CMPSCI 250: Introduction to Computation

Lecture #12: Divisibility and Primes
David Mix Barrington
17 February 2012

Divisibility and Primes

- Introduction to Number Theory
- An Application: Hashing With Open Addressing
- Do Incredibly Large Naturals Even Exist?
- Primes and Prime Factorization
- The Sieve of Eratosthenes
- Congruences and Congruence Classes

Introduction to Number Theory

- We've defined the **natural numbers** to be the non-negative integers {0, 1, 2, 3,...}. **Number theory** is the branch of mathematics that deals with the naturals.

- We'll define properties of the naturals using quantifiers, starting from basic predicates like $x = y$, $x \leq y$, $x + y = z$, and $x \cdot y = z$. We will give *definitions* of the naturals and these predicates and prove the properties from them.

- Because **counting** is a fundamental human activity, and the naturals are an abstraction of counting, number theory has a long history. We'll see results originally proved in ancient Greece and in medieval China. But there are easily stated questions in number theory to which no one knows the answer.

- Remember that naturals, and integers in general, are different from `int`s.

An Application: Hashing With Open Addressing

- In CMPSCI 187 we studied **hashing**, where a large address space is mapped into a smaller space called a **hash table**. The mapping from address space to hash table cannot be one-to-one, and we have a problem if it fails to be one-to-one on the address values that we actually use. A **collision** is when two relevant addresses are mapped to the same hash address.

- One way of computing a hash address is to divide the original address by the size s of the hash table and let the remainder, in the range from 0 to s - 1, be the hash address.

- One way to deal with collisions, called **open addressing**, has us look at new hash addresses if the first hash address h is full -- we look at h + k, h + 2k, h + 3k,... until we find an empty space in the table.

- We might not want k = 1. Will we still eventually find an open space if one exists?

Do Incredibly Large Naturals Even Exist?

- Some questions of number theory involve ridiculously large naturals. For example, the **Goldbach Conjecture** says that every even natural greater than 2 is the sum of two prime numbers. It is known that if this fails, it fails on a very large number (greater than 10^{18} according to Wikipedia). One paper in theoretical computer science treats all input sizes up to exp*(20) (a tower of twenty two-to-the operations) as a special case.

- If naturals exist in order to count sets, what about naturals that are too big to denote any set of material objects in the universe? Or numbers so big that no computer could ever name them? We say in mathematics that given any property of naturals, either a natural with that property exists or it doesn't.

- Logicians have shown that given any proof system for number theory, there must be statements that are *true, but not provable in the system.* There is some question about what it means for an unprovable statement to be true.

Primes and Prime Factorization

- We'll begin now with the foundations of number theory. The first division, of one natural **dividing** another, was in Monday's lecture. We defined the division relation D so that D(x, y) means $\exists z$: $x \cdot z = y$.

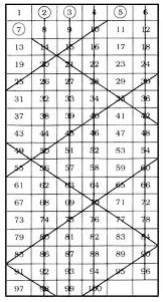
- A **prime number** is a natural, greater than 1, that is divided only by itself and 1. In symbols, we say $P(x) \leftrightarrow (x > 1) \wedge \forall y$: $D(y, x) \rightarrow (y = 1 \vee y = x)$. Numbers greater than 1 that are not prime are called **composite** -- a composite x can be written as $y \cdot z$ where both y and z are greater than 1. By convention, we say that 0 and 1 are neither prime nor composite.

- A composite number can be **factored**, and its factors can also be factored if they are composite. Operating recursively, we can write any positive natural as a product of prime numbers -- its **prime factorization**. In today's discussion we'll practice doing this by hand.

The Sieve of Eratosthenes

- To test whether a natural is prime, we can use **trial division**, seeing whether it has any divisor between 1 and itself. A useful trick to test n for primality is that if n has no divisors in the range from 2 through its **square root**, it is prime.

- The ancient Greeks developed a system to simultaneously test all the numbers in a given range for primality. In the picture, we have listed all the numbers from 1 though 100. We identify 2 as prime and cross out all its multiples. We do the same for 3, 5 and 7. The next prime, 11, is bigger than the square root of 100, so we don't need to check it. 25 of these 100 naturals are prime. They get rarer as we go on.



- Note that after 2 and 3, every prime is one more or one less than a multiple of 6.

Image from Ivars Peterson, *The Mathematical Tourist*

Congruences and Congruence Classes

- We have one more major definition in number theory. Recall that the parity relation P, where P(x, y) means that x and y are both odd or both even, is an equivalence relation. We can write this using the Java `%` operation, in which `x % y` is the remainder when y is divided by x. P(x, y) is true if and only if `x % 2 == y % 2`. Equivalently, P(x, y) is true if 2 divides x - y (or else y - x, whichever is a natural).

- If P(x, y) is true we also say that x and y are **congruent modulo 2**. In general x and y are **congruent modulo k** if `x % k == y % k`, or equivalently if k divides x - y or y - x. For example, 3 and 17 are congruent modulo 7.

- Congruence modulo k is an equivalence relation, and we refer to the equivalence classes of this relation as the **congruence classes modulo k**. Periodic processes in the real world or in computing can be modeled with the system of **modular arithmetic** we will begin studying next week.