Finite Automata and Random Sequence

C. P. Schnorr and H. Stimm

Institut für Angewandte mathematick
der Johann Wolfgang Goethe-Universiät
D-6000 Frankfurt 1, Robert-Mayer-Str. 10
Bundesrepublik Deutschland

Abstract. We consider the behaviour of finite automata on infinite binary sequence and study the class of random tests which can be carried out by finite automata. We give several equivalent characterizations of those infinite binary sequences which are random relative to finite automata. These characterizations are based on the concepts of selection rules, martingales and invariance properties defined by finite automata.

1 Introduction

Stimulated by Kolmogorov and and Chaitin, several attempts have been made recently, by Martin-Löf, Loveland, and Schnorr, to conceive the notion of infinite binary random sequences on the basis of the theory of algorithms. In Schnorr [13] an intuitive concept of the effective random test is proposed. The hitherto investigated, different approaches to this one has proven to be equivalent. It thus provides an robust definition of the notion of randomness. We say a sequence is random if it can pass any effective random test.

Random sequences can not necessarily be calculated effectively, i. e., they are not recursive. The problem, which is important for the application, to calculate the possible "random" sequences can therefore only be investigated in the sense that random sequences can be approximated by recursive sequences. Recursive sequences that behave "random" are called pseudo-random sequences. From a pseudo-random sequence one would expect that it at least fulfills the"most important" probability laws. The assumption is that the algorithmically simple random tests correspond to particularly important probability laws. Among the algorithm, those represented by finite automata are to be regarded as particularly simple. This raises the question of whether there is an excellent class of randomness which can be tested by finite automata.

To answer this question, we restricted some of the approaches described in [13] to the implementation of effective random tests and to consider only those that can be carried out through finite automata. We consider the following random concepts from [13].

(I) The principle of dichotomy of game system,
(II) The principle of unpredictability and
(III) Invariability characteristics of the random sequence

In each of these cases the so-called Bernoulli sequences (also called normal, non-reactive sequences or admissible numbers) fulfill all random tests which can be carried out by finite automata. Thus, precisely the Bernoulli sequences are classified as random against finite automata.

We restrict ourselves here to the case of uniform distribution over a finite alphabet. A transfer to finite probability spaces with arbitrary probabilities is obvious. The results of this work generalize the results of the investigations by Agafonoff [1] on the selection rules generated by finite automata.

2 Bernoulli sequences and ergodic states of finite automata

Let X be a finite set, let X^* (X^∞) be the set of finite (infinite) sequences over X. Let $\Lambda \in X^*$ be the empty sequence. $|u|$ designate the length of $u \in X^*$. For $u \in X^*$ ($z \in X^\infty$) and $i \in \mathbb{N}$ ($\mathbb{N}$ is the set of nonnegative integers), $u(i)$ ($z(i)$) is the initial sequence of length i of u or z. For $u \in X^*$, $y \in X^* \cup X^\infty$, denote $uy \in X^* \cup X^\infty$ the concatenation of u and y. This naturally results in a product $AB \subseteq X^* \cup X^\infty$ of set $A \subseteq X^*$ and $B \subseteq X^* \cup X^\infty$. $\chi_A : B \to \{0, 1\}$ be the characteristic function (indicator function) of set $A \subseteq B$. $\|A\|$ denotes the number of elements of set A. In the following $\|X\| = p$. We assume $p \geq 2$ by default.

Definition 2.1. *We say $z \in X^\infty$ is* ***Bernoulli sequence*** *(for uniform distribution) if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{X^*u}(z(i)) = p^{-|u|}$$

for all $u \in X^*$. Let $\mathcal{B} \subseteq X^\infty$ be the set of Bernoulli sequence.

That is the relative frequency of u occurring as a subsequence of z, is just the probability of u. Bernoulli sequences we defined here are also studied by Borel, Copeland, Reichenbach and Popper under the names of normal number, admissible numbers, and non-reactive sequences.

Let $\bar{\mu}$ denote the product probability measure on X^∞ of the uniform distribution on X. Then we have the following result from standard theory of probabilistic theory.

Corollary 2.2.

$$\bar{\mu}(\mathcal{B}) = 1.$$

Now let us establish a first connection between Bernoulli sequences and finite automata.

Definition 2.3. *A finite automaton is a 4-tuple $\mathcal{U} = (X, S, \delta, s_1)$, where X and S are finite sets, namely the input alphabet X and state set S. $s_1 \in S$ is the initial state. $\delta : X \times S \to S$ is the transfer function.*

If the automaton reads the input symbol $x \in X$ in state $s \in S$, it goes into the state $\delta(x, s)$. We describe the behavior of the automaton on a sequence recursively:

$$\delta^*(\Lambda, s) = s$$
$$\delta^*(ux, s) = \delta(x, \delta^*(u, s))$$

Now let us prove proposition 2.5, which later makes it possible to restrict ourselves to strictly connected automata by testing finite automata. Some preliminary considerations are necessary.

Let $\mathcal{U}$ be a finite automaton, and $s, t \in S$. s is **transformable** to $t \iff (\exists u \in X^*)\delta^*(u, s) = t$. (or we said t is reachable from s). .

We denote this as $s \to t$. The relation $\to$ is reflexive and transitive. $\to$ can be restricted to an equivalence relation $\leftrightarrow$ as follows:

$$s \leftrightarrow t \iff s \to t \wedge t \to s.$$

For $s \in S$, we denote $\bar{s} \subseteq S$ as the equivalence class of s of the equivalence relation $\leftrightarrow$. Denote $S_{/\leftrightarrow}$ to be the set of equivalence classes. We say $\mathcal{U}$ is strong connected if S has exactly one equivalence class.

We can define the following partial order in $S_{/\leftrightarrow}$:

$$\bar{s} > \bar{t} \iff (\exists u \in \bar{s})(\exists v \in \bar{t})u \to v$$

By the definition of $\leftrightarrow$, this is the same as:

$$\bar{s} > \bar{t} \iff (\forall u \in \bar{s})(\forall v \in \bar{t})u \to v$$

Since S is finite, there is a minimal element in S. the minimal element in $S_{/\leftrightarrow}$ has the following ergodic properties:

Corollary 2.4. *Let $\bar{s} \in S_{/\leftrightarrow}$ minimal. Then:*

$$(\forall a \in \bar{s})(\forall u \in X^*)\delta^*(u, a) \in \bar{s}.$$

Proof. Assume there are $a \in \bar{s}$ and $u \in X^*$ with $\delta(u, a) \notin \bar{s}$. Let $t := \delta^*(u, a)$. We then know $\bar{t} \neq \bar{s}$ and $\bar{s} > \bar{t}$. This is in conflict with the minimality of $\bar{s}$. □

The minimal elements of $S_{/\leftrightarrow}$ are called **ergodic set** in the theory of Markov chain. We say elements in such sets are **ergordic**. This notion is important in the following statement.

Proposition 2.5. *Let z be a Bernoulli sequence and $\mathcal{U}$ be a finite automaton. Then there is an n for each state s such that $\delta^*(z(n), s)$ is an ergodic state.*

The proof of the proposition is based on:

Lemma 2.6. *For every finite automaton there is a $u \in X^*$, so that $\delta^*(u, s)$ is an ergordic state for all states s.*

Proof. Let $S = \{s_1, \cdots, s_m\}$ be the state set of the automaton. We prove by induction on i such that $\forall i \leq m$ there is a $u^{(i)}$ such that for all $j \leq i$, $\delta^*(u^{(i)}, s_j)$ is ergodic.

Case $\mathbf{i = 1}$: Let $\bar{t} \in S_{/\leftrightarrow}$ be a minimal element such that $\bar{s}_1 > \bar{t}$. Then $s_1 \to t$. Hence there exists $u^{(1)} \in X^*$ so that $\delta^*(u^{(1)}, s_1) = t$. Then $\delta^*(\underline{u^{(1)}, s_1)}$ is ergodic.

From $\mathbf{i}$ to $\mathbf{i+1}$: Let $\bar{t} \in S_{/\leftrightarrow}$ a minimal element with $\overline{\delta^*(u^{(i)}, s_{i+1})} > \bar{t}$. Then $\delta^*(u^{(i)}, s_{i+1}) > t$. Choose $v \in X^*$ so that $\delta^*(v, \delta^*(u^{(i)}, s_{i+1})) = t$. Set $u^{(i+1)} := u^{(i)}v$. Then $\delta^*(u^{(i+1)}, s_{i+1}) = t$. Since $t \in S_{/\leftrightarrow}$ is minimal, it follows that it is ergodic. Note that for all $j \leq i$, $\delta^*(u^{(i+1)}, s_j)$ is also ergodic, since by induction hypothesis, $\delta^*(u^{(i)}, s_j)$ with $j \leq i$ is ergodic, and ergodic states can only transfer to ergodic states. □

Proof of Proposition 2.5. Let z be a Bernoulli sequence and $\mathcal{U}$ a finite automaton. Let $u \in X^*$ fulfills Lemma 2.6. Then u occurs as a subsequence of z. Choose n so that $z(n) \in X^*u$. Then it follows from 2.6 that $\delta^*(z(n), s)$ is ergodic.

3 Finite automata generated asset function

We want to consider games (in naive sense) about sequence $u \in X^*$. The result of a game over the sequence u is a non-negative real number $V(u)$, which indicated the capital of the player after he played over the sequence. Overall, a function

$$V : X^* \to \mathbb{R}^+, \quad \text{where } \mathbb{R}^+ \text{ is the set of non-negative real number,}$$

is generated by a fixed "game strategy" of the player, which he follows at play over a sequence $u \in X^*$. A game is generally regarded as "fair" if the function V satisfies the Martingale Property (M):

$$V(u) = \frac{1}{p} \sum_{x \in X} V(ux) \quad (u \in X^*). \tag{M}$$

In this case the chance to win is just as rough as the opportunity to lose. In the following, whenever we say *asset function*, we mean functions that satisfy property (M). Or in today's terminology, they are **martingale**.

Intuition of the asset function: at the beginning of the game the gambler has capital $V(\Lambda) \in R^+$. After playing over the sequence $u \in X^*$, he bets an amount of $B_x(u) \in R^+$ on $x \in X$ being the next symbol coming after u. The amount of his betting on all possible symbols should not exceed the capital he has. That is, the **betting strategy** $B_x : X^* \to \mathbb{R}^+$ should meet the **betting constraint** (E):

$$\sum_{x \in X} B_x(u) \leq V(u) \quad (u \in X^*). \tag{E}$$

Then the capital $V(ux)$ after this game round is determined by the payoff condition (A):

$$V(ux) = V(u) + \sum_{y \in X} (p\delta_{x,y} - 1)B_y(u), \tag{A}$$

where $\delta_{x,y}$ is the Kronecker Symbol. That is, for the bet on y, the player receives $(p-1)$-fold on the occurrence of y. It is easy to see that the conditions of (A) satisfy the martingale condition (M). It follows from (E) that V is not negative. Conversely, it it not difficult to show that every asset function $V : X^* \to \mathbb{R}^+$ can be generated by suitable function $B_x : X^* \to \mathbb{R}^+$ with the property E.

Now let's investigate the asset functions obtained from finite automata $\mathcal{U} = (X, S, \delta, s_1)$. These are the ones for which the quotient $V(ux)/V(u)$depends only on the instantaneous state $\delta^*(u, s_1)$ of the automaton and on the input symbol x. We then view these quotients as output of the automaton.

Formally, we define the function $V_{\mathcal{U}} : X^* \to \mathbb{R}^+$ that induce by an automaton $\mathcal{U} = (X, S, \delta, s_1)$ with an output function $\lambda : X \times S \to \mathbb{R}^+$ as:

$$\begin{aligned} V_{\mathcal{U}}(\varLambda) &= 1 \\ V_{\mathcal{U}}(ux) &= V_{\mathcal{U}}(u)\lambda(x, \delta^*(u, s_1)) \quad (u \in X^*, x \in X) \\ &\text{(Multiplication in } \mathbb{R}) \end{aligned} \tag{3.1}$$

$V_{\mathcal{U}}$ is a martingale if and only if:

$$\sum_{x \in X} \lambda(x, s) = p \quad (s \in S)$$

We then say such martingales are generated by finite automata.

Now we define it in terms of betting strategies. Let $\mathcal{U} = (X, S, \delta, s_1)$ be a finite automaton and $\overline{\lambda} : X \times S \to \mathbb{R}^+$ be an output function. Then one can define $B_x : X^* \to \mathbb{R}^+$ $(x \in X)$ and a function $\overline{V}_{\mathcal{U}} : X^* \to \mathbb{R}^+$ as follows:

$$\begin{aligned} \overline{V}_{\mathcal{U}}(\varLambda) &= 1 \\ B_x(u) &= \overline{V}_{\mathcal{U}}(u)\overline{\lambda}(x, \delta^*(u, s_1)) \\ \overline{V}_{\mathcal{U}}(ux) &\underset{(1)}{=} \overline{V}_{\mathcal{U}}(u) + \sum_{x \in X} (p\delta_{x,y} - 1)B_y(u) \\ &= \overline{V}_{\mathcal{U}}(1 + \sum_{x \in X} (p\delta_{x,y} - 1)\overline{\lambda}(y, \delta^*(u, s_1))) \quad (u \in X^*, x \in X). \end{aligned} \tag{3.2}$$

The betting constraint (E) is fullfiled exactly when

$$\sum_{x \in X} \overline{\lambda}(x, s) \leq 1 \quad (s \in S).$$

In this case, $\overline{V}_{\mathcal{U}}$ is not negative and thus a martingale. Note that (3.2) say the fraction of capital $B_x(u)/\overline{V}_{\mathcal{U}}(u)$ that bets on x, depends only on the instantaneous sate $\delta^*(u, s_1)$ and x. In the above definition, equation (3.1) and (3.2) is just the payoff condition (A).

Due to (3.2) (last formula), however, the quotient $\overline{V}_{\mathcal{U}}(ux)/\overline{V}_{\mathcal{U}}(u)$ also depends only on the state $\delta^*(u, s_1)$ and x. This means that the martingale we defined in (3.2) can also be defined in the way of (3.1), if we choose a suitable output function. To do this, let $x \in X$ and $s \in S$ and:

$$\lambda(x, s) := 1 + \sum_{y \in X} (p\delta_{x,y} - 1)\overline{\lambda}(y, s).$$

Conversely, let $V_{\mathcal{U}}$ be induced by $\mathcal{U}$, and the output function λ be as in (3.1). Now let $\overline{\lambda}(x, s) := \lambda(x, s)/p$. We show by induction that for the function B_x, $\overline{V}_{\mathcal{U}}$ the following are true:

$$\overline{V}_{\mathcal{U}}(u) = V_{\mathcal{U}}(u), \quad B_x(u) = \frac{1}{p} V_{\mathcal{U}}(ux) \quad (u \in X^*, x \in X).$$

For $x = \Lambda$, that is clear. We go from u to ux. It follows

$$\begin{aligned}\overline{V}_{\mathcal{U}}(ux) &= \overline{V}_{\mathcal{U}}(u) + \sum_{y \in X} (p\delta_{x,y} - 1) B_y(u) \\ &= V_{\mathcal{U}}(u) + \sum_{y \in X} (p\delta_{x,y} - 1) p^{-1} V_{\mathcal{U}}(uy) \\ &= V_{\mathcal{U}}(u) + pp^{-1} V_{\mathcal{U}}(ux) - p^{-1} \sum_{y \in X} V_{\mathcal{U}}(uy) \\ &= V_{\mathcal{U}}(ux).\end{aligned}$$

Also we have:

$$\begin{aligned}B_x(uy) &= \overline{V}_{\mathcal{U}}(uy)\overline{\lambda}(x, \delta^*(uy, s_1)), && \text{by (3.2)} \\ &= V_{\mathcal{U}}(uy)p^{-1}\lambda(x, \delta^*(uy, s_1)) \\ &= p^{-1} V_{\mathcal{U}}(uyx). && \text{by (3.1)}\end{aligned}$$

Overall, we have

Proposition 3.1. *The class of martingales defined by (3.1) and (3.2) based on a finite automaton $\mathcal{U}$, are the same.*

In the following, we use martingale defined by (3.1). (Added by translator: Instead of using the term "Bernoulli sequence", we use **"normal sequence"** or **"normal number"**; we also call a martingale generated by finite automata a **"finite state gambler"** for convenience.

4 Characterization of Normal Sequence by Finite State Gambler

We view a finite state gambler $V_{\mathcal{U}} : X^* \to \mathbb{R}^+$ as a random test. A sequence $z \in X^\infty$ is to be rejected by this test as not random, if $V_{\mathcal{U}}$ unbounded increases along z, that is, if

$$\limsup_{n \to \infty} V_{\mathcal{U}}(z(n)) = \infty.$$

This requirement results from the principle of dichotomy of game system. The following first essential result of this work shows that normal numbers and non-normal number differ considerably with respect to finite state gamblers.

Notation: the quantifier $\forall^{\infty} n$ means "for all but finitely many n"; and $\exists^{\infty} n$ means "for infinitely many n".

Proposition 4.1. *(a) Let $z \in X^{\infty}$ be a normal number and $V_{\mathcal{U}} : X^* \to \mathbb{R}^+$ be a finite state gambler. Then either (1) or (2) applies.*

$$(\forall^{\infty} n \in \mathbb{N}) : V_{\mathcal{U}}(z(n)) = V_{\mathcal{U}}(z(n+1)). \qquad \textit{(1)}$$

there is an r with $0 < r < 1$, such that,

$$(\forall^{\infty} n \in \mathbb{N}) : V_{\mathcal{U}}(z(n)) \leq r^n. \qquad \textit{(2)}$$

(b) Let $z \in X^{\infty}$ be a non-normal sequence, then there is a finite state gambler $V_{\mathcal{U}}$ and $r > 1$, such that

$$(\exists^{\infty} n \in \mathbb{N}) : V_{\mathcal{U}}(z(n)) > r^n.$$

Part (a) states that the asset function $V_{\mathcal{U}}$ on a normal number z either becomes constant after a finite number of steps or decreases exponentially. In order to prove this, we show that the quotient $V_{\mathcal{U}}(z(n+1))/V_{\mathcal{U}}(z(n))$ takes the value $1+\epsilon$ as well as $1-\epsilon$ approximately as frequently. Since $(1+\epsilon)(1-\epsilon) = (1-\epsilon^2) < 1$, the assertion follows. Part (b) says that, on a non-normal sequence, a suitable $V_{\mathcal{U}}$ can grow exponentially fast.

To prove this we need two lemma.

Lemma 4.2. *Let $\mathcal{U} = (X, S, \delta, s_1)$ be a finite, strongly connected automaton. Let $x \in X$, $t \in S$ be fixed. Then there is a $u \in X^*$, such that*

$$(\forall s \in S)(\exists i < |u|) : \delta^*(u(i), s) = t \wedge u(i+1) = u(i)x$$

Proof. Let $S = s_1, \cdots, s_m$. By induction on j, we show that for every $j \leq m$ there is a $u^{(j)}$ such that,

$$(\forall k < j)(\exists i < |u^{(j)}|) : \delta^*(u^{(j)}(i), s_k) = t \wedge u^{(j)}(i+1) = u^{(j)}(i)x.$$

Case $\mathbf{j = 1}$. Select $v \in X^*$ so that $\delta^*(v, s_1) = t$ and let $u^{(1)} := vx$.

From j to j+1: Let $r = \delta^*(u^{(i)})$. Select $v \in X^*$ such that $\delta^*(v, r) = t$ and let $u^{(i+1)} := u^{(i)}vx$. the induction assertion is readily verified. □

To prove the second lemma, we need some definitions and propositions from the theory of Markov chain, all taken from the book [4] *Finite Markov Chains with Stationary Transition Probabilities* authored by Kay Lai Chung.

Let $\mathcal{U} = (X, S, \delta, s_1)$ be a strongly connected finite automaton, $S = \{s_1, \cdots, s_m\}$ the state set.

Let $s_q, s_r \in S$ be fixed. We define the random variables $Z_n : X^{\infty} \to S$, where $(n \in \mathbb{N})$ as:

$$Z_n(z) := \delta^*(z(n), s_q).$$

Then Z_n form a Markov chain with initial distribution

$$(p_i)_{i=1,\cdots,m} = (0, \cdots, 0, 1, 0, \cdots, 0)$$

(it is one at the qth place) and transition matrix

$$p_{ij} \coloneqq \frac{1}{p} \| \{x \in X : \delta(x, s_i) = s_j\} \|.$$

For the n-step transition matrix $(p_{ij}^{(n)})$, we have:

$$(P_{ij}^{(n)}) = (p_{ij})^n.$$

It easy to see that S, the state set of the Markov chain, is a positively recurring class. Therefore the Cesaro limit exists:

$$\pi_j \coloneqq \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)}$$

independent of i ([4],§6 Theorem 4 and its corollary and §5 Theorem 6).

We define a function $f : S \to \{0, 1\}$ as follows:

$$f(s) \coloneqq \begin{cases} 1 & \text{if } s = s_r \\ 0 & \text{else.} \end{cases}$$

Then $f(Z_0) + f(Z_1) + \cdots + f(Z_n)$ indicates how often the Markov chain visited state s_r in the first n steps. It then applies that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n} f(Z_k) = \pi_r, \qquad \overline{\mu}\text{-almost sure}$$

([4],§15 Theorem 2). According to the definition of the random variable Z_n, since this is true for every $s_q, s_r \in S$, we have:

(4.3) $$(\forall s_i, s_j \in S) : \overline{\mu}\Big\{ z \in X^\infty : \lim_{n \to \infty} \frac{1}{n} \| \{k \le n : \delta^*(z(k), s_i) = s_j\} \| = \pi_j \Big\} = 1.$$

Starting from random variable $\overline{Z}_n : X^\infty \to S \times X$ $(n \in \mathbb{N})$ defined by

$$\overline{Z}_n(z) \coloneqq (\delta^*(z(n, s_q)), z_{n+1})$$

one obtains by analogous procedure (with the same π_j)

(4.4) $(\forall s_i, s_j \in S)(\forall x \in X)$:
$$\overline{\mu}\Big\{ z \in X^\infty : \lim_{n \to \infty} \frac{1}{n} \| \{k \le n : \delta^*(z(k), s_i) = s_j \wedge z_{k+1} = x\} \| = \frac{\pi_j}{p} \Big\} = 1$$

With the help of these statements, we now ready to prove

Lemma. 4.5. *Let $\mathcal{U} = (X, S, \delta, s_1)$ be a strongly connected finite automaton with $S = \{s_1, \cdots, s_m\}$ be its state set. Then there are numbers $\pi_j > 0$ $(j = 1, \cdots, m)$, such that for all normal sequence z, all $s_i \in S$ and all $x \in X$*

$$\frac{\pi_j}{p} = \lim_{n\to\infty} \frac{1}{n} \|\{k \leq n : \delta^*(z(k), s_i) = s_j \wedge z_{k+1} = x\}\|$$

Obviously, the π_j's form a probability distribution on S.

Proof. Let π_j $(j = 1, \cdots, m)$ be as in (4.4). Due to (4.4), for each $\varepsilon, \varepsilon' > 0$:

(4.6) $(\forall^\infty k \in \mathbb{N})(\forall s_i, s_j \in S)(\forall x \in X)$:

$$p^{-k} \left\| \left\{ u \in X^k : \left| \frac{1}{k} \|q < k : \delta^*(u(q), s_i) = s_j \wedge u_{q+1} = x\| - \frac{\pi_j}{p} \right| < \varepsilon \right\} \right\| > 1 - \varepsilon^{'}$$

Furthermore, for very normal sequence z, every $k \in \mathbb{N}$ and every $\varepsilon'' > 0$:

(4.7) $$(\forall^\infty n \in \mathbb{N})(\forall u \in X^k) : \left| \frac{1}{n} \sum_{q=1}^{n} \chi_{X^* u}(z(kq)) - p^{-k}) \right| < \varepsilon^{''} \quad .$$

For $\varepsilon, \varepsilon^{'} > 0$, pick k large enough so that (4.6) holds. For k and $\varepsilon^{''} > 0$ pick n large enough so that (4.7) holds. Then it follows immediately

$$\left| \frac{1}{nk} \left\| \left\{ q < nk : \delta^*(z(q), s_i) = s_j \wedge z_{q+1} = x \right\} \right\| - \frac{\pi_j}{p} \right| < \varepsilon + \varepsilon^{'} + \varepsilon^{''}.$$

Since the last inequality holds for arbitrary $\varepsilon, \varepsilon^{'}, \varepsilon^{''} > 0$ for sufficiently large k and n, the convergence expression in (4.5) follows. Since $\mathcal{U}$ is strongly connected, $\pi_j > 0$. □

Proof of Proposition 4.1(a): We want to investigate the limiting behavior of finite state gambler. First we show that, without loss of generality, one can restrict oneself to consider only strong connected finite automata. Let normal sequence z be an input, then a finite state automaton $\mathcal{U} = (X, S, \delta, s_1)$ arrives in an ergodic state t in finitely many (some n) steps (by Proposition 2.5).

The limiting behavior of a $V_{\mathcal{U}}$ on z generated by U is, in the case of $V_{\mathcal{U}}(z(n)) \neq 0$, the same as the finite state gambler $V_{\overline{\mathcal{U}}}$ on the normal sequence $z_{n+1}z_{n+2}z_{n+3}\cdots$, where $V_{\overline{\mathcal{U}}}$ is the corresponding finite state machine induced by $\mathcal{U}$ and defined as $\overline{\mathcal{U}} = (X, \overline{t}, \delta_{/\overline{t}}, t)$, in which the state set is precisely the equivalence class of t, the transition function is just δ restricted on $\overline{t}$, and the initial state is simply t.

Let $V_{\mathcal{U}}$ be a finite state gamble (defined as in (3.1)), with $\mathcal{U} = (X, S, \delta, s_1)$ strongly connected and $S = \{s_1, \cdots, s_m\}$:

$$V_{\mathcal{U}}(ux) = V_{\mathcal{U}}(u)\lambda(x, \delta^*(u, s_1))$$

$$\sum_{x\in X} \lambda(x, s) = p, \quad (s \in S).$$

We distinguish two cases:

$$(\forall s \in S)(\forall x \in X) : \lambda(x, s) = 1. \tag{1}$$

In this case, $V_{\mathcal{U}}$ is constant. These martingale arise when the player places the same bets on all $x \in X$.

$$(\forall t \in S)(\exists y \in X) : \lambda(y, t) < 1. \tag{2}$$

For a state $s \in S$ we define

$$U_s := \prod_{x \in X} \lambda(x, s).$$

The following assertion is then verified

(4.8) $0 \leq U_s \leq 1$
$U_s = 1 \iff (\forall x \in X) : \lambda(x, s) = 1.$

With standard methods one can show that the expression $\prod_{x \in X} \lambda(x, s)$ assumes its maximum 1 exactly at the point $\lambda(x, s) = 1$ $(x \in X)$ under the constraints $\sum_{x \in X} \lambda(x, s) = p$ and $\lambda(x, s) \geq 0$.

Dur to Lemma (4.5), for each normal sequence z it follows that

$$\lim_{n \to \infty} V_{\mathcal{U}}(z(n)) \Big/ \Big(\prod_{j=1}^{m} U_{s_j}^{\pi_j / p} \Big)^n = 1.$$

Dur to (4.8), if for all $s \in S$, $U_s = 1$, then for $\forall x \in X, \forall s \in X$, $\lambda(x, s) = 1$, this means the gambler stops gambling. Otherwise there must be some $s \in S$, such that $U_s < 1$, that is, there exist $\lambda(y, s) < 1$. This implies

$$\prod_{j=1}^{m} U_{s_j}^{\pi_j / p} < 1.$$

So there is an r with

$$(\forall n \in \mathbb{N}) : V_{\mathcal{U}}(z(n)) < r^n.$$

Proof of Proposition 4.1(b):

Proof. If $z \in X^\infty$ is not a normal sequence, then there is a $u \in X^*$, and $y \in X$ and a $\delta > 0$, so that

(4.9) $$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{X^* u}(z(i)) = p^{-|u|}$$

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{X^* uy}(z(i)) = p^{-|u|}(p^{-1} + \delta).$$

We define a martingale V, such that $V(w(n+1)) > V(w(n))$ whenever $w(n+1) \in X^*uy$. Pick an arbitrary number α such that $-1 < \alpha < 1$, we define V with $V(\Lambda) = 1$, and

$$V(w(n+1))/V(w(n)) = \begin{cases} 1 & \text{if } w(n) \notin X^*u \\ 1+\alpha & \text{if } w(n+1) \in X^*uy \\ 1-\alpha/(p-1) & \text{else.} \end{cases}$$

It is easy to see that V can be implement as a finite state gambler. Let

$$r(\alpha) := \Big(1+\alpha\Big)^{p^{-|u|}(p^{-1}+\delta)} \Big(1-\alpha/(p-1)\Big)^{p^{-|u|}(1-p^{-1}-\delta)}.$$

Then by (4.9) it applies that

$$\limsup_{n\to\infty} V(z(n)) r(\alpha)^{-n} = 1.$$

To prove our assertion, it suffices to show that α can be chosen such that $r(\alpha) > 1$. Note that $r(0) = 1$, it suffices to show

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\, r(\alpha)\big|_{a=0} > 0.$$

It applies:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\alpha}\, r(\alpha)\big|_{a=0} &= p^{-|u|}(p^{-1}+\delta) - p^{-|u|}(1-p^{-1}-\delta)/(p-1) \\ &= \delta p^{-|u|+1}/(p-1) > 0, \quad \text{since } \delta > 0. \end{aligned}$$

□

5 Characterization of Normal Sequences by Predictability

In this section, we will consider normal sequence in terms of prediction by finite automata, in which normal sequences behave like ideal random sequence.

Let $z \in X^\infty$. Then we interpret an output $x \in X$ of a finite automata after inputting the initial sequence $z(i)$ as the prediction of the next letter $z(i+1)$.

The following theorem states that the ratio of correct predictions is can be greater than $1/p$ if and only if the input sequence is not normal.

Proposition 5.1. *A sequence $z \in X^\infty$ is not normal if and only if there are finite automata $\mathcal{U} = (X, S, \delta, s_1)$ and an output function $\lambda : X \times S \to X$ with*

$$\limsup_{n\to\infty} \frac{1}{n} \left\| \Big\{ i \le n : \lambda\Big(z_i, \delta^*\big(z(i), s_1\big)\Big) = z_{i+1} \Big\} \right\| > \frac{1}{p}.$$

Proof. “ $\Leftarrow$:″ Let $z \in X^\infty$, $\mathcal{U} = (X, S, \delta, s_1)$ be a finite automaton with output function $\lambda : X \times S \to X$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \left\| \Big\{ i \le n : \lambda\Big(z_i, \delta^*\big(z(i), s_1\big)\Big) = z_{i+1} \Big\} \right\| > \frac{1}{p}.$$

Then there is a state $t \in S$, a $\delta > 0$ and an infinite set $M \subseteq \mathbb{N}$, so that for all $n \in M$:

$$\begin{aligned} &\frac{1}{n} \left\| \Big\{ i \le n : \delta^*\big(z(i), s_1\big) = t \wedge \lambda(z_i, t) = z_{i+1} \Big\} \right\| \\ &> \frac{1}{p}\frac{1}{n} \left\| \Big\{ i \le n : \delta^*\big(z(i), s_1\big) = t \Big\} \right\| + \delta \end{aligned}$$

For a fixed α with $-1 < \alpha < 1$, we define the following martingale V with $V(\varLambda) = 1$ and

$$V\big(w(n+1)\big)/V\big(w(n)\big) = \begin{cases} 1 & \delta^*(w(n)) \ne t \\ 1+\alpha & \delta^*(w(n)) = t \text{ and } \lambda(w_n, t) = w_{n+1} \\ 1 - \alpha/(p-1) & \textit{otherwise} \end{cases}$$

One can verify that V is a finite state gambler as define in (3.1). We now set

$$a := \limsup_{\substack{n \,\to\, \infty \\ n \,\in\, M}} \frac{1}{n} \left\| \big\{ i \le n : \delta^*\big(z(i), s_1\big) = t \big\} \right\|$$

and

$$r(\alpha) := \big(1+\alpha\big)^{ap^{-1}+\delta} \big(1 - \alpha/(p-1)\big)^{a(1-p^{-1})-\delta}.$$

Then

$$\limsup_{n\to\infty} V\big(z(n)\big)/r(\alpha)^n \ge 1.$$

as shown in the proof of (4.1)(b), α can be chosen such that $r(\alpha) > 1$. Hence z is not a normal sequence by Proposition 4.1.

“ $\Rightarrow$:″ If $z \in X^\infty$ is not a normal sequence, then there are $u \in X^*$, $y \in X$ and $\delta > 0$ as in (4.9). Then there exist $t \in X$ and $\delta^{'} \ge 0$,so that,

$$\textbf{(5.2)} \quad (\exists^\infty n \in \mathbb{N}) : \frac{1}{n} \sum_{i=1}^{n} \chi_{X^*uy}\big(z(i)\big) > p^{-|u|}\Big(\frac{1}{p} + \frac{\delta}{2}\Big)$$
$$\wedge\ \frac{1}{n} \sum_{i=1}^{n} \chi_{X^*ut}\big(z(i)\big) < p^{-|u|}\Big(\frac{1}{p} - \delta^{'}\Big).$$

We construct an automaton $\mathcal{U} = (X, S, \delta, s_1)$ and an output function $\lambda : X \times S \to X$, so that the correct prediction ratio is greater than $1/p$.

We define a finite automaton $\mathcal{U} = (X, S, \delta, s_1)$ as

$$S := \bigcup_{i=0}^{|u|} = \Big\{ v \in X^* : |v| \le |u| \Big\}, \quad s_1 := \lambda,$$

$$\delta(x,v) := \begin{cases} vx & if \quad |v| < |u| \\ w & otherwise,\ with\ vx \in Xw. \end{cases}$$

and the output function $\lambda : X \times S \to X$ as:

$$\lambda(x,v) := \begin{cases} t & if \quad v \neq u \\ y & if \quad v = u. \end{cases}$$

We discuss two cases.
Case 1 :

$$|u| = 0, \quad i.e. \quad u = \Lambda \tag{1}$$

In this case

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} &\left\| \left\{ i \le n : \lambda(z_i, \delta^*(z(i), s_1)) = z_{i+1} \right\} \right\| \\ &= \limsup_{n\to\infty} \frac{1}{n} \left\| \left\{ i \le n : z_{i+1} = y \right\} \right\| \\ &= 1/p + \delta > 1/p. \end{aligned}$$

Case 2 :

$$|u| > 0. \tag{2}$$

Then we have

(5.3) $$\begin{aligned} 1/p &= \lim_{n\to\infty} \frac{1}{n} \left\| \left\{ i \le n : z_{i+1} = t \right\} \right\| \\ &= \lim_{n\to\infty} \frac{1}{n} \Big(\sum_{i=1}^{n} \sum_{v \in X^{|u|}} \chi_{X^*vt}\big(z(i)\big) \Big). \end{aligned}$$

The automaton $\mathcal{U}$ would thus make exactly $1/p$ correct prediction on the average, if it only outputs t. If u is a substring of z, however, it output y. Therefore, by (4.9), (5.2) and (5.3):

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} &\left\| \left\{ i \le n : \lambda\big(z_i, \delta^*(z(i), s_1)\big) = z_{i+1} \right\} \right\| \\ &= 1/p - p^{-|u|}(1/p - \delta^{'}) + p^{-|u|}(1/p + \delta/2) \\ &= 1/p + \delta^{'} + \delta/2 > 1/p. \end{aligned}$$

□

6 Invariance Properties Defined by Finite Automata

For $z \in X^\infty$ be a "random" sequence, we expect the following: If we pick in some way (by means of a "selection rule") letters from z, we pick so many such that a new infinite sequence is formed. Then in this sequence all elements of X occur equally frequently. That is, it satisfies the strong law of large numbers.

We consider selection rules which are generated by finite automata, in which the selection or non-selection of the next letter depends only on the instantaneous state of an automaton. We can obviously describe this by an output function $\lambda : X \times S \to X^*$, which satisfies the following additional conditions:

$$\textbf{(6.1)} \quad \begin{array}{rll} \textit{either} & (\forall x \in X) : \lambda(x,s) = \Lambda & \text{(no selection in } s\text{)} \\ \textit{or} & (\forall x \in X) : \lambda(x,s) = x & \text{(select the letter } x \text{ in } s\text{)} \end{array}$$

By assigning to each infinite sequence z (every finite sequence u) the sequence $\Phi_{\mathcal{U}}(z)$ (or $\Phi_{\mathcal{U}}(u)$) selected from it, we obtain a uniquely determined partial function $\overline{\phi}_{\mathcal{U}}$. In what follows, we say $\Phi_{\mathcal{U}}(z)$ is defined if z is an infinite sequence then $\Phi_{\mathcal{U}}(z)$ also an infinite sequence. Otherwise we say $\Phi_{\mathcal{U}}(z)$ is not define, or z is not in the domain of $\Phi_{\mathcal{U}}$.

Agafonov [1] has shown that the chosen sequence satisfies the strong laws of large number for all selection rules generated by finite automatons, if only the input sequence is normal and the selected sequence is infinite.

We want to generalize this concept by dropping the essential limitation of possible output functions. Surely one must not allow all output functions $\lambda : X \times S \to X^*$, because with $\lambda(x,s) = y$ $(y \in X)$ fixed, then $\overline{\phi}_{\mathcal{U}}$ is a constant, which assigns each sequence $z \in X^\infty$ the sequence $yyy\cdots$, which certainly does not meet the strong law of the large numbers. We show that it is sufficient to require "measure-bounded" for $\overline{\phi}_{\mathcal{U}}$, in which there is a $k \in \mathbb{N} \setminus \{0\}$, so that for all measurable sets $M \subseteq X^\infty$:

$$\overline{\mu}\Big(\overline{\phi}_{\mathcal{U}}^{-1}(M)\Big) \leq k\overline{\mu}(M).$$

Certainly, all functions $\overline{\phi}_{\mathcal{U}}$ generated by selection rules are measure-bounded (see also Doob); in addition, there are still a large number of additional measure-bounded functions. As an example, we introduce the permutations: an automaton reads in each of the n members, permutes them, and outputs them again. The following automaton performs this:

Let $\pi : X^n \to X^n$ ($n \in \mathbb{N}$ be fixed) be a permutation. $\mathcal{U} = (X, S, \delta, s_1)$ be defined by

$$S := \bigcup_{i=0}^{n} = \{u \in X^* : |u| \leq n\}, \quad s_1 := \Lambda$$

$$\delta(x,v) := \begin{cases} vx & \text{if } |v| < n \\ x & \text{if } |v| = n \end{cases}$$

and the output function $\lambda : X \times S \to X^*$

$$\lambda(x,v) := \begin{cases} \Lambda & \text{if } |v| < n \\ \pi(v) & \text{if } |v| = n. \end{cases}$$

It then applies

Proposition. 6.2.

(a) If $z \in X^{\infty}$ is a normal sequence and $\overline{\phi}_{\mathcal{U}}$ is a selecting function generated by a finite automaton $\mathcal{U}$ with an output function $\lambda : X \times S \to X^{*}$, then z is either not in the domain of $\overline{\phi}_{\mathcal{U}}$ or $\overline{\phi}_{\mathcal{U}}(z)$ follows the strong law of large number:

$$\lim_{n\to\infty} \frac{1}{n} \big\| \{ i \leq n : (\overline{\phi}_{\mathcal{U}}(z))_i = x \} \big\| = \frac{1}{p} \quad (x \in X).$$

(b) if $z \in X^{\infty}$ is not a normal sequence, then there is a computably continuous, measure-preserving function $\overline{\phi}_{\mathcal{U}} : X^{\infty} \to X^{\infty}$ generated by a finite state automaton so that $\overline{\phi}_{\mathcal{U}}(z)$ does not satisfy the strong law of large numbers.

Thus the partial, measure-bounded functions generated by finite automata preserve normality.

$\overline{\phi}_{\mathcal{U}}$ is measure-preserving if for all measurable set $M \subseteq X^{\infty}$

$$\overline{\mu}\big(\overline{\phi}_{\mathcal{U}}^{-1}(M)\big) = \overline{\mu}(M)$$

and computably continuous if there is a totally recursive function $h : \mathbb{N} \to \mathbb{N}$, so that for all z from the definition range of $\overline{\phi}_{\mathcal{U}}$, if for any $i \in \mathbb{N}$, we have

$$n \geq h(i) \implies \big|\overline{\phi}_{\mathcal{U}}\big(z(n)\big)\big| \geq i.$$

Proof. (a) Let $z \in X^{\infty}$ be in the domain of $\overline{\phi}_{\mathcal{U}}$. As shown in (4.1), we can only consider strongly connected machines. According to Lemma (4.5), each $x \in X$ has a fixed limiting frequency in each state $s \in S$. Thus the frequency of x in $\overline{\phi}_{\mathcal{U}}(z)$ has a limit. Let this be K. If $K \neq \frac{1}{p}$, then $\overline{\phi}_{\mathcal{U}}(z)$ is not a normal sequence for every normal sequence z. That is, the set of normal sequences (of measure 1) result a set of the measure 0 after selecting. $\overline{\phi}_{\mathcal{U}}(z)$ is therefore not measure-bounded, a contradiction to the assumption. Thus follows the assertion.

(b) The inversion is not difficult to prove (see Agafonoff [1]).

□

References

1. Agafonoff, V. N.: Normal sequence and finite automata. Soviet Math. Dokl. 9, 324-325 (1968) (engl. Übersetzung).
2. Böhling, K. H., Indermark, K.: Endliche automaten I. Mannheim: BI 1969.
3. Chaitin, G. J.: On the length of programs for computing finite binary sequences. Journ. ACM 13, 547-569 (1966).
4. Chung, K. L.: Markov chains with stationary transition probabilities, 2. edition. Berlin-Göttingen-Heidelberg: Springer 1967.
5. Copeland, A. H.: Admissible numbers in the theory of probability. Amer, Journ. Math. 50, 535-552 (1928).
6. Doob, J. L.: Note on probability. Annals of Math. 37, 363-367 (1936).

7. Hotz, G., Walter, H.: Automatentheorie und formale Sprachen. Mannheim: BI 1969.
8. Kolmogoroff, A. N.: Drei Zugänge zur Definition des Begriffs, *Infomationsgehalt* [russisch]. Prob. peredaci informacii 1, 3-11 (1965).
9. Loveland, D. W.: A variant of the Komogoroff concept of complexity. Inform. Control 15, 510-526 (1969).
10. Martin-Löf, P.: The definition of Random sequences. Inform. Control 6, 602-619 (1966).
11. Popper, K.: Logik der Forschung zur Erkenntnistheorie der modernen naturwissenschaft. Schiften zur wissenschaftlichen Weltauffassung, Bd. 9. Wien: Springer 1935.
12. Reichenbach, H.: Wahrscheinlichkeitslehre. Leiden: Sijthoff 1935.
13. Schnorr, C. P.: Zufälligkeit und Wahrscheinlichkeit. Lecture Notes. Berlin-Heidelberg-New York: Springer 1971.