ITCS 531: Number Theory 3 - Primality testing

Rob Egrot

$\mathbb{Z}_n$

- In the previous class we saw that arithmetic modulo n 'makes sense'.
- I.e. we can define operations of addition, subtraction and multiplication on equivalence classes modulo n for all $n \in \mathbb{N} \setminus \{0\}$.
- The definition below is used to pick out the number system defined by looking at integers mod n.

Definition 1 ($\mathbb{Z}_n$)

If $n \in \mathbb{N} \setminus \{0\}$ then $\mathbb{Z}_n$ is the set of integers mod n.

Multiplicative inverses

- In standard arithmetic over $\mathbb{R}$, every number except 0 has an inverse under multiplication.
- That is, for all $x \in \mathbb{R} \setminus \{0\}$ there is $y \in \mathbb{R} \setminus \{0\}$ with $xy = 1$.
- We write x^{-1} or $\frac{1}{x}$ for the multiplicative inverse of x.
- In the integers $\mathbb{Z}$, only the numbers 1 and -1 have an inverse.
- In $\mathbb{Z}_n$ this is not usually true.

Inverses in $\mathbb{Z}_n$

Definition 2 (Modular multiplicative inverse)

For $a \in \mathbb{Z}$ we define $b \in \mathbb{Z}$ to be the multiplicative inverse, or just the *inverse*, of $a \mod n$ if $ab \equiv_n 1$.

- We write a^{-1} for the multiplicative inverse of a (when it exists!).
- Soon we will prove a result that tells us exactly when integers have an inverse mod n.

Coprimality

Definition 3 (Coprime)

Integers a and b are *coprime* if their highest common factor (**HCF**) is 1

Lemma 4

Let $a, b, c \in \mathbb{Z}$*, let* $a|bc$*, and let* a *and* b *be coprime. Then* $a|c$*.*

Proof.

This is exercise 3.1. $\square$

- Now we have all we need to prove the first important result of the class.

When do inverses exist in $\mathbb{Z}_n$?

Proposition 5

- *Let $a \in \mathbb{Z}$ and let $n \in \mathbb{N} \setminus \{0\}$.*
- *Then a has multiplicative inverse $\mod n$ if and only if a and n are coprime.*
- *Moreover, the multiplicative inverse of $a \mod n$ is unique in $\mathbb{Z}_n$, whenever it exists.*

Proof of proposition 5 part 1

- This proof has three parts.
  - We must show that *if* a and n are coprime, *then* a has an inverse in $\mathbb{Z}_n$.
  - Also *if* a has an inverse in $\mathbb{Z}_n$, *then* a and n are coprime.
  - Finally, if b and c are both inverses to a mod n, then $b \equiv_n c$.
- Suppose a and n are coprime.
- Since a and n are coprime, it follows from Bézout's identity that there are $x, y \in \mathbb{Z}$ with $xa + yn = 1$.
- So $xa - 1 = -yn$, but this means that $xa \equiv_n 1$ by definition.
- So a has an inverse in $\mathbb{Z}_n$ as required.

Proof of proposition 5 part 2

- Now suppose that a has an inverse, and call it x.
- Then we have $xa \equiv_n 1$.
- I.e. there is y with $xa - 1 = yn$.
- We can rewrite this as $xa - yn = 1$.
- Suppose $d|a$ and $d|n$.
- Then $d|(ax - yn)$, and so $d|1$.
- The only way this can be true is if $d = \pm 1$.
- This means $HCF(a, n) = 1$, and so a and n are coprime.

Proof of proposition 5 part 3

- Finally, if an inverse to a exists then we have just shown that (a, n) must be coprime.
- Let $ab \equiv_n 1$ and $ac \equiv_n 1$.
- Then there are $k, l \in \mathbb{Z}$ with $ab - 1 = kn$ and $ac - 1 = ln$.
- So $a(b - c) = (k - l)n$.
- We obviously have $a|a(b - c)$.
- So by lemma 4 we must have $a|(k - l)$.
- So $b - c = \frac{k-l}{a}n$, and $\frac{k-l}{a} \in \mathbb{Z}$.
- So $b \equiv_n c$.

Investigating prime factorization

- Now we know the basics of modular arithmetic, we can start to seriously study prime numbers and prime factorizations.
- The difficulty of finding the prime factors of large numbers is the basis for much of modern cryptography (i.e. RSA).
- An old result about prime numbers known as *Fermat's little theorem* will be important.
- This neat theorem gives us a kind of detector for numbers which are not prime (i.e. composite numbers).
- With some ingenuity this can be turned into a powerful probabilistic method for testing whether a number is prime.
- First we will need another small technical lemma.

Injective multiplication

Lemma 6
Let $a \in \mathbb{Z} \setminus \{0\}$ and $n \in \mathbb{N} \setminus \{0\}$ be coprime. Then, for all $b, c \in \mathbb{Z}$, if $ab \equiv_n ac$, we have $b \equiv_n c$.

Proof.

- Since a and n are coprime, by proposition 5 we know a has a multiplicative inverse $a^{-1} \pmod{n}$.
- So $a^{-1}ab \equiv_n a^{-1}ac$.
- And so $b \equiv_n c$ by definition of the inverse.

$\square$

Fermat's little theorem

Theorem 7 (Fermat's little theorem)

If p is prime then $a^{p-1} \equiv_p 1$ whenever a and p are coprime.

Proof.

- By lemma 6 we have

$$\{1, 2, 3, \ldots, p-1\} = \{a \mod p, 2a \mod p, 3a \mod p, \ldots, (p-1)a \mod p\}.$$

- So, by multiplying

$$(p-1)! \equiv_p a^{p-1}(p-1)!. \qquad (\dagger)$$

- Now, since p is prime, it follows that p cannot divide $(p-1)!$.
- So p and $(p-1)!$ are coprime.
- By proposition 5 it follows that $(p-1)!$ has an inverse modulo p.
- Multiplying $(\dagger)$ by this inverse gives $a^{p-1} \equiv_p 1$ as required.

$\square$

Primality testing with Fermat's little theorem

- Fermat's little theorem gives us an efficient way we can test whether a number is prime.
- Given $n \in \mathbb{N}$ we pick a with $1 < a < n$, then calculate a^{n-1} mod n.
- If this is not 1 then n is not prime, by Fermat's little theorem.
  - As if n is prime then a would automatically be coprime with n.
- However, $a^{n-1} \equiv_n 1$ does not imply that n is prime.
- This is because Fermat's little theorem only tells us that *if* p is prime *then* $a^{p-1} \equiv_p 1$.
- It doesn't say that if p is *not* prime then $a^{p-1} \not\equiv_p 1$.
- For example, $341 = 11 \times 31$, but $2^{340} \equiv_{341} 1$.

Evidence for primality

- Passing Fermat's test does give us evidence that a number is prime, due to the following result.

Lemma 8

Let $n \in \mathbb{N}$ and suppose there is $1 \leq a < n$ such that a is coprime with n and $a^{n-1} \not\equiv_n 1$. Then the modular inequality $b^{n-1} \not\equiv_n 1$ must hold for at least half the natural numbers b less than n.

Proof.

- Suppose $b < n$ and b passes Fermat's test (i.e. $b^{n-1} \equiv_n 1$).
- Then ab fails Fermat's test, because $(ab)^{n-1} = a^{n-1}b^{n-1} \equiv_n a^{n-1} \not\equiv 1$.
- Moreover, if $ab \equiv_n ac$ for $1 < b, c < n$ then $b = c$ (lemma 6).
- So every element that passes Fermat's test has a partner that doesn't, and these partners are all distinct.
- So there are at least as many elements that fail as that pass.

$\square$

Probabilistic primality testing

- If there is at least one value a that is coprime with n with $a^{n-1} \not\equiv_n 1$, this gives us a good test for determining whether a number n is prime:
  - We repeat Fermat's test k times with different random numbers a with $1 < a < n$.
  - If the test fails for any a we conclude with certainty that n is not prime (by the little theorem).
  - If every test is passed we conclude that the probability that n is not prime must be at most $\frac{1}{2^k}$.
  - Because, if n is not prime, every a provides at least a 50% chance of making n fail the test.
  - So, for high confidence just choose large k.
  - This test is always correct when it says a number is composite, but it occasionally says a number is prime when it is not.

Problems with Carmichael numbers

- There is a small problem with this.
- Lemma 8 relies on the existence of at least one a that is coprime with n and fails Fermat's test (i.e. $a^{n-1} \not\equiv_n 1$).
- Unfortunately, there are composite numbers where every coprime a passes Fermat's test.
- These numbers are called *Carmichael numbers*.
- The smallest Carmichael number is 561.
- This is not prime as $561 = 3 \times 11 \times 17$, but for every $1 < a < 561$ that is coprime to 561 we have $a^{560} \equiv_{561} 1$.
- So our probability calculation from before is not correct.
- There are an infinite number of Carmichael numbers, but they are rare, so Fermat's test works most of the time.
- There are also more advanced methods, like the Rabin-Miller test.

Roots of polynomials in $\mathbb{Z}_n$

- Remember that a polynomial with variable x and degree n is a function
$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$
where $a_0, \ldots, a_n$ are fixed parameters.
- A polynomial over $\mathbb{R}$ can have, at most, the same number of real roots as its degree
  - A *root* of a single variable function f is a value x such that $f(x) = 0$.
- The Fundamental Theorem of Algebra says that polynomial over $\mathbb{R}$ has exactly the number of complex roots as its degree, but this is not in the scope of this course.
- We will show soon that the limit on the number of roots of a polynomial we have just described also applies to polynomials over $\mathbb{Z}_p$, when p is prime.

A special case of polynomial division

- We will use this following lemma.

Lemma 9
If $x, y \in \mathbb{R}$ *then*

$$\frac{x^n - y^n}{x - y} = x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \ldots + xy^{n-2} + y^{n-1}.$$

Proof.
Direct calculation of

$$(x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \ldots + xy^{n-2} + y^{n-1})$$

shows it is equal to $x^n - y^n$. $\square$

- I.e. the polynomial $x - y$ divides the polynomial $x^n - y^n$.
- Note that this works even if $x = y$.
- For polynomials, potential division by zero makes sense.

Polynomials over $\mathbb{Z}_p$

- Let f be a polynomial over $\mathbb{Z}$ of degree n. I.e.

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

where $a_i \in \mathbb{Z}$ for all i and $a_n \neq 0$.

- Let p be a prime number.

- We define $f_p(x) = a'_0 + a'_1 x + a'_2 x^2 + \ldots + a'_n x^n$ where each $a'_i = a_i \mod p$ for all i.

- So f_p is f converted to being a polynomial over $\mathbb{Z}_p$.

- E.g. if $f(x) = 8 + 14x + 3x^2$, then $f_5(x) = 3 + 4x + 3x^2$.

Lagrange's theorem on polynomial roots

Theorem 10 (Lagrange)

Let p be prime, let $f(x) = a_0 + a_1x + \ldots + a_mx^m$ be a polynomial over $\mathbb{Z}$, and let f_p be as above. Suppose the degree of f_p is n. Then, unless every coefficient of f_p is zero, f_p has at most n distinct roots modulo p.

Lagrange's theorem on polynomial roots - proof 1

- You don't need to remember this proof.
- First note that the degree of f_p must be less than or equal to the degree of f, i.e. we must have $n \leq m$.
- We induct on n, the degree of f_p.
- Remember we're trying to show f_p is either zero or has at most n roots (mod p).
- The result is clearly true when $n = 1$, because here we have $f_p = a_0' + a_1' x$, and the root occurs when $x \equiv_p -a_0' a_1'^{-1}$.

Lagrange's theorem on polynomial roots - proof 2

- Suppose now that the result is true for all $n \leq k$.
- Let the degree of f_p be $k+1$.
- Suppose that f_p has a root b modulo p. I.e. $f_n(b) \equiv_p 0$.
- If such a root does not exist then we are done, as $0 \leq n$.
- Consider the polynomial

  $$f_p(x)-f_p(b) = a'_1(x-b)+a'_2(x^2-b^2)+\ldots+a'_{k+1}(x^{k+1}-b^{k+1}).$$

- By lemma 9, $(x-b)$ divides (x^l-b^l) for all $1 \leq l \leq k+1$, so we can define a polynomial $g(x) = \frac{f_p(x)-f_p(b)}{x-b}$ over $\mathbb{Z}_p$.

Lagrange's theorem on polynomial roots - proof 3

- By definition of g we have $f_p(x) - f_p(b) = (x-b)g(x)$.
- Moreover, g has degree at most k.
- Let c be a root of $f_p(x)$ modulo p.
- Then, setting $x = c$ we get $0 \equiv_p (c-b)g(c)$, as b is also a root of f_p mod p.
- I.e. $p|(c-b)g(c)$.
- Since p is prime this means either
  a) $p|(b-c)$, which happens if and only $c \equiv_p b$, or
  b) $p|g(c)$, in which case c is a root of $g(x)$ modulo p.
- But, by the inductive hypothesis, there are at most k roots of g modulo p.
- So there at most $k+1$ roots of f_p modulo p.