1 Primality

We are studying the complexity of two very fundamental, and intimately related, computational problems:

PRIMALITY Given an integer x, is it a prime?

FACTORING Given an integer x, what are its prime factors?

Obviously, PRIMALITY cannot be harder than FACTORING, since, if we knew how to factor, we would definitely know how to test for primality. What is surprising and fundamental —and the basis of modern cryptography— is that PRIMALITY *is easy while* FACTORING *is hard!*

As we know, PRIMALITY can be trivially solved in $O(x)$ time —in fact, we need only test factors up to $\sqrt{x}$. But, of course, these are both exponential algorithms —exponential in the number n of bits of x, which is the more accurate and meaningful measure of the size of the problem (seen this way, the running times of the algorithms become $O(2^n)$ and $O(2^{n/2})$, respectively). In fact, pursuing this line (testing fewer and fewer factors) will get us nowhere: Since FACTORING is hard, our only hope for finding a fast PRIMALITY algorithm is to look for an algorithm that decides whether n is prime without discovering a factor of n in case the answer is "no."

We describe such an algorithm next. This algorithm is based on the following fact about exponentiation modulo a prime:

Theorem 11.1: (Fermat's Little Theorem.) *If p is prime, then for all $a \neq 0 \bmod p$ we have $a^{p-1} = 1 \bmod p$.*

Proof: Consider the set of all nonzero numbers modulo p, $\Phi = \{1, 2, \ldots, p-1\}$. Now if we pick an a in this set, and multiply all these numbers by a, modulo p, we get another set, $\Phi_a = \{a \cdot 1, a \cdot 2, \ldots a \cdot (p-1)\}$, all $\bmod p$. We claim that all $p-1$ numbers in Φ_a are distinct, and therefore $\Phi_a = \Phi$. In proof, if $a \cdot i = a \cdot j \bmod p$, then, by multiplying both sides by $a^{-1} \bmod p$ (since p is prime and $a \neq 0 \bmod p$ we know that a has an inverse) we get $i = j$. Therefore, the products $\Pi_{x \in \Phi} x$ and $\Pi_{x \in \Phi_a} x$ are equal, that is,

$$(a \cdot 1) \cdot (a \cdot 2) \cdots (a \cdot (p-1)) = 1 \cdot 2 \cdots (p-1) \bmod m.$$

Now, multiplying both sides of this equation by $1^{-1} \bmod p$, then $2^{-1} \bmod p$, and so on, all the way to $(p-1)^{-1} \bmod p$, we get Theorem 1. $\square$

Theorem 11.1 suggests a test of primality for p: Take a number $a \neq 0 \bmod p$ and raise it to the $(p-1)$st power modulo p. If the result is not 1, then we know that p is not prime. But what if $a^{p-1} = 1 \bmod p$? Can we be sure that p is prime? Not really. There will always be a's that satisfy this equation (1 and $p-1$ being only the most obvious choices). The *converse* of Theorem 11.1 is not true. All we can prove is this:

Theorem 11.2: If x is *not* a prime, *and if x is not a Carmichael number,* then for *most* $a \neq 0 \bmod x$ $a^{x-1} \neq 1 \bmod x$.

Theorem 11.2 (which it would be a little of a detour to prove now) is a weak converse of Theorem 11.1: It says that *if* x is composite, and *if* x does not happen to be among a set of extremely rare exceptions called

Carmichael numbers,[1] then the primality test suggested by Fermat's Little Theorem will indeed expose the fact that x is not a prime with probability at least 50%.

The above discussion suggests the following *randomized algorithm* for primality:

```
algorithm prime(x)
repeat K times:
    pick an integer a between 1 and x at random
    if a^{x-1} ≠ 1 mod x then return(``x is not a prime'') }
return(``with probability at least 1 - 2^{-K},
        x is either a prime or a Carmichael number'')
```

The claimed probability of correctness is easy to prove: If x is neither prime nor a Carmichael number, then Theorem 11.2 says that each exponentiation will expose this with probability at least $\frac{1}{2}$. Since all these K trials are independent (that is, we choose our a each time without looking at the a's we chose before), the probability that all K failed to expose x is 2^{-K}. Thus, if all exponentiations give 1, the probability that x is prime or Carmichael is at least $1-2^{-K}$. By taking $K=100$, (and remembering that Carmichael numbers are extremely rare), we establish primality to a degree of confidence ($.999\ldots$ up to thirty nines) that surpasses all other aspects of life and computation. The test takes $O(K\cdot n^3)$ steps, where n is the number of bits of x, since it consists of K exponentiations.

Incidentally, there is a slightly more elaborate randomized algorithm that also detects Carmichael numbers, so there is a *polynomial randomized algorithm* for PRIMALITY. We will not cover this more elaborate algorithm in this course.

Not only are primes easy to detect, but they are also relatively ambudant:

Theorem 11.3: (The Prime Number Theorem.) *The number of prime numbers between* 1 *and x is "about" $\frac{x}{\ln x}$. (In other words, if I take a random number with D decimal digits, the chances of it being prime is just a little less than one in* 2*D. "One in about twenty social security numbers is a prime.")*

The Prime Number Theorem is one of the most fundamental facts in Mathematics, and very hard to prove. But, together with the previous algorithm, it enables us to easily find large primes (try a few, and pick one that tests well), and this is a key ingredient of the RSA algorithm. Another ingredient is the following variant of Theorem 11.1:

Theorem 11.4: If p and q are primes, then for all $a\neq 0 \bmod p,q$ we have $a^{(p-1)\cdot(q-1)} = 1 \bmod p\cdot q$.

Theorem 11.4's proof is exactly the same as Theorem 1's, except that we consider the set of all numbers among $1,2,\ldots,p\cdot q-1$ *that are relatively prime to $p\cdot q$.* Notice that there are $(p-1)\cdot(q-1)$ such numbers —check it out, one in every p of the numbers between 0 and $p\cdot q$ are divisible by p, and one every q by q, and $pq(1-\frac{1}{p})(1-\frac{1}{q}) = (p-1)(q-1)$.

Example. Let us take $p=3$ and $q=5$. From among all numbers modulo $p\cdot q=15$, namely $\{0,1,2,\ldots,14\}$, one-third are divisible by 3 (namely, $\{0,3,6,9,12\}$), and of the remaining 10, one-fifth are divisible by 5 ($\{5,10\}$). All $(p-1)\cdot(q-1)=8$ of the remaining numbers ($\Phi=\{1,2,4,7,8,11,13,14\}$) are prime to 15, and therefore they all have an inverse modulo 15. This enables the proof of Theorem 11.1, by taking a to be

[1]A number c is a Carmichael number if it is not a prime, and still for all prime divisors d of c with $d>1$ it so happens that $d-1$ divides $c-1$. The smallest Carmichael number is $561=3\cdot 11\cdot 17$ (notice how indeed $3-1, 11-1$, and $17-1$ all divide $561-1$). If c is a Carmichael number and a is relatively prime to c, then $a^{c-1}=1 \bmod c$. Can you prove it?

any one among these 8 numbers.