BCH Codes

Shan-Yuan Ho

April 24, 2010

In the last lecture, we showed that the encoding polynomial $p(x)$ for a single error correcting code can be any primitive polynomial, say of degree r. A primitive polynomial is a prime (non-factorable) polynomial such that every polynomial of degree at most $r-1$ is the remainder of some monomial, x^j, upon dividing by $p(x)$. With such a code and encoded words of length 2^r-1, all code words have no remainder when divided by $p(k)$, so that the remainder of a code word altered by one error, is the remainder of the error alone. The monomial corresponding to that error can be found from a table of the remainders of monomials $\pmod{p(x)}$.

Multiple error correcting polynomial codes were invented by mathematicians Bose, Ray-Chaudhuri, and Hocquenghem in the 1950's. These codes are called BCH codes in their honor. Although BCH codes can be defined over any field, we will again, for simplicity, restrict to the binary field and study binary BCH codes.

1 The BCH Code

Denote messages, generators (encoding polynomials), codewords, and received messages by $m(x), p(x), c(x)$, respectively. These are represented as sequences corresponding to the coefficients of a polynomial, where we take the convention of writing the coefficients from lowest to highest degree.

Recall from the last section that polynomial codes are obtained by multiplying message polynomials by encoding polynomials. Thus,

$$c(x) = m(x)p(x) = \sum_{i=1}^{n} c_i x^i \tag{1}$$

which is also respresented by $c = (c_1, c_2, \ldots, c_n)$ for $c_i \in GF(2)$.

A binary BCH code is defined as follows. Let $p(x)$ be a primitive polynomial of degree r with coefficients in the binary field. If $c(x)$ is a non-zero polynomial such that $c(x) = c(x^3) = c(x^5) = \cdots = c(x^{2t-1}) = 0 \pmod{p(x)}$ for all t such that $1 \leq (2t-1) \leq 2^r - 1$, then $c(x)$ is a t-error correcting code of length $n = 2^r - 1$.

A received sequence can have at most t errors to guarantee correct decoding. Let e_i, $1 \leq e_i \leq n$ and $1 \leq i \leq t$, represent the location of the i-th bit in error. Then the error monomials are x^{e_i} and the error polynomial $e(x) = \sum_{i=1}^{k} x^{e_i}$ is their sum. The received polynomial is then

$$r(x) = c(x) + e(x) \tag{2}$$

Suppose we have a t-error correcting BCH code. Then the remainder of the received ploynomial $r(x) \pmod{p(x)}$ is equal to the sum of the remainders of the error monomials. Furthermore, if we evaluate the received polynomial at a higher power of x and then divide by $p(x)$, then this is equal to the error polynomial evaluated at the higher power and taking its remainder. That is, Rem$[r(x^j)]$=Rem$[e(x^j)]$ for $1 \leq j \leq 2t-1$. Recall that if $p(x)$ is primitive, there is a bijective map from $x^{e_i} \rightarrow \text{Rem}[x^{e_i}]$, for all e_i, $1 \leq e_i \leq n$. Thus, if the error monomial or the remainder value after dividing the error monomial x^{e_i} by $p(x)$ is known, then the position of the error is known. For multiple errors, we only have the remainder of the sums of error monomials. To determine each bit position in error, we need to extract each monomial from this information.

In the following section, we show (i) how to find the encoding polynomial and (ii) how to determine the bit positions in error from the remainders of the received polynomial evaluated at higher powers.

2 The Generator Polynomial

For a t-error correcting code, the generator polynomial $Q(x)$ of standard BCH codes has the form

$$Q(x) = p(x)p_3(x)p_5(x)\cdots p_{2t-1}(x) \tag{3}$$

where $p(x)$ is a primitve polynomial and all the polynomials $p_3(x^3)$, $p_5(x^5)$, $\cdots$, $p_{2t-1}(x^{(2t-1)})$ must be divisible by $p(x)$, i.e.,

$$p_3(x^3) = p_5(x^5)\cdots p_{2t-1}(x^{2t-1}) = 0 \pmod{p(x)} \tag{4}$$

We have shown that this form is sufficient and demonstrated how to find it in the previous lecture notes on polynomial codes. We show it here again with an example. For primitive $p(x) = 1 + x + x^4$, to find $p_3(x)$, note that $(x^6 + x^9 + x^{12}) = 1 + x^3 \pmod{p(x)}$. Then, $(1 + x^3 + x^6 + x^9 + x^{12}) = 0 \pmod{p(x)}$. Let $y = x^3$ and let $p_3(y) = 1 + y + y^2 + y^3 + y^4$. It follows that $p_3(x^3) = 0 \pmod{p(x)}$.

Thus, $Q(x) = p(x)p_3(x) = (1+x+x^4)(1+x+x^2+x^3+x^4)$ is a generator polynomial for a 2-error correcting code. Using a similar procedure, we find $p_5(x) = 1+x+x^2$. Therefore, $Q(x) = p(x)p_3(x)p_5(x) = (1 + x + x^4)(1 + x + x^2 + x^3 + x^4)(1 + x + x^2)$ is the generator polynomial for a 3-error correcting BCH code.

3 The Error Locator Polynomial and the Elementary Symmetric Functions

Define $t_i = \sum_{j=1}^{t} x^{ie_j}$ to be the i-th power sums of the error monomials. Since $t_2 = t_1^2$ in the binary field, the even t_i provide no new information. The t_i's are called the *power sum symmetric functions*, and we will use these to find the bit error locations.

Define the *error locator polynomial* $E(y)$ of degree t such that it is equal to zero when evaluated at the error monomials and nonzero otherwise.

$$\begin{aligned} E(y) &= (y - x^{e_1})(y - x^{e_2}) \cdots (y - x^{e_k}) \\ &= y^k + s_1 y^{k-1} + s_2 y^{k-2} + \cdots s_{k-1} y + s_k \\ &= 0 \end{aligned} \tag{5}$$

Then, from the fundamental theorem of algebra, for a t-error correcting code, all the roots of the error locator polynomial are the error monomials, x^{e_j}, $1 \leq e_i \leq n$ and $1 \leq i \leq t$. We need to find a relationship between the coefficients, s_j's of the error locator polynomial and the odd power sum symmetric functions, t_j's.

The *k-th elementary symmetric function of d elements* is defined as the sum of the products of k different elements from among the d elements, combined in all possible ways. For example, the 2nd elementary function of a, b, c, d is $ab+ac+ad+bc+bd+cd$.

There is a linear relationship between the elementary symmetric functions and the odd power sums, t_i, of the error monomials. The coefficients, s_j, $1 \leq j \leq k$ of the error locator polynomial are related to the t_i's in the following for k odd and $s_0 = t_0 = 1$,

$$\sum_{i=1}^{k} s_i t_{k-1} = s_k + s_{k-1} t_1 + \cdots + s_1 t_{k-1} + t_k = 0 \tag{6}$$

To see this for a k-error correcting code, note that the error locator polynomial in equation 5 evaluated at each error monomial $y = x^{e_i}$ must be satisfied. That is, for each j, $1 \leq j \leq t$

$$x^{te_j} + s_1 e^{(t-1)e_j} + \cdots + s_{t-1} x^{e_j} + s_t = 0 \tag{7}$$

Sum equation 5 over all the error monomials. Then, t is the coefficient of s_t and t_{t-j} is the coefficient of s_j, except that the coefficient of s_t is t. For any $t' > t$, multiply equation 6 by $y^{t'-t}$, and follow the same proof – evaluate at each error monomial and sum. The relationship between s_i and t_i in equation 6 follows.

For the case of a two error correcting code over a binary field, the standard encoding polynomial is $Q(x) = p(x)p_3(x)$. The error locator polynomial will have degree 2, and its coefficients determined from equation 6 are given by $s_1 = t_1$ and $s_2 = t_1^2 + \frac{t_3}{t_1}$.

4 Example: 3 Error Correcting BCH Code

The encoding polynomial is of the form $Q(x) = p(x)p_3(x)p_5(x)$. Suppose we use $p(x) = 1 + x + x^4$, then $Q(x) = (1 + x + x^4)(1 + x + x^2 + x^3 + x^4)(1 + x + x^2)$. Suppose the received sequence is 101000110110010, then

$$\begin{aligned} r(x) &= 1 + x^2 + x^6 + x^7 + x^9 + x^{10} + x^{13} \\ r(x^3) &= 1 + x^6 + x^{18} + x^{21} + x^{27} + x^{30} + x^{39} \\ &= 1 + x^6 + x^3 + x^6 + x^{12} + x^0 + x^9 \\ &= x^{13} \\ r(x^5) &= 1 + x^2 + x^6 + x^7 + x^9 + x^{10} + x^{13} \\ &= 1 + x^{10} + x^{30} + x^{35} + x^{45} + x^{50} + x^{65} \\ &= 0 \end{aligned} \tag{8}$$

We divide by $p(x) = 1+x+x^4$ for all 3 received sequences above. Note that $r(x^5) = 0$ means there are only 2 errors in our received sequence. Since $r(x^3) = x^{13}$, it is already a monomial. We only need to add the remainders of the monomials of $r(x)$. 1000+0010+0011+1101+0101+1110+1011 = 0100 means $r(x) = x \pmod{p(x)}$.

$$\text{Rem}[r(x)] = t_1 = x \tag{9}$$

$$\text{Rem}[r(x^3)] = t_3 = x^{13} \tag{10}$$

$$\text{Rem}[r(x^5)] = t_5 = 0$$

From the $s - t$ relations in equation 6, we have

$$s_1 + t_1 = 0 \tag{11}$$

$$s_3 + s_2t_1 + s_1t_2 + t_3 = 0 \tag{12}$$

$$s_3t_2 + s_2t_3 + s_1t_4 + t_5 = 0 \tag{13}$$

Substituting equations 9 into equations 11 and solving, we get we get $s_1 = x$, $s_2 = x^7$, and $s_3 = 0$.

From the elementary symmetric functions,

$$s_1 = x^{e_1} + x^{e_2} + x^{e_3} = x \tag{14}$$

$$s_2 = x^{e_1}x^{e_2} + x^{e_2}x^{e_3} + x^{e_3}x^{e_1} = x^{e_1+e_2} + x^{e_2+e_3} + x^{e_3+e_1} = x^7 \tag{15}$$

$$s_3 = x^{e_1}x^{e_2}x^{e_3} = x^{e_1+e_2+e_3} = 0 \tag{16}$$

the error locator polynomial becomes

$$
\begin{aligned}
E(y) &= (y - x^{e_1})(y - x^{e_2})(y - x^{e_3}) && (17)\\
&= y^3 + s_1 y^2 + s_2 y + s_3 && (18)\\
&= y^3 + xy^2 + x^7 y = y^2 + xy + x^7 = 0
\end{aligned}
$$

We now substitute in all the monomials $y = x^j$, $0 \leq j \leq 2^r$ to see which monomials are solutions to the error locator polynomial. We find that x^2 and x^5 are roots. The correct sequence is 100001110110010 and the message is 11010.