 follows trivially. Q E D

Lemma 2: Let $\{a_n\}$,$\{c_n\}$ and $\{d_n\}$ be nonnegative number sequences such that $\{d_n\}$ is bounded. Then

$$\sum_{n=1}^{\infty} a_n = \infty \quad \text{and} \quad \lim_n d_n\cdot\sum_{i=1}^{n} c_i = \infty$$

if and only if there exists a sequence $\{k_n\}$ of integers with $1\le k_n\le n$ for all n and

$$\lim_n \sum_{i=1}^{k_n} a_i = \infty \quad \text{and} \quad \lim_n d_n\cdot\sum_{i=k_n+1}^{n} c_i = \infty.$$

Proof: (if part) By hypothesis,

$$d_n\cdot\sum_{i=1}^{n} c_i \ge d_n\cdot\sum_{i=k_n+1}^{n} c_i \to \infty \text{ as } n\to\infty$$

And

$$\sum_{i=1}^{n} a_i \ge \sum_{i=1}^{k_n} a_i \to \infty \text{ as } n\to\infty.$$

(only if part) We need only find a sequence $\{k_n\}$ of integers with $k_n\to\infty$ as $n\to\infty$ and

$$d_n\cdot\sum_{i=k_n+1}^{n} c_i \to \infty \text{ as } n\to\infty$$

because $\sum_{n=1}^{\infty} a_n = \infty$ and $\lim_n k_n = \infty$ imply that

$$\lim_n \sum_{i=1}^{k_n} a_i = \infty.$$

Now, if $\sum_{n=1}^{\infty} c_n < \infty$, then $d_n\cdot\sum_{i=1}^{n} c_i \to\infty$ as $n\to\infty$ implies that $\lim_n d_n = \infty$, but this contradicts the hypothesis that $\{d_n\}$is bounded. So we can assume that $\sum_{n=1}^{\infty} c_n = \infty$.

Let k_n be the largest integer such that

$$\sum_{i=k_n+1}^{n} c_i > \sum_{i=1}^{n} c_i/2.$$

It is clear that k_n is monotonically nondecreasing and that if $k_n \not\to \infty$ as $n\to\infty$, then $k_n\to K<\infty$, and, in fact, $k_n=K$ for all n large enough. Because k_n is largest, we have for all n large enough:

$$\sum_{i=K+2}^{n} c_i \le \sum_{i=1}^{n} c_i/2$$

so that

$$\sum_{i=K+2}^{n} c_i \le \sum_{i=1}^{K+1} c_i$$

for all n large enough. But this would imply that

$$\sum_{n=1}^{\infty} c_n \le 2 \cdot \sum_{n=1}^{K+1} c_n < \infty$$

which is a contradiction with $\sum_{n=1}^{\infty} c_n = \infty$.

Therefore, $\lim_n k_n = \infty$ and

$$d_n \cdot \sum_{i=k_n+1}^{n} c_i \ge (d_n/2) \cdot \sum_{i=1}^{n} c_i \to \infty \text{ as } n \to \infty .$$

Q E D

Lemma 3: If $\{c_n\}$ and $\{d_n\}$ are nonnegative number sequences with $\lim_n c_n = \infty$, then

$$\lim_n c_n \cdot e^{-\lambda \cdot d_n} = 0 \quad \text{for all } \lambda > 0$$

if and only if

$$\lim_n d_n / \log c_n = \infty$$

Proof: (if part) Assume, without loss of generality that $c_n > 1$ for all n. Given $\lambda > 0$ find an integer N such that for all $n \ge N$: $d_n > (2/\lambda) \cdot \log c_n$. Then

$$c_n \cdot e^{-\lambda \cdot d_n} \le c_n \cdot e^{-2 \log c_n} = 1/c_n \to 0 \text{ as } n \to \infty .$$

(only if part) Assume, without loss of generality, that $c_n > 1$ for all n. Suppose that $d_n / \log c_n \not\to \infty$ as $n \to \infty$. Then there exists a constant $M < \infty$ and a subsequence $\{n'\}$ such that $d_{n'} / \log c_{n'} \le M$ for all n'. Thus,

$$c_{n'} \cdot e^{-d_{n'}/M} \ge c_{n'} \cdot e^{-\log c_{n'}} = 1$$

for all n'. Therefore, $c_n \cdot e^{-d_n/M}$ does not converge to 0 as $n \to \infty$, contradicting the hypothesis. Q E D

Lemma 4: Let $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ be nonnegative number sequences.

(i) If $r > 0$, then

$$\lim_n a_n / b_n^r = 0 \quad \text{and} \quad \lim_n b_n c_n = 0$$

for some sequence $\{b_n\}$ if and only if $\lim_n a_n c_n^r = 0$.

(ii) Let $\lim_n a_n = \infty$. Then

$$\lim_n a_n e^{-\lambda b_n} = 0 \text{ for all } \lambda > 0 \text{ and } \lim_n b_n c_n = 0$$

for some sequence $\{b_n\}$ if and only if

$$\lim_n a_n \cdot e^{-\lambda / c_n} = 0 \text{ for all } \lambda > 0$$

which, on its turn, is equivalent to the condition:

$$\lim_n c_n \log a_n = 0$$

Proof: (i,"if" part) Let $b_n^{r+1} = a_n / c_n$. Then

$$b_n c_n = a_n / b_n^r = a_n^{1 - r/(r+1)} \cdot c_n^{r/(r+1)} = (a_n c_n^r)^{1/(r+1)} \to 0 \text{ as } n \to \infty .$$

(i,"only if" part) Trivially, $a_n c_n^r = (a_n / b_n^r) \cdot (b_n c_n)^r \to 0$ as $n \to \infty$.

(ii,"if" part) We remark that $c_n \log a_n \to 0$ as $n \to \infty$ if and only if $a_n e^{-\lambda / c_n} \to 0$ as $n \to \infty$ for all $\lambda > 0$ by lemma 3. Choose

$$b_n = \log a_n / (c_n \log a_n)^{1/2}$$

and note that $b_n c_n = (c_n \log a_n)^{1/2}$ and $b_n / \log a_n = (c_n \log a_n)^{-1/2}$. The theorem follows if we note that, by lemma 3, $a_n e^{-\lambda b_n} \to 0$ as $n \to \infty$ for all $\lambda > 0$ in view of $b_n / \log a_n \to \infty$ and $a_n \to \infty$ as $n \to \infty$.

(ii,"only if" part) Trivially, employing lemma 3 again,

$$c_n \log a_n = (b_n c_n) / (b_n / \log a_n) \to 0 \text{ as } n \to \infty .$$

Q E D

Lemma 5: Let $t > 1$, $0 \le C < \infty$, and let $\mathcal{G}_{t,C}$ be the class of random variables Y with $E\{Y\} = 0$ and $\sup_{Y \in \mathcal{G}_{t,C}} E\{|Y|^t\} \le C$. Let $Y_1, Y_2, \ldots, Y_n$ be iid random variables that are distributed as Y. Then for each $\varepsilon > 0$ there exists a constant K depending upon ε, t and C such that

$$\sup_{Y \in \mathcal{G}_{t,C}} P\{ \bigcup_{k=n}^{\infty} \{ |k^{-1} \cdot \sum_{i=1}^{k} Y_i| \ge \varepsilon \} \} \le K / n^{t-1} , \quad (42)$$

Let $\mathcal{J}$ be the class of random variables Y with $E\{Y\} = 0$ and with the property that for every $\delta > 0$ there exists a $c(\delta) > 0$ with

$$\sup_{Y \in \mathcal{J}} E\{e^{\lambda Y}\} \le e^{|\lambda|^{\delta}}$$

$$,\text{all } \lambda \in [-c(\delta),+c(\delta)]\,. \qquad (43)$$

Then for each $\varepsilon>0$ there exist numbers $B>0$ and $M>0$ (both depending upon ε and the function c) such that

$$\sup_{Y\in\mathcal{J}} P\{\bigcup_{k=n}^{\infty}\{|k^{-1}.\sum_{i=1}^{k} Y_i|\geq\varepsilon\}\}\leq M.e^{-Bn}\,. \qquad (44)$$

Proof: (42) is obtained by inspecting the proof of [11, theorem 1]. To show (44), let $\varepsilon>0$ be arbitrary. We will extend the proof of a theorem of [12]. In particular, we show that (44) holds with $B=c(\varepsilon/2).\varepsilon/2$ and $M=2/(1-e^{-B})$. Let Y be a random variable from $\mathcal{J}$ and note that by Chebyshev's inequality and by (43),

$$P\{|k^{-1}.\sum_{i=1}^{k} Y_i|\geq\varepsilon\}\leq e^{-|\lambda|\varepsilon k}.$$

$$((E\{e^{\lambda Y}\})^k+(E\{e^{-\lambda Y}\})^k)$$

$$\leq 2e^{-|\lambda|\varepsilon k}.e^{|\lambda|k\varepsilon/2} \quad \text{for all } \lambda\in[-c(\varepsilon/2),+c(\varepsilon/2)]$$

$$\leq 2e^{-k.c(\varepsilon/2).\varepsilon/2} \quad \text{by choice of } \lambda$$

$$= 2e^{-Bk}.$$

Further, $P\{\bigcup_{k=n}^{\infty}\{|k^{-1}.\sum_{i=1}^{k} Y_i|\geq\varepsilon\}\}\leq$

$$\sum_{k=n}^{\infty} P\{|k^{-1}.\sum_{i=1}^{k} Y_i|\geq\varepsilon\}$$

$$\leq\sum_{k=n}^{\infty} 2e^{-Bk}\leq(2/(1-e^{-B})).e^{-Bn}=M.e^{-Bn}$$

Q E D

Proof of theorem 1: Let $\varepsilon>0$ be arbitrary. Then, if $\chi_{\{.\}}$ denotes the indicator function of $\{.\}$,

$$P\{M'_n>q_{min}+\varepsilon\mid(H_{n-1},Z_{n-1})\}=$$

$$\alpha_n.P\{Q(W_{L_n})>q_{min}+\varepsilon\mid(H_{n-1},Z_{n-1})\}$$

$$+(1-\alpha_n).P\{Q(W_{I_n})>q_{min}+\varepsilon\mid(H_{n-1},Z_{n-1})\}$$

$$\leq\alpha_n+1-\beta_n+\beta_n.\chi_{\{Q(X_{n-1})>q_{min}+\varepsilon\}}$$

and

$$P\{M'_n>q_{min}+\varepsilon\}\leq\alpha_n+1-\beta_n+P\{Q(X_{n-1})>q_{min}+\varepsilon\}$$

which tends to 0 as $n\to\infty$ in view of (22) and (23). If (20) holds, then $|M'_n|\leq K_1<\infty$ for all n so that $M'_n\vee q_{min}\to q_{min}$ in L_r for all $r>0$. Also, under (20), $M_n\vee q_{min}\to q_{min}$ in L_r for all $r>0$ if and only if $M_n\vee q_{min}>q_{min}$ in probability as $n\to\infty$. But note that $E\{M_n\}=E\{M'_n\}$ so that the "in probability" part of theorem 1 is proved. For the second part of the theorem, we remark that

$$P\{\bigcup_{k=n}^{\infty}\{M_k>q_{min}+\varepsilon\}\}\leq$$

$$P\{\bigcup_{k=n}^{\infty}\{\alpha_k+1-\beta_k>\varepsilon/2K_1\}\}$$

$$+P\{\bigcup_{k=n}^{\infty}\{Q(X_{k-1})>q_{min}+\varepsilon/2\}\}$$

$$=P\{\bigcup_{k=n}^{\infty}\{Q(X_{k-1})>q_{min}+\varepsilon/2\}\}$$

for all n large enough in view of (22). Thus, if $Q_n\vee q_{min}\to q_{min}$ WP1 as $n\to\infty$, then $M_n\vee q_{min}\to q_{min}$ WP1 as $n\to\infty$. Q E D

Proof of theorem 2: Let $\alpha_0\overset{\Delta}{=}1$ and note that the sequence $\{1/b_n.\sum_{i=0}^{n}\alpha_i\}$ is bounded. From (25), (26) and lemma 2 find a sequence $\{k_n\}$ of integers with

$$\lim_n k_n=\infty\,;\ \lim_n\sum_{i=1}^{k_n}\alpha_i\eta_i=\infty\,;$$

$$\lim_n(1/b_n.\sum_{i=0}^{n}\alpha_i).\sum_{i=k_n+1}^{n}\gamma_i.(1-\alpha_i)=\infty$$

Let $\varepsilon>0$ be arbitrary and define the following events:

$$A_{n1}=\{|L_n-\sum_{i=0}^{n}\alpha_i|\leq\sum_{i=0}^{n}\alpha_i/2\}$$

$$A_{n2}=\{|L_{k_n}-\sum_{i=0}^{k_n}\alpha_i|\leq\sum_{i=0}^{k_n}\alpha_i/2\}$$

$$A_{n3}=\bigcap_{i=1}^{L_{k_n}}\{N_i^n\geq b_n\}$$

$$A_{n4}=\bigcup_{i=1}^{L_{k_n}}\{Q(W_i)\leq q_{min}+\varepsilon/2\}$$

$$A_{n5}=\bigcap_{i=1}^{L_n}\{\{\{|\bar{Y}_i^n-Q(W_i)|<\varepsilon/4\}\cap\{N_i^n\geq b_n\}\}\cup\{N_i^n<b_n\}\}$$

$$A_{n0} = \{Q(X_n) \le q_{min} + \epsilon\}$$

Noting that $A_{n1} \cap A_{n2} \cap A_{n3} \cap A_{n4} \cap A_{n5} \subseteq A_{n0}$, we have :

$$P\{A_{n0}^c\} \le P\{A_{n1}^c\} + P\{A_{n2}^c\} + P\{A_{n1} \cap A_{n2} \cap A_{n3}^c\} + P\{A_{n4}^c\} + P\{A_{n1} \cap A_{n5}^c\} \tag{45}$$

where $(.)^c$ denotes the complement of a set.

By lemma 1,

$$P\{A_{n1}^c\} \le 2\exp\{-\sum_{i=0,}^{n} \alpha_i / 10\} \tag{46}$$

$$P\{A_{n2}^c\} \le 2\exp\{-\sum_{i=0,}^{n} \alpha_i / 10\} \tag{47}$$

If X denotes a random vector with the distribution function G in $\mathbb{R}^m$, then we know that $P\{Q(X) \le q_{min} + \epsilon/2\} = \xi > 0$. Then, using lemma 1 again:

$$\begin{aligned}
P\{A_{n4}^c\} &\le P\{\bigcap_{i=1,}^{L_{k_n}} \{Q(W_i) > q_{min} + \epsilon/2\}\} \\
&\le P\{\sum_{i=1,}^{k_n} \chi_{\{U_i'=U_i=1\}} < \sum_{i=1,}^{k_n} \alpha_i\gamma_i/2\} \\
&+ P\{\sum_{i=1,}^{k_n} \chi_{\{U_i'=U_i=1\}} \ge \sum_{i=1,}^{k_n} \alpha_i\gamma_i/2 ; \\
&\bigcap_{i=1,}^{L_{k_n}} \{\{Q(W_i) > q_{min} + \epsilon/2\} \cap \{U_{T_i}=1\}\} \cup \{U_{T_i}=0\}\}\} \\
&\le \exp\{-\sum_{i=1,}^{k_n} \alpha_i\gamma_i/10\} + (P\{Q(X) > q_{min} + \epsilon/2\})^{\sum_{i=1,}^{k_n} \alpha_i\gamma_i/2} \\
&\le 2\exp\{-\mathrm{Min}(1/10 ; \xi/2) \cdot \sum_{i=1,}^{k_n} \alpha_i\gamma_i\}
\end{aligned} \tag{48}$$

Further,

$$\begin{aligned}
P\{A_{n1} \cap A_{n5}^c\} &\le P\{\bigcup_{i=1,}^{L_n} \{\{|\overline{Y}_i^n - Q(W_i)| \ge \epsilon/4\} \cap \{N_i^n \ge b_n\}\} ; L_n \le (3/2) \cdot \sum_{i=0,}^{n} \alpha_i\} \\
&\le (3/2) \cdot \sum_{i=0,}^{n} \alpha_i \cdot \sup_{x \in B} P\{\bigcup_{l=b_n,}^{\infty} \{|\overline{Y}_{(x,l)} - Q(x)| \ge \epsilon/4\}\}
\end{aligned}$$

where $\overline{Y}_{(x,l)}$ is the average of l iid random variables all having distribution function F_x in $\mathbb{R}^1$ (and mean Q(x),obviously). By lemma 5, we can upper bound the last term by

$$(3/2) \cdot \sum_{i=0,}^{n} \alpha_i \cdot g(\mathcal{E}, \epsilon, b_n) \tag{49}$$

where

$g(\mathcal{E}, \epsilon, b_n) =$
- 0 if $\mathcal{E}$ is $\emptyset$
- K_2/b_n^{t-1} if $\mathcal{E}$ is $\mathcal{E}_t$ for some $t > 1$ ($K_2 > 0$ depends upon $\mathcal{E}$ and ϵ)
- $K_3 e^{-K_4 b_n}$ if $\mathcal{E}$ is χ where $K_3 > 0$ and $K_4 > 0$ are constants depending upon $\mathcal{E}$ and ϵ.

We also have that

$$P\{A_{n1} \cap A_{n2} \cap A_{n3}^c\} \le (3/2) \cdot \sum_{i=0,}^{k_n} \alpha_i \cdot P\{1 + \sum_{i=k_n+1,}^{n} \chi_{\{U_i'=V_i=0\}} < b_n ; \quad L_n \le (3/2) \cdot \sum_{i=0,}^{n} \alpha_i\}. \tag{50}$$

Note that $E\{\chi_{\{U_i'=V_i=0\}} | L_{i-1}\} = \gamma_i(1-\alpha_i)/L_{i-1}$ with probability one. As we pointed out, we have for all n large enough:

$$b_n < (2/30) \cdot \sum_{i=k_n+1,}^{n} \gamma_i(1-\alpha_i) \Big/ \sum_{i=0,}^{n} \alpha_i$$

Therefore, (50) is for all n large enough upper bounded by $(3/2) \cdot \sum_{i=0,}^{k_n} \alpha_i \cdot P\{\sum_{i=k_n+1,}^{n} (\chi_{\{U_i'=V_i=0\}} - E\{\chi_{\{U_i'=V_i=0\}} | L_{i-1}\}) \ldots (3/2) \cdot \sum_{i=0,}^{n} \alpha_i\}$

$$< -(1/2).(\sum_{i=k_n+1,}^{n} \gamma_i(1-\alpha_i))/((3/2)\sum_{i=0,}^{n} \alpha_i) \mid L_n \le$$

$$(3/2).\sum_{i=0,}^{n} \alpha_i \}$$

$$\le (3/2)\sum_{i=0,}^{k_n} \alpha_i.\exp\{-\sum_{i=k_n+1,}^{n}$$

$$\gamma_i(1-\alpha_i)/((30/2).\sum_{i=0,}^{n} \alpha_i)\}$$

$$\le (3/2)\sum_{i=0,}^{k_n} \alpha_i . e^{-b_n}. \qquad (51)$$

We remark that for all the environments considered in theorem 2, $\lim_n P\{A_{n1} A_{n5}^c\} = 0$. Further, (25) implies that $P\{A_{n1}^c\}+P\{A_{n2}^c\}+P\{A_{n4}^c\} \to 0$ as $n\to\infty$ in view of (46-48). Finally, by (51), for all n large enough,

$$P\{A_{n1}\cap A_{n2}\cap A_{n3}^c\} \le (3/2).\sum_{i=0,}^{k_n} \alpha_i .e^{-b_n} \to 0$$

as $n\to\infty$

in view of $k_n \le n$ and (28) (where we use the fact that (27) implies (28)) for environments that are $\mathcal{K}$ or $\mathcal{E}_t$ for some t>1. Thus, for these environments, $\lim_n P\{A_{n0}^c\} = 0$ for all $\epsilon>0$ in view of (45). If the environment is $\mathcal{D}$, notice that $P\{A_{n0}^c\} \le P\{A_{n4}^c\} \le 2\exp\{-\text{Min}(1/10;\ \epsilon/2).$ $.\sum_{i=1,}^{k_n} \alpha_i\eta_i\} \to 0$ as $n\to\infty$ in view of (25-26). However, with $b_n=1$ and $k_n=n$, the condition (26) is not needed. Q E D

<u>Proof of theorem 3</u>: Consider first environments that are $\mathcal{E}_t$ for some t>1. By lemma 4 and (30), we can find a sequence $\{b_n\}$ of integers with $b_n \ge 1$,

$$\lim_n (\sum_{i=1,}^{k_n} \alpha_i)/b_n^{t-1} = 0 \qquad (52)$$

and

$$\lim_n b_n.(\sum_{i=0,}^{n} \alpha_i)/(\sum_{i=k_n+1,}^{n} \gamma_i(1-\alpha_i)) = 0. \qquad (53)$$

Let $\epsilon>0$ be arbitrary and define $A_{n0}, A_{n1}, A_{n2}, A_{n3}$ and A_{n4} as in the proof of theorem 2. Let further

$$A_{n6} = \bigcap_{i=1,}^{L_{k_n}} \{|\bar{Y}_i^n - Q(W_i)| < \epsilon/4\} \qquad (54)$$

and note that $A_{n1}\cap A_{n2}\cap A_{n3}\cap A_{n4}\cap A_{n6} \subseteq A_{n0}$.

Therefore,

$$P\{A_{n0}^c\} \le P\{A_{n1}^c\}+P\{A_{n2}^c\}+P\{A_{n1}\cap A_{n2}\cap A_{n3}^c\}+ P\{A_{n4}^c\}+P\{A_{n2}\cap A_{n3}\cap A_{n6}^c\}. \qquad (55)$$

We recall from the proof of theorem 2 that

$$P\{A_{n1}^c\} + P\{A_{n2}^c\} \le 4\exp\{-\sum_{i=0,}^{k_n} \alpha_i/10\} \qquad (56)$$

$$P\{A_{n4}^c\} \le 2\exp\{-\text{Min}(1/10;\mathcal{E}/2).\sum_{i=1,}^{k_n} \alpha_i\eta_i\} \qquad (57)$$

and, in view of (53), for all n large enough,

$$P\{A_{n1}\cap A_{n2}\cap A_{n3}^c\} \le (3/2).\sum_{i=0,}^{k_n} \alpha_i .e^{-b_n} \qquad (58)$$

Next, using an argument as in theorem 2,

$$P\{A_{n2}\cap A_{n3}\cap A_{n6}^c\} \le (3/2).\sum_{i=0.}^{k_n} \alpha_i .$$
$$\sup_{x\in B} P\{\bigcup_{l=b_n,}^{\infty} \{|\bar{Y}_{(x,l)} - Q(x)| \ge \epsilon/4\}\}$$
$$\le (3/2).(\sum_{i=0,}^{k_n} \alpha_i).g(\mathcal{E},\epsilon,b_n) \le \qquad (59)$$
$$(3K_2/2).(\sum_{i=0,}^{k_n} \alpha_i)/b_n^{t-1}$$

where $g(.,.,.)$ is defined in (49) and K_2 is a positive constant depending upon $\mathcal{E}$ and ϵ. It is not hard to see that $\lim_n P\{A_{n0}^c\} = 0$ in view of (55), (56-59), (29), (25) and (52) (where we use the fact that (52) implies that the right-hand side of (58) tends to 0 as $n\to\infty$).

If $\mathcal{E}$ is deterministic, then $P\{A_{n0}^c\} \leq P\{A_{n4}^c\}$, which can be bounded as in (57). Clearly, $\lim_n P\{A_{n0}^c\} = 0$ in view of (25) and (29).

If $\mathcal{E}$ is $\mathcal{K}$, then, by lemma 4 and (31), it is possible to find a sequence $\{b_n\}$ of integers with $b_n \geq 1$, (53) and

$$\lim_n \left(\sum_{i=1}^{k_n} \alpha_i\right) . e^{-\lambda b_n} = 0 \quad \text{for all } \lambda > 0. \qquad (60)$$

All the terms on the right hand side of (55) are bounded as for $\mathcal{E}_t$ type environments (see (56-58)) with the exception that for some constants $K_3 > 0$ and $K_4 > 0$ (depending upon $\mathcal{E}$ and ε):

$$P\{A_{n2} \cap A_{n3} \cap A_{n6}^c\} \leq (3K_3/2) . \left(\sum_{i=0}^{k_n} \alpha_i\right) . e^{-K_4 b_n} . \qquad (61)$$

Again, it is not hard to see that $\lim_n P\{A_{n0}^c\} = 0$ in view of (55-58), (60) and (61). Theorem 3 then follows from the arbitrariness of ε. Q E D

9. REFERENCES

1. Ya. Z. Tsypkin, A.S. Poznyak: "Finite learning automata." , Engineering Cybernetics, Vol. 10, No. 3, pp. 479-490, 1972.

2. I.J. Shapiro, K.S. Narendra: "Use of stochastic automata for parameter self-optimization with multimodal performance criteria.", IEEE Transactions on Syst., Sci. and Cybernetics, Vol. SSC-5, No. 4, pp. 352-360, 1969.

3. R. Viswanathan, K.S. Narendra: "Application of stochastic automata models to learning systems with multimodal performance criteria.", Becton Center, Yale Univ., New Haven, Conn., Techn. Rept. CT-40, 1972.

4. G.J. McMurtry, K.S. Fu: "A variable structure automaton used as a multimodal searching technique.", IEEE Transactions on Autom.Contr., Vol. AC-11, No. 3, pp. 379-387, 1966.

5. L.P. Devroye: "A class of optimal performance directed probabilistic automata.", IEEE Transactions on Syst., Man and Cybernetics, Vol. SMC-6, No. 11, pp. 777-784. 1976.

6. G.J. McMurtry: "Adaptive optimization procedures.", in: "Adaptive, learning and pattern recognition systems." J.M. Mendel and K.S. Fu Ed., Ac. Press, N.Y., 1970.

7. L.D. Cockrell, K.S. Fu: "On search techniques in adaptive systems.", Purdue Univ., Lafayette, Ind., Techn. Rept. TR-EE-70-01, 1970.

8. L.P. Devroye: "On the convergence of statistical search.", IEEE Transactions on Syst., Man and Cybernetics, Vol. SMC-6, No. 1, pp.46-56, 1976.

9. R.A. Jarvis: "Optimization strategies in adaptive control: a selective survey.", IEEE Transactions on Syst., Man and Cybernetics, Vol. SMC-5, No. 1, pp. 83-94, 1975.

10. M. Loeve: "Probability theory.", 3d ed., Van Nostrand, Princeton, N.J., 1968.

11. T.J. Wagner: "On the rate of convergence for the law of large numbers.", Ann.Math.Stat., Vol. 40, No. 6, pp. 2195-2197, 1969.

12. L. Baum, M. Katz, R.R. Read: "Exponential convergence rates for the law of large numbers.", Trans.Amer.Math.Soc., Vol. 102, pp. 187-199, 1962.

13. D.K. Fuk, S.V. Nagaev: "Probability inequalities for sums of independent random variables.", Theory of Probab. and its Applications, Vol. 16, No. 4, pp. 643-660, 1971.

14. L. Baum, M. Katz: "Convergence rates in the law of large numbers.", Trans.Amer.Math.Soc., Vol. 121, pp. 108-123, 1965.