Submitted to Association for Computing
Machinery as a short communication

UCRL-19429
Preprint

c.2

ON FINDING PRIME NUMBERS

RECEIVED
LAWRENCE
RADIATION LABORATORY

MAR 5 1970

LIBRARY AND
DOCUMENTS SECTION

Barrie Pardoe

November 1969

AEC Contract No. W-7405-eng-48

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 5545

UCRL-19429 cy. 2

LAWRENCE RADIATION LABORATORY
UNIVERSITY of CALIFORNIA BERKELEY

32

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

ON FINDING PRIME NUMBERS*

Barrie Pardoe

Lawrence Radiation Laboratory
University of California
Berkeley, California 94720

November 1969

1. Introduction

Singleton's comments[1] on applying Eratosthenes' Sieve to the sequence of positive integers with multiples of 2, 3 and 5 omitted prompt me to outline the following treatment of the sequence with multiples of 2 and 3 only omitted. The terms of this sequence can be expressed in simple form and some elementary but interesting results can be derived. In what follows, the method resulting from the treatment is given first, its efficiency when machine coded for a CDC 6600 is assessed second, and its derivation is given third.

The following notation is used throughout:

$\left[p/q\right]$ integral part of p/q,

$R(p/q) = p - q \cdot \left[p/q\right]$,

$p|q$ p is a divider of q,

(p, Q) highest common factor of p and q,

$\Longrightarrow$ implies,

iff if and only if.

2. Method

2.1. The sequence of integers with multiples of 2 and 3 omitted is given by

$$s_n = 6 \cdot \left[\frac{n+1}{2}\right] + (-1)^n , \quad n = 1, 2, 3, \cdots .$$

2.2. Given prime $p = s_n$ and $p \mid s_{n'}$ where $n' = n + 2rp$, $r = 0, 1, 2, \cdots$, then

$$p \mid s_{n''} \text{ iff } n'' = n' + m_1 \text{ or } n'' = n' + m_1 + m_2 ,$$

where $m_1 + m_2 = 2p,$

$$m_1 = 2 \cdot \left\{ 2n + R(n/2) \right\} + 1,$$

$$m_2 = 2n + 1.$$

2.3. If $p^2 = s_{n''}$, then $p \mid s_{n'' + m'}$

where,

for n even, $n'' = 3n^2 + 2n$, $m = m_1 = 4n + 1$, $m_2 = 2n + 1$,

for n odd, $n'' = 3n^2 + 4n + 1$, $m = m_2 = 2n + 1$, $m_1 = 4n + 3$.

3. Efficiency of the Method

These expressions were coded in machine language for the CDC 6600. Each term in the sequence was represented by a bit in an array. Given n for a prime p, masks were generated of length m_1 and m_2 and, starting with the term in the sequence corresponding to p^2, bits in the array corresponding to multiples of p were turned off by alternating the masks m_1 and m_2 up to the end of the array.

For a given array, the primes were printed out by flagging integers as to whether they are composite or not. Only integers of the form $10r + 1, 3, 7, 9$ need be flagged, and four such flags can be represented by a hexadecimal character (0-9, A-F $\equiv$ 0000-0101, 0110-1111). In this manner, flags for 50 000 integers could be printed on one page (60 lines of 132 characters) with sufficient information to provide easy table look up and a prime count.

The time taken for execution of the program was very close to linear as far as was tested. Terms in the sequence were scanned at the rate of 10^5/sec, and terms were scanned and the flagged integers printed at 31 000/sec.

The 788 062 primes $\leq$ 12 000 029 were found in 40 sec and printed in 90 sec.

The program passed through the array once and was therefore limited to the size of the array($66\,667_{10}$ words of 60 bits in the example just cited), but the program could easily be modified to pass through a smaller array repeatedly, the next higher part of the sequence being mapped onto the array each time. This modification would increase the length of the sequence that could be scanned and reduce the memory requirements of the program.* By keeping a list of primes up to p, the search for primes less than p^2 can be made. Thus an array of length N of 60-bit words would permit a search for all primes up to $(180N)^2$. Taking N=1000 would give all primes up to $3.2 \cdot 10^{10}$.

4. Derivation of Method

The sequence of positive integers omitting multiples of 2 and 3 is

$$s_n = 6 \cdot \left[\frac{n+1}{2}\right] + (-1)^n, \qquad n = 1, 2, 3, \cdots,$$

from which

$$s_{n+m} - s_n = 6 \cdot \left\{\left[\frac{n+m+1}{2}\right] - \left[\frac{n+1}{2}\right]\right\} + (-1)^n \cdot \left\{(-1)^m - 1\right\}.$$

$$m > 0$$

*This method has already been used by Singleton (Ref. 1).

If m even, $m \geqslant 2$,

$$s_{n+m} - s_n = 3m.$$

If $p \mid s_n$, $p \mid s_{n+m}$ if $p \mid 3m$,

$p > 3 \implies p \mid m$,

$m = 2p, 4p, 6p, \cdots$.

Thus $m = 2p$ is the lowest even number of terms of the sequence, the difference of which is divisible by p.

If m odd,

$$s_{n+m} - s_n = 6 \cdot \left[\frac{m}{2}\right] + 3 + (-1)^n.$$

Again, if $p \mid s_n$, $p \mid s_{n+m}$ iff

$$6 \cdot \left[\frac{m}{2}\right] + a(n) = q \cdot p, \quad a(n) = 3 + (-1)^n, \quad q = 1, 2, 3, \ldots,$$

which has a solution[2] in q, m as $(6, p) = 1$ and $1 \mid a(n)$.

Further, if m_1 is the least m satisfying (1), $m_1 < 2p$ as the existence of $m = m_1 > 2p$ satisfying (1) implies as $m = m_1' < 2p$ satisfying (1). Further, there is only one value of m, $0 < m < 2p$, such that $p \mid s_{n+m}$ as m is odd and two or more values of m would demand some m to be even and $< 2p$.

This being so, m_1 can now be determined.

Let

$$m_1 + m_2 = 2p, \quad p = s_n \text{ for some } n.$$

Both m_1, m_2 are odd.

$$p \mid s_{n+m_1} \implies 6\cdot\left[\frac{m_1}{2}\right] + a(n) = q\cdot p\,,\ q \geqslant 1 \tag{2}$$

$$\implies 6\cdot\left[\frac{m_2}{2}\right] + 6 - a(n) = (6-q)\cdot p \tag{3}$$

From (2), (3):

$$q \geqslant 1\,,\ m_1 > 0\,,\ a(n) \text{ even} \implies q \text{ even},$$

$$4 \geqslant a(n) = 3 + (-1)^n \implies 6 - q > 0;$$

$$\therefore\ q = 2, 4 \implies 6 - q = 4,\ 2.$$

Take $q = 2$, either $6\cdot\left[\frac{m_1}{2}\right] + a(n) = 2p$

or $6\cdot\left[\frac{m_2}{2}\right] + 6 - a(n) = 2p.$

Consider $6\cdot\left[\frac{m}{2}\right] + R = 2p, \qquad 0 \leqslant R < 6\,.$ (4)

If, for a solution to (4), $R = a(n)$, $m_1 = m$,

but if, on the other hand, $R = 6 - a(n)$, $m_2 = m$.

From (4):

$$6\cdot\left[\frac{m}{2}\right] + R = 12\cdot\left[\frac{n+1}{2}\right] + 2(-1)^n,$$

so if n even, $R = 2 = 6 - a(n)$ and $m = 2n + 1$;

if n odd, $R = 4 = 6 - a(n)$ and $\left[\frac{m}{2}\right] = 2\left[\frac{n+1}{2}\right] - 1 \implies m = 2n + 1$;

$$\therefore\ m_2 = 2n + 1,$$

$$m_1 = 2\cdot\left\{2n + R(n/2)\right\} + 1.$$

This proves the result stated earlier, that if

$p = s_n$ then $p \mid s_{n'}$, where $n' = n + 2rp$, $n' = n + 2rp + m_1$,

Clearly, from the sequence s_n, p^2 is the first composite term which is a multiple of p to be deleted if all multiples of lesser primes have been deleted.

To determine n' such that $s_{n'} = p^2$, where $n' = n + m$,

$$s_n' - s_n = p^2 - p = 36 \cdot \left[\frac{n+1}{2}\right]^2 + 12(-1)^n \cdot \left[\frac{n+1}{2}\right] + 1 - 6 \cdot \left[\frac{n+1}{2}\right] - (-1)^n$$

$$= 6 \cdot \left[\frac{m}{2}\right] + a(n, m), \quad a(n, m) = 0, \quad m \text{ even}$$

$$= a(n), \quad m \text{ odd},$$

from which

$$n \text{ even} \implies a(n, m) = 0 \implies m \text{ even} \implies m = n(3n + 1),$$

$$n \text{ odd} \implies a(n, m) = 2 \implies m \text{ odd} \implies m = 3n(n+1) + 1.$$

Hence if n even, $n' = 3n^2 + 2n$,

and terms $s_{n''}$, $n'' = n' + m_1$, $n'' = n' + m_1 + m_2 = n' + 2p$

are multiples of p,

whereas if n odd, $n' = 3n^2 + 4n + 1$,

and terms $s_{n''}$, $n'' = n' + m_2$, $n'' = n' + m_2 + m_1 = n' + 2p$

are multiples of p.

FOOTNOTES AND REFERENCES

*Work performed under auspices of the U. S. Atomic Energy Commission.

1. Singleton, R. C., Algorithm 357: An Efficient Prime Number Generator. Comm. ACM 12 (Oct. 1969), 563.

2. Sierpinski, W., "Elementary Theory of Numbers," Publications of the Inst. of Maths., Polish Academy of Science, 1964, Theorem 15, p. 27.

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

TECHNICAL INFORMATION DIVISION
LAWRENCE RADIATION LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720