1 Chinese Remainder Theorem and Digital Fingerprints

The Chinese remainder theorem

Suppose we have a system of simultaneous equations, like maybe this one:

$$\begin{aligned} x &\equiv 2 \pmod 5 \\ x &\equiv 5 \pmod 7 \end{aligned}$$

What can we say about x? Well, notice that one solution is $x = 12$; $x = 12$ satisfies both equations. This is not the only solution: for instance, $x = 12 + 35$ also works, as does $x = 12 + 70$, $x = 12 + 105$, and so on. Evidently adding any multiple of 35 to any solution gives another valid solution, so we might as well summarize this state of affairs by saying that $x \equiv 12 \pmod{35}$ is one solution of the above system of equations.

What about other solutions? For this example, there are no other solutions; every solution is of the form $x \equiv 12 \pmod{35}$. Why not? Well, suppose x and x' are two valid solutions. From the first equation, we know that $x \equiv 2 \pmod 5$ and $x' \equiv 2 \pmod 5$, so we must have $x \equiv x' \pmod 5$. Similarly $x \equiv x' \pmod 7$. But the former means that 5 is a divisor of $x - x'$, and the latter means that 7 is a divisor of $x - x'$, so $x - x'$ must be a multiple of 35 (here we have used that $\gcd(5,7) = 1$), which in turn means that $x \equiv x' \pmod{35}$. In other words, all solutions are the same modulo 35: or, equivalently, if all we care about is $x \bmod 35$, the solution is unique.

You can check that the same would be true if we replaced the numbers $5,7,2,5$ above by any others. The only thing we used is that $\gcd(5,7) = 1$.

Here is the generalization:

Theorem 10.1: *(The Chinese remainder theorem.) Let m,n be relatively prime, and let a,b be arbitrary. Then the pair of equations $x \equiv a \pmod m$, $x \equiv b \pmod n$ have a unique solution for $x \bmod mn$.*

Moreover, the solution x can be computed efficiently

Proof: Let us first showing how to find a solution to the pair of equations. We begin by showing how to find two numbers u and v with the property that $u \equiv 1 \bmod m$ and $u \equiv 0 \bmod n$, while $v \equiv 0 \bmod m$ and $v \equiv 1 \bmod n$. Once we have these numbers it is easy to construct x, since we can choose $x = au + bv$. Why does this work?

Let us see how to construct u. The second condition on u states that it must be a multiple of n, say kn. How do we find k? The first condition says that $kn \equiv 1 \bmod m$. This means $k \equiv n^{-1} \bmod m$, which must exist since $gcd(m,n) = 1$. The construction of v is similar. By the previous discussion we can now find an x that satisfies both equations.

To show that x is unique $\bmod mn$, suppose there is an x' that satisfies both equations. Then $x \equiv x' \bmod m$ which implies that $m|(x-x')$. Similarly $n|(x-x')$. Since m and n are relatively prime, this means that $mn|(x-x')$. Thus $x \equiv x' \bmod mn$. $\square$

The Chinese remainder theorem is often useful when doing modular arithmetic with a composite modulus; if we want to compute some unknown value modulo mn, a standard trick is to compute it modulo m, compute it modulo n, and then deduce its value mn using the Chinese remainder theorem (CRT).

1.1 Digital Fingerprinting

Suppose we wish to test a file on the moon (on our spacecraft) is uncorrupted, by checking it against the reference file on earth. We could transmit the entire file and check, but bandwidth to the moon is very expensive. Fingerprinting provides a nice way of doing this. Suppose the file on earth is an n bit string x and the file on the moon is y. We can interpret x and y as numbers $0 \leq x, y < 2^n$ in the obvious way.

Now we select a random prime p chosen at random from all primes up to n^3. Send $F_p(x) = x \bmod p$ as the fingerprint of the file, together with the prime p that was chosen. On the moon, the fingerprint $F_p(y)$ is calculated and if the fingerprints don't match then there the file must have been corrupted. Of course it is possible that $x \neq y$ and we happened to choose p such that the fingerprints happened to match up. We will show below that this is extremely unlikely. But before doing so, let us observe that $F_p(x)$ has length $4logn$ bits, and so is much shorter than x.

Now let us assume that $x \neq y$. What is the probability that the fingerprints of x and y will match up when we choose p at random? The prime number theorem tells us that there are at least m/lnm primes less than or equal to m. Thus there are $n^3/lnn^3 = n^3/3lnn$ primes among which p is chosen at random. Now we claim that if $x \neq y$ then the fingerprints $F_q(x) = F_q(y)$ for at most n primes q. These are the bad choices of primes for p. But since p was chosen at random from among $n^3/3lnn$ primes, at most n of which are bad, it follows that the probability that p was bad is at most $n/(n^3/3lnn) << 1/n$. So the error probability is pretty low, and the algorithm is efficient: $(nlogn)$ time to compute the fingerprint and $O(logn)$ bits to transmit.

To finish the analysis, let's prove the claim that if $x \neq y$ then the fingerprints $F_q(x) = F_q(y)$ for at most n primes q. Proof by contradiction. Suppose that there are more than n primes q such that $x \equiv y \bmod q$. Then if M is the product of these primes, then by the Chinese remainder theorem, $x \equiv y \bmod M$. Moreover $M > 2^n$ since it is the product of more than n numbers each of which is at least 2. Since $0 \leq x, y < 2^n - 1$, $x \equiv y \bmod M$ implies that $x = y$. Contradiction. Therefore the claim follows.