Computing Science 204 (B1)

The Sieve of Eratosthenes

Saturday, July 22

Before the advent of the computer, one of the most efficient methods of constructing tables of primes was the *sieving process*, invented by the Greek mathematician Eratosthenes (276-194 B.C.). The work involved in the process is simplified by the following lemma.

Lemma. If an integer n is composite, then it has a prime factor p with $p \leq \sqrt{n}$.

proof. If this were not the case, then we could write $n = ab$ where $a > \sqrt{n}$ and $b > \sqrt{n}$. This would imply that

$$n = ab > \sqrt{n}\sqrt{n} = n,$$

which is a contradiction.

$\square$

We illustrate the method by constructing a table containing all primes less that 130. We begin by listing all the integers from 2 to 129 (since 1 is not a prime it is not listed).

The first number in our list, 2, must be a prime and no multiple of 2 except 2 itself can be prime. We remove all multiples of 2 (except $2 \cdot 1$) from our list. The next remaining number, 3, must be a prime, so we delete all the multiples of 3. We now delete the multiples of 5, the multiples of 7, and the multiples of 11. Because the largest prime less than $\sqrt{130}$ is 11, then from the lemma, all the remaining numbers must be primes. In the table below, rather than delete multiples of 2, 3, 5, 7, 11, we have underlined the integers that were not deleted, thus, all primes less than 130 are all underlined.

		$\underline{2}$	$\underline{3}$	4	$\underline{5}$	6	$\underline{7}$	8	9
10	$\underline{11}$	12	$\underline{13}$	14	15	16	$\underline{17}$	18	$\underline{19}$
20	21	22	$\underline{23}$	24	25	26	27	28	$\underline{29}$
30	$\underline{31}$	32	33	34	35	36	$\underline{37}$	38	39
40	$\underline{41}$	42	$\underline{43}$	44	45	46	$\underline{47}$	48	49
50	51	52	$\underline{53}$	54	55	56	57	58	$\underline{59}$
60	$\underline{61}$	62	63	64	65	66	$\underline{67}$	68	69
70	$\underline{71}$	72	$\underline{73}$	74	75	76	77	78	$\underline{79}$
80	81	82	$\underline{83}$	84	85	86	87	88	$\underline{89}$
90	91	92	93	94	95	96	$\underline{97}$	98	99
100	$\underline{101}$	102	$\underline{103}$	104	105	106	$\underline{107}$	108	$\underline{109}$
110	111	112	$\underline{113}$	114	115	116	117	118	119
120	121	122	123	124	125	126	$\underline{127}$	128	129

The sieve of Eratosthenes can be used to derive a formula for the number of primes less than or equal to n if the primes less than or equal to $\sqrt{n}$ are known. First a standard definition.

Definition. If x is a real number, then $\pi(x)$ denotes the number of prime numbers less than or equal to x.

Theorem. If $p_1, p_2, \ldots, p_k$ are the primes less than $\sqrt{n}$, then

$$\begin{aligned}
\pi(n) = n - 1 + \pi(\sqrt{n}) &- \left\{ \left\lfloor \frac{n}{p_1} \right\rfloor + \left\lfloor \frac{n}{p_2} \right\rfloor + \cdots + \left\lfloor \frac{n}{p_k} \right\rfloor \right\} \\
&+ \left\{ \left\lfloor \frac{n}{p_1 p_2} \right\rfloor + \left\lfloor \frac{n}{p_1 p_3} \right\rfloor + \cdots + \left\lfloor \frac{n}{p_1 p_k} \right\rfloor + \left\lfloor \frac{n}{p_2 p_3} \right\rfloor + \cdots + \left\lfloor \frac{n}{p_{k-1} p_k} \right\rfloor \right\} \\
&- \left\{ \left\lfloor \frac{n}{p_1 p_2 p_3} \right\rfloor + \cdots + \left\lfloor \frac{n}{p_{k-2} p_{k-1} p_k} \right\rfloor \right\} + \cdots \\
&+ (-1)^k \left\lfloor \frac{n}{p_1 p_2 \cdots p_k} \right\rfloor
\end{aligned}$$

A rigorous proof of this formula follows immediately from the Principle of Inclusion and Exclusion. We will give an informal argument.

Recall the steps in the sieve of Eratosthenes. The number of elements in the original set is $n-1$. The number divisible by p_1 is $\left\lfloor \frac{n}{p_1} \right\rfloor$, the number divisible by p_2 is $\left\lfloor \frac{n}{p_2} \right\rfloor$. If we delete all the numbers divisible by p_1 (including p_1 itself) and then all the numbers divisible by p_2 (including p_2 itself), we have deleted

$$\left\lfloor \frac{n}{p_1} \right\rfloor + \left\lfloor \frac{n}{p_2} \right\rfloor - \left\lfloor \frac{n}{p_1 p_2} \right\rfloor$$

numbers. If we continue this line of reasoning, we obtain $n-1$ plus the complicated expression involving the greatest integer function for the number of elements remaining in the set. In the process, we have removed the prime numbers $p_1, p_2, \ldots, p_k$ from the list as well as their multiples. When we replace these $\pi(\sqrt{n})$ numbers, we get the correct result, namely, $\pi(n)$.

As an example of the use of this formula, we can calculate the number of primes less than or equal to 129. In this case

$$\pi(\sqrt{n}) = \pi(\sqrt{129}) = 5,$$

and the primes less than or equal to $\sqrt{129}$ are $2, 3, 5, 7, 11$. From the formula in the theorem, we have

$$\begin{aligned}
\pi(129) = 129 - 1 + 5 &- \left\lfloor \frac{129}{2} \right\rfloor - \left\lfloor \frac{129}{3} \right\rfloor - \left\lfloor \frac{129}{5} \right\rfloor - \left\lfloor \frac{129}{7} \right\rfloor - \left\lfloor \frac{129}{11} \right\rfloor \\
&+ \left\lfloor \frac{129}{2\cdot 3} \right\rfloor + \left\lfloor \frac{129}{2\cdot 5} \right\rfloor + \left\lfloor \frac{129}{2\cdot 7} \right\rfloor + \left\lfloor \frac{129}{2\cdot 11} \right\rfloor + \left\lfloor \frac{129}{3\cdot 5} \right\rfloor + \left\lfloor \frac{129}{3\cdot 7} \right\rfloor \\
&+ \left\lfloor \frac{129}{3\cdot 11} \right\rfloor + \left\lfloor \frac{129}{5\cdot 7} \right\rfloor + \left\lfloor \frac{129}{5\cdot 11} \right\rfloor + \left\lfloor \frac{129}{7\cdot 11} \right\rfloor \\
&- \left\lfloor \frac{129}{2\cdot 3\cdot 5} \right\rfloor - \left\lfloor \frac{129}{2\cdot 3\cdot 7} \right\rfloor - \left\lfloor \frac{129}{2\cdot 3\cdot 11} \right\rfloor - \left\lfloor \frac{129}{2\cdot 5\cdot 7} \right\rfloor - \left\lfloor \frac{129}{2\cdot 5\cdot 11} \right\rfloor \\
&- \left\lfloor \frac{129}{2\cdot 7\cdot 11} \right\rfloor - \left\lfloor \frac{129}{3\cdot 5\cdot 7} \right\rfloor - \left\lfloor \frac{129}{3\cdot 5\cdot 11} \right\rfloor - \left\lfloor \frac{129}{3\cdot 7\cdot 11} \right\rfloor - \left\lfloor \frac{129}{5\cdot 7\cdot 11} \right\rfloor \\
&+ \left\lfloor \frac{129}{2\cdot 3\cdot 5\cdot 7} \right\rfloor + \left\lfloor \frac{129}{2\cdot 3\cdot 5\cdot 11} \right\rfloor + \left\lfloor \frac{129}{2\cdot 3\cdot 7\cdot 11} \right\rfloor + \left\lfloor \frac{129}{2\cdot 5\cdot 7\cdot 11} \right\rfloor + \left\lfloor \frac{129}{3\cdot 5\cdot 7\cdot 11} \right\rfloor \\
&- \left\lfloor \frac{129}{2\cdot 3\cdot 5\cdot 7\cdot 11} \right\rfloor \\
&= 129 - 1 + 5 - 64 - 43 - 25 - 18 - 11 + 21 + 12 + 9 + 5 + 8 + 6 + 3 + 3 + 2 + 1 \\
&\quad - 4 - 3 - 1 - 1 - 1 - 0 - 1 - 0 - 0 - 0 + 0 + 0 + 0 + 0 + 0 - 0 \\
&= 31.
\end{aligned}$$

We can verify that this is correct by counting the number of primes in the table (the integers that are underlined).

Although this formula is very awkward to use, it is the only formula for the exact value of $\pi(n)$.