INFORMATION AND CONTROL 21, 171–184 (1972)

Some Criteria for Determining Recognizability of a Set*

WALTER S. BRAINERD

Department of Mathematical Statistics, Columbia University, New York 10027

AND

RONALD B. KNODE

Department of Mathematics, U.S. Naval Postgraduate School, Monterey, California 93940

Let a_n be the number of strings of length n in a set $A \subseteq \Sigma^*$, where Σ is a finite alphabet. Several criteria for determining that a set is not recognizable by a finite automaton are given, based solely on the sequence $\{a_n\}$. The sequence $\{a_n\}$ is also used to define a finitely addititive probability measure on all recognizable sets.

1. INTRODUCTION

The most common proof that a set is not recognizable by a finite automaton uses a fundamental theorem which says that if a sufficiently long string x is accepted by a particular automaton, the string can be factored as $x = uyv$, where, for each $n \geqslant 0$, uy^nv will also be accepted by the automaton. Using this theorem to prove unrecognizability of a set requires some knowledge of the way symbols are arranged in the strings of the set, in order to prove that such a factorization cannot always be made.

Minsky and Papert (1966) and Cobham (1966, 1969) have developed other criteria for recognizability. This paper discusses some properties of recognizable sets based solely on the sequence $\{a_n\}$, where a_n is the number of strings of length n in a set. The connection between finite automata and Markov chains suggested by Hartmanis and Stearns (1967) is explored, providing a finitely additive measure on all recognizable sets, which can be

* This research was supported in part by the National Science Foundation under Grant Number GJ-28403 at Columbia University.

characterized in terms of the sequence $\{a_n\}$. The sequence $\{a_n\}$ also determines the generating function $f(z) = \sum_{n=0}^{\infty} a_n z^n$. The main result of the paper indicates some connections between a recognizable set A, the measure of A, its generating function, the rate of growth of the sequence $\{a_n\}$, and some properties of the minimal automaton recognizing A. These results can be used to show that certain sets are not recognizable.

Consider a problem which motivated this investigation. Suppose $+$ is a symbol for a binary operation and x is a variable symbol. The well-formed expressions involving $+$ and x can be represented as binary trees with label $+$ at the branch points and x at the end points, as shown in Fig. 1.

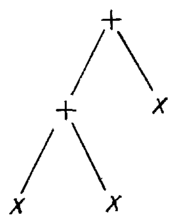

FIG. 1. Tree representation of $(x + x) + x$.

If any of the usual methods of writing these expressions is used, such as Polish prefix or infix with parentheses, the resulting set is context-free, but not recognizable by a finite automaton. See, for example, Brainerd (1969). Is there some way of writing the expressions so that the set of strings is recognizable?

There are $c_n = \binom{2n}{n}/(n + 1)$ trees with n $+$'s, which also have exactly $n + 1$ x's; see Knuth (1968). If A is a set of strings in which just the symbols $+$ and x appear and a_n is the number of strings of length n, then $a_{2n+1} = c_n$ and $a_{2n} = 0$. Theorem 4 will show, using this information alone, that the set A cannot be recognizable. Even if other symbols, such as parentheses, are used, then the set still cannot be recognizable, since recognizable sets are closed under the morphism which deletes the extra symbols; see Hopcroft and Ullman (1969).

Let $\Sigma = \{1, 2,..., k\}$ be a finite alphabet and let Σ^* denote the set of finite strings over Σ. Let $|x|$ denote the length of $x \in \Sigma^*$. We are primarily interested in the case where $\#\Sigma = k > 1$, i.e., Σ contains at least two letters. The case $\#\Sigma = 1$ will be discussed in Section 7.

All automata discussed will be *deterministic minimal finite* automata. If the machine is in state q, then qx will represent the unique state of the automaton after reading input $x \in \Sigma^*$. The initial state will always be q_1, and the set recognized by the automaton is $\{x \in \Sigma^* \mid q_1 x \in F\}$, where F is the set of *final* states.

A state q is a *dead* (or *sink*) state if $q \notin F$ and $q\sigma = q$ for each $\sigma \in \Sigma$. A minimal automaton can have at most one dead state.

There is a convenient bijection between Σ^* and the natural numbers given by

$$\nu(\sigma_n \cdots \sigma_1 \sigma_0) = \sum_{i=0}^{n} \sigma_i \times k^i$$

The null sting ϵ corresponds to zero, i.e., $\nu(\epsilon) = 0$. Note that this is not the usual k-ary notation which, due to leading zeros, does not define a bijection between $\{0, 1,..., k-1\}^*$ and the natural numbers.

The following definitions are slight modifications of those in Minsky and Papert (1966). For $n > 0$ and $A \subseteq \Sigma^*$, let $\alpha_n = \{x \in A \mid \nu(x) < n\}$. For the set Σ^*, $\alpha_n = n$, so, in general, α_n/n is the proportion of strings x in A with $\nu(x) < n$ and $\lim_{n\to\infty} \alpha_n/n$, if it exists, is a measure of the set A. Minsky and Papert proved, among other things, that if $\lim_{n\to\infty} \alpha_n/n = 0$, then a minimal automaton recognizing A must have a dead state. This will be generalized slightly by Theorem 4. A difficulty is that $\lim_{n\to\infty} \alpha_n/n$ does not exist even for some very simple recognizable sets. If $A = (\Sigma^2)^* = \{x \in \Sigma^* \mid |x| \text{ is even}\}$, then there are subsequences of α_n/n converging to $\frac{1}{3}$ and $\frac{2}{3}$, thus $\lim_{n\to\infty} \alpha_n/n$ does not exist. A measure will now be defined on all recognizable sets, which is equal to $\lim_{n\to\infty} \alpha_n/n$ whenever the limit exists.

2. A Measure for Recognizable Sets

Each deterministic finite automaton with s states determines an $s \times s$ stochastic matrix M, where $M_{ij} = \#\{\sigma \in \Sigma \mid q_i\sigma = q_j\}/k$. In other words a Markov chain is obtained by treating the input letters as being generated by independent Bernoulli trails with the probability of each letter equal to $1/k$. Thus, if the machine is in state q and an input letter is generated, the machine will enter state $q\sigma$, $1 \leqslant \sigma \leqslant k$, with probability $1/k$. The result of a computation will be uncertain due to the random input, but the machine itself still operates in a deterministic way.

The terminology used here is that of Doob (1953) and Kemeny and Snell (1960). A Markov chain induced by a minimal automaton will have some special properties. Let T be the set of transient states. If there are any transient states, then the initial state q_1 must be one of them. If there are no transient states, i.e., $T = \varnothing$, then M itself must be ergodic.

A measure $\mu(A)$ of any recognizable set may be calculated as the probability that the Markov chain induced by the minimal automaton accepting A is in a final state.

The following facts concerning stochastic matrices will be used [see, for example, Doob (1953)].

The value of $(M^n)_{ij}$ is the probability of being in state j after n steps, if initially in state i. Let $p_j^{(n)} = (M^n)_{1j}$ be the probability of being in state j after n steps, starting in the initial state. Note that

$$a_n/k^n = \sum_{j \in F} p_j^{(n)}. \tag{2.1}$$

If j is a transient state ($j \in T$), then

$$\lim_{n\to\infty} p_j^{(n)} = 0. \tag{2.2}$$

If j is not a transient state ($j \in \bar{T}$), then there are integers $t_j > m_j \geqslant 0$, such that

$$\lim_{n\to\infty} p_j^{(nt_j+m)} = \begin{cases} \pi_j > 0 & \text{if } \; m \equiv m_j (\operatorname{mod} t_j), \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

Hence the Cesaro limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} p_j^{(n)} = \pi_j/t_j . \tag{2.4}$$

The measure $\mu(A)$ is thus defined and characterized by the equations

$$\begin{aligned} \mu(A) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} (a_n/k^n) \qquad (2.5) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{N} \sum_{j \in F} p_j^{(n)} \\ &= \sum_{j \in F \cap \bar{T}} (\pi_j/t_j). \end{aligned}$$

Kemeny and Snell (1960) give methods for calculating $\mu(A)$. The essential step is the inversion of a matrix which will always be nonsingular.

THEOREM 1. *For each recognizable set* A, *the measure*

$$\mu(A) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{\infty} (a_n/k^n)$$

exists and satisfies the following properties:

a. $0 \leqslant \mu(A) \leqslant 1$

b. $\mu(A \cup B) = \mu(A) + \mu(B)$ *if* $A \cap B = \varnothing$, *i.e.,* μ *is finitely additive. More generally,* $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

c. $\mu(A) = 0$ *if* A *is finite, i.e.,* μ *is diffuse. More generally* $\mu(A \cup B) = \mu(B)$.

d. $\mu(\Sigma^* - A) = 1 - \mu(A)$.

Proof. For part (b), if $C = A \cup B$ and $D = A \cap B$, then $c_n = a_n + b_n - d_n$. For part (c), if A is finite, then all final states are transient states and so $\mu(A) = 0$ by (2.5). ∎

The measure μ has other intuitively appealing properties. For example, if $\#\Sigma = k$, then $\mu(\sigma \cdot \Sigma^*) = \mu(\Sigma^* \cdot \sigma) = 1/k$, for each letter σ. Also $\mu(\Sigma^i)^* = 1/i$. This last example is the one cited in Section 1 for which the Minsky–Papert measure does not exist. Tsichritzis (1969) discusses other interesting properties of measures on countable sets.

THEOREM 2. *If the Minsky–Papert measure* $\lim_{n\to\infty} (\alpha_n/n)$ *exists for a recognizable set* A, *then it is equal to* $\mu(A)$. *Indeed, if* $\alpha_n/n \to p$, *then* $a_n/k^n \to p$.

Proof. Let $c_n = \nu(1^n)$, the number whose k-ary representation is a string of n 1's. Then $a_n = \alpha_{c_{n+1}} - \alpha_{c_n}$, and so

$$\begin{aligned} a_n/k^n &= (\alpha_{c_{n+1}} - \alpha_{c_n})/k^n \\ &= (\alpha_{c_{n+1}}/c_{n+1}) \times (k^{n+1} - 1)/k^n(k-1) \\ &\quad - (\alpha_{c_n}/c_n) \times (k^n - 1)/k^n(k-1) \\ &\to p(k^{n+1} - k^n)/k^n(k-1) \\ &= p, \end{aligned}$$

since the subsequence $\{\alpha_{c_n}/c_n\}$ must also converge to p. ∎

3. The Rate of Growth of a_n

If Σ has k letters then $\lim_{n\to\infty} a_{n+1}/a_n \leqslant k$, for any set A. It might be conjectured that if the measure $\mu(A)$ of a recognizable set is zero, then the maximal rate of growth cannot be achieved. This will be shown to be true, but first, we will show that there are recognizable sets with $\mu(A) = 0$ such that $\lim_{n\to\infty} a_{n+1}/a_n$ is arbitrarily close to k.

For a given number s of states, there is an interesting automaton which, it is conjectured, has the maximum $\lim_{n\to\infty} a_{n+1}/a_n$ of all machines with s states such that $\mu(A) = 0$, i.e., of all machines having a dead state as the only ergodic state. For simplicity, let $k = 2$. This machine is nicely represented by the tree in Fig. 2.

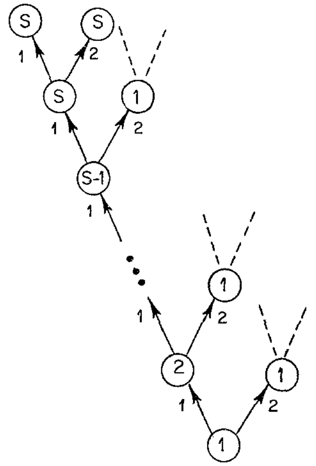

FIG. 2. The s-state Fibonacci machine.

All states are final except the dead state s. It is easily verified that for $0 \leqslant n < s$, $a_n = 2^n$ and for $n \geqslant s$, $a_n = a_{n-1} + a_{n-2} + \cdots + a_{n-(s-1)}$. Thus a_n is a generalized Fibonacci sequence, and for $s = 3$, it is the familiar sequence 1, 2, 3, 5, 8, 13,... . For this reason, call this machine the s-state Fibonacci machine. As in Alfred (1965), $x = \lim_{n\to\infty} a_{n+1}/a_n$ may be calculated by solving the equation $P_s(x) = x^{s-1} - \cdots - x - 1 = 0$. (For $s = 3$, $a_n = a_{n-1} + a_{n-2}$ implies $(a_n/a_{n-2}) \times (a_{n-1}/a_{n-1}) = (a_{n-1}/a_{n-2}) + (a_{n-2}/a_{n-2})$ and taking the limit of both sides yields $x \cdot x = x + 1$ or $x^2 - x - 1 = 0$, whose positive solution is the golden ratio $(1 + \sqrt{5})/2$.) Now $P_s(1) = 1 - (s - 1) < 0$, for $s > 2$, and $P_s(2) = 1$, which means that there is a solution between 1 and 2. Furthermore, for $x > 2$,

$$1 > \frac{1}{x-1} = \frac{1/x}{1-1/x} = \sum_{i=1}^{\infty} x^{-i} > \sum_{i=1}^{s-1} x^{-i}.$$

Thus

$$x^{s-1} > x^{s-1} \times \sum_{i=1}^{s-1} x^{-i} = \sum_{i=1}^{s-1} x^{s-i-1} = x^{s-2} + \cdots + x + 1.$$

Hence $P_s(x) > 0$ for all $x > 2$. Since all solutions of $P_s(x)$ are less than 2, that means $\lim_{n\to\infty} a_{n+1}/a_n < 2$.

THEOREM 3. *For every* δ, $0 < \delta < 1$, *there is a recognizable set* $A \subseteq \{1, 2\}^*$ *such that* $\mu(A) = 0$ *and* $\lim_{n\to\infty} a_{n+1}/a_n = r$, *where* $2 - \delta < r < 2$.

Proof.

$$\sum_{i=1}^{\infty} (2-\delta)^{-i} = \left(\frac{1}{2-\delta}\right) \Big/ \left(1 - \frac{1}{2-\delta}\right) = 1/(1-\delta) > 1.$$

Choose s large enough so that $\sum_{i=1}^{s-1} (2-\delta)^{-i} > 1$. Then

$$\sum_{i=1}^{s-1} (2-\delta)^{s-i-1} > (2-\delta)^{s-1}.$$

Thus

$$P_s(2-\delta) = (2-\delta)^{s-1} - (2-\delta)^{s-2} - \cdots - (2-\delta) - 1 < 0.$$

Since $P_s(2) > 0$, the solution $r = \lim_{n\to\infty} a_{n+1}/a_n$ for the set A accepted by the s-state Fibonacci machine must be between $2 - \delta$ and 2. The measure $\mu(A) = 0$, because all final states are transient states. ∎

4. GENERATING FUNCTIONS

The sequence $a_n = \#\{x \in A \mid |x| = n\}$ of a set A determines the generating function $f(z) = \sum_{n=0}^{\infty} a_n z^n$. Cobham (1966) and Kuich (1970) show that the generating function of any recognizable set must be rational. Let ρ_A be the radius of convergence of the power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$. If $A \subseteq B$, then $a_n \leqslant b_n$ and so $\rho_A \geqslant \rho_B$. For the set $\Sigma^* = \{1, 2,\ldots, k\}^*$, $f(z) = \sum_{n=0}^{\infty} k^n z^n$, hence $\rho_{\Sigma^*} = 1/k$. Thus, for any set $A \subseteq \Sigma^*$, $\rho_A \geqslant 1/k$.

In the next section, relationships between the value of ρ_A and other things discussed previously are proved. Some additional facts about generating functions are needed. Let $g(z) = \sum_{n=0}^{\infty} b_n z^n$ be the generating function for a set B. If $A \cup B$, AB, and A^* are unambiguous sets (i.e., A and B are disjoint; $x \in AB$ can be factored in only one way as $x = uv$, where $u \in A$, $v \in B$; and each $x \in A^*$ can be factored uniquely as $x = x_1 \cdots x_n$, where $x_1 \in A$, $1 \leqslant 1 \leqslant n$), then the generating functions for $A \cup B$, AB, and A^* are $f(z) + g(z)$, $f(z) \times g(z)$, and $1/[1 - f(z)]$, respectively; see Kuich (1970).

5. Criteria for Recognizability

THEOREM 4. *Let $A \subseteq \Sigma^*$, with $\#\Sigma = k$, be recognized by a finite automaton. Let Q be the states of the minimal automaton $\mathfrak{A}$ which recognizes A. Let $a_n = \{x \in A \mid |x| = n\}$, $\rho_A =$ the radius of convergence of $\sum_{n=0}^{\infty} a_n z^n$, and $\mu(A) = \lim_{N\to\infty} 1/N \sum_{n=1}^{N} (a_n/k^n)$. Then the following are equivalent:*

a. *$\mathfrak{A}$ has a dead state d accessible from each state $q \in Q$.*

a′. $\forall q\, \exists x (qx = d)$

a″. $\forall y\, \exists x\, \forall z (yxz \notin A)$

b. $\exists x\, \forall y\, \forall z (yxz \notin A)$

b′. $\exists x\, \forall q (qx = d)$

b″. $\exists x (\Sigma^* x \Sigma^* \cap A = \varnothing)$

c. $\rho_A > 1/k$

d. $a_n/k^n \to 0$

e. $\mu(A) = 0$

f. *There is no subsequence of $\{a_n\}$ of the form $\{a_{nt+m}\}_{n=0}^{\infty}$, $t > m \geqslant 0$, such that* $\lim_{n\to\infty} a_{(n+1)t+m}/a_{nt+m} = k^t$.

Proof. a, a′, and a″, as well as b, b′, and b″ are obviously equivalent. The chain of implications a′ ⇒ b′ ⇒ b″ ⇒ c ⇒ d ⇒ e ⇒ a ⇒ f ⇒ e will be proved.

a′ ⇒ b′: Let $Q = \{q_1, ..., q_s\}$ be the states of $\mathfrak{A}$. Assume $q_i x_i = d$, $1 \leqslant i \leqslant s$.

$$
\begin{aligned}
&\text{Let } q_2 x_1 = q_{i_2}, && \text{then } q_2 x_1 x_{i_2} = d\\
&\text{Let } q_3 x_1 x_{i_2} = q_{i_3}, && \text{then } q_3 x_1 x_{i_2} x_{i_3} = d\\
&\quad \cdots\\
&\text{Let } q_s x_1 x_{i_2} \cdots x_{i_{s-1}} = q_{i_s}, && \text{then } q_{i_s} x_1 x_{i_2} \cdots x_{i_s} = d\\
&\text{Let } x = x_1 x_{i_2} \cdots x_{i_s}, && \text{then } \forall q\ qx = d.
\end{aligned}
$$

b″ ⇒ c: The technique used in this proof was provided by S. Eilenberg. Assume $\Sigma^* w \Sigma^* \cap A = \varnothing$, where $w = \sigma_1 \sigma_2 \cdots \sigma_m$. For each $\sigma \in \Sigma$, let $\bar{\Sigma}_\sigma = \Sigma - \{\sigma\}$. Let $A' = \Sigma^* - \Sigma^* w \Sigma^* \supseteq A$, so that $\rho_A \geqslant \rho_{A'}$. Let

$$B = (\bar{\Sigma}_{\sigma_1}) \cup (\sigma_1 \bar{\Sigma}_{\sigma_2}) \cup (\sigma_1 \sigma_2 \bar{\Sigma}_{\sigma_3}) \cup \cdots \cup (\sigma_1 \sigma_2 \cdots \sigma_{m-1} \bar{\Sigma}_{\sigma_m})$$

and let $C = \{\epsilon, \sigma_1, \sigma_1\sigma_2, ..., \sigma_1\sigma_2 \cdots \sigma_{m-1}\}$. For example, if $w = 121$ and $k = 3$, then

$$B = \{2, 3, 11, 13, 122, 123\} \quad \text{and} \quad C = \{\epsilon, 1, 12\}.$$

Now $A' = \Sigma^* - \Sigma^* w \Sigma^* \subseteq BA \cup C$, since if $x \in A'$, then either $x \in C$ or x consists of a member of B followed by a string in which w does not occur.

LEMMA. *If $X \subseteq BX \cup C$ and $\epsilon \notin B$, then $X \subseteq B^*C$.*

Proof.

$$\begin{aligned} X &\subseteq BX \cup C \subseteq B(BX \cup C) \cup C = B^2X \cup BC \cup C \\ &\subseteq B^3X \cup B^2C \cup BC \cup C \subseteq \cdots. \end{aligned}$$

By induction $X \subseteq B^{n+1}A \cup B^nC \cup \cdots \cup BC \cup C$, for any $n \geqslant 1$. Let $x \in X$, $|x| = n$, then $x \notin B^{n+1}X$, since $y \in B^{n+1}$ implies $|y| \geqslant n + 1$. Thus $x \in B^nC \cup \cdots \cup BC \cup C \subseteq B^*C$. (This is part of the proof of the well-known result that if $\epsilon \notin B$, then the unique solution of $X = BX \cup C$ is $X = B^*C$.)

Since $A' \subseteq BA' \cup C$, $A' \subseteq B^*C$, by the lemma. Thus $\rho_A \geqslant \rho_{A'} \geqslant \rho_{B^*C}$. The generating function for B is $k'(z + z^2 + \cdots + z^m) = k'z(z^m - 1)/(z - 1)$, where $k' = k - 1 = \#\bar{\Sigma}_\sigma$, and the generating function for C is $1 + z + \cdots + z^{m-1} = (z^m - 1)/(z - 1)$. Thus $g(z) = h(z)/[1 - k'zh(z)]$, where $h(z) = (z^m - 1)/(z - 1) = (1 - z^m)/(1 - kz + k'z^{m+1})$. Let $D(z) = 1 - kz + k'z^{m+1}$, then $D(0) = 1$ and $D(1/k) = k'/k^{m+1} > 0$. The derivative $D'(z) = (m + 1)\, k'z^m - k$. For $0 \leqslant z \leqslant 1/k$,

$$D'(z) \leqslant (m + 1)\, k'/k^m - k < k[(m + 1)/k^m - 1] \leqslant 0.$$

Thus, since $D'(z) < 0$ for $0 \leqslant z \leqslant 1/k$, $D(z) > 0$, for $0 \leqslant z \leqslant 1/k$ and thus $\rho_A \geqslant \rho_{A'} \geqslant \rho_{B^*C} > 1/k$.

c $\Rightarrow$ d: If $\sum_{n=0}^{\infty} a_n(1/k)^n$ converges, then $a_n/k^n \to 0$.

d $\Rightarrow$ e: If $a_n/k^n \to 0$, then

$$\mu(A) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} a_n/k^n = \lim_{n \to \infty} a_n/k^n = 0.$$

e $\Rightarrow$ a: If $\mu(A) = 0$, then each final state of $\mathcal{Cl}$ must be a transient state by (2.5). Any two nontransient nonfinal states would be equivalent dead states; thus there is exactly one nontransient state (the dead state) which must be accessible from every state, since at least one nontransient state is accessible from each transient state in any Markov chain.

a $\Rightarrow$ f: Let x be a string of length $nt + m$ in A, where $n \geqslant 0$, $t > 0$, and $m \geqslant 0$ are arbitrary. Let i be the smallest integer such that

$$(n + 1)\, t + m > s = \#Q.$$

By assumption, there is a y such that $q_1 xy = d$, the dead state, and y may be picked so that $|\, y \,| < s = \#Q$. Thus $xyz \notin A$, for all z. Thus

$$\#\{w \in A \mid |\, xw \,| = (n + i)\, t + m\} \leqslant k^{it} - 1;$$

hence $a_{(n+i)t+m}/a_{nt+m} \leqslant k^{it} - 1$. Hence

$$\begin{aligned}
&(\lim_{n\to\infty} a_{(n+1)t+m}/a_{nt+m})^i \\
&\quad = \lim_{n\to\infty} a_{(n+1)t+m}/a_{nt+m} \times \lim_{n\to\infty} a_{(n+2)t+m}/a_{(n+1)t+m} \times \cdots \\
&\qquad \times \lim_{n\to\infty} a_{(n+i)t+m}/a_{(n+i+1)t+m} \\
&\quad = \lim_{n\to\infty} a_{(n+i)t+m}/a_{nt+m} < k^{it},
\end{aligned}$$

and so

$$\lim_{n\to\infty} a_{(n+1)t+m}/a_{nt+m} < k^t, \qquad \text{if the limit exists.}$$

It should be noted that this result does not follow from only the fact that the automaton has a dead state. It is crucial that the dead state be accessible from every state. For the set $A = 1\Sigma^*$, $\mu(A) = \frac{1}{2}$ and A is recognized by the minimal automaton in Fig. 3 which has a dead state.

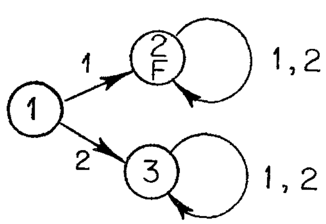

FIG. 3. An automaton with a dead state and $\mu(A) > 0$.

f $\Rightarrow$ e: The contrapositive is proved. Assume $\mu(A) > 0$. Then by (2.5) and (2.3), there is a final nontransient state $j \in F \cap \bar{T}$ and numbers $0 \leqslant m < t_j$ such that $\lim_{n\to\infty} p_j^{(nt_j+m)} = \pi_j > 0$. Let $t = \prod_{i\in F\cap\bar{T}} t_i$, where t_i is the number given by (2.3) for each final nontransient state. Then

$$L = \lim_{n\to\infty} a_{nt+m}/k^{nt+m} = \lim_{n\to\infty} \sum_{i\in F\cap\bar{T}} p_i^{(nt+m)}$$

which exists by (2.4) and the choice of t. Furthermore,

$$L \geqslant \lim_{n\to\infty} p_j^{(nt+m)} = \pi_j > 0.$$

Thus

$$\lim_{n\to\infty} a_{(n+1)t+m}/a_{nt+1}$$
$$= \lim_{n\to\infty} [a_{(n+1)t+m}/k^{(n+1)t+m}] \times [k^{nt+m}/a_{nt+m}] \times (k^t) = (L/L) \times k^t = k^t,$$

which contradicts condition f. ∎

6. Examples

Any set which satisfies one of the conditions in Theorem 4 but does not satisfy one of the other conditions must not be recognizable by a finite automaton. Several examples will now be given. In each example, assume $\Sigma = \{1, 2\}$.

Let $A = \{x \cdot \overleftarrow{x} \mid x \in \Sigma^*\}$, where $\overleftarrow{x}$ is the string x reversed. For this set $a_{2n} = 2^n$ and $a_{2n+1} = 0$. Thus $\mu(A) = \lim_{n\to\infty} 2^n/2^{2n} = 0$, satisfying condition e. However, A violates condition b″, since for each x, $\epsilon \cdot x \cdot \overleftarrow{x} \in A$.

Let $A = \{x \mid \nu(x) \text{ is a prime}\}$. For the set of primes $\alpha_n \sim n/\log n$ and so $\alpha_{2^n} \sim 2^n/n \log 2$. Thus,

$$a_n \sim 2^{n+1}/(n+1)\log 2 - 2^n/n \log 2,$$
$$a_{n+1}/a_n \sim (2^{n+2} - 2^{n+1})/(2^{n+1} - 2^n) = 2,$$

which contradicts condition f. However,

$$\begin{aligned} a_n/2^n &= [2/(n+1) - 1/n]/\log 2 \\ &= (n-1)/(n^2+n) \times \log 2 \\ &\to 0 \end{aligned}$$

and so A satisfies condition d.

It is of interest that the set of primes also does not satisfy b″; in fact, for each x, there are infinitely many y such that $\nu(xy)$ is prime; see Sierpinsky (1959).

Let A be the set of strings representing well-formed expressions in a binary operator and variable symbol. As was stated in the introduction,

$$a_{2n+1} = c_n = \frac{1}{n+1}\binom{2n}{n} \sim \frac{1}{n+1}\frac{\sqrt{4\pi n}\,(2n/e)^{2n}}{[\sqrt{2\pi n}\,(n/e)^n]^2}$$

$$\sim \frac{1}{n}\,\frac{2^{2n+1}\pi^{1/2}n^{2n+1/2}e^{-2n}}{2\pi n^{2n+1}e^{-2n}} = \frac{2^{2n}}{\sqrt{\pi n^3}}\,.$$

Thus

$$\lim_{n\to\infty}\frac{a_{2n+1}}{2^{2n+1}} = \lim_{n\to\infty}\frac{c_n}{2^{2n+1}} = \lim_{n\to\infty}\frac{1}{2\sqrt{\pi n^3}} = 0.$$

Since $a_{2n} = 0$, $\lim_{n\to\infty} a_n/2^n = 0 = \mu(A)$. On the other hand,

$$\lim_{n\to\infty}\frac{a_{2n+3}}{a_{2n+1}} = \lim_{n\to\infty}\frac{c_{n+1}}{c_n} = \lim_{n\to\infty}\frac{2^{2n+2}}{\sqrt{\pi(n+1)^3}}\;\frac{\sqrt{\pi n^3}}{2^{2n}} = 2^2 = 4.$$

Thus A contradicts condition f with $t = 2$ and $m = 1$, so that it cannot be a recognizable set.

There are two other interesting proofs that the expressions involving $+$ and x cannot be written as a recognizable set of strings. If $a_{2n+1} = c_n$ and $a_n = 0$, then $f(z) = [1 - \sqrt{(1-4z^2)}]/2z$, which is not rational. Hence, by Cobham (1966) and Kuich (1970), A cannot be recognizable. A proof in Brainerd (1969) uses the theory of runs and the fact that in each string of length $2n + 1$, there are n $+$'s and $n + 1$ x's.

Another closely related set is the set of balanced parentheses over the alphabet $\Sigma = \{(,)\}$. For this set $a_{2n} = c_n$ and $a_{2n+1} = 0$. This set is also context-free, but not recognizable by Theorem 4.

7. The Case of the One Letter Alphabet

Any deterministic automaton over $\Sigma = \{1\}$ must be of the form shown in Fig. 4.

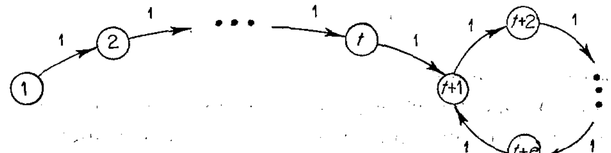

FIG. 4. An automaton accepting a subset of $\{1\}^*$.

The set of transient states is $T = \{1, 2,..., t\}$, $t \geqslant 0$, and $\bar{T} = \{t + 1,..., t + e\}$, $e \geqslant 1$ is a single ergodic set. The sequence $a_n/1^n = a_n$ is Cesaro summable to $\mu(A)$. The stationary probability for each nontransient state is $1/e = 1/\#\bar{T}$, hence $\mu(A) = \#(F \cap \bar{T})/\#\bar{T}$. In the one-letter case,

$$\alpha_n = \sum_{i=0}^{n-1} a_i \text{ ; hence } \lim_{n\to\infty} \alpha_n/n = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} a_n = \mu(A).$$

Thus the Minsky–Papert measure coincides with the measure $\mu(A)$ and always exists in this case. Since a_n is an ultimately periodic sequence of zeros and ones, each of the conditions in Theorem 4 asserts that A is finite and so are also equivalent in the case $k = 1$.

RECEIVED: August 2, 1971

REFERENCES

1. ALFRED, U., "An Introduction to Fibonacci Discovery," The Fibonacci Association, San Jose, CA, 1965.
2. BRAINERD, W. S. (1969), Semi-Thue systems and representations of trees, *in* "Conference Record of the 1969 Tenth Annual Symposium on Switching and Automata Theory," pp. 240–244.
3. COBHAM, A. (1966), Sets definable by finite automata, IBM Research Notes NC-415, NC-458, NC-577, and NC-665.
4. COBHAM, A. (1969), On the base-dependence of sets of numbers recognizable by finite automata, *Math. Systems Theory* **3**, 186–192.
5. DOOB, J. L. (1953), "Stochastic Processes," Wiley, New York.
6. FELLER, WILLIAM (1957), "An Introduction to Probability Theory and Its Applications," Wiley, New York.
7. HARRISON, M. A. (1966), "Introduction to Switching and Automata Theory," McGraw-Hill, New York.
8. HARTMANIS, J. AND R. E. STEARNS (1967), Sets of numbers defined by finite automata, *American Math. Monthly* **74**, 539–542.
9. HOPCROFT, J. E. AND J. D. ULLMAN (1969), "Formal Languages and Their Relation to Automata," Addison–Wesley, Cambridge, MA.
10. KEMENY, J. G. AND J. L. SNELL (1960), "Finite Markov Chains," Van Nostrand, Princeton, NJ.
11. KNODE, R. B. (1969), On the unrecognizability of sets, Master's thesis, Naval Postgraduate School.
12. KNUTH, D. E (1968), "The Art of Computer Programming: Fundamental Algorithms," Addison–Wesley, Cambridge, MA.
13. KUICH, W. (1970), On the entropy of context-free languages, *Information and Control* **16**, 173–200.

14. MINSKY, M. AND S. PAPERT (1966), Unrecognizable sets of numbers, *JACM* **13**, 281–286.
15. SIERPINSKY, W. (1959), Sur les nombres Premiers ayant des chiffres initiaux et finals donnés, *Acta Arith.* **5**, 265–266.
16. TSICHRITZIS, D. (1969), Measures on countable sets, Technical Report No. 8, Dept. of Computer Science, University of Toronto.

Printed in Belgium