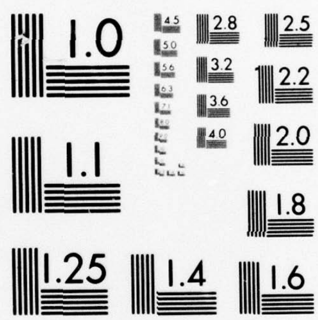
AD-A053 986 NORTH CAROLINA UNIV AT CHAPEL HILL DEPT OF STATISTICS F/G 9/4
TOWARD A THEORY OF PROBABILISTIC AUTOMATA WITH ENVIRONMENTS,(U)
1978 J GOULD, E J WEGMAN N00014-67-A-0321-0006

UNCLASSIFIED AFOSR-TR-78-0730 NL

| OF |
AD
A053986

END
DATE
FILMED
6-78
DDC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

B-5

TOWARD A THEORY OF PROBABILISTIC AUTOMATA WITH ENVIRONMENTS

JERREN GOULD
Claremont Graduate School
Claremont, California (U.S.A.)

EDWARD J. WEGMAN
University of North Carolina
Chapel Hill, North Carolina (U.S.A.)

See 1473 in back

The viewpoint is taken that every probabilistic automaton is situated within a sequence of environments which affects the initial state distribution and the transition function. There exists a probabilistic automaton within a deterministic environment sequence (ADE) which defines an event which is not a PCE, yet in certain nontrivial cases the behavior of an ADE can be simulated by a probabilistic automaton. For probabilistic automata within random environment sequences there is a mean equivalent canonical representation which eliminates the randomness due to probabilistic transition.

The traditional method to describe a system is to model the behavior as it stands. The system may not, however, exhibit identical behavior when placed within a distinct external situation. That is, the input and quintessential dynamics of the system remain the same, but the external factors have changed. These external factors are the environment within which the system operates. By the traditional approach a system which performs differently within distinct environments would have to be regarded as a collection of distinct systems. Consideration of automata within environments allows the various factors which may influence the behavior of the system to be incorporated in the model. Thus, by this more fundamental description systems which perform differently within distinct environments could be identical with the difference in behavior being attributed to the environments. Automata in media were investigated also in [8] and [10], but our concept is more general and includes this earlier work as a special case.

Many biological or social systems operate within environments which influence their behavior. Even mechanical or electrical systems, intended to be independent of external factors, may exhibit such behavior. If the environment may be controlled or the internal dynamics modified, then there may be methods of optimizing the behavior of the system.

Probabilistic automata (PA) are mathematical models for finite state systems which admit at

Work supported by Grant No. AFOSR-75-2840 and ONR Grant No. N00014-67-A-0321-0006, Task NR 042-269.

discrete intervals certain inputs and emit certain outputs. If the system is in state s_i and if the present input is σ, then the system may go into any state and the probability of going to state s_j depends only on s_i, s_j, and σ. The output of the system depends only on the state obtained. There is extensive literature on probabilistic automata. In this paper we shall follow the notations of [7] and [5]. In particular, the formulation given therein amounts to assuming that the set of outputs contains just two elements. Because of the restriction to two outputs, these automata can be viewed as defining (recognizing) sets of sequences of inputs (tapes); this point of view is adopted throughout this paper.

Probabilistic automata exhibit a behavior independent of environmental variation. It seems quite natural, however, to consider automata with stochastic behavior located within environments which affect their properties. Thus, the probabilistic transition function of the system is not only related to the present state and input, but also to the present configuration of the environment. Also, the initial state distribution and final output relation may depend on their repective environments.

This work surveys some lines of research in Gould and Wegman [3] and Gould [1], [2]. In this paper we develop a formulation of environments and of probabilistic automata within environments and answer some of the basic questions about them.

The sequence of environments may be derived from a deterministic rule; in such a case we have an automaton in deterministic environments (ADE). ADE are, in general, stronger than probabilistic automata. In an effort to find when an ADE and a PA have the same capability, we introduce the concept of simulation of an ADE by a PA (Theorem 2). Reduction of the environment set to certain finite structures is also considered.

Finally, we consider the case of probabilistic automata operating in random environments (ARE); we assume that the realization of any environment is governed by some probabilistic structure. Hence, the environment sequence is a stochastic process. We encounter the dual nature of the randomness involved in ARE. We have that the environments are random and that for each value the environment assumes, the probabilistic transitions are defined. The relative frequency of acceptance of a tape under a fixed random environment sequence is a consistent estimator of the expected acceptance probability. Finally, we find that there is a mean equivalent canonical representation which eliminates the randomness due to probabilistic state assignment and transitions. Thus, for any automaton in random environments we can find a finite automaton in random environments with finite environment set, deterministic assignment of the initial state, and deterministic transitions which has a state distribution equivalent in the mean for any input tape.

1. Probabilistic Automata

In this section we present some of the basic definitions and results from the theory of probabilistic automata which will be used in the succeeding sections.

Let Σ be the finite input set, the alphabet, and let Σ^* be the class of all finite sequences of elements of Σ. Let us also include Λ, the empty tape, in Σ^*. If $x = \sigma_1 \ldots \sigma_K$ is a tape, then the length $L(x)$ of x is $L(x)=K$. Note that $L(\Lambda)=0$. If x and y are tapes, then xy will denote the tape which is the concatenation of x and y.

Let M_n denote the set of all $n \times n$ stochastic matrices and V_n denote the set of all n-dimensional stochastic vectors.

Definition 1 A probabilistic automaton (PA) over the alphabet Σ is a system $A = (S, M, \pi_0, F)$, where $S = \{s_1, \ldots, s_n\}$ is a finite set (the set of internal states), M is a function $M: \Sigma \to M_n$ (the matrix transition function) such that $m_{ij}(\sigma)$ is the probability of changing to state s_j under input σ given that the system is in state s_i, $\pi_0 \in V_n$ (the initial state distribution), and $F \subset S$ (the set of acceptance states).

The function M can be extended to define the transition probabilities for going from state s_i to state s_j by a sequence $x \in \Sigma^*$. Let $M(\Lambda) = I_n$, the $n \times n$ identity. For $x=\sigma_1 \ldots \sigma_K$, we obtain $M(x)$ by the rule $M(x) = M(\sigma_1)M(\sigma_2)\ldots M(\sigma_K)$

Let $A = (S, M, \pi_0, F)$ be a PA over alphabet Σ. We define the state distribution of A after input tape x as $\pi(x) = \pi_0 M(x)$. A tape $x \in \Sigma^*$ is said to be accepted by A if a state in F is obtained after tape x is input. Let η^F be the n-dimensional column vector whose i^{th} component is 1 if $s_i \in F$ and 0 otherwise. The probability that tape x is accepted by A is defined as $p(x)$ and is calculated $p(x) = \pi(x)\eta^F = \pi_0 M(x)\eta^F$.

A PA may be used to define sets of tapes. These sets will not only depend on A but also on an additional parameter called the cut-point. Let A be a PA and λ be a real number, $\lambda \in [0,1)$. The set of all tapes defined by A with cut-point λ is $T(A,\lambda) = \{x | x \in \Sigma^*, \lambda < p(x)\}$. Also, we say $U \subset \Sigma^*$ is a probabilistic cut-point event (PCE) if $U = T(A,\lambda)$ for some PA A and some $\lambda \in [0,1)$. PCE is a synonym for stochastic language.

2. Automata in Deterministic Environments

We shall consider the aggregate of all configurations of the environment as an abstract set denoted by E. Presently, we do not make any stipulations as to the origin, form, structure, or cardinality of E. As each new symbol is input to the system, the configuration of the environment attends to the system and the probabilistic transition ensues as a function of the input symbol and the configuration of the environment. We also view the initial state distribution as a function of the environment.

Here we shall consider the case in which the sequence of environments is specified by a deterministic rule. That is, for the initial distribution and the subsequent transitions, we are given the precise condition of the environment. Let E^∞ denote the cartesian product of a countable number of copies of the environment set E. The environment sequence is a mapping $\underline{e}(K) = e_K \in E$ for K=0,1,... We also denote $\underline{e}$ by the sequence $(e_0, e_1, \ldots)$.

Definition 2 An automaton in deterministic environments (ADE) is a system $A = (\Sigma, S, G, \pi_0, F, E)$, where Σ is the finite input alphabet, $S=\{s_1, \ldots s_n\}$ is the finite internal state set, G is a mapping $G: \Sigma \times E \to M_n$ (the basic matrix transition function), $\pi_0: E \to V_n$ (the initial distribution function), $F \subset S$ (the set of acceptance states), and E is the set of environments.

The matrix transition function M is defined on $\Sigma^* \times E^\infty$ by the inductive recursion rule:

1. $M(\Lambda, \underline{e}) = I_n \quad \forall \underline{e} \in E^\infty$
2. if $x \in \Sigma^*$ such that $L(x)=K$ and $\underline{e} \in E^\infty$, then $M(x\sigma, \underline{e}) = M(x, \underline{e})G(\sigma, e_{K+1})$.

Hence, for all $x \in E^\infty$,

$$M(x,\underline{e}) = \prod_{i=1}^{K} G(\sigma_i, e_i),$$

where $x = \sigma_1 \ldots \sigma_K$ and $\underline{e} = (e_0, e_1, \ldots)$.

The state distribution of A within environment sequence $\underline{e}$ after input tape x is computed as with PA except the relevant quantities depend on the environments. Thus, we obtain $\pi_0(\underline{e}) = \pi_0(e_0)M(x,\underline{e})$. Similarly, $p(x,\underline{e}) = \pi(x,\underline{e})\eta^F$ is the probability that A within environment sequence $\underline{e}$ accepts tape x. Clearly, these functions do not depend on environments beyond the input symbols.

An ADE may also be used to define sets of tapes in a manner similar to that of PA except that the set defined will depend on the environment sequence.

Definition 3 Let A be an ADE and λ be a real number, $\lambda \in [0,1)$. The set of tapes $T(A,\underline{e},\lambda) = \{x | x \in \Sigma^*, \lambda < p(x,\underline{e})\}$ is called the set of tapes defined by A within environment sequence $\underline{e}$ with cut-point λ.

We now study the problem of whether it suffices to ignore the extension of ADE with the set of acceptance states related to the environment. Logically, we restrict the relationship to the environment concurrent with the terminal input symbol of any tape. Let us order the 2^n subsets of S and let $\phi_j(e)$ be the probability that F_j, the j^{th} subset of S, is the set of acceptance states when the environment concurrent with the terminal input symbol is e. Let $T(A,\underline{e},\lambda)$ be any set of tapes defined by an extended ADE $A=(\Sigma,S,G,\pi_0,\phi,E)$ within environment sequence $\underline{e}$ with cut-point λ and with the set of acceptance states related probabilistically to the terminal environment by ϕ.

Consider the ADE B with $n2^n$ states over the same alphabet, with the same environment set, and $\pi_0{}^B(e_0) = (\pi_0(e_0)\phi_1(e_0), \ldots, \pi_0(e_0)\phi_{2^n}(e_0))$

$$G^B(\sigma,e) = \begin{bmatrix} G(\sigma,e)\phi_1(e) & G(\sigma,e)\phi_2(e) \cdots G(\sigma,e)\phi_{2^n}(e) \\ G(\sigma,e)\phi_1(e) & G(\sigma,e)\phi_2(e) \cdots G(\sigma,e)\phi_{2^n}(e) \\ \vdots & \vdots \qquad\qquad \vdots \\ G(\sigma,e)\phi_1(e) & G(\sigma,e)\phi_2(e) \cdots G(\sigma,e)\phi_{2^n}(e) \end{bmatrix}$$

$$\eta^{F^B} = \begin{bmatrix} \eta^{F_1} \\ \vdots \\ \eta^{F_2} \\ \vdots \\ \vdots \\ \eta^{F_{2^n}} \end{bmatrix}, \text{ where } \pi_0 \text{ and G are defined in } A.$$

It is easily verified that for any $x \in \Sigma^*$ and $\underline{e} \in E^\infty$ the probability that x is accepted by A within environment sequence $\underline{e}$ is identical to the probability that x is accepted by B within environment sequence $\underline{e}$. So $T(A,\underline{e},\lambda) = T(B,\underline{e},\lambda)$ for all $\underline{e} \in E^\infty$ and all $\lambda \in [0,1)$. But B is an ADE as in definition 2; that is, B has a constant set of acceptance states. Hence, we have the following result.

Theorem 1 Every set of tapes which can be defined by an extended ADE with the set of acceptance states related probabilistically to the terminal environment can be defined by an ADE with a constant set of acceptance states.

Another possible extension is to allow the components of $\eta^{F(e)}$ to be arbitrary numbers in the interval [0,1]. However, for any $e \in E$, $\eta^{F(e)}$ is contained in the convex hull of $\eta^{F_1}, \eta^{F_2}, \ldots, \eta^{F_{2^n}}$. So $\eta^{F(e)}$ can be represented as

$$\eta^{F(e)} = \sum_{j=1}^{2^n} \phi_j(e)\eta^{F_j} \quad \text{with}$$

$$0 \le \phi_j(e) \le 1 \quad \text{and} \quad \sum_{j=1}^{2^n} \phi_j(e) = 1.$$

It follows as a corollary to theorem 1 that the class of sets of tapes defined by these automata is identical to the class of sets of tapes defined by ADE as in definition 2.

Probabilistic automata can be considered as a special case of ADE in two ways. First, for any PA $A = (S, M, \pi_0, F)$ over Σ we can define an ADE B with any nonempty environment set E such that $\pi_0(e) = \pi_0$ and $G(\sigma,e) = M(\sigma)\ \forall \sigma \in \Sigma, \forall e \in E$. Such an ADE is not influenced by the configuration of the environment. Hence, for any $\lambda \in [0,1)$, $T(A,\lambda) = T(B,\underline{e},\lambda)\ \forall \underline{e} \in E^\infty$. Also, for any PA A we can define an ADE with any nonempty environment set E such that for some $e_0, e \in E$ $\pi_0(e_0) = \pi_0$ and $G(\sigma,e) = M(\sigma) \forall \sigma \in \Sigma$. For the environment sequence $\underline{e} = (e_0,e,e,\ldots)$ we obtain $T(A,\lambda) = T(B,\underline{e}\ \lambda)$ for every $\lambda \in [0,1)$. Thus, every set of tapes definable by a PA can be trivially defined by an ADE in either of two ways.

Let $\Sigma = \{\sigma\}$. Paz [6] demonstrated that there exists $U \subset \Sigma^*$ which is not a PCE. Consider a two state ADE A over alphabet Σ, where

$$E = \{0,1\},\ \pi_0(e) = (1-e, e),\ G(\sigma,e) = \begin{bmatrix} 1-e & e \\ 1-e & e \end{bmatrix} \forall e \in E$$

and $\eta^F = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The elements of Σ^* are of the form σ^K for some $K \ge 0$. Let V be any subset of Σ^*; then $V = \{\sigma^{K_1}, \sigma^{K_2}, \ldots\}$.

For $\underline{e}(K) = \begin{cases} 1 & \text{if } K=K_i \text{ for some } i \\ 0 & \text{otherwise} \end{cases}$ it is clear that $V = T(A,\underline{e},0)$. Since V is arbitrary, every subset of this particular Σ^* is definable by A within some environmental sequence. But $U \subset \Sigma^*$ is not definable by a PA. Thus, the class of automata in deterministic environments produces a strictly larger class of definable sets.

3. Simulation

Let A be an ADE with finite environment set. We shall construct an effective procedure to find a PA which simulates the operation of A.

Definition 4 A PA A' over alphabet Σ' simulates the ADE $A = (\Sigma,S,G,\pi_0,F,\underline{E})$ if there is a relation R between Σ'^* and $\Sigma^* \times E^\infty$ such that

1. for any $(x,\underline{e}) \in \Sigma^* \times E^\infty$ there exists $x' \in \Sigma'^*$ such that $[x', (x,\underline{e})] \in R$;
2. if $[x',(x,\underline{e})] \in R$, then for each $\lambda \in [0,1)$, $x' \in T(A',\overline{\lambda})$ iff $x \in T(A,\underline{e},\lambda)$;
3. if for some $\lambda \in [0,1)$ $x' \in T(A',\lambda)$, then there exists $(x,\underline{e}) \in \Sigma^* \times E^\infty$ such that $[x', (x,\underline{e})] \in R$.

If a PA A' simulates the ADE A, given any input tape x for A within environment sequence $\underline{e}$, we have a rule R to find an input tape x' for A' such that x' has the same acceptance probability as x for A within some environment sequence $\underline{e}$. We say a tape $x' \in \Sigma'^*$ is admissible if and only if there is an input tape x for A within some environment sequence $\underline{e}$ from which the rule R will yield x'. By condition 3 above we see that tapes which are not admissible do not belong to any set defined by A': that is, if x' is not admissible, then $p'(x') = 0$.

For any tape $x \in \Sigma^*$ with $L(x) = K \geq 0$ only the environments $e_0,e_1,\ldots,e_K$ influence the acceptance probability. Hence, the relation R need only depend on x and $e_0,e_1,\ldots,e_K$.

We shall use the notation $\#(A)$ to denote the cardinality of the set A.

Theorem 2 Let $A = (\Sigma,S,G,\pi_0,F,E)$ be an ADE such that $\#(E) < \infty$. There exists a PA A' over some finite alphabet Σ' which simulates A.

Proof: Let $m = \#(\Sigma)$ and $\mu = \#(E)$. Without any loss of generality we shall let $\Sigma = \{1,2,\ldots m\}$ and $E = \{0,1,\ldots,\mu-1\}$. Consider $\Sigma' = \{0,1,\ldots, \mu(m+1)-1\}$. Let $\Sigma_0 = \Sigma \cup \{0\}$. We now define a mapping $g: \Sigma_0 \times E \to \Sigma'$ as follows:

$g(\sigma,e) = e + \sigma\mu \quad \forall e \in E$ and $\forall \sigma \in \Sigma_0$. We extend the definition of g to $(\Sigma_0 \times E)^*$, the set of all finite sequences of elements $\Sigma_0 \times E$, by component-wise application and concatenation of the results. The extension of g is one-to-one correspondence between $(\Sigma_0 \times E)^*$ and Σ'^*. If b is any nonempty element of $(\Sigma_0 \times E)^*$, then b is isomorphic to $(y, (e_0,\ldots,e_K))$, where for some $K \geq 0$, $y \in \Sigma_0^*$, $L(y) = K + 1$, and $e_0,\ldots,e_K \in E$. We define the relation R to be the set of all elements of the form:

$$[g(0x,(e_0,\ldots,e_K)),\ (x,\underline{e})]$$

where $x \in \Sigma^*$, $L(x) = K$, $\underline{e} = (e_0,\ldots,e_K,\ldots)$, and $K = 0,1,\ldots$ Clearly, for each $(x,\underline{e}) \in \Sigma^* \times E^\infty$ there exists a unique $x' = g(0x, (e_0,\ldots,e_K))$, where $K = L(x)$, such that $[x',(x,\underline{e})] \in R$.

Now we shall construct a PA $A' = (S',M',\pi_0',F')$ over Σ' to simulate A. Let $S' = S \cup \{s_{n+1}, s_{n+2}\}$. For any σ' such that $0 \leq \sigma' \leq \mu-1$ we define

$$M'(\sigma') = \begin{pmatrix} & & & 0 & 1 \\ & 0 & & \vdots & \vdots \\ & & & 0 & 1 \\ \pi_0^{(1)}(\sigma') & \cdots & \pi_0^{(n)}(\sigma') & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}$$

where $\pi_0^{(i)}(\sigma')$ is the $i^{\underline{th}}$ component of $\pi_0(\sigma')$. Note that $0 \leq \sigma' \leq \mu-1$ implies $\sigma' \in E$. For $\mu \leq \sigma' \leq \mu(m+1)-1$, define

$$M'(\sigma') = \begin{pmatrix} & 0 & 0 \\ G(g^{-1}(\sigma')) & \vdots & \vdots \\ & 0 & 0 \\ 0 \cdots 0 & 0 & 1 \\ 0 \cdots 0 & 0 & 1 \end{pmatrix}$$

Let $\pi_0' = (0,\ldots,0,1,0)$ and $F' = F$. Therefore,

$$\eta^{F'} = \begin{pmatrix} \eta^F \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Suppose $[x', (\Lambda,\underline{e})] \in R$. Therefore, $x' = g(0\Lambda,e_0) = g(0,e_0) = e_0$. So

$$p'(x') = \pi_0' M'(e_0)\eta^{F'} = (\pi_0(e_0),0,0)\eta^{F'}$$
$$= \pi_0(e_0)\eta^F = p(\Lambda,\underline{e}).$$

Thus, if $[x', (\Lambda,\underline{e})] \in R$, then for each $\lambda \in [0,1)$, $x' \in T(A',\lambda)$ iff $x \in T(A,\underline{e},\lambda)$.

Suppose $[x',(x,e)] \in R$, where $x \neq \Lambda$. Hence, $L(x) > 0$. By definition of R we have that $x' = g(0x, (e_0,\ldots,e_K))$, where $x = \sigma_1\ldots\sigma_K$ has length $K > 0$. So $x' = g(0,e_0)g(\sigma_1,e_1)\ldots g(\sigma_K, e_K) = e_0\sigma_1'\ldots\sigma'_K$, where $\mu \leq \sigma_j' \leq \mu(m+1)-1$ for $i = 1,\ldots K$. Thus, $p'(x') = \pi_0'M'(e_0\sigma_1'\ldots\sigma_K')\eta^{F'}$
$= \pi_0'M'(e_0)M'(\sigma_1'\ldots\sigma_K')\eta^{F'}$. But

$$M'(\sigma_1'\ldots\sigma_K') = \prod_{i=1}^{K} \begin{pmatrix} & 0 & 0 \\ & \vdots & \vdots \\ G(g^{-1}(\sigma')) & 0 & 0 \\ 0 \cdots 0 & 0 & 0 \\ 0 \cdots 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} & 0 & 0 \\ & \vdots & \vdots \\ M(x,\underline{e}) & 0 & 0 \\ 0 \cdots 0 & 0 & 1 \\ 0 \cdots 0 & 0 & 1 \end{pmatrix}$$

Hence, we obtain $p'(x') = \pi_0(e_0)M(x,\underline{e})\eta^F = p(x,\underline{e})$. So, in this case, we have $x' \in T(A',\overline{\lambda})$ iff $x \in [0,1)$.

We have verified that $\forall x \in \Sigma^*$ and $\forall \underline{e} \in E^\infty$ if $[x',(x,e)] \in R$, then for each $\lambda \in [0,1)$, $x' \in T(A',\lambda)$ iff $x \in T(A,\underline{e},\lambda)$.

Let $x' \in \Sigma'^*$ be any tape for A' such that there does not exist $x \in \Sigma^*$ and $\underline{e} \in E^\infty$ so that $[x', (x,\underline{e})] \in R$. The tape x' is not admissible. Yet $x' = g(b)$ for some $b \in (\Sigma_0 \times E)^*$.

Λ', the empty tape for A', is the image under g of the empty tape in $(\Sigma_0 \times E)^*$. Clearly, Λ' is not admissible and $p'(\Lambda') = \pi_0'\eta^{F'} = 0$. Hence, $\Lambda' \notin T(A',\lambda)$ for any $\lambda \in [0,1)$.

Let $x' \in \Sigma'^*$ be any nonempty tape which is not admissible. $x' = \sigma_0'\ldots\sigma_K'$ is the image under g of some element $b \in (\Sigma_0 \times E)^*$, where b has the form $(y, (e_0,\ldots,e_K))$ and $L(x') = L(y) = K + 1$. Note that $y \in \Sigma_0^*$. Let $y = \tau_0\tau_1\ldots\tau_K$. x' is not admissible iff $\tau_0 \neq 0$ or $\tau_i \notin \Sigma$ for some $i=1,\ldots K$. If $\tau_0 \neq 0$, then $\mu \leq \sigma_0' \leq \mu(m+1)-1$. So A' enters the absorbing state $s_{n+2} \in F'$ at the first transition. So $x' \notin T(A',\lambda)$ for any $\lambda \in [0,1)$. If $\tau_i \notin \Sigma$ for some $i = 1,\ldots,K$, then $\tau_i = 0$. Hence, $0 \leq \sigma_i' \leq \mu-1$ and A' enters state s_{n+2} at the (i+1)-th transition. Again, we have $x' \notin T(A',\lambda)$ for any $\lambda \in [0,1)$.

Thus, if for some $\lambda \in [0,1)$, $x' \in T(A',\lambda)$, then x' is admissible; that is, there exists $(x,\underline{e}) \in \Sigma^* \times E^\infty$ such that $[x', (x,\underline{e})] \in R$.

Consequently, A', as constructed, simulates A. □

The results of theorem 2 and the previous section seem to offer a paradox. We have found a set of tapes $U \subset \Sigma^*$ defined by an ADE A with finite environment set which is not a PCE. So for any $\lambda \in [0,1)$ there is no PA A_p such that $U = T(A_p,\lambda)$. But by theorem 2 there exists a PA A' over an expanded alphabet Σ' such that A' simulates A. Let $U' = \bigcup_{\underline{e}\in E^\infty} \{g(0x,\underline{e}) \mid x \in T(A,\underline{e},\lambda)\}$. $U' = T(A',\lambda)$ and, hence, is a PCE. We not only have expanded Σ to Σ' but also enriched the internal essence of the system with a structure that had been included in the environment. The problem for $U \subset \Sigma^*$ remains. $U \neq U'$ and U' has a different character within Σ'^*. If fact, U as a subset of Σ'^* is a set of inadmissible tapes and is not defined by A'. In general, from theorem 2 we obtain $T(A',\lambda) = \bigcup_{\underline{e}\in E^\infty}\{g(0x,\underline{e}) \mid x \in T(A,\underline{e},\lambda)\}$. Moreover, for any inadmissible tape x' for A' (this must include $T(A',\lambda)$) we can obtain a tape x and a class of environment sequences within which the probability of acceptance of x by A exceeds λ.

Given any arbitrary n-dimensional vector ξ, we define $|\xi| = \max |\xi_i|$. Also, for any $n \times n$ matrix Q, we define $|Q| = \max_{i,j} |q_{ij}|$.

For any ADE $A = (\Sigma,S,G,\pi_0,F,E)$ consider the pseudo-metric $d(e,e') = \max\left(|\pi_0(e)-\pi_0(e')|, \max_{\sigma\in\Sigma} |G(\sigma,e)-G(\sigma,e')|\right)$, where $e,e' \in E$. The environments e and e' have identical influences on the system iff $d(e,e') = 0$. *The environment set is irreducible iff d is a metric.*

Suppose E is reducible to a set $E\dagger$. That is, there is an onto mapping $h: E \to E\dagger \subset E$, where if $h(e) = e\dagger$, then $d(e,e\dagger) = 0$. The mapping h is unique only when $E\dagger$ is irreducible. h can be extended to $h: E^\infty \to (E\dagger)^\infty$ by component-wise application.

<u>Corollary</u> Let $A = (\Sigma,S,G,\pi_0,F,E)$ be any ADE, then $T(A,\underline{e},\lambda) = T(A,h(\underline{e}),\lambda)$ for all $\underline{e} \in E^\infty$ and $\lambda \in [0,1)$. Furthermore, if $\#(h(E)) < \infty$, then A can be simulated by a PA.

<u>Proof</u>: Since it is clear that $p(x,\underline{e}) = p(x,h(\underline{e}))$ $\forall x \in \Sigma^*$ and $\underline{e} \in E^\infty$, it must be that $T(A,\underline{e},\lambda) = T(A,h(\underline{e}),\lambda)$ for any $\lambda \in [0,1)$. For $\#(h(E)) < \infty$, by theorem 2 there exists a PA A' which simulates the ADE $A\dagger = (\Sigma,S,G,\pi_0,F,E\dagger)$. Thus, there exists a relation $R\dagger$ between Σ'^* and $(\Sigma \times E\dagger)^*$ satisfying definition 4. Define the relation R between Σ'^* and $(\Sigma \times E)^*$ by the rule:

$$[x',(x,\underline{e})] \in R \text{ iff } [x',(x,h(\underline{e}))] \in R\dagger .$$

It is easily verified that R is the relation to justify that A' simulates A. $\square$

Suppose we have an ADE $A = (\Sigma,S,G,\pi_0,F,E)$, where E and some binary operation form a semi-group. For each $\sigma \in \Sigma$, let the mapping $G(\sigma,\cdot): E \to M_n$ be a homorphism. Clearly, for each $\sigma \in \Sigma$ $G(\sigma) = \{G(\sigma,\underline{e}) \mid e \in E\}$ is a semi-group with the operation multiplication. If the semi-group of environments can be generated by a finite set, then A may be simulated by a PA.

<u>Theorem 3</u> Let $A = (\Sigma,S,G,\pi_0,F,E)$ be an ADE such that:

1. $E = S(F,\circ)$, where F is any finite set and $S(F,\circ)$ denotes the semi-group generated by F and the operation $\circ$;
2. for any $\sigma \in \Sigma$ the mapping $G(\sigma,\cdot): E \to M_n$ is a homomorphism;
3. $\pi_0(\cdot)$ is a constant.

Then there exists a PA A' which simulates A.

<u>Proof</u>: Let $x \in \Sigma^*$ and $\underline{e} = (e_0,e_1,\ldots) \in E^\infty$. Suppose $L(x) = K$. Each e_i, $i=1,\ldots,K$, has a finite decomposition $e_i = f_{i1} \circ \ldots \circ f_{ir_i}$.

Thus, $M(x,\underline{e}) = \prod_{i=1}^{K} G(\sigma_i,e_i) = \prod_{i=1}^{K}\prod_{j=1}^{r_2} G(\sigma_i f_{ij})$.

Let $x\dagger = \sigma_1^{r_1} \cdots \sigma_K^{r_K}$ and let $\underline{f}$ be any environment sequence with initial terms $\bar{f}, f_{11},\ldots f_{1r_1}, f_{21},\ldots f_{Kr_K}$, where $f \in F$ is arbitrary. Then $p(x,\underline{e}) = p(x\dagger,\underline{f})$. Hence, for all $\lambda \in [0,1)$, $x \in T(A,\underline{e},\lambda)$ iff $x\dagger \in T(A,\underline{f},\lambda)$. Clearly, for any decomposition of the environments the corresponding derived input tape will have identical acceptance properties.

The ADE $A\dagger = (\Sigma,S,G,\pi_0,F,F)$ has finite environment set and so we have a relation $R\dagger$ between Σ'^* and $\Sigma^* \times F^\infty$ whereby a PA A' simulates $A\dagger$. Define the relation R between Σ'^* and $\Sigma^* \times E^\infty$ as follows:

$[x',(x,\underline{e})] \in R$ iff $[x',(x\dagger,\underline{f})] \in R\dagger$ for any decomposition of $\underline{e}$ and its corresponding derived input tape $x\dagger$.

It is easily verified that R is the relation to justify that A' simulates A. $\square$

Continuous-time probabilistic automata as introduced by Knast [4] are seen to be a special case of the ADE if we consider the time interval for an input to a continuous-time PA to be the environment which attends to that input. Moreover, the environment set $E = [0,\infty)$ and the operation addition form a semi-group and the transition function is a homomorphism for each $\sigma \in \Sigma$. However, there is no finite set which generates $[0,\infty)$ by the operation of addition. Knast showed, however, that under certain conditions a continuous-time PA can be approximated by an ADE $A_h = (\Sigma,S,G,\pi_0,F,E_h)$, where π_0 is constant and $E_h = \{h,2h,\ldots\}$ for $h > 0$. *Clearly, $E_h = S(h)$ by the operation* addition and for any $\sigma \in \Sigma$ and $e \in E_h$, $G(\sigma,e) = (G(\sigma,h))^{e/h}$. Hence, the approximation may be simulated by a PA.

4. <u>Automata in Random Environments</u>

We shall now assume that the realization of the environment is governed by some probabilistic structure. Hence, the sequence of environment configurations is a stochastic process. This formulation is useful when we are only able to make certain probabilistic assumptions about the occurrence of any environment configuration or when the environment configuration can only be measured by statistical techniques.

Let $A = (\Sigma,S,G,\pi_0,F,E)$ be an ADE. Let (Ω,B,P) be a probability space. For each $j \in \{0,1,\ldots\}=J$, let z_j be a measurable function from (Ω,B) to (E,B'), where B' is a σ-field of subsets of E. The family of random variables $Z = \{z_j \mid j \in J\}$ is an environmental stochastic process (ESP) and z_j denotes the random variable for the configuration of the environment attending to j^{th} input symbol. Let us assume that the mappings $\pi_0: E \to V_n$ and $G(\sigma,\cdot): E \to M_n$ $\forall\sigma\in\Sigma$ are measurable; hence, the compositions $\pi_0 \circ z_0 = \pi_0(z_0)$ and $G(\sigma,\cdot)\circ z_j = G(\sigma,z_j), j\geq 1$ $\forall\sigma\in\Sigma$, are measurable; they are random stochastic vectors and random stochastic matrices respectively.

Definition 5 An automaton in random environments (ARE) is a system (A,Z), where A is an ADE and Z is an ESP.

Define $\underline{z}: \Omega \to E^{\infty}$ by $\underline{z}(\omega) = (z_0(\omega), z_1(\omega), \ldots)$. The mapping $\underline{z}$ is a measurable map from (Ω, B) to $(E^{\infty}, B'^{\infty})$, where B'^{∞} is the smallest σ-field generated by the measurable cylinders. For $x \in \Sigma^*$ such that $L(x) = K$ we define the random matrix transition function

$$M(x,\underline{z}(\omega)) = \prod_{i=1}^{K} G(\sigma_i, z_i(\omega)).$$

Similarly, we have $\pi(x,\underline{z}(\omega)) = \pi_0(z_0(\omega))M(x,\underline{z}(\omega))$ and $p(x,\underline{z}(\omega)) = \pi(x,\underline{z}(\omega))\eta^F$. Furthermore, $T(A,\underline{z},\lambda)$ for any $\lambda \in [0,1)$ is a set function of $\omega \in \Omega$ taking values in the set of all subsets Σ^*. As can be seen by theorem 1 there is no strict generalization in allowing the set of acceptance states to be related to the terminal environment.

Suppose $\underline{z}^{(1)}, \underline{z}^{(2)}, \ldots$ are independent identically distributed (IID) sequences of random environments. Thus, for any fixed, but arbitrary $x \in \Sigma^*$, $\{p(x,\underline{z}^{(i)})\}_1^{\infty}$ is a sequence of uniformly bounded IID random variables. By the strong law of large numbers we obtain

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} p(x,\underline{z}^{(i)}) = E\, p(x,\underline{z}) \quad \text{a.s.}$$

Definition 6 Let (A,Z) be an ARE and λ a real number, $\lambda \in [0,1)$. The set of tapes $ET(A,\underline{z},\lambda) = \{x \mid x \in \Sigma^*, \lambda < E\, p(x,\underline{z})\}$ is called the expected set of tapes defined by (A,Z) with cut-point λ.

Let us define the random variable $\psi(\theta) = \theta$ on the probability space $([0,1], B[0,1], \mu)$, where $B[0,1]$ is the relative Borel field and μ is Lebesgue measure. For any fixed, but arbitrary $x \in \Sigma^*$ and $\underline{e} \in E^{\infty}$, we define the random variable

$$I(x,\underline{e},\psi(\theta)) = \begin{cases} 1 & \text{if } \psi(\theta) \le p(x,\underline{e}) \\ 0 & \text{otherwise} \end{cases}.$$

$I(x,\underline{e},\psi)$ is measurable and $\mu\{\theta \mid I(x,\underline{e},\psi(\theta)) = 1\} = p(x,\underline{e})$. Thus, $I(x,\underline{e},\psi)$, which has the same relevant probability structure as an indicator that A within environment sequence $\underline{e}$ is in a state in F after input x, has expectation $E\, I(x,\underline{e},\psi) = p(x,\underline{e})$. So for an ADE the relative frequency of acceptance estimates the acceptance probability.

Now consider the composition mapping $I(x,\underline{z},\psi)$ defined on the product space $[0,1] \times \Omega$, where

$$I(x,\underline{z}(\omega),\psi(\theta)) = \begin{cases} 1 & \text{if } \psi(\theta) \le p(x,\underline{z}(\omega)) \\ 0 & \text{otherwise} \end{cases}$$

We now obtain $E\, I(x,\underline{z},\psi)$ with respect to the product measure

$$EI(x,\underline{z},\psi) = \int_{[0,1]\times\Omega} I(x,\underline{z}(\omega),\psi(\theta))\, d(\mu\times P).$$

By Fubini's theorem

$$EI(x,\underline{z},\psi) = \int_{\Omega}\int_{[0,1]} I(x,\underline{z}(\omega),\psi(\theta))\, d\mu(\theta)\, dP(\omega)$$

$$= \int_{\Omega} p(x,\underline{z}(\omega))\, dP(\omega)$$

$$= Ep(x,\underline{z}),$$

where the last expectation is taken with respect to (Ω, B, P). Supposing $\{(\underline{z}^{(i)}, \psi^{(i)})\}_1^{\infty}$ are independent copies of $(\underline{z},\psi)$, then $\{I(x,\underline{z}^{(i)},\psi^{(i)})\}_1^{\infty}$ is a sequence of random variables IID according to the product measure $\mu \times P$. Thus, the relative frequency of acceptance of x by A within IID environment sequences is a strong consistent unbiased estimator of $Ep(x,\underline{z})$. Furthermore, when there is no additional information as to the realizations of the environment processes, the marginal distribution of the frequency of acceptance of x based on N repetitions is a binomial random variable with parameters N and $Ep(x,\underline{z})$. Accordingly, we are able to statistically decide whether $x \in ET(A,\underline{z},\lambda)$ for all $\lambda \ne Ep(x,\underline{z})$.

Let $(\bar{A},Z)$ be an ARE, where $\bar{A} = (\Sigma, S, G, \pi_0, F, E)$ and Z is a ν state homogeneous Markov chain with initial distribution α and transition matrix Q. Consider the PA B over Σ with $n\nu$ states and

$$\pi_0 = (\pi_0(1)\alpha_1, \ldots, \pi_0(\nu)\alpha_\nu)$$

$$M(\sigma) = \begin{bmatrix} G(\sigma,1)q_{11} & \cdots & G(\sigma,\nu)q_{1\nu} \\ \vdots & & \vdots \\ G(\sigma,1)q_{\nu 1} & \cdots & G(\sigma,\nu)q_{\nu\nu} \end{bmatrix}$$

$$\eta^{F^B} = \begin{bmatrix} \eta^F \\ \vdots \\ \eta^F \end{bmatrix}.$$

It is left to the reader to verify that $Ep(x,\underline{z}) = p^B(x)\ \forall x \in \Sigma^*$. Hence, if (A,Z) is an ARE with Z as a finite homogeneous Markov chain, then $ET(A,\underline{z},\lambda)$ is a PCE for all $\lambda \in [0,1)$. In other words, if a PA over a singleton alphabet generates the random environment sequence, then the expected set of tapes defined for any cut-point is a PCE. This generalizes the result of Turakainen [9] that if the generator is a finite automaton over a singleton alphabet, then the set defined for any cut-point is a PCE.

If the environment random variables within the sequence are independent, we find

$$Ep(x,\underline{z}) = (E\pi_0(z_0))\left(\prod_{i=1}^{K} EG(\sigma_i, z_i)\right)\eta^F$$

for any $x \in \Sigma^*$ and such a result defines an ADE. Furthermore, if there are only a finite number of distinct distribution functions corresponding to the environmental random variables, then the ADE can be simulated by a PA A'. In the particular case when $\underline{z}$ is a sequence of IID random variables, we obtain $T(A',\lambda) = ET(A,\underline{z},\lambda)\ \forall\lambda \in [0,1)$.

5. A Canonical Representation for Automata in Random Environments

In section 4 we encountered the dual nature of the randomness involved in automata in random environments. The environment sequence is random and for each value the environment sequence assumes, the probabilistic state transitions are defined. Hence, the state of the system after an input tape within a random environment sequence is a function defined on a product probability space. We shall show that for any ARE (A,Z) there is an ARE (A_D,Y) over the same alphabet whose transition probabilities are equivalent in the mean, but the state of (A_D,Y) obtained after an input within a random environment sequence is only related to the value the random environment sequence obtains. The essence of the machine within A_D is deterministic. To accomplish this we need that the initial distribution function and the basic matrix transition take on values which, in addition to being stochastic vectors and matrices, respectively, have components that are either 0 or 1. Thus, the only randomness in the system (A_D, Y) is the randomness of the environment sequence. We shall also construct (A_D, Y) to have a finite environment set.

Let $D(M_n)$ denote the subset of M_n such that if $M \in D(M_n)$, then $m_{ij} = 0$ or 1. Also, let $D(V_n)$ be the subset of V_n such that if $v \in D(V_n)$, then $v_i = 0$ or 1.

<u>Theorem 4</u> Let (A,Z) be an ARE, where $A = (\Sigma, S, G, \pi_0, F, E)$ is an ADE and $Z = \{z_j \mid j \in J\}$ is an ESP defined on (Ω, B, P) taking values in E. There exists an ARE (A_D, Y), where $A_D = (\Sigma, S, H, \rho_0, F, E')$ is an ADE and $Y = \{y_j \mid j \in J\}$ is an ESP defined on some (Ω', B', P') taking values in E' such that (A_D, Y) satisfies the following properties:

1. $H: \Sigma \times \Omega' \to D(M_n)$
2. $\rho_0: \Omega' \to D(V_n)$
3. $\#(E') < \infty$
4. $E\pi(x,\underline{z}) = E\rho(x,\underline{y})\ \forall x \in \Sigma^*$, where $\underline{y} = (y_0, y_1, \ldots)$
5. $ET(A,\underline{z},\lambda) = ET(A,\underline{y},\lambda)\ \forall \lambda \in [0,1)$.

Here we are using π and ρ to denote the state distributions of (A,Z) and (A_D, Y), respectively. Recall $n = \#(S)$. Notice that knowledge of the input x and the environment sequence $\underline{y}(\omega')$ determines the state of A_D obtained.

<u>Proof</u>: For any $\sigma \in \Sigma$ and $e \in E$ we can decompose $G(\sigma,e)$ into a convex linear combination of $p = n^2 - n + 1$ elements of $D(M_n)$. Thus,

$$\sum_{i=1}^{p} \alpha_{i\sigma}(e) B_{i\sigma}(e) = G(\sigma,e)$$

for each $\sigma \in \Sigma$ and $e \in E$, where $B_{i\sigma}(e) \in D(M_n)$, $\alpha_{i\sigma}(e) \ge 0$ and

$\sum_{i=1}^{p} \alpha_{i\sigma}(e) = 1$. So for each $e \in E$ we have generated the following arrays of numbers and matrices:

$$\begin{Bmatrix} \alpha_{1\sigma_1}(e) & \cdots & \alpha_{p\sigma_1}(e) \\ \vdots & & \vdots \\ \alpha_{1\sigma_m}(e) & \cdots & \alpha_{p\sigma_m}(e) \end{Bmatrix} \qquad \begin{Bmatrix} B_{1\sigma_1}(e) & \cdots & B_{p\sigma_1}(e) \\ \vdots & & \vdots \\ B_{1\sigma_m}(e) & \cdots & B_{p\sigma_m}(e) \end{Bmatrix}$$

where $m = \#(\Sigma)$. For any $i_1, \ldots, i_m = 1, \ldots, p$ let $a_e(i_1, \ldots, i_m) = \prod_{j=1}^{m} \alpha_{i_j \sigma_j}(e)$. Now consider $A(e) = \{a_e(i_1, \ldots, i_m) \mid i_1, \ldots, i_m = 1, \ldots, p\}^L$ as a list of numbers. A <u>list</u> differs from a set in that redundancy is preserved. Now for any $e \in E$, clearly $a_e(i_1 \ldots, i_m) \ge 0$ and $\sum_{i_1} \cdots \sum_{i_m} a_e(i_1, \ldots i_m) = 1$.

For each $\sigma_K \in \Sigma$, we partition $A(e)$ into p sublists $A_{1\sigma_K}(e), \ldots, A_{p\sigma_K}(e)$, where

$$A_{i\sigma_K}(e) = \{a_e(i_1, \ldots, i_{K-1}, i, i_{K+1}, \ldots i_m) \mid i_1, \ldots, i_{K-1}, i_{K+1}, \ldots, i_m = 1 \ldots, p\}^L .$$

We obtain the sum of the elements of the list $A_{i\sigma_K}(e)$,

$$\sum_{i_1} \cdots \sum_{i_{K-1}} \sum_{i_{K+1}} \cdots \sum_{i_m} a_e(i_1 \ldots, i_{K-1}, i, i_{K+1}, \ldots, i_m)$$
$$= \alpha_{i\sigma_K}(e) \sum_{i_1} \cdots \sum_{i_{K-1}} \sum_{i_{K+1}} \cdots \sum_{i_m} \prod_{j \ne K} \alpha_{i_j \sigma_j}(e)$$
$$= \alpha_{i\sigma_K}(e) .$$

These are p^m elements in the list $A(e)$. We now construct the list $A'(e) = \{a'_j(e) \mid i=1, \ldots, p^m\}$

For $i = 1 + \sum_{j=1}^{m} (i_j - 1)p^{j-1}$, set $a'_i(e) = a_e(i_1, \ldots, i_m)$. This uniquely describes each element of $A'(e)$ and $A'(e) = A(e)$. For each $\sigma_K \in \Sigma$, we partition $A'(e)$ into p sublists $A'_{1\sigma_K}(e), \ldots, A'_{p\sigma_K}(e)$, where $a'_i(e) \in A'_{j\sigma_K}(e)$ iff $a_e(i_1, \ldots, i_{K-1}, j, i_{K+1}, \ldots, i_m) \in A_{j\sigma_K}(e)$ and

$i = 1 + (j-1)p^{K-1} + \sum_{r \ne K} (i_r - 1)p^{r-1}$. Actually, $A'_{j\sigma_K}(e) = A_{j\sigma_K}(e)$. Thus, we have reordered $A(e)$ and maintained the properties of the partitioning.

We now construct the list $B'(e) = \{B_{i\sigma}(e) \mid \sigma \in \Sigma, i=1, \ldots, p^m\}^L$ by the rule:

1. Set i=1 and K=1
2. $a'_i(e)$ must belong to one of the sublists $A'_{1\sigma_K}(e), \ldots, A'_{p\sigma_K}(e)$, say $A'_{j\sigma_K}$. Set $B'_{i\sigma_K}(e) = B_{j\sigma_K}$.
3. Delete one occurrence of $a'_i(e)$ from the list $A'_{j\sigma_K}(e)$, but let the amended list retain the name $A'_{j\sigma_K}(e)$.
4. Increase the value of i by 1.
5. If $i \le p^m$, go to step 2. If not, go to step 6.
6. If $K < m$, then set i=1, increase the value of K by 1, and go to step 2. If not, stop.

This algorithm must necessarily terminate in mp^m iterations with the sublists empty.

We obtain for each $e \in E$.

1. $a'_i(e) \ge 0$. $\forall i = 1, \ldots, p^m$, $\sum_{i=1}^{p^m} a'_i(e) = 1$
2. For each $\sigma \in \Sigma$, $A'(e)$ can be partitioned into p sublists $A'_{1\sigma}(e), \ldots, A'_{p\sigma}(e)$ such that $\sum_{a'_i(e) \in A'_{j\sigma}} a'_i(e) = \alpha_{j\sigma}(e)$.
3. $B'_{i\sigma}(e) \in D(M_n)$ $\forall i = 1, \ldots, p^m$ $\forall \sigma \in \Sigma$ and $B_{i\sigma}(e) = B_{j\sigma}(e)$ i is an index of an element in $A'_{j\sigma}(e)$.

Also for any $\sigma \in \Sigma$ and $e \in E$.

$$\sum_{i=1}^{p^m} a'_i(e) B'_{i\sigma}(e) = \sum_{j=1}^{p} \sum_{a'_i(e) \in A'_{j\sigma}(e)} B_{j\sigma} a'_i(e)$$
$$= \sum_{j=1}^{p} B_{j\sigma}(e) \sum_{a'_i \in A'_{j\sigma}(e)} a'_i(e)$$
$$= \sum_{j=1}^{p} B_{j\sigma}(e) \alpha_{j\sigma}(e) = G(\sigma,e).$$

Hence for each $e \in E$ we have the following arrangement:

$$\begin{matrix} a'_1(e) & B'_{1\sigma_1}(e) & \cdots & B'_{1\sigma_m}(e) \\ \vdots & & & \\ a'_{p^m}(e) & B'_{p^m\sigma_1}(e) & \cdots & B'_{p^m\sigma_m}(e) \end{matrix}$$

Since $D_{n,m} = D(M_n) \times \cdots \times D(M_n)$, the Cartesian product of $D(M_n)$ with itself m times, is a finite set with n^{nm} distinct elements, we can order it in a systematic way. Let D_K be the K^{th} element of $D_{n,m}$ by our ordering. For each $e \in E$ and $K = 1, \ldots, n^{nm}$ define $I_K(e) = \{i \mid (B'_{i\sigma_1}(e), \ldots, B_{i\sigma_m}(e)) = D_K\}$. Note that $I_1(e), \ldots, I_{n^{nm}}(e)$ is a partition of the set $\{1, 2, \ldots, p^m\}$. Thus, we set

$$\beta_K(e) = \begin{cases} \sum_{i \in I_K(e)} a'_i(e) & \text{if } I_K(e) \ne \emptyset \\ 0 & \text{if } I_K(e) = \emptyset \end{cases}$$

Hence, $\beta_K(e) \geq 0$ for $K=1,\ldots,n^{nm}$ and for all $e \in E$. Also, for each $e \in E$,

$$\sum_{K=1}^{n^{nm}} \beta_K(e) = \sum_{K=1}^{n^{nm}} \sum_{i \in I_K(e)} a'_i(e) = \sum_{i=1}^{p^m} a'_i(e) = 1 \quad \text{and}$$

$$\sum_{K=1}^{n^{nm}} \beta_K(e) D_K = \sum_{K=1}^{n^{nm}} \sum_{i \in I_K(e)} a'_i(e)(B'_{i\sigma_1}(e),\ldots,B'_{i\sigma_m}(e))$$

$$= \sum_{i=1}^{p^m} a'_i(e)(B'_{i\sigma_1}(e),\ldots,B'_{i\sigma_m}(e))$$

$$= (G(\sigma_1,e),\ldots,G(\sigma_m,e)).$$

Now we are ready to define (A_D,Y). Retain Σ, S, and F from (A,Z). Let $E' = \{1,\ldots,n^{nm}\}$. Define $H(\sigma_i,K)$ to be the i^{th} matrix in D_K. For $K=1,\ldots,n$, let $\rho_o(k)$ be an n-dimensional row vector of all zeroes except for a 1 in the k^{th} component. For $n+1 \leq K \leq n^{nm}$ let $\rho_o(K)$ be any arbitrary element of $D(\gamma_n)$.

Define $\gamma_K(e) = \begin{cases} \pi_o^{(K)}(e) & \text{for } K=1,\ldots,n \\ 0 & \text{for } K=n+1,\ldots,n^{nm}, \end{cases}$

where $\pi^{(K)}(e)$ is the k^{th} component of $\pi_o(e)$. By Kolmogorov's Existence Theorem we define the process $Y = \{y_j \mid j \in J\}$ by the finite dimensional distributions

$$P\{y_{j_1}=K_1,\ldots,y_{j_r}=K_r\} = \int_\Omega \beta_{K_1}(z_{j_1})\ldots\beta_{K_r}(z_{j_r})dP,$$

where $j_1,\ldots,j_r \in J$ are all distinct. *If same* $j_i=0$, replace $\beta_{K_i}(z_{j_i})$ by $\gamma_{K_i}(z_o)$.

Let $x \in \Sigma^*$ be an input tape of arbitrary length, say $x = \tau_1\ldots\tau_r$, $\tau_i \in \Sigma$ $i=1,\ldots,r$.

By construction

$$E\rho(x,y) = \sum_{K_o}\cdots\sum_{K_r}\rho_o(K_o)H(\tau_1,K_1)\cdots H(\tau_r,K_r) \; P(y_o=K_o,\ldots,y_r=K_r)$$

$$= \sum_{K_o}\cdots\sum_{K_r}\rho_o(K_o)H(\tau_1,K_1)\ldots H(\tau_r,K_r)\int_\Omega \gamma_{K_o}(z_o)\beta_{K_1}(z_1)\ldots\beta_{K_r}(z_r)dP$$

$$= \int_\Omega \sum_{K_o}\cdots\sum_{K_r}\rho_o(K_o)H(\tau_1,K_1)\cdots H(\tau_r,K_r)\,\gamma_{K_o}(z_o)\beta_{K_1}(z_1)\ldots\beta_{K_r}(z_r)dP$$

$$= \int_\Omega \Big(\sum_{K_o}\rho_o(K_o)\gamma_{K_o}(z_o)\Big)\Big(\sum_{K_1}\beta_{K_1}(z_1)H(\tau_1,K_1)\Big)\ldots\Big(\sum_{K_r}\beta_{K_r}(z_r)H(\tau_r,K_r)\Big)dP$$

$$= \int_\Omega \Pi_o(z_o)G(\tau_1,z_1)\ldots G(\tau_r,z_r)dP$$

$$= \int_\Omega \Pi(x,\underline{z})dP = E\,\Pi(x,\underline{z})$$

where all sums are from 1 to n^{nm}.

Clearly, it follows that $ET(A,\underline{Z},\lambda)=ET(A_D,\underline{Y},\lambda)$ $\forall\lambda \in [0,1)$ □

As a special case we see that a PA can be viewed as a finite automaton subject to the influences of an IID random environment sequence.

REFERENCES

[1] Gould, J. (1975a), Automata in environments: II. stability results, manuscript.

[2] Gould, J. (1975b), Automata in environments: III. random environments, manuscript.

[3] Gould, J. and Wegman, E.J. (1975), Automata in environments: I. basic concepts, manuscript

[4] Knast, R. (1969), Continuous-time probabilistic automata, *Inform. Contr.*, 15, 333-352.

[5] Paz, A. (1966), Some aspects of probabilistic automata, *Inform. Contr.*, 9, 26-60.

[6] Paz, A. (1970), *Probabilistic Automata*, Academic Press, New York & London.

[7] Rabin, M.O. (1963), Probabilistic automata, *Inform. Contr.*, 6, 230-245.

[8] Tsetlin, M.L. (1961), On the behavior of finite automata in random media, *Automat. Remote Control*, 22, 1210-1219.

[9] Turakainen, P. (1969), On time-variant probabilistic automata with monitors, *Ann. Univ. Turku. Ser.A I.*, 130, 3-11.

[10] Varshavskii, V.I. and Varontsova, I.P. (1963), On the behavior of stochastic automata with a variable structure, *Automat. Remote Control*, 24, 327-333.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 18 AFOSR-TR- 78-0730	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): 6 TOWARD A THEORY OF PROBABILISTIC AUTOMATA WITH ENVIRONMENTS,	5. TYPE OF REPORT & PERIOD COVERED: Interim
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): 10 Jerren Gould ~~and~~ Edward J. Wegman	8. CONTRACT OR GRANT NUMBER(s): 15 N00014-67-A-0321-0006; AFOSR-75-2840
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of North Carolina, Department of Statistics, Chapel Hill, NC 27514	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 16 2304/A5 17 A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332	12. REPORT DATE: 11 1978 12 9p.
	13. NUMBER OF PAGES: 8
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The viewpoint is taken that every probabilistic automaton is situated within a sequence of environments which affects the initial state distribution and the transition function. There exists a probabilistic automaton within a deterministic environment sequence (ADE) which defines an event which is not a PCE, yet in certain nontrivial cases the behavior of an ADE can be simulated by a probabilistic automaton. For probabilistic automata within random environment sequences there is a mean equivalent canonical representation which eliminates the randomness due to probabilistic transition.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

182 850 JOB