CS 70 FALL 2007 — DISCUSSION #4

ASSANE GUEYE, LUQMAN HODGKINSON, AND VAHAB POURNAGHSHBAND

1. Administrivia

(1) Course Information
- Reminder: The first midterm will be held on **Wednesday October 3rd** 6pm-8pm in 10 Evans

(2) Discussion Information
- The third homework is graded and will be distributed in section.

2. Polynomials on the Reals

Briefly, recall the following polynomial basics.

Definition 1. A *polynomial of degree d* on the reals is a function $p(x) = a_0 + a_1x^1 + a_2x^2 + \ldots + a_dx^d$, where the input variable x and the $d+1$ constants $a_0, \ldots, a_d$ are all real numbers, and additionally $a_d \neq 0$. r is a *root* of polynomial $p(x)$ if $p(r) = 0$.

Theorem 2. *Over the reals:*

(1) A degree d polynomial has at most d roots.

(2) For any $(x_1, y_1), \ldots, (x_{d+1}, y_{d+1}) \in \mathbb{R}^2$ there exists a unique polynomial $p(x)$ of degree at most d such that $p(x_i) = y_i$, for each $1 \leq i \leq d+1$.

Exercise 1. Find (and prove) an upper-bound on the number of times two degree d polynomials can intersect. What if the polynomials' degrees differ?

3. Polynomial Interpolation on the Reals

Property 2 (see Theorem 2) says that any set of $d+1$ points $(x_1, y_1), \ldots, (x_{d+1}, y_{d+1}) \in \mathbb{R}^2$ can be interpolated by a polynomial of degree at most d. But how can we efficiently perform such an interpolation? In lecture we saw that the Lagrange interpolation method achieves this feat.

Method 3. The Lagrange interpolation procedure:

i. $q_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{d+1} (x - x_j)$ is a degree d polynomial satisfying $q_i(x_j) = 0$ for all $j \neq i$ and $q_i(x_i)$ is some non-zero constant;

ii. $\Delta_i(x) = \frac{q_i(x)}{q_i(x_i)}$ is a degree d polynomial equal to 1 at x_i and 0 on the x_j with $j \neq i$;

iii. $y_i\Delta_i(x)$ is a degree d polynomial equal to y_i at x_i and 0 on the x_j with $j \neq i$; and

iv. $p(x) = \sum_{i=1}^{d+1} y_i\Delta_i(x)$ is a polynomial of degree at most d that satisfies $p(x_i) = y_i$ for each $1 \leq i \leq d+1$ (i.e. witnessing Property 2 as desired).

Date: October 9, 2007.

Exercise 2. Use the Lagrange interpolation method to determine the polynomial of degree at most 3 that fits the points $(-1, 2), (0, 1), (1, 2), (2, 5)$. What is the (exact) degree of this polynomial?

4. From Reals to Fields (e.g. $\mathbb{F}_m$)

5. Secret Sharing

Recall from class the following application of Lagrange interpolation on $\mathbb{F}_m$. A GSI wishes to distribute secret $s \in \mathbb{Z}$ among n CS70 students $1, \ldots, n$ so that at least k of these students must get together in order to reconstruct s from each of their pieces of information.

Protocol 4. The secret sharing protocol:

i. The GSI and students agree on a prime $q > n, s$.
ii. The GSI picks (in secret) any $k-1$ degree polynomial $P(x)$ on $\mathbb{F}_q$ such that $P(0) = s$.
iii. The GSI distributes $P(i)$ to student i, for each $1 \leq i \leq n$.
iv. Any group of k students can get together and construct the (at most) $k-1$ degree Lagrange polynomial $L(x)$ that fits their respective $P(i)$ values.
v. Property 2 ensures that $L = P$ and so that $L(0) = P(0) = s$.

Exercise 3. Suppose (!) you're a CS70 student, and your GSI has distributed a secret s to 10 students including yourself and your neighbor. The GSI picked a polynomial $P(x)$ of degree 2 (and so s/he hopes that no fewer than $k = 3$ students could reconstruct s) modular $q = 11$. Suppose the two of you are told that $P(6) \equiv 7 \mod 11$ and that $P(7) \equiv 5 \mod 11$. What can you say about s?

Exercise 4. What if you make friends with another student who tells you that $P(8) \equiv 7 \mod 11$? If possible, determine s.

6. Error Correcting Codes - Erasure Errors or Known Error Positions

In lecture, we learned an error-correcting scheme that allows us to correct k errors by adding k more characters to the transmitted message. This ability is quite useful in communications where we know we lost characters (such as noise or random disconnections on a line where it is very clear what noise is, etc...).

Let us review how this process works. Reconstruct the following statements in the alphabet $A = 0$, $I = 1$, $N = 2$, $S = 3$, and $T = 4$, knowing that the message size is 3 (the number of acceptable erasures is 1). Note that you are solving for a quadratic polynomial modulo 5.

(1) S_II
(2) _TIS
(3) SI_I

Exercise 5. Figure out the dropped letter at each level.
Concatenate the 3 character words together. Do you see the message?

7. Error Correcting Codes - General Errors or Unknown Error Positions

In modern day communications, most data is transmitted in the digital domain. This change makes it hard to determine what is noise or an omission in the data, so the error correcting scheme must be improved. Specifically, the error correcting scheme must tell you where the errors occured and allow you to decode the original message. In fact, CDs and other storage devices contain a large amount of redundant data.

The main idea is to add k more characters to the message such that message may be decoded. To follow this procedure, we add another polynomial $E(i) = (x - e_1) * (x - e_2)...(x - e_k)$ whose zeros are at the positions of the errors and then solve $Q(x) = P(x) * E(x)$. Note that E will have k coefficients as well, and that is why we need to extend the size of the message sent to be $n + 2k$ characters long.

Exercise 6. Suppose you received 42045 through a noisy channel that changes a packet in every five packets. What's the length of the initial message? Figure out the initial message, knowing that you are working on GF(7).