RCL-March 13, 2019
Jon Crawford – May 27, 2019

Homework 1: Sieve of Eratosthenes
50 Points

Consider the algorithm named Sieve of Eratosthenes:

Sieve

Input: an integer $n > 1$.

Let A be an array of Boolean values, indexed by integers 2 to $\sqrt{\ }\ n$, initially all set to true.

for $i = 2,3,4,...,$ not exceeding $\overline{\ \ n}$:

if $A[i]$ is true:

for $j = i^2, i^2 + i, i^2 + 2i, i^2 + 3i, ...,$ not exceeding n: $A[j]$:= false.

Output: all i such that $A[i]$ is true. Here is C++

```
void Sieve(fsu::BitVector& bv)
// in: bv is a bit vector
// out: bv[k] is true iff k is prime
{
  size_t max = bv.Size();
  size_t sqrtmax = ceil(sqrt(max));
  // 1: initialize bv
  bv.Set();
  bv.Unset(0);
  bv.Unset(1);  // 0 and 1 is not prime
  for (size_t i = 4; i < max; i += 2)
    bv.Unset(i);

  for (size_t i = 3; i < sqrtmax; i += 2) // we can skip over the even numbers
  {
    if (bv[i]) // i is prime
      for (size_t j = i*i; j < max; j+= i) // clear all multiples of i
        bv.Unset(j);
  }
}
```

Along with Euclid's algorithm to find the greatest common divisor of two positive integers, the Sieve of Eratosthenes is one of the oldest computational algorithms, dating back at least to 300 BC [1], [2]. The Sieve is still an important tool today, so much so that it has been improved and optimized in many ways.

One modern version may help discover new non-Mersenne primes [3]. (The current record prime numbers are all Mersenne primes, that is, of the from $2^p - 1$ for some much smaller prime p. These discoveries used techniques that work only for the rare Mersenne primes. See [4].) Your task is to prove:

Proposition. The Sieve of Eratosthenes algorithm stated above has runtime complexity ≤O($n$log$\log n$).

This may seem a daunting assignment at first. However, please read all about the Sieve in Wikipedia and any other open sources on the web. You may assume some facts about prime numbers and logarithms as well, including:

Prime Number Theorem. Let $\pi(x)$ denote the number of primes less than or equal to x. Then:

$$\pi(x) \sim \frac{x}{\log x}$$

where log$x$ is the natural logarithm of x. [5]

and the result of Mertens (1874) that *pre-dates* the prime number theorem by 22 years: **Mertens' Lemma.** Take $\sum_{p \le n}$ to mean the sum over all prime numbers $p \le n$. Then

$$\sum_{p \le n} \frac{1}{p} \sim \log\log n + M$$

where $M \approx 0.261497...$ is the so-called *Meissel-Mertens constant*. [6]

To prove the Proposition, first eliminate the optimizations to obtain a simplified version of Sieve (see below). Show in passing that the total loop count of Sieve(n) is ≤ that of Simpler Sieve(n) and then count the steps in the loops. In particular, show that the number of steps in the inner loop executing at the prime p is $\le \frac{n}{p}$. Then find the total count of both loops, do a little algebraic re-arranging, and apply Mertens' Lemma. (You supply full details ... this is just a hint.)

Simpler Sieve (Eratosthenes' original verson) Input: an
integer $n > 1$.
Let A be an array of Boolean values, indexed by integers 2 to n, initially all set to true.
for i = 2,3,4,..., not exceeding n:
    if $A[i]$ is true:
        for $j = i + i, i + 2i, i + 3i, ...,$ not exceeding n:
            $A[j]$:= false.
Output: all i such that $A[i]$ is true

The first step is to eliminate the optimizations in order to obtain SimpleSieve(bv):

```
void SimpleSieve(fsu::BitVector& bv)
{
  size_t max = bv.Size();
  bv.Set();
  bv.Unset(0);
  bv.Unset(1);
  for(size_t i = 2; i < max; ++i) // 2 to n
  {
    if(bv[i])
    {
      for(size_t j = i + i; j < max; j += i)
      {
        bv.Unset(j);
      }
    }
  }
}
```

Then, consider in passing that SimpleSieve(bv) will make more iterations based on the fact that it does not first set all the even numbers to false and skip over them in the loop. It also iterates j from a size of i+i on the inner loop rather than the size of i*i used in Sieve(bv). Therefore, the loop count of Sieve(n) ≤ SimpleSieve(n), this is clear.

Now, since the loop for j is incrementing at j = i+i, this means that i is equivalent to the p because it is representing the prime numbers being incremented at since the inner loop only executes when bv[i] = true. With max representing n. [8] Then, $p \text{ is } \leq \frac{n}{p}$ due to the fact that the number of times that i increments is some number of times x from i + i, i + 2i, i + 3i … to i + xi but not exceeding n. Clearly i + xi ≤ n means xi ≤ n and surely x ≤ $\frac{n}{i}$ since i is positive values 2 to n. [9] Proving that $p \text{ is } \leq \frac{n}{p}$ as the inner loop only fires on primes. Taking the Prime Number theorem, and Mertens' Lemma coupled with this value. That means that since the stated loop execution at p ≤ $\frac{n}{p}$ can be equated to that of the sum of $\frac{n}{2}, \frac{n}{3}, \frac{n}{5}$ … since p represents primes.

Or, $\sum \frac{n}{i}$ where i equals p, and from p ≤ n. The law's of Algebra state that $\frac{n}{i}$ = n ($\frac{1}{i}$). [10]

Giving n times the Mertens' Lemma left hand side when n is factored out, or n * $\sum_{p \leq n} \frac{1}{p}$. Which in turn equals n times the Mertens' Lemma right hand side or, n * $\log \log n + M$. Then, using Calculus and the concept of a bound as above, the + M can be absorbed. Now, since Sieve(n) ≤ SimpleSieve(n), then the Sieve has runtime complexity ≤ O(n log log n).

References

[1] According to Wikipedia, the earliest known reference to the sieve is in Nicomachus of Gerasa's Introduction to Arithmetic (translated from ancient Greek in [2]) which describes it and attributes it to Eratosthenes of Cyrene, a Greek mathematician.

[2] Hoche, Richard, ed. (1866), Nicomachi Geraseni Pythagorei Introductionis arithmeticae libri II, Leipzig: B.G. Teubner, p. 31

[3] See Science Alert: https://www.sciencealert.com/an-ancient-greek-algorithm-could-be-thekey-to-finding-new-prime-numbers

[4] See GIMPS: https://en.wikipedia.org/wiki/Great Internet _Mersenne Prime Search

[5] See the Wikipedia entry: https://en.wikipedia.org/wiki/Prime number theorem

[6] F. Mertens. **J. reine angew. Math. 78** (1874), 46 *Ein Beitrag zur analytischen Zahlentheorie*

[7] Mertens Theorems Wikipedia entry: https://en.wikipedia.org/wiki/Mertens%27_theorems

[8] N.A, n.d, Sieve of Erastothenes. Retrieved May 31, 2019 from, https://www.algolist.net/Algorithms/Number_theoretic/Sieve_of_Eratosthenes

[9] Krishan, Kumar, (n.d.) GeeksForGeeks, Retrieved June 2, 2019 from, https://www.geeksforgeeks.org/sieve-of-eratosthenes/

[10] N.A, n.d, Sieve of Erastothenes. Retrieved May 31, 2019 from, https://cp-algorithms.com/algebra/sieve-of-eratosthenes.html