15-503/859P: Cryptography **Lecturer:** Manuel Blum
Topic: Ending Zero Knowledge; Beginning Number Theory **Date:** 8 Feb 2006
Scribe: Brandon Kuczenski

5.1 Overview

- Finish Off Zero Knowledge
- Get Into Number Theory
  - Extended GCD (Greatest Common Denominator)
  - Continued Fractions (a tabular way to compute GCD)

5.2 Manuel's Research Agenda

A couple of times when Manuel was in college, a professor would tell him and his peers something about their research agenda, something they thought was interesting, and more often than not they turned out to be right. He will now endeavor to do the same for us.

"What I would like from Zero Knowledge"

We don't know any short, simple zero-knowledge proof that a graph does *not* have a Hamilton Cycle. Of course, we don't have any way *at all* to show that a graph does not have such a cycle—that would be tantamount to showing that Hamilton Cycle is in coNP. But we do have short, simple proofs for certain special cases.

Manuel would like to be able to convert that into a short, simple, straightforward *Zero Knowledge* proof. For cases in which a simple proof exists, is there also a simple way to do it in Zero Knowledge?

For example, the Rubik's Cube. A graph representing a Rubik's Cube contains 27 nodes corresponding to the cells in the 3x3x3 cube, and an edge connecting each pair of nodes for which the corresponding Rubik's Cube cells share a face. It turns out that this graph does not have a Hamilton Cycle, and there is a simple proof (though he did not elucidate it to us). But to convert this proof into Zero Knowledge is cumbersome.

Another example: prove that there exists an irrational number α such that α^α is rational. This proof is very, very short, and it's not very hard. But how can it be done in Zero Knowledge? In other words, how can you prove the theorem without exposing to the verifier any clue about how it's proved?

Logic Cards

Imagine a deck of cards, on each of which is printed a number and some kind of a logical statement, like $(A \cdot B) \Rightarrow C$. These cards are dealt out to different people. Say you have the above card. You wait until somebody with a lower number than you tells you that "A is true;" and somebody else with a lower number tells you that "B is true;" and another person with a lower number tells you that the implication itself, $(A \cdot B) \Rightarrow C$, is true. Then you are directed to somebody with a higher number and tell him or her with confidence that, based on the information you've been given, C is true.

Imagine there is a symbolic statement of the proof of the Rubik's Cube theorem in which the symbols have all been rearranged and obfuscated to the point that you have no idea to what A, B, or C refer. Then your possession of the knowledge of the $(A \cdot B) \Rightarrow C$ implication will still give you no idea how the Rubik's Cube theorem was proved, despite your participation in the proof.

This is but a brief paraphrase of the technique, but it can indeed be carried out in a rigorous, Zero Knowledge way. However, there are far too many cards. The Zero Knowledge protocol would allow the

Verifier to look at any card, convince herself that it is valid, and then the deck would be destroyed. A newly randomized deck of cards could be produced as many times as desired to convince the Verifier the proof is valid.

5.3 HOL-Light

These types of formal systems for representing mathematical proofs rigorously as chains of logical inferences exist. One noteworthy one is HOL-Light, in which HOL stands for "Higher-Order Logic." It is a Formal Proof Verifier, which means it uses established logical rules of inference to represent, and then verify, mathematical proofs.

HOL-Light is being used by Tom Hales to prove Kepler's Conjecture, colloquially known as the "Grocer's Problem" or the "Fruit Stacking Problem." Kepler's conjecture, now thought to be proven by Hales, was that the grocer's technique is amost space-efficient way to stack spheres, which could be thought of as generalized fruit[1].

This conjecture had been an open question for a very long time, but Hales developed a proof and sent it off to be published. The proof was divided up amongst experts who considered it, worked on it, taught courses on it, and after five or so years, the reviewers came to the publisher and announced they were 99% certain the proof was correct.

Of course, in Mathematics, 99% is not enough. So Hales set about converting his proof into the language of HOL-Light. Once the process is complete, the HOL-Light program will be able to run through the proof, line-by-line, and announce whether it is logically sound. Unfortunately, at this point it takes Prof. Hales one day to convert one line of his proof into approximately one page of HOL-Light, a rate at which the conversion will be complete in about 300 years.

Because there is a HOL-Light, there is a guaranteed way to take any mathematical theorem that has a proof and convert it to a Zero Knowledge proof.

Unfortunately, HOL-Light is cumbersome. It would be highly desirable to find a way to generate a short, simple zero knowledge proof from a short, simple mathematical proof.

5.4 Number Theory: Greatest Common Divisor (GCD)

Theorem 1 (Extended GCD) *Let $a, b \in \mathbb{Z}^+$. Then $\exists d \in \mathbb{Z}^+$ which is the largest integer that divides a and b. Furthermore, $\exists u, v : ua - vb = d$ with $0 \leq u < b, 0 \leq v < a$.*

Of course, the interesting part of the theorem is the u and v. For one, they show that d is the largest by way of the equation $ua - vb = d$. If a larger number s divided both a and b, then it could be factored out of the right hand side, and then d would be represented as the product of two integers, one of which is larger than d, a contradiction.

5.5 Continued Fractions: another approach

Let us write π as a continued fraction.

$$\pi = 3.1415926\ldots = 3 + .1415926\ldots = 3 + \frac{1}{7.06251\ldots} = 3 + \frac{1}{7 + .06251\ldots} = 3 + \frac{1}{7 + \frac{1}{15.992\ldots}}\ldots \quad (1)$$

Of course, writing nested fractions like that takes up two-dimensional space. As a space-saving measure, we write continued fractions like this:

$$\pi = 3 + 1/(7 + 1/(15 + 1/(1 + 1/(292 + \ldots)))) \quad (2)$$

[1] In fact, there are at least two different ways to stack spheres which both have the same optimal volumetric efficiency, a little more than 74%.

In general, a decimal number can be represented by a series of *partial quotients*:

$$q = q_0 + 1/(q_1 + 1/(q_2 + 1/(q_3 + \ldots))) \tag{3}$$

or even more compactly as

$$q = [q_0; q_1, q_2, q_3, \ldots] \tag{4}$$

where the more q's you add, the better your approximation.

You can write the resultant approximation as a ratio of integers P_n and Q_n, the results of computing through the nth q. For π,

$$\frac{P_0}{Q_0} = 3 = \frac{3}{1}; \tag{5}$$

$$\frac{P_1}{Q_1} = 3 + \frac{1}{7} = \frac{22}{7}; \tag{6}$$

$$\frac{P_2}{Q_2} = 3 + 1/(7 + 1/(15)) = \frac{333}{106} \tag{7}$$

The fractions P_k/Q_k are known as approximates.

To compute these conventionally, we start with the most-nested (e.g. the rightmost) fraction and work our way left. Question: how can we work incrementally, from left to right?

Answer: by looking at small examples and extrapolating, we can develop a model we can prove inductively.

Example

Consider the continuing fraction $1+1/(1+1/(1+1/(1+\ldots)))$, written as $[1; 1, 1, 1, 1, \ldots]$. The approximates are:

$$\frac{1}{1}, \frac{2}{1}, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \ldots \tag{8}$$

Now, how about $[1; 1, 2, 1, 3, 1, \ldots]$. The approximates look like:

$$\frac{1}{1}, \frac{2}{1}, \frac{5}{3}, \frac{7}{4}, \ldots \tag{9}$$

If we're really smart, we will infer something that looks like:

$$\frac{P_{k+1}}{Q_{k+1}} = \frac{P_{k-1} + P_k q_{k+1}}{Q_{k-1} + Q_k q_{k+1}} \tag{10}$$

which we can prove by induction. If we take P_0, Q_0 to be the integer part of the fraction and P_1, Q_1 to be $1 + P_0 q_1, q_1$, then we can show the base ($k = 1$) case simply.

For the inductive case, assume P_k, Q_k can be written as $P_{k-2} + P_{k-1} q_k,\ Q_{k-2} + Q_{k-1} q_k$. Then

$$\frac{P_{k+1}}{Q_{k+1}} = \frac{P_{k-2} + P_{k-1}\left(q_k + \frac{1}{q_{k+1}}\right)}{Q_{k-2} + Q_{k-1}\left(q_k + \frac{1}{q_{k+1}}\right)} = \frac{P_{k-2} + P_{k-1} q_k \ + P_{k-1}\frac{1}{q_{k+1}}}{Q_{k-2} + Q_{k-1} q_k \ + Q_{k-1}\frac{1}{q_{k+1}}} = \frac{\mathbf{P_k} + P_{k-1}\frac{1}{q_{k+1}}}{\mathbf{Q_k} + Q_{k-1}\frac{1}{q_{k+1}}} = \frac{P_{k-1} + P_k q_{k+1}}{Q_{k-1} + Q_k q_{k+1}} \tag{11}$$

5.6 Continued Fractions and GCD

A further theorem:

Theorem 2 (Cross Product of Continued Fraction Approximates) *Given approximates to a continued fraction* $\frac{P_0}{Q_0}, \frac{P_1}{Q_1}, \ldots \frac{P_i}{Q_i}$, *for any* i, $P_{i+1}Q_i - P_iQ_{i+1} = \pm 1$.

Theorem 3 (Continued Fractions for Integer Ratios) *If* a, b *are integers, they can be represented exactly by a finite continued fraction. In other words,* $\frac{a}{b} = \frac{P_k}{Q_k}$ *for some finite* k *approximates of some continued fraction.*

If P_k and Q_k in the above theorem are relatively prime, then the greatest common divisor of a and b can be written:

$$\gcd(a, b) = \frac{a}{P_k} = \frac{b}{Q_k} = d \tag{12}$$

By then combining the above with the Cross Product theorem, multiplying by d on both sides, we get

$$\begin{aligned} d \cdot (P_kQ_{k-1} - P_{k-1}Q_k) &= \pm 1 \cdot d & (13)\\ P_kdQ_{k-1} - P_{k-1}Q_kd &= \pm d & (14)\\ aQ_{k-1} - P_{k-1}b &= \pm d & (15) \end{aligned}$$

which is precisely Extended GCD.

5.7 An Application

Encryption of ZAP mail (Larry Roberts)

Choose a prime (carefully). If the prime is carefully chosen, then the quotient of any smaller integer divided by it will contain a repeating decimal pattern whose period is the prime divisor. Not all primes work. For example, $1/11 = 0.090909\ldots$. But consider the prime 97 and the integer 31.

$$\frac{31}{97} = 0.31958\mathbf{76288}65979381443298969072\ldots$$

This example does not repeat for 96 decimal places.

To break it...

In any message, there will be chunks that are blank, in other words, where the pad itself shows through[2]. Because of the nature of the repeating decimal pattern, we can extract the prime given any short sequence of digits in the expansion. For example, take the segment 76288, shown in bold above.

Create a continued fraction representation of the segment:

$$0.76288 = 1/1.310822\ldots = 1/1 + 0.310822\ldots = 1/1 + 1/3 + 0.217273\ldots = \ldots \tag{16}$$

Run this to its conclusion and the resulting continued fraction terminates with $[0; 1, 3, 4, 1, 1, 1, 1, 15, 2]$. Constructing the approximates to the fraction, we get $P_k = [0, 1, 3, 13, 16, 29, 45, 74, 1115, 2384]$ and $Q_k = [1, 1, 4, 17, 21, 38, 59, 97, 1514, 3125]$

Then any prime encountered in the values for Q_k is a candidate for that denominator prime.[3] Each can be tested as a decryption candidate.

[2] Manuel did not explain the exact mechanism of encoding the message with the pad.

[3] If the length in bits of the encryption key is known, this can further narrow the search.