Based on K. H. Rosen: Discrete Mathematics and its Applications.

Lecture 16: Prime numbers. Fundamental Theorem of Arithmetic (FTA). Section 4.3

1 Number Theory: FTA and gcd of numbers

1.1 Prime numbers and the FTA

Definition 1. An integer p greater than 1 is called **prime** if the only positive factors of p are 1 and p. A positive integer that is greater than 1 and is not prime is called **composite**.

Theorem 2. *(THE FUNDAMENTAL THEOREM OF ARITHMETIC) Every integer greater than* 1 *can be written uniquely as a prime or as the product of two or more primes where the prime factors are written in order of nondecreasing size.*

Theorem 3. *If n is a composite integer, then n has a prime divisor less than or equal to $\sqrt{n}$.*

Proof. In n is a composite numbers, this means, it factors as a product $n = ab$, where $a > 1$ and $b > 1$. If $a \leq \sqrt{n}$ we finish the proof, otherwise if $a > \sqrt{n}$, then

$$b = \frac{n}{a} \leq \frac{n}{\sqrt{n}} = \sqrt{n}$$

will do the work for us. □

Theorem 4. *There are infinitely many primes.*

Proof. We will prove this theorem using a proof by contradiction. We assume that there are only finitely many primes listed as $p_1, p_2, \ldots, p_n$. Consider the number

$$Q = p_1 p_2 \ldots p_n + 1.$$

By the fundamental theorem of arithmetic, Q has a prime factor p. However, none of the primes p_j divides Q. By properties of the division if p_j divides Q and also divides the product $p_1 p_2 \ldots p_n$, then it must divide the difference $Q - p_1 p_2 \ldots p_n = 1$, which is not possible. As conclusion, we have a prime p that is not in our finite list (which was suppose to contain all the primes) and this shows that we do not have a finite list containing all primes. □

The above result says that there are infinitely many primes. One could ask, how likely are we to find them, say in the interval $[0, x]$ (for a real number x)?

Theorem 5. *(THE PRIME NUMBER THEOREM) The ratio of the number of primes not exceeding x and $\frac{x}{\ln(x)}$ approaches 1 as x grows without bound. (Here $\ln(x)$ is the natural logarithm of x.)*

Now, this result says that probability of selecting a prime number when a positive integer is selected in the interval $[0, n]$ is approximately $\frac{\frac{n}{n}}{\ln(n)} = \frac{1}{\ln(n)}$. In other words the n-th prime p_n is approximately of the order $n \ln(n)$.

1.2 Greatest Common divisor

Definition 6. Let a and b be integers, not both zero. The largest integer d such that $d \mid a$ and $d \mid b$ is called the **greatest common divisor** of a and b. The greatest common divisor of a and b is denoted by $\gcd(a, b)$.

The following lemma is the basis for Euclidean algorithm

Lemma 7. *Let $a = bq+r$, where $a, b, q,$ and r are integers. Then $\gcd(a, b) = \gcd(b, r)$. In particular when $a, b > 0$ and q, r are the quotient of the remainder of the division of a by b, we have*

$$\gcd(a, b) = \gcd(b, r).$$

The Euclidean Algorithm to find $\gcd(a, b)$ for $a \geq b > 0$:

1. Put $r_0 = a$ and $r_1 = b$.
2. Compute r_2 from $r_0 = r_1 q_1 + r_2$ with $0 \leq r_2 < r_1 = b$.
3. Compute r_3 from $r_1 = r_2 q_2 + r_3$ with $0 \leq r_3 < r_2$.
4. Compute, in general from r_{k+2} from $r_k = q_{k+1} r_{k+1} + r_{k+2}$, we obtain $r_{k+2} < r_{k+1}$.

Since $r_0 = a \geq b = r_1 > r_2 > \cdots > r_{k+1} > r_{k+2} \cdots \geq 0$, the process eventually terminates with some $r_{n+1} = 0$ and $r_{n-1} = r_n q_n$. The lemma above is saying that

$$\gcd(r_0 = a, r_1 = b) = \gcd(r_2, r_1) = \cdots = \gcd(r_n, r_{n-1}) = \gcd(r_n, 0) = r_n.$$

The Euclid's algorithm is replacing in each step the smaller number of (r_k, r_{k+1}) with the remainder of the division of r_k and r_{k+1} and works instead with the new pair (r_{k+1}, r_{k+2}) of smaller numbers to find the gcd.

procedure gcd(a, b : positive integers)
$x = a$
$y = b$
while $y \neq 0$
  $r = x$ mod y
  $x = y$
  $y = r$
return x ($x = \text{gcd}(a, b)$)

At the same we find the gcd of two numbers, we can express that gcd as a linear combination of the original numbers using integer coefficients. The result can be stated:

Theorem 8. *(BÉZOUT'S THEOREM) If a and b are positive integers, then there exist integers s and t such that* $\text{gcd}(a, b) = sa + tb$.

An extended Euclidean algorithm is an algorithm that finds $\text{gcd}(a, b)$ and, at the same time, finds integers t and s such that $\text{gcd}(a, b) = sa + tb$. In the following table the subsequent q, r, s, t are obtained using, for $n \geq 1$, the formulas:

$$q_{n+1} = r_n // r_{n-1} \qquad r_{n+1} = r_n \% r_{n-1} = r_{n-1} - q r_n$$

$$s_{n+1} = s_{r-1} - q s_r \qquad t_{n+1} = t_{r-1} - q t_r$$

n	0	1	2	3	4	5	6
q			5	4	1	1	2
r	240	46	10	6	4	2	0
s	1	0	1	-4	5	-9	
t	0	1	-5	21	-26	47	

For example to get the third column ($n = 2$), we divide 240 by 46 obtaining quotient $q = 5$ and remainder $r = 10$. On the other hand $s_2 = s_0 - qs_1 = 1 - 0(5) = 1$ and $t_2 = t_0 - qt_1 = 0 - 1(5) = -5$. In this way we obtain the second column $(q, r, s, t) = (5, 10, 1, -5)$ and continue the the process.
When we encounter $r_n = 0$, the triple (gcd, s, t) is in the previous column as $(r_{n-1}, s_{n-1}, t_{n-1})$.
In our example above, $\text{gcd}(240, 46) = 2$ and $2 = (-9)(240) + (47)(46)$.