Polynomials in GF(2) and CRCs

This chapter covers arithmetic with polynomials that have coefficients from GF(2). These operations are the basis of many useful telecommunications functions including computation of cyclic redundancy checks (CRCs).

After this chapter you should be able to: represent a sequence of bits as a polynomial with coefficients from GF(2), compute the result of multiplying a polynomial by x^n, compute the result of dividing two polynomials, compute the value of a CRC given the message and generator polynomials, and determine if a CRC computation indicates an error has occurred. You should be able to determine if a CRC is guaranteed to detect a particular error sequence.

GF(2)

A Galois field, denoted as $GF(q)$, is a set of integers and two operations that have certain properties. One of the properties is closure – the result of any operation on two elements of the field is also in the field.

For example, $GF(2)$ includes two integers (0 and 1) and the addition and multiplication operations are defined as addition and multiplication with the result taken modulo-2.

Exercise 1: Write the addition and multiplication tables for $GF(2)$. What logic function can be used to implement modulo-2 addition? Modulo-2 multiplication?

Exercise 2: What are the possible values of the results if we used values 0 and 1 but the regular definitions of addition and multiplication? Would this be a field?

Representing Codewords as Polynomials

Thus far we've represented codewords as sequences of bits. We can also represent codewords as polynomials with coefficients from $GF(2)$. For example, the polynomial:

$$1x^3 + 0x^2 + 1x^1 + 1x^0 = x^3 + x^1 + 1$$

can be used to represent the codeword 1011.

Exercise 3: What is the polynomial representation of the codeword 01101?

Polynomials are used to describe codes because many properties of codes can be derived from the mathematical properties of polynomials.

Note that it is the coefficients of the polynomial that are important. The polynomial itself is never evaluated and the variable x that appears in these polynomials is just a dummy variable. These polynomials can thus also be viewed as binary numbers or bit strings where the order of each term indicates the bit position.

Polynomial Arithmetic

We can add, subtract, multiply and divide polynomials with coefficients in GF(2). These operations are the basis for many useful communication-related functions including convolutional codes for FEC (Forward Error Correction), CRCs (Cyclic Redundancy Checks), and PRBS (Pseudo-Random Bit Sequence) generators.

Exercise 4: What is the result of multiplying $x^2 + 1$ by $x^3 + x$ if the coefficients are regular integers? If the coefficients are values in $GF(2)$? Which result can be represented as a bit sequence?

Digital Implementation of Polynomial Arithmetic

Arithmetic on polynomials with $GF(2)$ coefficients can be implemented with simple digital logic circuits. Flip-flops, organized as shift registers, store the bits of the message (coefficients equal to 1 or 0) and XOR and AND gates are used to compute modulo-2 addition and multiplication. The bits corresponding to codeword(s)/message(s) can be input and output sequentially, bit by bit, into the polynomial arithmetic circuits.

It's much simpler to do arithmetic using polynomials in GF(2) than using regular integers because we do not need to compute carries when computing results.

Cyclic Redundancy Checks

A Cyclic Redundancy Check (CRC) is a code used to detect errors in a sequence of k data bits. A "codeword" of n bits is transmitted for each k data bits. The length of the CRC is thus $n - k$:
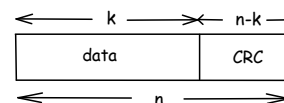
The algorithm used to compute the CRC is as follows:

The data to be transmitted, treated as a polynomial, is multiplied by the polynomial x^{n-k}. This increases the order of each term by $n - k$ (or equivalently, appends $n - k$ zero bits). This new polynomial, $M(x)$, is divided by a *generator polynomial*, $G(x)$[1]. The result is a quotient and a remainder:

$$\frac{M(x)}{G(x)} = Q(x) \text{ remainder } R(x)$$

We then replace the last $n - k$ bits of $M(x)$ (which were zero due to us having multiplied by x^{n-k}) with $R(x)$. This is equivalent to adding (or subtracting since polynomial addition and subtraction are the same for coefficients in GF(2)) $R(x)$ from $M(x)$. This ensures that the new polynomial will be divisible by $G(x)$.

Note that $n - k$ is one less than the number of terms in $G(x)$ since the remainder is always less then the divisor. If we number the terms by the order of x, then the highest order term in the remainder will be x^{n-k}.

The receiver carries out the same polynomial division operation on the combination of the message bits and CRC. If the remainder is not zero then the received polynomial is not a multiple of $G(x)$ and so at least one of the bits must have changed and an error has been detected.

Detecting Added/Deleted Zero Bits

We can add or remove any number of leading zeros coefficients to $M(x)$ without affecting its value or the CRC. To allow the CRC to detect missing/added leading zero bits, some implementations require that some initial data bits (typically the first $n - k$) be complemented before computing the CRC.

Similarly, appending or deleting zeros to the end of the message will also result in a zero remainder. We can avoid this problem by complementing the CRC before sending it. This generates a non-zero remainder but the value will be a specific value (the same for all messages) if there are no errors.

Another way to detect missing/added leading/trailing zero bits is to include the length of the message in the CRC computation.

Computing the CRC

Computing the CRC requires polynomial division. The process involves repeated subtraction of the generator polynomial from the message polynomial.

[1] Generator polynomials "generate" other codewords, in this case the CRC.

Unlike regular division, to compute the CRC we only need to compute the remainder.

Exercise 5: If the generator polynomial is $G(x) = x^3 + x + 1$ and the data to be protected is 1001, what are $n - k, M(x)$ and the CRC? Check your result. Invert the last bit of the CRC and compute the remainder again.

Checking the CRC

At the receiver the same circuit can be used to divide the received message and the appended remainder polynomial by the generator polynomial. If the remainder is zero then the received polynomial must be a multiple of the generator polynomial. This is always the case when we subtract the remainder $R(x)$ from the message polynomial. Therefore if the remainder in the SR is non-zero then there must have been an error.

CRC Error Detection Performance

CRC error detection will fail only if the error pattern is a multiple of $G(x)$.

If all the errors are located within an "error burst" of length $n - k$ then the error pattern cannot be a multiple of $G(x)$ and is guaranteed to be detected. However, the CRC will also detect most longer bursts since they are unlikely to be a multiple of $G(x)$.

Exercise 6: Is a 32-bit CRC guaranteed to detect 30 consecutive errors? How about 30 errors evenly distributed within the message?

A common situation is where the received bits are completely random (e.g. noise being detected as data). In this case the probability of not detecting an error is the probability that a random sequence of $n - k$ bits matches the required checksum.

Exercise 7: What is the probability that a CRC of length $n - k$ bits will be the correct CRC for a randomly-chosen codeword? Assuming random data, what is the undetected error probability for a 16-bit CRC? For a 32-bit CRC?

Standard CRC Generator Polynomials

There are several CRC generator polynomials in common use. The most common lengths are 16 and 32 bits since these are multiples of 8 bits. All(?) IEEE 802 standards use the same 32-bit CRC polynomial typically called "CRC-32". The ITU has defined a 16-bit CRC generator polynomial ("CRC-16-CCITT") that is also used in various standards.