ON THE RECOGNITION OF PRIMES BY AUTOMATA

J. Hartmanis

H. Shank

TR 68-1

On the Recognition of Primes by Automata

J. HARTMANIS AND H. SHANK

Cornell University,[*] *Ithaca, New York*

ABSTRACT. A study of the problem of recognizing the set of primes by automata is presented. A simple algebraic condition is derived which shows that neither the set of primes nor any infinite subset of the set of primes can be accepted by a pushdown or finite automaton.

In view of this result an interesting open problem is to determine the "weakest" automaton which can accept the set of primes. It is shown that the linearly bounded automaton can accept the set of primes, and it is conjectured that no automaton whose memory grows less rapidly can recognize the set of primes. One of the results shows that if this conjecture is true, it cannot be proved by the use of arguments about the distribution of primes, as described by the Prime Number Theorem. Some relations are established between two classical conjectures in number theory and the minimal rate of memory growth of automata which can recognize the set of primes.

KEY WORDS AND PHRASES: automata, context-free languages, memory requiroments, primes, regular sets, Turing machines

CR CATEGORIES: 5.22, 5.29, 5.30

1. *Introduction*

Several interesting results and problems in automata theory have originated in attempts to characterize the computational power of automata which generate or recognize sets of numbers that have been studied extensively in mathematics. Results of this type dealing with the algebraic and transcendental numbers, as well as with primes and perfect squares, are described in [1–7]. Among the sets of integers studied in mathematics, there are several that have well-known density properties. One such set is the set of primes whose density is described by the celebrated

PRIME NUMBER THEOREM. *If $\pi(n)$ denotes the number of primes not larger than n, then*

$$\lim_{n\to\infty} \pi(n)\cdot\frac{\ln n}{n} = 1.$$

Thus to show that a certain automaton cannot accept the set of primes it is sufficient to show that it cannot accept any set of numbers with the above density. This approach has recently been used by Minsky and Papert [5], who derived several properties about the density of sets of numbers accepted by finite automata and using these results showed that a finite automaton cannot accept the set of primes.

In this paper a simple algebraic condition is derived which shows that no infinite set of primes can be accepted by a pushdown automaton. This answers a problem raised by Minsky and Papert [5] and strengthens one of their results. The same result using different techniques was derived independently by Schützenberger [7].

[*] Department of Computer Science. This research has been supported in part by National Science Foundation Grant GP 6120.

Journal of the Association for Computing Machinery, Vol. 15, No. 3, July 1968, pp. 382–389.

An interesting open problem is to determine how much memory is required to recognize the set of primes. We show that a linearly bounded automaton can recognize the set of primes and conjecture that no automaton whose memory grows more slowly can recognize the set of primes. One of our results shows that if this conjecture is true, it cannot be proven by the use of density arguments. We also establish some relationships between some classic conjectures in number theory and the rate of memory growth of automata which can recognize the set of primes.

2. *Derivation of Algebraic Result*

In this section we derive a divisibility result about sets of numbers which have a periodic part in their representation.

Let α_i denote a binary string and $\underline{\alpha_i}$ the integer represented by the binary string α_i. Let α_i^q denote the binary string obtained by concatenating q times, $q > 0$, the string α_i; and let $l(\alpha_i)$ denote the length (number of digits) of the string α_i.

In the following proof we make use of a special case of Fermat's Theorem, which states:

If p is a prime and $p > 2$, then $2^{p-1} \equiv 1 \pmod p$.

THEOREM 1. *Let p be a prime with $p > 2$ and let α_1, α_2, α_3 be binary strings (α_1 starting with a one) such that $2^{l(\alpha_2)} \not\equiv 1 \pmod p$. Then $\underline{\alpha_1\alpha_2^{p-1}\alpha_3} \equiv \underline{\alpha_1\alpha_3} \pmod p$.*

PROOF. Observe that

$$\begin{aligned}\underline{\alpha_1\alpha_2^{p-1}\alpha_3} &= \underline{\alpha_3} + 2^{l(\alpha_3)}\underline{\alpha_2^{p-1}} + 2^{l(\alpha_3)+(p-1)l(\alpha_2)}\underline{\alpha_1}\\ &= \underline{\alpha_3} + 2^{l(\alpha_3)}[\underline{\alpha_2} + 2^{l(\alpha_2)}\underline{\alpha_2} + \cdots + 2^{(p-2)l(\alpha_2)}\underline{\alpha_2}] + 2^{l(\alpha_3)+(p-1)l(\alpha_2)}\underline{\alpha_1}\\ &= \underline{\alpha_3} + 2^{l(\alpha_3)}\underline{\alpha_2}\left[\frac{2^{l(\alpha_2)(p-1)} - 1}{2^{l(\alpha_2)} - 1}\right] + 2^{l(\alpha_3)+(p-1)l(\alpha_2)}\underline{\alpha_1}.\end{aligned}$$

By Fermat's theorem we obtain that $2^{l(\alpha_2)(p-1)} \equiv 1 \pmod p$, and therefore in the last equation the middle term is congruent to zero and a factor of the last term is congruent to one modulo p. Thus $\underline{\alpha_1\alpha_2^{p-1}\alpha_3} \equiv \underline{\alpha_3} + 2^{l(\alpha_3)}\underline{\alpha_1} \equiv \underline{\alpha_1\alpha_3} \pmod p$, as was to be shown.

Using the same technique and expanding two terms in a geometric series to show that they are zero, we obtain our next result.

COROLLARY 1. *Let p be a prime larger than two and let the binary strings α_1, α_2, α_3, α_4, α_5 be such that $2^{l(\alpha_2)} \not\equiv 1 \pmod p$ and $2^{l(\alpha_4)} \not\equiv 1 \pmod p$. Then*

$$\underline{\alpha_1\alpha_2^{p-1}\alpha_3\alpha_4^{p-1}\alpha_5} \equiv \underline{\alpha_1\alpha_3\alpha_5} \pmod p.$$

We now restate our results in a form which can be used directly to show that no infinite subset of the set of primes can be accepted by a pushdown or finite automaton.

COROLLARY 2. *If $p = \underline{\alpha_1\alpha_2^r\alpha_3}$ is a prime larger than two and $2^{l(\alpha_2)} \not\equiv 1 \pmod p$, then*

$$\underline{\alpha_1\alpha_2^r\alpha_2^{p-1}\alpha_3} \equiv \underline{\alpha_1\alpha_2^r\alpha_3} \equiv 0 \pmod p.$$

COROLLARY 3. *If $p = \underline{\alpha_1\alpha_2^k\alpha_3\alpha_4^k\alpha_5}$ is a prime larger than two, and $2^{l(\alpha_2)} \not\equiv 1 \pmod p$ and $2^{l(\alpha_4)} \not\equiv 1 \pmod p$, then*

$$\underline{\alpha_1\alpha_2^k\alpha_2^{p-1}\alpha_3\alpha_4^k\alpha_4^{p-1}\alpha_5} \equiv \underline{\alpha_1\alpha_2^k\alpha_3\alpha_4^k\alpha_5} \equiv 0 \pmod p.$$

3. *Application to Automata Theory*

It is known [8] that for every regular set R there exists a positive integer q such that w in R and $l(w) \geq q$ implies that $w = w_1w_2w_3$ with $w \neq \Lambda$ and $w_1w_2^kw_3$ is in R, for $k = 1, 2, \cdots$.

Similarly, it is known [8] that for every context-free language L there exists a q such that w in L and $l(w) \geq q$ implies that $w = w_1w_2w_3w_4w_5$ with $w_2 \neq \Lambda$ or $w_4 \neq \Lambda$ and $w_1w_2^kw_3w_4^kw_5$ is in L, for $k = 1, 2, \cdots$. Furthermore, we recall that the sets accepted by finite-state machines are regular sets and vice versa. Similarly, it is known that context-free languages are characterized as the sets of strings accepted by (nondeterministic) pushdown automata.

We now combine these results with our corollaries to show that no pushdown automaton and, therefore, no finite automaton can accept an infinite subset of the set of primes.

THEOREM 2. *If P is an infinite subset of* $1(0 + 1)^*$ *(the set of all finite binary strings starting with one) and $w \in P$ implies that $\underline{w}$ is a prime, then P is not a context-free language.*

PROOF. Let B, $B \subseteq 1(0 + 1)^*$, be an infinite context-free language. Then there exists w in B such that $w = w_1w_2w_3w_4w_5$ with $w_2 \neq \Lambda$ or $w_4 \neq \Lambda$, and for $n = 1, 2, \cdots$, $w_1w_2^nw_3w_4^nw_5$ is in B. We assume that $w_2 \neq \Lambda$ and $w_4 \neq \Lambda$; the other cases are handled similarly by using Corollary 2 instead of Corollary 3. Let k be a positive integer such that $2^{l(w_2)} < \underline{w_1w_2^kw_3w_4^kw_5}$ and $2^{l(w_4)} < \underline{w_1w_2^kw_3w_4^kw_5}$, which insures that $2^{l(w_2)} \not\equiv 1 \pmod{\underline{w_1w_2^kw_2w_4^kw_5}}$ and $2^{l(w_4)} \not\equiv 1 \pmod{\underline{w_1w_2^kw_3w_4^kw_5}}$. Thus if $p = \underline{w_1w_2^kw_3w_4^kw_5}$ is a prime, Corollary 3 implies that

$$q = \underline{w_1w_2^{k+p-1}w_3w_4^{k+p-1}w_5} \equiv 0 \pmod{p}.$$

Therefore q is divisible by p and since q and p are in B we see that B cannot be an infinite subset of the set of primes, as was to be shown.

4. *Memory Requirements*

Since neither the finite automata nor the pushdown automata can recognize the set of primes, it remains an interesting problem to find the "weakest" automaton which can recognize the set of primes.

A natural measure of the computational power of an automaton is the amount of memory it uses. To make this concept precise, we now recall some definitions about memory-limited computations from [9].

Consider a *two-tape Turing machine* M; one of the tapes is a read-only input tape on which, at the start of the computation, is written down between endmarkers the input string, and the reading head is placed on the first digit of the input string; the machine can move the reading head in either direction on this tape but cannot write on it. On the other tape, referred to as working tape, the Turing machine can read and write, and at the start of the computation this tape is blank.

A *machine operation* consists of reading the tape symbols under the reading heads, overprinting the symbol on the working tape, moving the tapes independently one square to the right or left or no move, and changing the internal state.

A set $A \subseteq I^*$ (I^* denotes the set of all finite strings over I) is *accepted* by M if, and only if, M started on w, w in I^*, stops in an accepting state if w is in A and stops in a rejecting state if w is not in A.

Let $L(n)$ be a computable function from nonnegative integers into nonnegative integers. Then we say that a set A is *accepted* by M *on* $L(n)$ *tape* if, and only if, M accepts A and every input w of length n, $n = l(w)$, is rejected or accepted by M without using more than $L(n)$ tape squares of the working tape.

The next result is not new and it states that if the available memory grows linearly with the memory required to represent the input, then the automaton can recognize the set of primes.

THEOREM 3. *The set of primes can be recognized on* $L(n) = n$*-tape.*

PROOF. (The proof is only outlined.) To construct the desired (deterministic-linear-bounded) automaton, let M have two separate tape tracks on its working tape and let M at the start of the computation mark off $l(w)$ tape squares on its working tape for input w. Then M writes successively on the upper track of its working tape the integers $i = 2, 3, \cdots, \underline{w}$; and on its lower track for each integer i, M writes successively the multiples ik, $k = 2, 3, \cdots$, until either $ik = \underline{w}$ or the marked-off working tape is used up. In the first case $\underline{w}$ is rejected; otherwise the process is repeated for $i + 1$. The computation stops and w is accepted when $i = \underline{w}$ is reached. It is seen that M operates on $L(n) = n$-tape and accepts the set of primes, as was to be shown.

From previous work [9] it is known that if M recognizes a set A on $L(n)$-tape and

$$\lim_{n\to\infty} \frac{L(n)}{\log \log n} = 0,$$

then the set A is regular. Furthermore, it is known that there exist nonregular sets which are recognizable on $L(n) = [\log \log n]$-tape. (The symbol $[x]$ denotes the largest integer less than or equal to x, and $\log n$ denotes $\log_2 n$.)

Since Theorem 2 showed that the set of primes is not a regular set, it is concluded that the growth of the minimal amount of tape to recognize the set of primes (or any infinite subset of the set of primes) must be at least as rapid as $L(n) = [\log \log n]$. At the same time it seems impossible that the set of primes could be recognized by an automaton which uses the minimal amount of tape for the recognition of nonregular sets. After a considerable amount of experimenting and "intuitive reasoning" the authors conjecture that the linearly bounded automaton is the "weakest" automaton which can recognize the set of primes. Stated more precisely:

CONJECTURE. *If A is recognized by M on $L(n)$-tape and*

$$lim_{n\to\infty} \frac{L(n)}{n} = 0,$$

then A is not the set of primes.

The authors also conjecture that there are infinite subsets of the set of primes which are easier to recognize than the full set of primes, but have no conjectures about the minimal rate of memory growth for the recognition of such sets.

Next it is shown that if the set of primes contains infinitely many primes which have in their binary representation many consecutive ones or zeros, then a lower bound can be given on the amount of memory required for their recognition. The first result relates this problem to an unsolved problem in number theory.

THEOREM 4. *If there are infinitely many primes of the form $p = 2^{2^k} + 1$, then the set of primes cannot be recognized with $L(n)$-tape such that*

$$\mathit{lim}_{n\to\infty} \frac{L(n)}{\mathit{log}\ n} = 0.$$

PROOF. We refer to the combination of the state of M, the tape pattern written on the working tape, and the position of the read-write head on the working tape as the *total state* of the machine. It is seen that for tape length $L(n)$ the machine with S states and Q different tape symbols can have no more than $S \cdot L(n) \cdot Q^{L(n)}$ total states, where the factor S accounts for the possible state of M, $L(n)$ gives the number of possible positions for the reading head, and $Q^{L(n)}$ describes the number of different patterns which can be written on the $L(n)$ tape squares of the working tape. For large $L(n)$ there is a q, $q > Q$, such that $S \cdot L(n) \cdot Q^{L(n)} < q^{L(n)}$. Because of the limit condition

$$\lim_{n\to\infty} \frac{L(n)}{\log n} = 0,$$

there is an N such that $n > N$ implies that $S \cdot L(n) \cdot Q^{L(n)} < q^{L(n)} < n$. This implies that on an input string $w = 10^{2^k-1}1$, $l(w) \geq n$, representing the number $2^{2^k} + 1 = 10^{2^k-1}1$, the machine must be at least twice in the same total state every time it crosses the run of zeros. If on one such crossing, the total state repeats itself after $t(t \geq 1)$ zeros, then the machine will be in the same total state after crossing the sequences $10^{2^k-1+rt}1$, $r = 0, 1, 2, \cdots$. Since this is true for each crossing, we see that the machine is not able to distinguish between the sequences $w = 10^{2^k-1}1$ and $w_r = 10^{2^k-1+(2^k-1)!r}1$, $r = 1, 2, 3, \cdots$, since after each traversing of these input strings, the total state of the machine is the same and thus w is accepted if and only if w_r is accepted.

Now, using Theorem 1 we see that the set accepted by M on $L(n)$-tape, with

$$\lim_{n\to\infty} \frac{L(n)}{\log n} = 0,$$

cannot be the set of primes, if there are infinitely many primes of the form $2^{2^k} + 1$, as was to be shown.

By the same technique the next result can be obtained.

COROLLARY 4. *If there are infinitely many primes of the form $2^p - 1$, where p is a prime, then the set of primes cannot be recognized with $L(n)$-tape such that*

$$\mathit{lim}_{n\to\infty} \frac{L(n)}{\mathit{log}\ n} = 0.$$

The same reasoning can be used to show that if there are infinitely many primes whose binary representation has no more than k zeros (or k ones), then the set of primes cannot be recognized on tape $L(n)$ with

$$\lim_{n\to\infty} \frac{L(n)}{\log n} = 0.$$

Unfortunately, it is not known whether any of the above-mentioned properties holds for the set of primes, and the first two properties are classic conjectures in number theory.

5. *Density Considerations*

In this section the distribution of numbers recognized by automata is considered, and it is shown that these are computationally "weak" automata which can accept sets of numbers with the same distribution as the set of primes (as given by the Prime Number Theorem). This shows that if the conjecture is true it cannot be proven by the use of arguments about the distribution of primes.

For a subset A of $1(0 + 1)^*$ let $\pi_A(n)$ denote the number of elements w in A such that $\underline{w} \leq n$.

To prove the next theorem we need the following result from analysis.

LEMMA. *If $a_1, a_2, a_3, \cdots$, is a sequence of positive numbers such that*

$$\lim_{n\to\infty} \frac{a_n}{a_{n-1}} = \alpha > 1,$$

then

$$\lim_{n\to\infty} \frac{1}{a_{n+1}} \sum_{i=1}^{n} a_i = \frac{1}{\alpha - 1}.$$

THEOREM 5. *The set $A = \{w \mid w \in 1(0 + 1)^*$ and $\underline{w} = [l(w) - 1]k$ for some integer $k\}$ is recognizable with $L(n) = [\log n]$-tape and*

$$\lim_{\underline{w}\to\infty} \pi_A(\underline{w}) \frac{\log_2 \underline{w}}{\underline{w}} = 1.$$

PROOF. It is seen that $l(w) = [\log \underline{w}] + 1$, and a straightforward algorithm shows that the set A is recognizable with tape bound $L(n) = [\log n]$. (This can be done by carrying out division of $\underline{w}$ by $l(w) - 1$; this process can be performed on $L(n) = [\log n]$-tape.)

Next we compute $\pi_A(\underline{w})$. Observe that

$$\lim_{\underline{w}\to\infty} \frac{l(w) - 1}{\log \underline{w}} = 1$$

and denote $[\pi_A(2^k - 1) \log (2^k)]/2^k$ by Θ_k. Then

$$\Theta_k = \frac{k}{2^k} \sum_{i=1}^{k} [\pi_A(2^i - 1) - \pi_A(2^{i-1} - 1)],$$

and therefore

$$\Theta_k \sim \frac{k}{2^k} \sum_{i=1}^{k} \frac{2^{i-1}}{i - 1}.$$

Thus, using the lemma with $\alpha = 2$, we conclude that

$$\lim_{k\to\infty} \Theta_k = 1.$$

For $n = 2^k + r$, $0 \leq r < 2^k$, we have

$$\pi_A(n) - \pi_A(2^k - 1) \sim \frac{r+1}{k},$$

and therefore

$$\frac{\pi_A(n)\log n}{n} \sim \left(\frac{r+1}{k} + \frac{2^k}{k}\Theta_k\right)[k + \log(1 + r2^{-k})](2^k + r)^{-1}$$

$$= \frac{2^k\Theta_k + r + 1}{2^k + r}[1 + \log(1 + r2^{-k})^{1/k}]$$

$$= \left[1 + \frac{2^{-k} + \Theta_k - 1}{1 + r2^{-k}}\right][1 + \log(1 + r2^{-k})^{1/k}].$$

From the last expression we see that

$$\lim_{n\to\infty} \frac{\pi_A(n)\log n}{n} = 1,$$

as was to be shown.

Note 1. The above density result differs from the density for prime numbers by the factor $\log_2 e$; due to the different bases for logarithms,

$$\lim_{n\to\infty} \frac{\pi(n)}{\pi_A(n)} = \log_2 e.$$

This factor can be eliminated by a somewhat more complicated construction of the set A and the machine M which recognizes A: for each input w, M computes an approximation e_w of $\log_e 2$, using no more than $\log l(w)$-tape, and then accepts w if, and only if, $\underline{w}$ is divisible by $[e_w l(w)]$. Since

$$\lim_{\underline{w}\to\infty} \frac{[e_w l(w)]}{\ln \underline{w}} = 1,$$

we see that on $L(n) = [\log n]$-tape we can recognize a set of integers with the same density as the set of primes.

Note 2. By an even more complicated process, which utilizes ideas developed in [9] about $L(n) = [\log\log n]$-recognizable sets and which uses properly encoded markers on the input tape to help check the division process, a machine M can be constructed which works on $L(n) = [\log\log n]$-tape and accepts a set of integers with the same density as the set of primes.

From Theorem 5 and the Notes it is seen that if the Conjecture is true, it cannot be proven by arguments about the density of the prime numbers as described by the Prime Number Theorem.

6. *Conclusion*

In this paper several results are derived about the recognition of the set of primes by automata. Still the most interesting problem about the "weakest" automaton which can recognize the set of primes remains open. At the same time it seems that in the problem of recognizing the set of primes by automata we have an interesting interaction between a classical branch of mathematics and automata theory. It is our hope that further work may reveal deeper connections and contribute to both fields.

ACKNOWLEDGMENT. The authors would like to thank Professors K. M. Brown, J. E. Hopcroft, and R. J. Walker for many helpful discussions about the recognition of primes.

REFERENCES

1. COBHAM, A. The recognition problem for the set of perfect squares. Proc. Seventh Annual Symposium on Switching and Automata Theory, IEEE, New York, Oct. 1966, pp. 78–87.
2. ——. Time and memory capacity bounds for machines which recognize squares or palindromes. IBM Res. Rep. RC-1621, May 31, 1966.
3. HARTMANIS, J., AND STEARNS, R. E. On the computational complexity of algorithms. *Trans. Amer. Math. Soc. 117*, 5 (May 1965), 285-306.
4. —— AND ——. Sets of numbers recognized by finite automata. *Amer. Math. Mon. 74*, 5 (May 1967), 539-542.
5. MINSKY, M., AND PAPERT, S. Unrecognizable sets of numbers. *J. ACM 13*, 2 (April 1966), 281-286.
6. RITCHIE, R. W. Finite automata and the set of squares. *J. ACM 10*, 4 (Oct. 1963), 528-531.
7. SCHÜTZENBERGER, M. P. A remark on acceptable sets of numbers. *J. ACM 15*, 2 (April 1968), 300-303.
8. HARRISON, M. A. *Introduction to Switching and Automata Theory*. McGraw-Hill, New York, 1965.
9. STEARNS, R. E., HARTMANIS, J., AND LEWIS, P. M. Hierarchies of memory limited computations. Proc. Sixth Annual Symposium on Switching Circuit Theory and Logical Design, IEEE, New York, Oct. 1965, pp. 179–190.

RECEIVED AUGUST, 1967; REVISED DECEMBER, 1967