Math 189H Joy of Numbers Activity Log

Tuesday, December 6, 2011

Paul Erdös: "A mathematician is a device for turning coffee into theorems." [It was suggested that 'coffee' should be replaced with 'caffeine'.]

African proverb: "If you think you are too small to make a difference, try sleeping with a mosquito in the room."

Today started with two presentations. With a coin flip determining the order of presentation, we first heard from Logan and Will on the Chinese Remainder Theorem, which is a result first discovered in the 12th century B.C.E. for solving simultaneous systems of congruence equations:

$z \equiv r_1 \bmod n_1$
to
$z \equiv r_k \bmod n_k$

The method discovered by Qin Jiushao was, setting $N = n_1 n_2 \cdots n_k$, to solve k equations one at a time, and then combine their answers into a solution to the original problem. His method will work if the moduli n_i are all relatively prime to one another. Setting $x_i = N/n_i$, if we solve $x_i g_i \equiv r_i \bmod n_i$ (which we can, since x_i and n_i are relatively prime), then his solution was given by $z = g_1 x_1 + \cdots + g_k x_k$, since every term in the sum except $g_i x_i$ is a multiple of n_i so does not affect the remainder ot z on division by n_i (!). Logan and Will illustrated this solution with several examples motivated by eggs in a basket and partgoers breaking into groups!

This was followed by Merideth and Chad discussing the Gregorian calendar, which was introduced by Pope Gregory XIII in 1582 as a system for assigning dates that more accurately fit with the solar year. (The previous Julian calendar fell a day ahead of the solar year every 128 years, which over time had shifted dates by 10 days!) By reverse-engineering the Gregorian calendar, we can build a formula (using modular arithmetic and our favorite 'floor' function) to compute the day of the week that any particular date fell or will fall on (after 1582?!). To make leap year easier to deal with, the formula numbers the month m with March starting as 1, we then need the day of the month k, the century c, and the year within the century (i.e., the year modulo 100) d. Then, with Sunday being 0 and numbering forward, the day of the week can be computed as

$$k + \lfloor \frac{13m-1}{5} \rfloor + d + \lfloor \frac{d}{4} \rfloor + \lfloor \frac{c}{4} \rfloor - 2c \text{ modulo } 7$$

Merideth and Chad illustrated this with several of our birth dates (which I will not reproduce here, both because I don't remember them (other than my own) and to reduce the risk of identity theft!).

After the presentations, we turned our attention to some further refinements of Euler's Theorem. We have learned that if p is prime, then for every a other than a multiple of p we have $a^{p-1} \underset{p}{\equiv} 1$. [This result was discovered a bit before Euler's more general result, by Pierre de Fermat, and is usually known as 'Fermat's Little Theorem'.] But we have ourselves encounter composite numbers n having a's with $a^{n-1} \underset{n}{\equiv} 1$, the first being $n = 91$ with $a = 2$, and more spectacularly some Carmichael numbers. This result in the end is also not really a test for primality so much as providing proof of non-primality; if $a^{n-1} \not\equiv 1 \bmod n$ but $\gcd(a, n) = 1$, then n cannot be prime. On the other hand, a Carmichael number can never be proven composite using this test. We'd like to do better!

One way to improve things is built upon another observation we've made (or rather, you made, on one of your exams?); if p is prime and $a^2 \underset{p}{\equiv} 1$, then it must be the case that either $a \underset{p}{\equiv} 1$ or $a \underset{p}{\equiv} -1$ (just as with ordinary arithmetic!). But what happens if n is not prime? How many solutions to $a^2 \underset{n}{\equiv} 1$ can we have (with $1 \leq a \leq n$)? Since as things stand to answer this question for a particular n we would have to test every a from 1 to n, and that sounded tedious, we once again let Maple 15 do the heavy lifting, and tested an assortment of values for n. What we found was (this is simulated data, the Maple worksheet chose 6-digit numbers at random!)

for $n = 496389 = 3 \cdot 165463$, there were 4 solutions,
for $n = 157167 = 3^3 \cdot 5821$, there were 4 solutions,
for $n = 105651 = 3^3 \cdot 7 \cdot 13 \cdot 43$, there were 16 solutions,
for $n = 331362 = 2 \cdot 3^2 \cdot 41 \cdot 449$, there were 8 solutions,
for $n = 593086 = 2 \cdot 13 \cdot 22811$, there were 4 solutions,
for $n = 514121 = 19 \cdot 27059$, there were 4 solutions,
for $n = 802175 = 5^2 \cdot 11 \cdot 2917$, there were 8 solutions, and
for $n = 836304 = 2^4 \cdot 3 \cdot 7 \cdot 19 \cdot 131$, there were 64 solutions!

Based on this (somewhat limited) evidence, we decided that even numbers behaved weirdly [but on the other hand, the only numbers we would ever be testing for primality would be odd ones!]. For the odd numbers, we definitely detected a pattern; the number of a with square congruent to 1 modulo n appears to be 2 raised to the number of distinct prime factors of n. We can in fact, borrowing the work of Logan and Will, see why this might be so; dissecting a particular (small) example, $n = 187 = 11 \cdot 17$, the numbers squaring to 1 are $1, 67, 120$, and $186 \equiv -1$. 1 and -1 are no surprise, perhaps, but what is so special about 67 and 120? Looked at through the lens of the factors of 187, we noted something: 67 is $\equiv 1 \bmod 11$, and is $\equiv -1 \bmod 17$, while 120 is $\equiv -1 \bmod 11$ and $\equiv 1 \bmod 17$. So the solutions to $a^2 \equiv 1 \bmod 11 \cdot 17$ turn out to be the solutions to the system(s) of equations

$a \equiv \pm 1 \bmod 11$
$a \equiv \pm 1 \bmod 17$

for each of the $4 = 2^2$ possible choices of pairs of signs ±. This holds more generally; if $n = p_1^{k_1} \cdots p_r^{k_r}$ (is odd!), then using Logan and Will's technique to solve the system $a \equiv \pm 1 \bmod p_i^{k_i}$ for any choice of signs ±, then that solution will automatically satisfy $a^2 \equiv 1 \bmod p_i^{k_i}$ for every i, so $p_i^{k_i} | a^2 - 1$, so $n = p_1^{k_1} \cdots p_r^{k_r} | a^2 - 1$ (since the prime powers are all relatively prime to one another!), so $a^2 \equiv 1 \bmod n$. This gives our 2^r solutions to $a^2 \underset{n}{\equiv} 1$ (!). [Tuesday's class essentially ended here.]

The point we need to take away from this is that there are more solutions to $a^2 \underset{n}{\equiv} 1$ when n is (odd and) composite than when n is prime. This fact was turned into a better (non-)primality test than Euler's Theorem turns out to be, by Miller and Rabin, in the mid-1970's. [Yes, we have gotten that close to research-level mathematics!] Their idea was to use the fact that passing the Euler Theorem 'test', $a^{n-1} \underset{n}{\equiv} 1$, provided a further test, since (if n is odd [otherwise why are you testing for primality?!] then) $n-1 = 2k$ is even, we have $a^{n-1} = a^{2k} = (a^k)^2$, so, if n is prime, it must be the case that $a^k \underset{n}{\equiv} 1$ or $a^k \underset{n}{\equiv} -1$. This is a further test! But even more, if $a^k \underset{n}{\equiv} 1$ and $k = 2\ell$ is still even, then it must be that $a^\ell \underset{n}{\equiv} 1$ or $a^k\ell \underset{n}{\equiv} -1$ (if n is prime). This really creates a series of tests for the primality of n, so long as we can extract further factors of 2 from $n-1$. In its most streamlined form, the *Miller-Rabin Primality Test* says:

If n is prime, and $n-1 = 2^k m$ with m odd, then looking at the sequence of numbers $a^m, a^{2m}, a^{2^2m}, \dots a^{2^km} = a^{n-1}$, mod n, either $a^m \underset{n}{\equiv} 1$ or one of the terms is $\underset{n}{\equiv} -1$ (and is then followed by at least one 1(!)). Turning this around, if the last term in the sequence is not $\underset{n}{\equiv} 1$, or the term that immediately precedes the first term that is 1 is not $\underset{n}{\equiv} -1$, then n cannot be prime.

But Miller and Rabin went one step further. If a number n 'passes' the Miller-Rabin test for some base a (meaning the test makes it look like it is prime), we will call n a strong pseudoprime to the base a. What they also showed that if n is not prime, then it can be a strong pseudoprime to at most 1/4th of the bases relatively prime to n. That means that roughly 3/4ths of the a's between 1 and n must be witnesses to the non-primality of n. Which are pretty good odds!

Every 'practical' primality test you are likely to encounter (for example, the one built into Maple 15) uses this property of the Miller-Rabin test. It more or less randomly chooses some number s of bases a_i (s=100? 1000?), incidentlly checks that each is relatively prime to n (otherwise it has accidentally stumbled across a factor, which is a really good witness to the non-primality of n), then, starting with $n-1 = 2^k m$ with m odd, computes $a^m, (a^m)^2, (a^m)^4, \dots (a^m)^{2^k} = a^{n-1}$ in turn, mod n. If this sequence starts with 1, it stops and moves on the to the next base a. Otherwise, if it encounters a 1 without first meeting a -1, it stops and declares n to be composite with a as a 'witness'. And if it never encounters a 1, then $a^{n-1} \not\equiv 1$ mod n, so again it is composite with a as a witness. And if it makes it all the way to the end and meets a 1, it moves on to the next base. If it gets all of the way through its choice of bases without ever meeting a witness, it declares n to be an 'industrial grade' prime (as most people term them); by Miller-Rabin, the odds that n is not prime are likely on the order of $4^{-s} \approx 10^{-.6s}$, which for a large value of s is a very small number! If, for example, it is smaller than $1/n$, then we are essentially saying that there is probably only one composite number the same size as the number we are working with that would have been mistakenly identified as not composite. And since all of this machinery is relying on (fast) exponentiation, we can crank out very-highly-probably primes in the numbers (no pun intended) that modern commerce really requires.