CS 70 Discrete Mathematics and Probability Theory
Fall 2021 DIS 5B

1 Berlekamp-Welch Warm Up

Let $P(i)$, a polynomial applied to the input i, be the original encoded polynomial before sent, and let r_i be the received info for the input i which may or may not be corrupted.

(a) When does $r_i = P(i)$? When does r_i not equal $P(i)$?

(b) If you want to send a length-n message, what should the degree of $P(x)$ be? Why?

(c) If there are at most k erasure errors, how many packets should you send? If there are at most k general errors, how many packets should you send? (We will see the reason for this later.) Now we will only consider general errors.

(d) What do the roots of the error polynomial $E(x)$ represent? Does the receiver know the roots of $E(x)$? If there are at most k errors, what is the maximum degree of $E(x)$? Using the information about the degree of $P(x)$ and $E(x)$, what is the degree of $Q(x) = P(x)E(x)$?

(e) Why is the equation $Q(i) = P(i)E(i) = r_iE(i)$ always true? (Consider what happens when $P(i) = r_i$, and what happens when $P(i)$ does not equal r_i.)

(f) In the polynomials $Q(x)$ and $E(x)$, how many total unknown coefficients are there? (These are the variables you must solve for. Think about the degree of the polynomials.) When you receive packets, how many equations do you have? Do you have enough equations to solve for all of the unknowns? (Think about the answer to the earlier question - does it make sense now why we send as many packets as we do?)

(g) If you have $Q(x)$ and $E(x)$, how does one recover $P(x)$? If you know $P(x)$, how can you recover the original message?

Solution:

(a) $r_i = P(i)$ when the received packet is correct. r_i does not equal $P(i)$ the received packet is corrupted.

(b) P has degree $n-1$ since n points would determine a degree $n-1$ polynomial.

(c) We send $n+k$ packets when we have k erasures and $n+2k$ packets for k general errors.

(d) The roots of error polynomial $E(x)$ represent the locations of corrupted packets. The receiver does not know the roots of $E(x)$. $E(x)$ is a polynomial that the receiver needs to compute in order to obtain $P(x)$. If there are at most k errors, then the maximum degree of $E(x)$ is k. The maximum degree of Q is $(n-1)+(k) = n+k-1$ since the degree of P is $n-1$ and the degree of E is at most k.

(e) If there is no error at point i, $P(i) = r_i$ and then multiplying each side by $E(i)$ gives $P(i)E(i) = r_iE(i)$. If there is an error at point i, then $E(i) = 0$, which means $P(i)E(i) = r_iE(i) = 0$.

(f) The maximum degree of $Q(x)$ is $n+k-1$, so the number of unknowns is $n+k$. The maximum degree of $E(x)$ is k, which would mean there would be $k+1$ unknowns. However, we know that the coefficient of x^k is 1 in $E(x)$, so the number of unknowns is k.

The total number of unknowns is $(n+k)+(k) = n+2k$

There are $n+2k$ equations, which is enough to solve for $n+2k$ unknowns.

(g) We can compute $P(x)$ using the equation: $P(x) = Q(x)/E(x)$. To recover the message, we compute $P(i)$ for $1 \leq i \leq n$.

2 Berlekamp-Welch Algorithm with Fewer Errors

In class we derived how the Berlekamp-Welch algorithm can be used to correct k general errors, given $n+2k$ points transmitted. In real life, it is usually difficult to determine the number of errors that will occur. What if we have less than k errors? This is a follow up to the exercise posed in the notes.

Suppose Alice wants to send 1 message to Bob and wants to guard against 1 general error. She decides to encode the message with $P(x) = 4$ (on GF(7)) such that $P(0) = 4$ is the message she want to send. She then sends $P(0), P(1), P(2) = (4,4,4)$ to Bob.

(a) Suppose Bob receives the message $(4,5,4)$. Without performing Gaussian elimination explicitly, find $E(x)$ and $Q(x)$.

(b) Now, suppose there were no general errors and Bob receives the original message $(4,4,4)$. Show that the $Q(x), E(x)$ that you found in part (a) still satisfies $Q(i) = r_iE(i)$ for all $i = 0,1,2$.

(c) Verify that $E(x) = x$, $Q(x) = 4x$ is another possible set of polynomials that satisfies $Q(i) = r_iE(i)$ for all $i = 0,1,2$.

(d) Suppose you're actually trying to decode the received message $(4,4,4)$. Based on what you showed in the previous two parts, what will happen during row reduction when you try to solve for the unknowns?

(e) Prove that in general, no matter what the solution of $Q(x)$ and $E(x)$ are though, the recovered $P(x)$ will always be the same.

Solution:

(a) $E(x) = x-1$ and $Q(x) = P(x)E(x) = 4x-4$.

(b) This is true because there were no errors, so $P(i) = r_i$ for $i = 0,1,2$.

(c) Since $Q(x) = P(x)E(x)$ and $P(i) = r_i$ for $i = 0,1,2$, we must have $Q(i) = r_iE(i)$ for all $i = 0,1,2$.

(d) There are multiple solutions to the system of equations.

(e) Suppose we got two solutions $Q'(x),E'(x)$ and $Q(x),E(x)$. Since they are both solutions, by definition, we have $Q'(i) = r_iE'(i)$ and $Q(i) = r_iE(i)$ for $1 \leq i \leq n+2k$. Therefore, $Q'(i)E(i) = Q(i)E'(i) = r_iE(i)E'(i)$. However, $Q'(x)E(x) - Q(x)E'(x)$ is a degree $n+2k-1$ polynomial, which is 0 at $n+2k$ points. Thus, $Q'(x)E(x) = Q(x)E'(x)$ for all x, so we arrive at

$$\frac{Q'(x)}{E'(x)} = \frac{Q(x)}{E(x)}.$$

This proves that the final solution for $P(x)$ is the same.

3 Secret Sharing with Spies

An officer stored an important letter in her safe. In case she becomes unreachable in battle, she decides to share the password (which is a number) with her troops. However, everyone knows that there are 3 spies among the troops, but no one knows who they are except for the three spies themselves. The 3 spies can coordinate with each other and they will either lie and make people not able to open the safe, or will open the safe themselves if they can. Therefore, the officer would like a scheme to share the password that satisfies the following conditions:

- When M of them get together, they are guaranteed to be able to open the safe even if they have spies among them.
- The 3 spies must not be able to open the safe all by themselves.

Please help the officer to design a scheme to share her password. What is the scheme? What is the smallest M? Show your work and argue why your scheme works and any smaller M couldn't work. (The troops only have one chance to open the safe; if they fail the safe will self-destruct.)

Solution:

The key insight is to realize that both polynomial-based secret-sharing and polynomial-based error correction work on the basis of evaluating an underlying polynomial at many points and then trying to recover that polynomial. Hence they can be easily combined.

Suppose the password is s. The officer can construct a polynomial $P(x)$ such that $s = P(0)$ and share $(i,P(i))$ to the i-th person in her troops. Then the problem is: what should the degree of $P(x)$ be and what is the smallest M?

First, the degree of polynomial d should not be less than 3. It is because when $d < 3$, the 3 spies can decide the polynomial $P(x)$ uniquely. Thus, n will be at least 4 symbols.

Let's choose a polynomial $P(x)$ of degree 3 such that $s = P(0)$. We now view the 3 spies as 3 general errors. Then the smallest $M = 10$ since n is at least 4 symbols and we have $k = 3$ general errors, leading us to a "codeword" of $4 + 2 \cdot 3 = 10$ symbols (or people in our case). Even though the 3 spies are among the 10 people and try to lie on their numbers, the 10 people can still be able to correct the $k = 3$ general errors by the Berlekamp-Welch algorithm and find the correct $P(x)$.

Alternative solution:

Another valid approach is making $P(x)$ of degree $M - 1$ and adding 6 public points to deal with 3 general errors from the spies. In other words, in addition to their own point $(i, P(i))$, everyone also knows the values of 6 more points, $(t+1, P(t+1)), (t+2, P(t+2)), \ldots, (t+6, P(t+6))$, where t is the number of the troops. The spies have access to total of $3 + 6 = 9$ points so the degree $M - 1$ must be at least 9 to prevent the spies from opening the safe by themselves. Therefore, the minimum M is 10.