An Automaton in the Nonstationary Random Environment

HIDEKAZU TSUJI, MASAHARU MIZUMOTO, JUNICHI TOYODA,
AND KOHKICHI TANAKA
Faculty of Engineering Science, Osaka University, Toyonaka, Osaka, Japan 560

Communicated by John M. Richardson

ABSTRACT

The purpose of this paper is to consider the behavior of a finite automaton in a nonstationary random environment. The behavior of finite deterministic automata in stationary random environments was considered by Tsetlin. In this paper, the probabilistic automaton is introduced as a random environment in order to generalize the stationary random environment. The interaction between the probabilistic automaton and the two-state deterministic automaton is considered in the case where the probabilistic automaton has two inputs and two states and, besides, is completely isolated by the 0th approximation. And the limiting state probability distribution of this finite automaton is also obtained. Moreover, it is shown that, if the probabilistic automaton is completely isolated by the $(0, k)$th approximation and satisfies some conditions, then the finite automaton can behave expediently against the probabilistic automaton.

1. INTRODUCTION

Automata which behave intelligently in random environments were formulated, at first, by Tsetlin [1]. He considered the behavior of a finite deterministic automaton in a random environment and showed that if a linear strategy is applied to the state transition of the automaton, then, under certain conditions the automaton shows asymptotically optimal behavior.

Based on Tsetlin's model, many studies have been done. For example, Varshavskii and Vorentsova [2] extended the Tsetlin's model and used a probabilistic automaton with a variable structure. Krylov [3] modified this Tsetlin's model by using a pseudoprobabilistic automaton and showed that asymptotic optimality is preserved. Fu and Li [4] also proposed the different type of the deterministic automaton which shows asymptotically optimal behavior. However, a nonstationary random environment whose penalty probability varies, is hardly considered. Chandrasekaran and Shen [5] only investigated an optimal behavior of a probabilistic automaton in a periodic environment.

In this paper, probabilistic automata are employed as nonstationary random environments. The behavior of a finite deterministic automaton with two states in a nonstationary random environment is considered. At first, when a probabilistic automaton is completely isolated by the 0th approximation in the sense that Yasui and Yajima [6] defined, the limiting state probability distribution of this finite automaton in this environment (or the probabilistic automaton) is obtained. If the probabilistic automaton satisfies more conditions than stated above, the finite automaton behaves expediently against the probabilistic automaton in the sense of the Tsetlin's model.

2. THE INTERACTION BETWEEN A FINITE AUTOMATON AND A PROBABILISTIC AUTOMATON

We shall briefly review the interaction between a finite automaton and a random environment introduced by Tsetlin [1].

He considered the function of finite automata responding to their actions within an environment in a random fashion as shown in Fig. 1. Suppose the

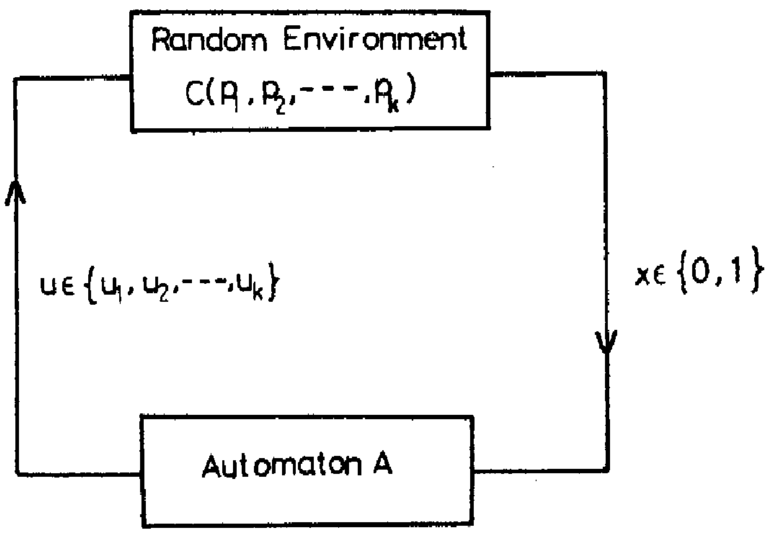

Fig. 1. Tsetlin's model.

random environment C is characterized by $C(p_1, p_2, \ldots, p_k)$ where $0 \leqslant p_i \leqslant 1$ for $i = 1, 2, \ldots, k$. The input x of an automaton A can take two values, that is, $x = 1$ (penalty) and $x = 0$ (nonpenalty), and the output u of A is the action taken by A. In one experiment, if the automaton A takes action u_i, $i = 1, 2, \ldots, k$, then the next input x of A is as follows

$$x = \begin{cases} 1, & \text{with probability } p_i, \\ 0, & \text{with probability } 1 - p_i. \end{cases}$$

The automaton operating in the random environment aims to minimize the expectation of penalty, or the limiting penalty ratio $M(A, C)$ given as

$$\lim_{n\to\infty} \frac{\text{number of penalties in first } n \text{ trials}}{n}$$

Then, after Tsetlin, a linear strategy was used in the state transition of the automaton A, and it was shown that if the random environment is stationary, that is, each p_i is unknown but fixed, the automaton A shows *expedient behavior*; that is, the expectation of penalty imposed on the automaton A is less than that in case where each u_i is put into the environment with the same probability $1/k$, thus the limiting penalty ratio may be given as

$$M(A,C) < \frac{1}{k}(p_1 + p_2 + \cdots + p_k).$$

Furthermore, it was shown that under such a condition as

$$\min(p_1, p_2, \ldots, p_k) \leqslant \tfrac{1}{2}$$

the automaton A shows asymptotically optimality; that is, $M(A, C)$ decreases and approaches to $\min(p_1, p_2, \ldots, p_k)$ as the number of the states of A is increasing. Let $\min(p_1, p_2, \ldots, p_k)$ be p_i, then u_i is the optimal output of the automaton A.

As described above, the output probability distribution of random environments in Tsetlin's model depends only on the input at that instant. However, generally speaking, the output of an environment depends not only on the input at that instant but also on preceding inputs. In other words, environments have internal states memorizing past behaviors. Therefore, in this paper, probabilistic automata are employed as random environments in order to extend the random environment defined by Tsetlin. In this case, it is supposed for simplicity that probabilistic automata (or random environments) have two inputs, two outputs and two states defined as follows.

Definition 1. *Environment C* is a probabilistic automaton given as follows

$$C = (U, T, X, \bar{\pi}_0, \{P(u) | u \in U\}, g)$$

where

(1) $U = \{u_1, u_2\}$, set of inputs;
(2) $T = \{t_1, t_2\}$, set of internal states;
(3) $X = \{0, 1\}$, set of outputs;
(4) $\bar{\pi}_0 = (1, 0)$, 2-dimensional probabilistic row vector (initial state probability distribution);

(5) $P(u)$, probability transition matrix of order 2 such that

$$P(u_1) = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}, \quad P(u_2) = \begin{pmatrix} 1-c & c \\ d & 1-d \end{pmatrix};$$

(6) $g(\cdot)$: $g(t_1) = 0, g(t_2) = 1$, output function.

Definition 2. *The transitional probability matrix for an input string* $u^* = u^{(1)}u^{(2)} \cdots u^{(m)}$ is defined as follows

$$P(u^*) = P(u^{(1)})P(u^{(2)}) \cdots P(u^{(m)}).$$

The state probability row vector (state probability distribution) at the instant of $m + 1$, is given by

$$\bar{\pi}(m+1) = \bar{\pi}_0 P(u^*).$$

Remark. In a special case where

$$a + b = 1, \quad c + d = 1,$$

the environment C defined in Def. 1 becomes the random environment in the Tsetlin's model.

We shall next discuss the interaction between a two-state deterministic automaton and a random environment C defined in Def. 1.

At first, we shall explain the probabilistic automaton which is completely isolated by the 0th approximation. Suppose that the state t_2 of the probabilistic automaton C corresponds to the output 1 (penalty) and the initial state is the state t_1. Then the penalty probability may be given by the (1, 2)-element of the transition probability matrix caused by an input sequence. At time t, the penalty probability caused by an input $u_i \in U (i = 1, 2)$ is given as $p_i(t)$.

Definition 3. The probabilistic automaton is said to be *completely isolated by the 0th approximation,* if the following conditions are satisfied

$$|p_1(t) - \alpha_0| < \epsilon, \quad |p_2(t) - \beta_0| < \epsilon, \quad |\alpha_0 - \beta_0| = 2\epsilon, \tag{2.1}$$

where

$$\alpha_0 = \frac{a}{a+b}, \quad \beta_0 = \frac{c}{c+d}. \tag{2.2}$$

This definition is shown in Fig. 2.

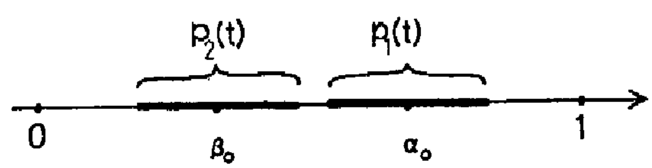

Fig. 2. Penalty probability of a completely isolated probabilistic automaton by the 0th approximation.

THEOREM 1. *(Yasui and Yajima [6]) A necessary and sufficient condition for the probabilistic automaton C being completely isolated by the 0th approximation is*

$$\|H\| \neq 0, \text{ and } h = \frac{\delta}{1-\delta}\frac{\max\{\|H\|, \|A_2\|, \|B_2\|\}}{\|H\|} \leqslant 1$$

where

$$\|H\| = \frac{|bc - ad|}{(a+b)(c+d)}, \quad \|A_2\| = \frac{a}{a+b}, \quad \|B_2\| = \frac{c}{c+d},$$

$$\delta = \max\{|1 - a - b|, \ |1 - c - d|\}.$$

When a probabilistic automaton C is completely isolated by the 0th approximation, the combination between a finite automaton and a probabilistic automaton may be regarded as the one between a finite auotmaton and a nonstationary random environment $C(p_1(t), p_2(t))$ which satisfies Eqs. (2.1) and (2.2) as shown in Fig. 3.

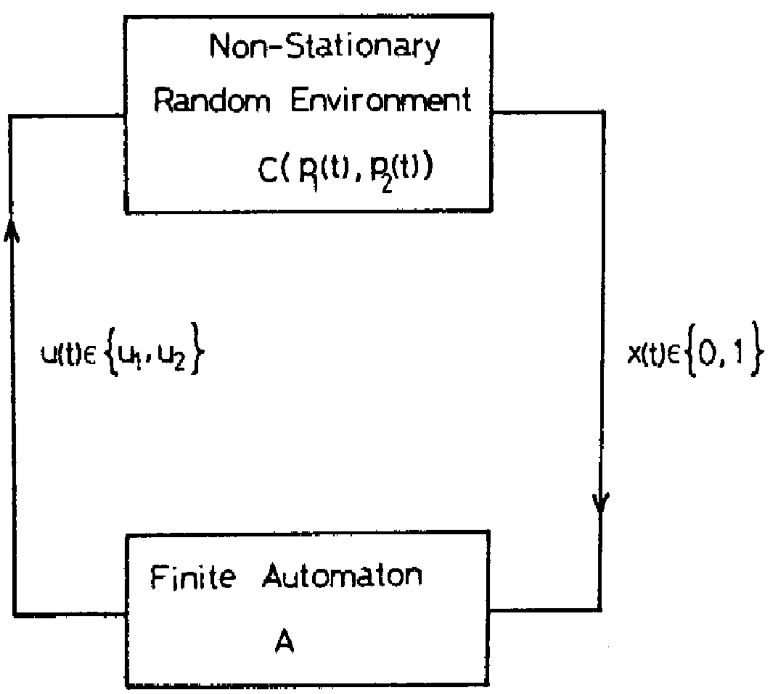

Fig. 3. Interaction between a nonstationary random environment and a finite automaton.

For state transitions of the automaton A, a linear strategy is used as shown in Fig. 4. By using Tsetlin's method of the analysis, the state transition of the automaton A is determined by the following probability matrix at time t,

$$P(t) = \begin{pmatrix} q_1(t) & p_1(t) \\ p_2(t) & q_2(t) \end{pmatrix}, \qquad p_i(t) + q_i(t) = 1, \text{for } i = 1, 2.$$

Fig. 4. State transition of a finite automaton A.

As mentioned above, if a random environment is completely isolated by the 0th approximation, the behavior of the finite automaton A can be expressed by using the nonstationary state transition probability matrix $P(t)$.

3. THE LIMITING PROBABILITY DISTRIBUTION OF THE FINITE AUTOMATON

Based on the result of the previous chapter, nonstationary automata with one input are defined and the limiting state probability distribution is considered in the present chapter.

Definition 4. *A nonstationary automaton A with one input* is a system

$$A = (S, P, \pi)$$

where

(1) $S = \{s_1, s_2\}$; set of internal states
(2) $P(t)$; state transition probability matrix at time t such that

$$P = P(t) = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \tag{3.1}$$

where

$$\alpha = \alpha(t), \quad \beta = \beta(t), \quad 0 \leqslant \alpha, \quad \beta \leqslant 1$$

$$|\alpha - \alpha_0| < \epsilon, \quad |\beta - \beta_0| < \epsilon$$

$$\epsilon = |\alpha_0 - \beta_0|/2 \tag{3.2}$$

(3) $\pi(t)$; state row vector at time t such that

$$\pi = \pi(t) = (r, 1-r). \tag{3.3}$$

where

$$r = r(t) \qquad 0 \leqslant r(t) \leqslant 1$$

$$\pi(t+1) = \pi(t)P(t). \tag{3.4}$$

The definition of the nonstationary automaton is over. Let us investigate the behavior of this automaton. At first, the case where α and β are fixed is considered.

The limiting state probability vector is given by

$$\lim_{n\to\infty} \pi(n) = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right). \tag{3.5}$$

Let $f(\alpha, \beta)$ be

$$f(\alpha, \beta) = \frac{\beta}{\alpha+\beta}, \tag{3.6}$$

where α and β satisfy Eq. (3.2), then we have

$$\max_{\alpha,\beta} f(\alpha, \beta) = f(\alpha_0 - \epsilon,\ \beta_0 + \epsilon) = \frac{\beta_0 + \epsilon}{\alpha_0 + \beta_0}, \tag{3.7}$$

$$\min_{\alpha,\beta} f(\alpha, \beta) = f(\alpha_0 + \epsilon,\ \beta_0 - \epsilon) = \frac{\beta_0 - \epsilon}{\alpha_0 + \beta_0}. \tag{3.8}$$

When the limiting state probability is equal to the value given by Eqs. (3.7) and (3.8), we represent the corresponding state transition matrix as P_ϵ and $P_{-\epsilon}$, respectively, i.e.

$$P_\epsilon = \begin{pmatrix} 1 - \alpha_0 + \epsilon & \alpha_0 - \epsilon \\ \beta_0 + \epsilon & 1 - \beta_0 - \epsilon \end{pmatrix}, \tag{3.9}$$

$$P_{-\epsilon} = \begin{pmatrix} 1 - \alpha_0 - \epsilon & \alpha_0 + \epsilon \\ \beta_0 - \epsilon & 1 - \beta_0 + \epsilon \end{pmatrix}. \tag{3.10}$$

We shall next consider the case where α and β are time varying, that is, the automaton is nonstationary. Let us give the following two lemmas.

LEMMA 1. *At an arbitrary time k, the following inequality about the transition matrix $P(k)$ holds;*

$$|\pi(k)P_{-\epsilon}| \leqslant |\pi(k)P(k)| \leqslant |\pi(k)P_\epsilon| \tag{3.11}$$

where, $|\pi|$ is designated as the 1st element of the vector.

Proof.

$$|\pi(k)P(k)| = \left|(r, 1-r)\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}\right|$$

$$= r(1-\alpha) + (1-r)\beta \tag{3.12}$$

where α and β satisfy Eq. (3.2). Then $|\pi(k)P(k)|$ has a maximum value when α and β are

$$\alpha(k) = \alpha_0 - \epsilon, \quad \beta(k) = \beta_0 + \epsilon,$$

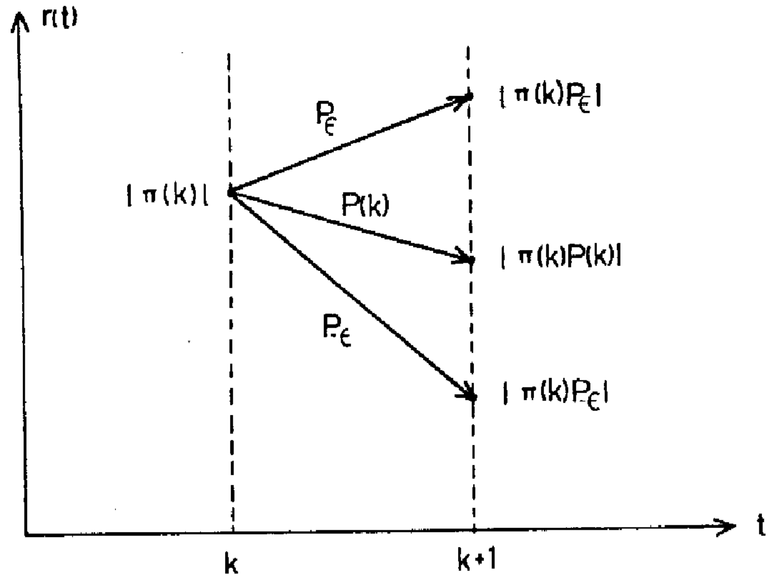

Fig. 5 Explanation of Lemma 1.

respectively. Thus,

$$P(k) = P_\epsilon.$$

Similarly, $|\pi(k)P(k)|$ has a minimum value if the following equation holds.

$$P(k) = P_{-\epsilon}. \blacksquare$$

LEMMA 2. *Let the state vector at an arbitrary time k be* π_1 or π_2, *and the transition matrix be* $P(k)$. *If the following equation holds*

$$\pi_1 = (r_1, 1 - r_1), \quad \pi_2 = (r_2, 1 - r_2), \quad r_1 > r_2$$

then, we have

(1) $0 \leqslant |\pi_1 P(k)| - |\pi_2 P(k)| < |\pi_1| - |\pi_2|$ (3.13)

if $1 - \alpha(k) - \beta(k) \geqslant 0$

(2) $0 < |\pi_2 P(k)| - |\pi_1 P(k)| < |\pi_1| - |\pi_2|$ (3.14)

if $1 - \alpha(k) - \beta(k) < 0$

Proof.

$$|\pi_1 P(k)| - |\pi_2 P(k)| = \left|(r_1, 1 - r_1)\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}\right| - \left|(r_2, 1 - r_2)\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}\right|,$$

$$= (r_1 - r_2)(1 - \alpha - \beta),$$

$$= (|\pi_1| - |\pi_2|)(1 - \alpha - \beta),$$

where

$$r_1 > r_2, \quad |1 - \alpha - \beta| < 1.$$

The lemma was thus proved. $\blacksquare$

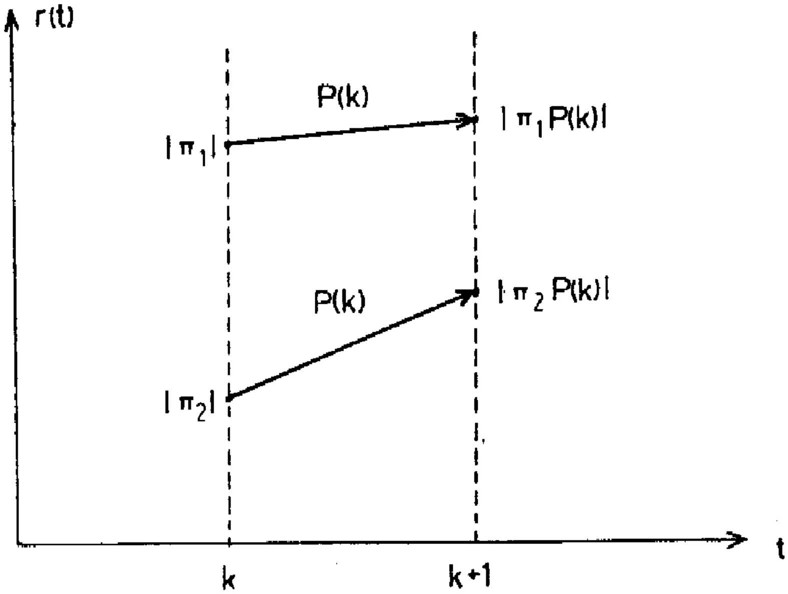

Fig. 6. Explanation of Lemma 2–(1).

Lemma 1 and Lemma 2 are illustrated in Figs. 6 and 7 respectively. Let us next show that the limiting value of the state probability of the state t_1 is within a certain extent.

THEOREM 2. *If* $1 - \alpha_0 - \beta_0 \geqslant 0$ *and* $n \to \infty$, *then*

$$\frac{\beta_0 - \epsilon}{\alpha_0 + \beta_0} \leqslant r(n) \leqslant \frac{\beta_0 + \epsilon}{\alpha_0 + \beta_0}. \tag{3.15}$$

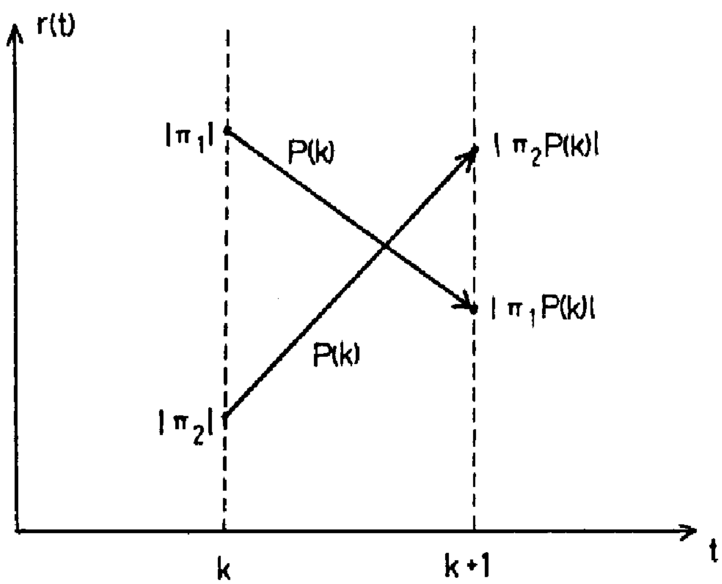

Fig. 7. Explanation of Lemma 2–(2).

Proof. (i) Using mathematical induction, we can show that the following (3.16) holds for a certain positive integer k and an arbitrary positive integer n,

$$\begin{aligned}|\pi(k)P_{-\epsilon}^{n}| &\leqslant |\pi(k)P(k)P(k+1)\cdots P(k+n-1)\\ &\leqslant |\pi(k)P_{\epsilon}^{n}|. \end{aligned} \tag{3.16}$$

By Lemma 1, (3.16) holds for $n = 1$. Let us next assume that Eq. (3.16) holds for $n = n_0$, that is, the following holds.

$$\begin{aligned}|\pi(k)P_{-\epsilon}^{n_0}| &\leqslant |\pi(k)P(k)P(k+1)\cdots P(k+n_0-1)|,\\ &\leqslant |\pi(k)P_{\epsilon}^{n_0}|. \end{aligned} \tag{3.17}$$

We can rewrite (3.17) as (3.18), i.e.

$$|\pi_{-\epsilon}| \leqslant |\pi(k+n_0)| \leqslant |\pi_{\epsilon}| \tag{3.18}$$

where

$$\pi_{-\epsilon} = \pi(k)P_{-\epsilon}^{n_0}, \qquad \pi_{\epsilon} = \pi(k)P_{\epsilon}^{n_0}. \tag{3.19}$$

From the assumption $1 - \alpha_0 - \beta_0 > 0$ and Lemma 2 (1), we have

$$|\pi(k+n_0)P_{\epsilon}| \leqslant |\pi_{\epsilon} P_{\epsilon}|. \tag{3.20}$$

Moreover, from Lemma 1 we obtain

$$|\pi(k+n_0)P_{-\epsilon}| \leqslant |\pi(k+n_0)P(k+n_0)| \leqslant |\pi(k+n_0)P_{\epsilon}|. \tag{3.21}$$

From (3.19), (3.20), and (3.21), we have

$$|\pi_{-\epsilon} P_{-\epsilon}| \leqslant |\pi(k+n_0)P(k+n_0)| \leqslant |\pi_{\epsilon}P_{\epsilon}|, \tag{3.22}$$

that is,

$$\begin{aligned}|\pi(k)P_{-\epsilon}^{n_0+1}| &\leqslant |\pi(k)P(k)P(k+1)\cdots P(k+n_0)|,\\ &\leqslant |\pi(k)P^{n_0+1}|. \end{aligned} \tag{3.23}$$

From the expression of the above equation (3.23), it is shown that (3.16) holds for $n = n_0 + 1$. Hence, we obtain that (3.16) holds for an arbitrary n by using mathematical induction. The above equations are summarized in Fig. 8.

(ii) If $n \to \infty$, then from (3.7), (3.8), (3.9), and (3.10), we can find

$$|\pi(k)P_{-\epsilon}^{n}| \to \frac{\beta_0 - \epsilon}{\alpha_0 + \beta_0}, \tag{3.24}$$

and

$$|\pi(k)P_{\epsilon}^{n}| \to \frac{\beta_0 + \epsilon}{\alpha_0 + \beta_0}. \tag{3.25}$$

Moreover, the following holds

$$|\pi(k)P(k)P(k+1)\cdots P(k+n)| = r(k+n+1). \tag{3.26}$$

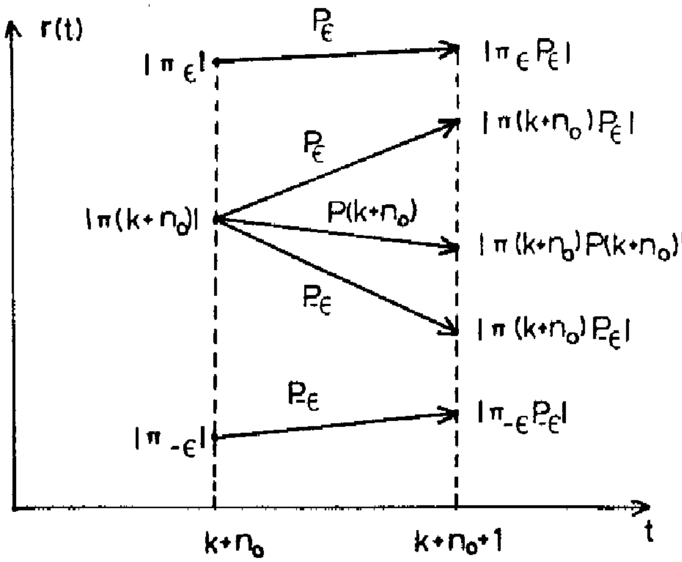

Fig. 8. Explanation of the proof of Theorem 2.

Thus, if $n \to \infty$, from (3.16), (3.24), (3.25), and (3.26), we conclude

$$\frac{\beta_0 - \epsilon}{\alpha_0 + \beta_0} \leqslant r(n) \leqslant \frac{\beta_0 + \epsilon}{\alpha_0 + \beta_0}. \blacksquare$$

THEOREM 3. *If* $1 - \alpha_0 - \beta_0 < 0$ *and* $n \to \infty$, *the expectation* $M(r(n))$ *of the state probability* $r(n)$ *given by*

$$M(r(n)) = \{r(1) + r(2) + \cdots + r(n)\}/n$$

becomes as follows,

$$\frac{\beta_0 - \epsilon}{\alpha_0 + \beta_0} \leqslant M(r(n)) \leqslant \frac{\beta_0 + \epsilon}{\alpha_0 + \beta_0}. \tag{3.27}$$

Proof. (i) Suppose the following (3.28) holds for a certain positive integer n_i, where $n_i < n_{i+1}$ for $i = 1, 2, \ldots,$

$$|\pi(1)P_\epsilon^{n_i+2}| < |\pi(1)P(1)P(2)\cdots P(n_i+2)|, \tag{3.28}$$

we have the following (3.29), (3.30), and (3.31)

$$|\pi(1)P(1)P(2)\cdots P(n_i+1)| < |\pi(1)P_\epsilon^{n_i+1}|, \tag{3.29}$$

$$|\pi(1)P(1)P(2)\cdots P(n_i+3)| < |\pi(1)P_\epsilon^{n_i+3}|, \tag{3.30}$$

and

$$|\pi(n_i+2)| + |\pi(n_i+3)| < |\pi(1)P_\epsilon^{n_i+1}| + |\pi(1)P_\epsilon^{n_i+2}|. \tag{3.31}$$

It is trivial that we can obtain these three equations from Lemma 1, Lemma 2 (2) and (3.28). The above equations are summarized in Fig. 9.

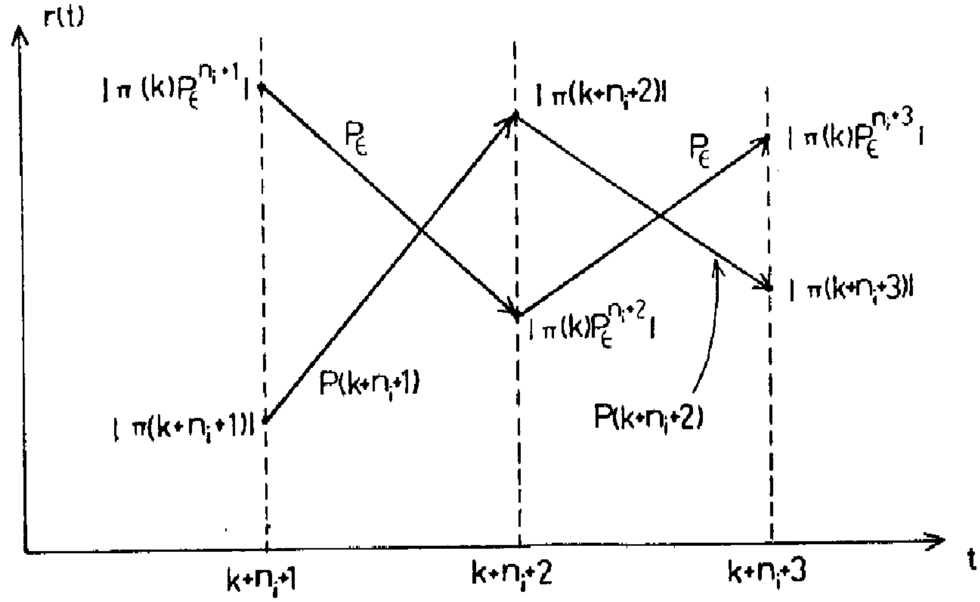

Fig. 9. Explanation of the proof of Theorem 3.

(ii) From (3.28), (3.29), (3.30), and (3.31), we obtain the following inequalities.

$$|\pi(2)| < |\pi(1)P_\epsilon|,$$

$$\vdots$$

$$|\pi(n_1+2)| + |\pi(n_1+3)| < |\pi(1)P_\epsilon^{n_1+1}| + |\pi(1)P_\epsilon^{n_1+2}|,$$

$$|\pi(n_1+4)| < |\pi(1)P_\epsilon^{n_1+3}|,$$

$$\cdots\vdots\cdots$$

$$|\pi(n_i+2)| + |\pi(n_i+3)| < |\pi(1)P_\epsilon^{n_i+1}| + |\pi(1)P_\epsilon^{n_i+2}|,$$

$$|\pi(n_i+4)| < |\pi(1)P_\epsilon^{n_i+3}|,$$

$$\cdots\vdots\cdots$$

$$|\pi(n)| < |\pi(1)P_\epsilon^{n-1}|.$$

Summing up all these inequalities, then

$$|\pi(1)| + |\pi(2)| + \cdots + |\pi(n)| < |\pi(1)| + |\pi(1)P_\epsilon| + \cdots + |\pi(1)P_\epsilon^{n-1}|. \quad (3.32)$$

If $n \to \infty$, from Eq. (3.25) we have

$$\{|\pi(1)| + |\pi(1)P_\epsilon| + \cdots + |\pi(1)P_\epsilon^{n-1}|\}/n \to \frac{\beta_0 + \epsilon}{\alpha_0 + \beta_0}. \quad (3.33)$$

Hence, if $n \to \infty$, from Eqs. (3.32) and (3.33) we have

$$M(r(n)) \leqslant \frac{\beta_0 + \epsilon}{\alpha_0 + \beta_0}. \quad (3.34)$$

Similarly, if $n \to \infty$, then

$$\frac{\beta_0 - \epsilon}{\alpha_0 + \beta_0} \leqslant M(r(n)). \quad \blacksquare$$

Corollary 1. *If $n \to \infty$, the following holds*

$$\frac{\beta_0 - \epsilon}{\alpha_0 + \beta_0} \leqslant M(r(n)) \leqslant \frac{\beta_0 + \epsilon}{\alpha_0 + \beta_0}.$$

Proof. This proof is easily obtained from Theorem 2 and Theorem 3.

4. EXPEDIENT BEHAVIOR OF FINITE AUTOMATA

Based on the limited state probability distribution of the finite automaton given in the previous chapter, the limiting penalty ratio represented as

$$\lim_{n \to \infty} \frac{\text{number of penalties in first } n \text{ trials}}{n}$$

is obtained. Comparing this limiting penalty ratio with the one obtained in the case where inputs are put into the environment with the same probability, it is shown that the finite automaton can behave expediently.

In the previous chapter, the expected state probability of the state s_1 of the nonstationary probabilistic automaton A has been obtained. From this result, the expected state probability $M(1 - r(n))$ of the state s_2 is

$$\frac{\alpha_0 - \epsilon}{\alpha_0 + \beta_0} \leqslant M(1 - r(n)) \leqslant \frac{\alpha_0 + \epsilon}{\alpha_0 + \beta_0} \tag{4.1}$$

as $n \to \infty$. Since α_0, β_0 and ϵ satisfy (3.2), we have $p_1(t) > p_2(t)$ if $\alpha_0 > \beta_0$ is assumed. Therefore, in the sense of Tsetlin's model, the optimal output of the finite automaton is u_2. Furthermore, if $n \to \infty$, from (4.1) and Corollary 1, we have

$$M(r(n)) < M(1 - r(n)). \tag{4.2}$$

In other words, the automaton A takes an optimal output with a larger probability than that associated with a nonoptimal one. In the case of Tsetlin's random environment, a finite automaton behaves expediently if and only if the finite automaton takes an optimal output with a larger probability. However, in the case of the nonstationary random environment as in this paper, we cannot conclude easily that a finite automaton behaves expediently. Therefore, in this case, we shall call that a finite automaton behaves pseudo-expediently.

Let us then consider a true expedient behavior. At first, penalty probability at enough large time in the case where each input is put into the probabilistic automaton with the same probability, is obtained.

The state transition probability $P(u^*) = P(u^{(1)})P(u^{(2)}) \cdots P(u^{(m)})$ of the input sequence $u^* = u^{(1)}u^{(2)} \cdots u^{(m)}$ is represented by $C^{(1)}C^{(2)} \cdots C^{(m)}$ as follows. When the jth input $u^{(j)}$ of the input sequence is $u_i (i = 1, 2)$, $C^{(j)}$ becomes $P(u_i)$, where

$$P(u_1) = A, \quad P(u_2) = B.$$

Let us represent the kth fundamental matrix of the matrix $C^{(j)}$ as $C_k^{(j)} (k = 1, 2)$, where $C_k^{(j)} = A_k$ when $C^{(j)} = A$, and $C_k^{(j)} = B_k$ when $C^{(j)} = B$.

LEMMA 3 (Yasui and Yajima [6]). *A product of the stochastic matrices* $C^{(1)}C^{(2)} \cdots C^{(m)}$ *can be represented as follows*

$$C^{(1)}C^{(2)} \cdots C^{(m)} = C_1^{(m)} + \nu_2^{(m)} C_1^{(m-1)} C_2^{(m)} + \nu_2^{(m-1)} \nu_2^{(m)} C_1^{(m-2)} C_2^{(m-1)} + \cdots + \nu_2^{(2)} \nu_2^{(3)} \cdots \nu_2^{(m)} C_1^{(1)} C_2^{(2)} + \nu_2^{(1)} \nu_2^{(2)} \cdots \nu_2^{(m)} C_2^{(1)}, \tag{4.3}$$

where

$$A_1 A_2 = \overline{0}, \quad B_1 B_2 = \overline{0}^1,$$

$$H = A_1 B_2 = -B_1 A_2 = \frac{|bc - ad|}{(a+b)(c+b)} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \tag{4.4}$$

$$\nu_2^{(j)} = \begin{cases} 1 - a - b = \lambda, & \text{if } C^{(j)} = A, \\ 1 - c - d = \mu, & \text{if } C^{(j)} = B, \end{cases} \tag{4.5}$$

$$A_1 = \begin{pmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \dfrac{b}{a+b} & \dfrac{a}{a+b} \end{pmatrix}, \quad B_1 = \begin{pmatrix} \dfrac{d}{c+d} & \dfrac{c}{c+d} \\ \dfrac{d}{c+d} & \dfrac{c}{c+d} \end{pmatrix}. \tag{4.6}$$

Definition 5. Let z be any input sequence of length m, then *the expected matrix* $M(p(z))$ *of the transition matrix* $P(z)$ is defined as follows

$$|M(P(z))|_{i,j} = \sum_{z \in U_m^*} p_z |P(z)|_{i,j}, \quad i, j = 1, 2$$

where

U_m^*, set of all input sequences of length m;
p_z, a priori probability of z;
$|D|_{i,j}$, the (i, j)-element of the 2×2 matrix D.

THEOREM 4. *If at an arbitrary time j, the input* $u^{(j)}$ *of the probabilistic automaton C becomes* u_1 *(or* u_2*) with probability* 1/2, *and* $\lambda = \mu$, *then the expected*

[1] $\overline{0}$ is 2×2 zero matrix

matrix $M(C^{(1)}C^{(2)}\cdots)$ *of the state transition matrix* $C^{(1)}C^{(2)}\cdots$ *for the infinite input sequence* $u^{(1)}u^{(2)}\cdots$ *is*

$$M[C^{(1)}C^{(2)}\cdots] = \frac{A_1 + B_1}{2}. \tag{4.7}$$

Proof. Let the length of an input sequence $u^{(1)}u^{(2)}\cdots u^{(m)}$ be m. Using Lemma 3, $C^{(1)}C^{(2)}\cdots C^{(m)}$ is given by Eq. (4.3). Let us now consider the expected matrix of each term in Eq. (4.3). The kth term ($k = 2, 3, \ldots, m$) in Eq. (4.3) is written as

$$\nu_2^{(m-k+2)}\nu_2^{(m-k+3)}\cdots\nu_2^{(m-1)}\nu_2^{(m)}C_1^{(m-k+1)}C_2^{(m-k+2)}.$$

Since both $C^{(m-k+1)}$ and $C^{(m-k+2)}$ are either A or B with probability 1/2 from the assumption, $C_1^{(m-k+1)}C_2^{(m-k+2)}$ is equal to one of A_1A_2, A_1B_2, B_1A_2, and B_1B_2 with probability 1/4. From Eq. (4.4) the expected matrix of the kth term is

$$M(\nu_2^{(m-k+2)}\cdots\nu_2^{(m)}C_1^{(m-k+1)}C_2^{(m-k+2)}) = \lambda^k(A_1B_2/4 + B_1A_2/4) = \overline{0}.$$

The expected matrices of the 1st term and the $(m+1)$th term are, respectively,

$$M(C_1^{(m)}) = \frac{A_1 + B_1}{2},$$

and

$$M(\nu_2^{(1)}\nu_2^{(2)}\cdots\nu_2^{(m)}C_2^{(1)}) = \lambda^m\frac{A_2 + B_2}{2}.$$

Hence, the expected matrix of the state transition probability matrix for the input sequence $u^{(1)}u^{(2)}\cdots u^{(m)}$ is

$$M(C^{(1)}C^{(2)}\cdots C^{(m)}) = \frac{A_1 + B_1}{2} + \lambda^m\frac{A_2 + B_2}{2}.$$

When $|\lambda| < 1$ and an input sequence is infinite, i.e. $n \to \infty$, we have

$$\lambda^m\frac{A_2 + B_2}{2} \to 0. \blacksquare$$

By Theorem 4 and (2.2), it can be shown that if either u_1 or u_2 is put into the probabilistic automaton with probability 1/2 and $\lambda = \mu$ (or $a + b = c + d$), then the penalty probability $R(C, t)$ at enough large time t is $(\alpha_0 + \beta_0)/2$.

Extending the concept of being completely isolated by the 0th approximation defined by Yasui and Yajima [6], we can define the concept of being completely isolated by the $(0, k)$th approximation. Based on this concept, the expedient behavior is considered as follows.

Definition 6. A probabilistic automaton C is said to be completely isolated by the $(0, k)$th approximation, if the following conditions are satisfied

$$|p_1(t) - \alpha_0| < \epsilon, \qquad |p_2(t) - \beta_0| < \epsilon,$$
$$|\alpha_0 - \beta_0| \geqslant k\epsilon, \qquad k \geqslant 2, \tag{4.8}$$

where α_0, β_0, $p_1(t)$ and $p_2(t)$ are the same as those in Def. 3, respectively.

LEMMA 4. *A necessary and sufficient condition for the probabilistic automaton being completely isolated by the* $(0, k)$th *approximation is*

(1) $\|H\| \neq 0$

(2) $$h' = \frac{k\delta}{1-\delta} \cdot \frac{\max\{\|H\|, \|A_2\|, \|B_2\|\}}{\|H\|} \leqslant 1 \tag{4.9}$$

where δ, $\|H\|$, $\|A_2\|$, and $\|B_2\|$ *are the same as those in Theorem 1, respectively.*

Proof. From the proof of Lemma 2 in the Yasui and Yajimas' paper [6], we have

$$|\alpha_0 - \beta_0| = \|H\| \geqslant k\epsilon = k\frac{\delta}{1-\delta}\max\{\|H\|, \|A_2\|, \|B_2\|\}. \ \blacksquare$$

LEMMA 5. *If a probabilistic automaton is completely isolated by the* $(0, k)$th *approximation, this automaton is completely isolated by the* 0th *approximation.*

Proof. It is trivial from Defs. 3 and 6.

Next, let us compare the penalty probability in the case where the finite automaton and the probabilistic automaton have a mutual interaction, with the penalty probability in the case where inputs are put into the probabilistic automaton with the same probability.

LEMMA 6. *If two input sequences,* $u^* = u^{(1)}u^{(2)} \cdots u^{(n)}$ *and* $u^{*\prime} = u^{(1)\prime}u^{(2)\prime} \cdots u^{(n)\prime}$, *satisfy the following conditions*

$$u^{(j)} = u^{(j)\prime}, \quad \text{for } j = 1, 2, \ldots, n-1,$$
$$u^{(n)} = u_1, \quad u^{(n)\prime} = u_2, \quad \lambda = \mu,$$

then, the matrices P_1 *of* u^* *and* P_2 *of* $u^{*\prime}$ *satisfy the following*

$$P_1 - P_2 = A_1 - B_1 - \lambda H. \tag{4.10}$$

Proof. From Lemma 3 and the assumption, we obtain

$$P_1 - P_2 = C_1^{(n)} - C_1^{(n)\prime} + \lambda[C_1^{(n-1)}C_2^{(n)} - C_1^{(n-1)\prime}C_2^{(n)\prime}].$$

Next, we consider the two cases of the input at time $n - 1$ as follows:

(i) If $u^{(n-1)} = u^{(n-1)'} = u_1$, then

$$C_1^{(n-1)} C_2^{(n)} = A_1 A_2 = \overline{0},$$
$$C_1^{(n-1)'} C_2^{(n)'} = A_1 B_2 = H.$$

(ii) If $u^{(n-1)} = u^{(n-1)'} = u_2$, then

$$C_1^{(n-1)} C_2^{(n)} = B_1 A_2 = -H,$$
$$C_1^{(n-1)'} C_2^{(n)'} = B_1 B_2 = \overline{0}.$$

From (i) and (ii), Eq. (4.10) is obtained. ∎

THEOREM 5. *If the probabilistic automaton C is completely isolated by the* (0, k)th *approximation and the following conditions are satisfied*

(1) $\lambda = \mu \geqslant 0$,

(2) $1 - \alpha_0 - \beta_0 \geqslant 0$,

(3) $\dfrac{\alpha_0}{k+1} \leqslant \beta_0 \leqslant \dfrac{(k-4)\alpha_0}{k}$,

then the finite automaton A behaves expediently against the probabilistic automaton C.

Proof. By Lemma 5 and the assumption of being completely isolated by the (0, k)th approximation, the probabilistic automaton C is completely isolated by the 0th approximation. From this result along with the condition (2) and Theorem 2, the state probability $r(t)$ of the state s_1 at enough large time t is

$$r(t) = \frac{\beta_0 - \epsilon_1(t)}{\alpha_0 + \beta_0}, \qquad -\epsilon < \epsilon_1(t) < \epsilon.$$

Since the outputs u_1 and u_2 of the finite automaton A correspond to the states s_1 and s_2, respectively, the probability $M(A, C, t)$ with which the finite automaton A receives penalty at enough large time t is as follows;

$$M(A, C, t) = \frac{\beta_0 - \epsilon_1(t)}{\alpha_0 + \beta_0}\,\alpha(t) + \frac{\alpha_0 + \epsilon_1(t)}{\alpha_0 + \beta_0}\,\beta(t). \tag{4.11}$$

Next, let $R(C, t)$ be penalty-probability at enough large time t in the case where the inputs are put into the probabilistic automaton with the same probability. Then, from the condition (1) and Theorem 4, we have

$$R(C, t) = \frac{\alpha_0 + \beta_0}{2}.$$

Now, compare $M(A, C, t)$ with $R(C, t)$,

$$R(C,t) - M(A,C,t) = \frac{\alpha_0 + \beta_0}{2} - \left\{ \frac{\beta_0 - \epsilon_1(t)}{\alpha_0 + \beta_0} \alpha(t) + \frac{\alpha_0 + \epsilon_1(t)}{\alpha_0 + \beta_0} \beta(t) \right\},$$

$$= \frac{F}{2(\alpha_0 + \beta_0)}$$

where

$$F = (\alpha_0 + \beta_0)^2 - 2\{\alpha(t)\beta_0 + \alpha_0\beta(t) + (\beta(t) - \alpha(t))\epsilon_1(t)\}.$$

From Lemma 6, we have

$$\alpha(t) - \beta(t) = \alpha_0 - \beta_0 - \lambda|H|,$$

where $|H|$ is the $(1, 2)$-element of the 2×2 matrix H. Moreover, define

$$\beta(t) = \beta_0 + \epsilon_2(t), \qquad -\epsilon < \epsilon_2(t) < \epsilon,$$

then

$$F = (\alpha_0 - \beta_0)^2 - 2\{\lambda|H|(\epsilon_1(t) - \beta_0) + \epsilon_2(t)(\alpha_0 + \beta_0) + \epsilon_1(t)(-\alpha_0 + \beta_0)\}.$$

We can assume $\alpha_0 > \beta_0$ without loss of generality. Then, from (2.2) and (4.4), we can find easily

$$|H| = \alpha_0 - \beta_0 > 0.$$

From this result and the condition (1), we have

$$F > (\alpha_0 - \beta_0)^2 - 2\{\lambda|H|(\epsilon - \beta_0) + \epsilon(\alpha_0 + \beta_0) - \epsilon(-\alpha_0 + \beta_0)\}$$

$$= (\alpha_0 - \beta_0)^2 - 2\{\lambda|H|(\epsilon - \beta_0) + 2\epsilon\alpha_0\}.$$

From the assumption that the probabilistic automaton C is completely isolated by the $(0, k)$th approximation, i.e. $|H| = \alpha_0 - \beta_0 = k\epsilon$, we have

$$F > k^2\epsilon^2 - 2\{\lambda k\epsilon(\epsilon - \beta_0) + 2\epsilon\alpha_0\}$$

$$= \epsilon\{k^2\epsilon - 2[\lambda k(\epsilon - \beta_0) + 2\alpha_0]\}.$$

Moreover, from the conditions (1) and (3), we can find

$$F > \epsilon(k^2\epsilon - 4\alpha_0) > 0.$$

Hence, we conclude

$$R(C,t) > M(A,C,t).$$

Now, let the limiting penalty ratios corresponding to $M(A, C, t)$ and $R(C, t)$ be $M(A, C)$ and $R(C)$, respectively, then it is easily seen that the following holds.

$$R(C) > M(A,C).$$

This result shows that the finite automaton behaves expediently. ∎

Now, we consider one example of the probabilistic automaton which satisfies the conditions given in Theorem 5.

EXAMPLE.

$$A = \begin{pmatrix} 0.6 & 0.4 \\ 0.55 & 0.45 \end{pmatrix}, \quad B = \begin{pmatrix} 0.8 & 0.2 \\ 0.75 & 0.25 \end{pmatrix}$$

Then we obtain

$$\lambda = \mu = 0.05 > 0$$

$$\|H\| = \frac{4}{19}, \quad \|A_2\| = \alpha_0 = \frac{8}{19}, \quad \|B_2\| = \beta_0 = \frac{4}{19}$$

$$h' = \frac{2k}{19} \leqslant 1, \quad \text{i.e.}\ k \leqslant \frac{19}{2}$$

$$1 - \alpha_0 - \beta_0 = \frac{7}{19} > 0.$$

Assuming $k = 9$, then we have

$$\frac{\alpha_0}{k+1} = \frac{8}{190}, \quad \beta_0 = \frac{4}{19}, \quad \frac{(k-4)\alpha_0}{k} = \frac{40}{171}, \quad \frac{8}{190} < \frac{4}{19} < \frac{40}{171}.$$

Hence, the conditions in Theorem 8 are satisfied.

5. CONCLUSION

The behavior of the finite automaton in the nonstationary random environment, that is, the interaction between the finite automaton and the probabilistic automaton is considered based on Tsetlin's model in the case where the automaton has two inputs and two states. However, for the case where the number of the states is n larger than 2, the analysis may seem to be much more complicated and it will be the problem to be investigated in the future.

The authors thank Mr. Yoneharu Fujita for his discussions about the present study.

REFERENCES

1. M. L. Tsetlin, On the behavior of finite automata in random media, *Automat. Remote Control* **22**, 1210–1219 (1961).
2. V. I. Varshavskii and I. P. Vorentsova, On the behavior of stochastic automata with a variable structure, *Automat. Remote Control* **24**, 327–333 (1963).
3. V. Yu. Krylov, On one automaton that is asymptotically optimal in random medium, *Automat. Remote Control* **24**, 1114–1116 (1963).

4. K. S. Fu and T. J. Li, Formulation of learning automata and automata games, *Information Sciences* 1, 237–256 (1969).
5. B. Chandrasekaran and D. W. Shen, Adaptation of stochastic automata in nonstationary environments, *Proc. Nat'l Electronics Conf.* 23, 39–44 (1967).
6. T. Yasui and S. Yajima, Two-state two-symbol probabilistic automata, *Information and Control* 16, 203–224 (1970).

Received July, 1971