Error Detection and Correction

ICNE532C - Computer Networks

S.Venkatesan

Indian Institute of Information Technology, Allahabad

September 27, 2016

Sections

- Types of errors, the concept of redundancy, and distinguishes between error detection and correction.
- Block Coding - Error Detection and Hamming Distance
- Cyclic Codes
- Checksums
- Forward Error Correction

Types of Errors

- Single Bit error - Only one bit of a given data unit is changed.
- Burst error - 2 or more bits in the data unit is changed.

Redundancy

- The central concept in detecting or correcting errors is redundancy.
- We need to send some extra bits with our data.
- These redundant bits are added by the sender and removed by the receiver.
- This allows the receiver to detect or correct corrupted bits.

Detection vs Correction

- The correction of errors is more difficult than the detection.
- In error correction, we need to know the exact number of bits that are corrupted and, more importantly, their location in the message.
- If we need to correct a single error in an 8-bit data unit, we need to consider eight possible error locations; if we need to correct two errors in a data unit of the same size, we need to consider 28 possibilities.

Coding

- **Block Coding**
- Convolution Coding

Block Coding

- Message will be divided in to blocks, each of k bits, called **datawords**
- We add r redundant bits to each block to make the length $n = k + r$
- The resulting n bits are called as **codewords**
- With k bits we can create 2^k datawords; with n bits we can create 2^n codewords.
- Since $n > k$, the number of possible codewords is larger than the number of possible datawords.
- Block coding process is one-to-one; the same dataword is always encoded as the same codeword. This means that we have $2^n - 2^k$ codewords that are not used.

Error Detection

- The reciever has (or can find) a list of valid codewords.
- The original codeword has changed to an invalid one.

Dataword	**Codeword**
00	000
01	011
10	101
11	110

Hamming Distance

- It is the number of bit differences in two words.
- Hamming distance is the number of corrupted bits during transmission.
- Eg. If the codeword 00000 is sent and 01101 is received, 3 bits are in error and the Hamming distance between the two is **d**(00000,01101) = 3
- If the Hamming distance between the sent and the received codeword is not zero, the codeword has been corrupted during transmission.

Minimum Hamming Distance for Error Detection

- It is the smallest Hamming distance between all possible pairs of codewords.
- If s errors occur during transmission, the Hamming distance between the sent codeword and received codeword is s.
- If our system is to detect up to s errors, the minimum distance between the valid codes must be $(s + 1)$, so that the received codeword doest not match a valid codeword.
- In other words, if the minimum distance between all valid codewords is $(s + 1)$, the received codeword cannot be erroneously mistaken for another codeword. Error will be detected.

Linear Block Codes

- It is a code in which the XOR (addition modulo-2) of two valid codewords creates another valid codeword.
- *Minimum distance for Liner Block Codes*: The minimum Hamming distance is the number of 1s in the nonzero valid codeword with the smallest number of 1s.

Parity-Check Code

- It is the most familiar error-detecting and linear block code.
- In this code, a k-bit dataword is changed to an n-bit codeword where $n = k + 1$. The extra bit is called as parity bit.
- *Minimum distance for Liner Block Codes*: The minimum Hamming distance is the number of 1s in the nonzero valid codeword with the smallest number of 1s.
- The minimum Hamming distance for this category is $d_{min} = 2$, which means that the code is a single-bit error-detecting code.

Dataword	**Codeword**
0000	0000**0**
0001	0001**1**
0111	0111**1**

Parity-Check Code - Techniques

- Parity bit $r_0 = a_3 + a_2 + a_1 + a_0 (modulo - 2)$
- If the number of 1s is even, the result is 0; if the number of 1s is odd, the result is 1. In both cases, the total number of 1s in the codeword is even.
- Syndrome bit $s_0 = a_3 + a_2 + a_1 + a_0 + q_0 (modulo - 2)$
- The syndrome is 0 when the number of 1s in the received codeword is even; otherwise, it is 1.
- The syndrome is passed to the decision logic analyzer. If the syndrome is 0, there is no detectable error in the received codeword; the data portion of the received codeword is accepted; if the syndrome is 1, the data portion of the received codeword is discarded.
- Two dimensional parity code.

Cyclic Codes

- It is the special linear block code with one extra property.
- If a codeword is cyclically shifted (rotated), the result is another codeword.
- For example, if 1011000 is a codeword and we cyclically left-shift, then 0110001 is also a codeword.
- *Cyclic redundancy check (CRC)*, which is used in networks such as LANs and WANs

Dataword	**Codeword**
0000	0000**000**
0001	0001**001**
0111	0111**010**

CRC encoder and decoder

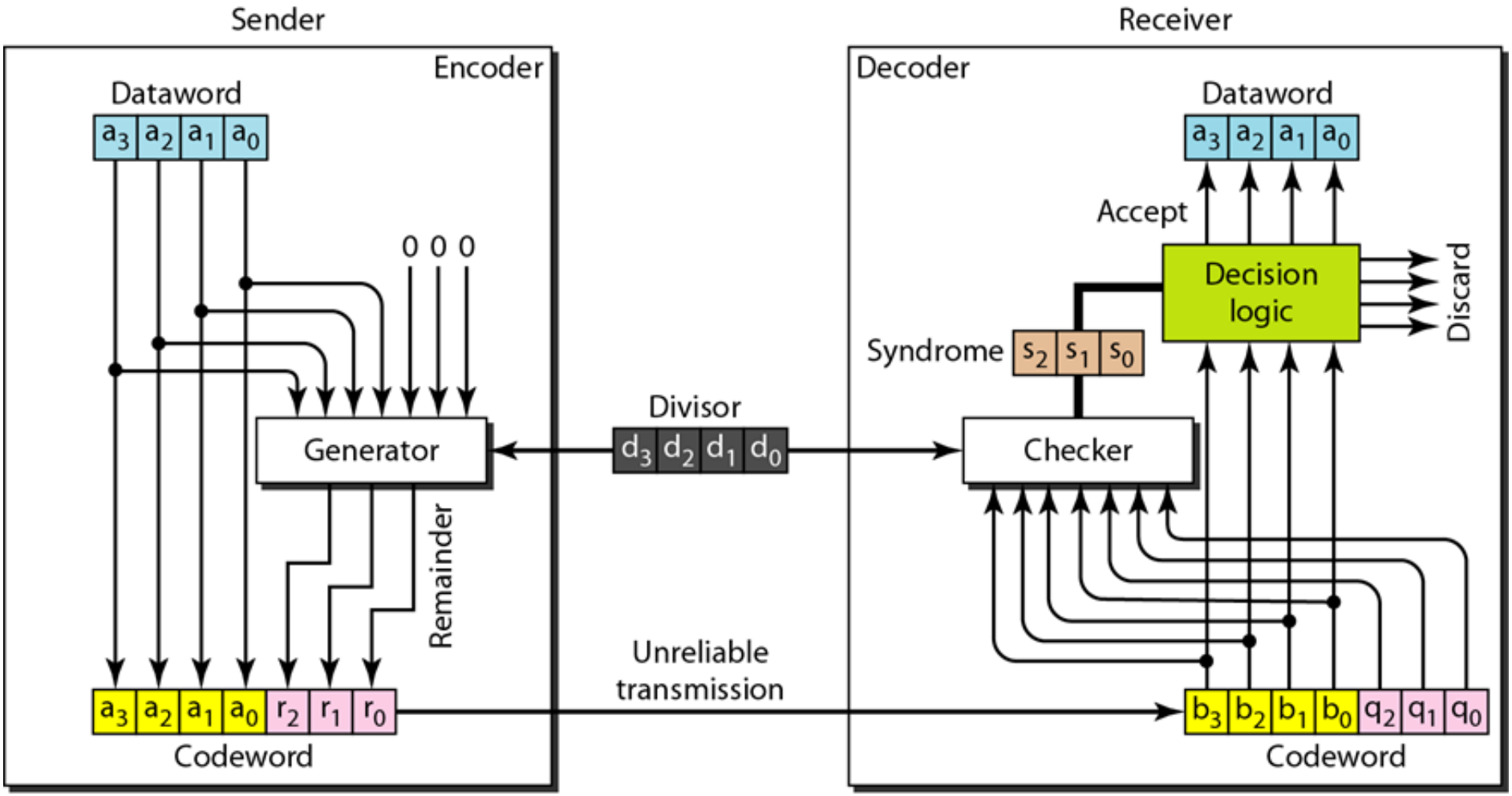
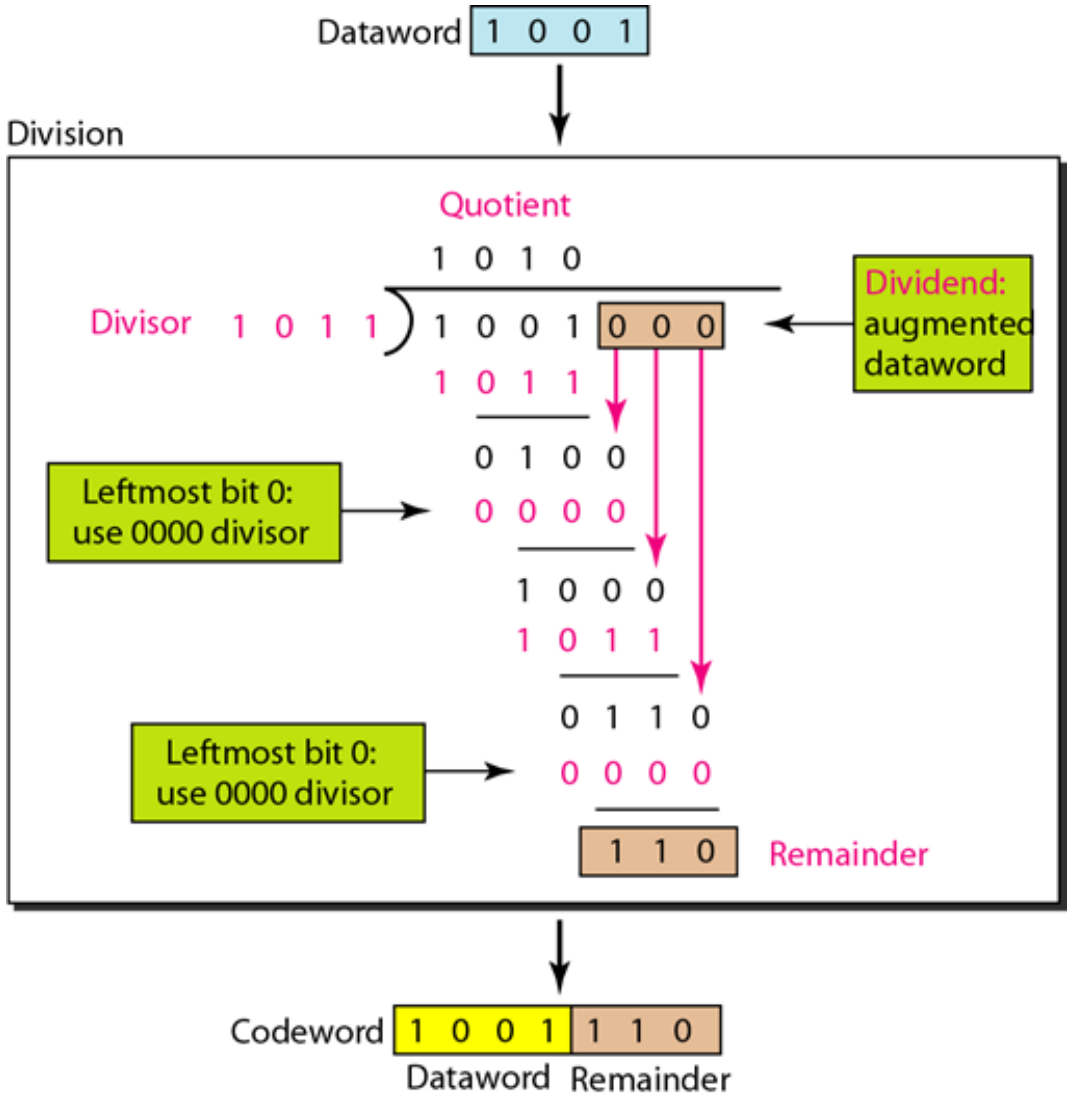
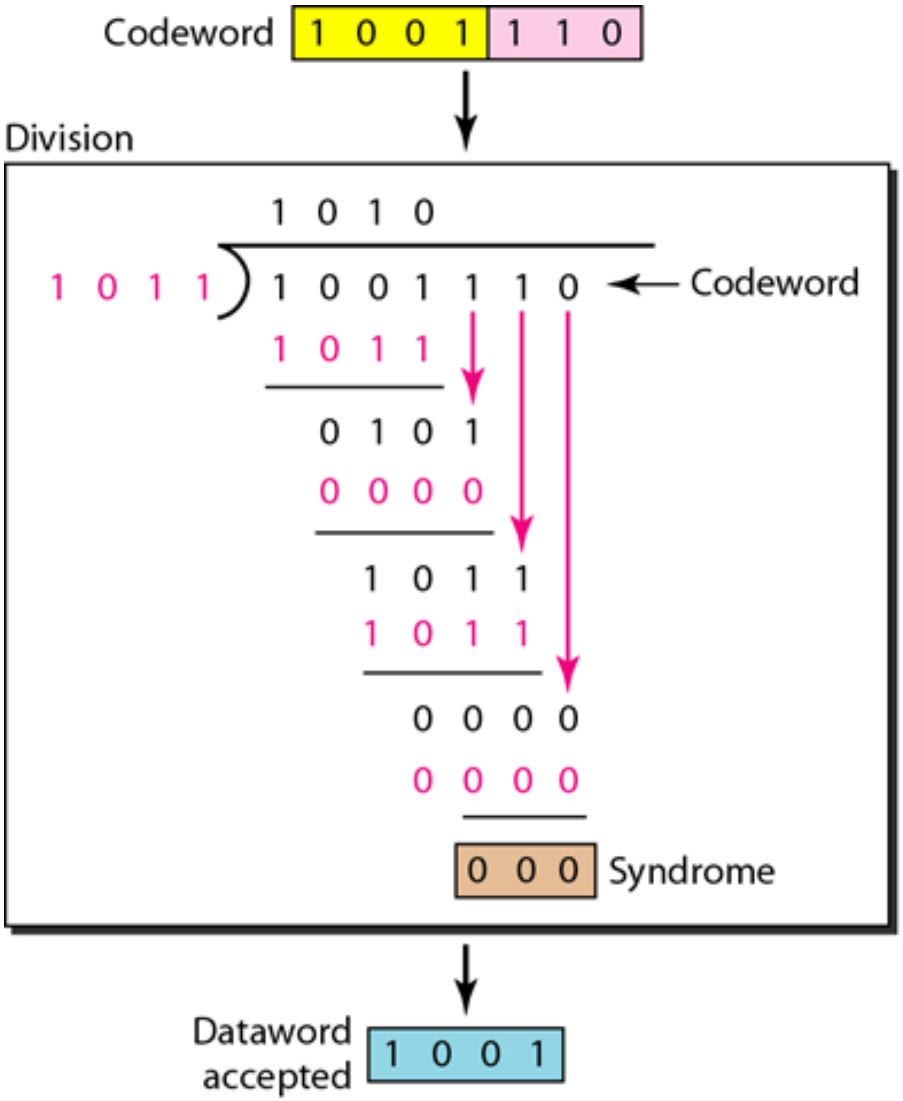


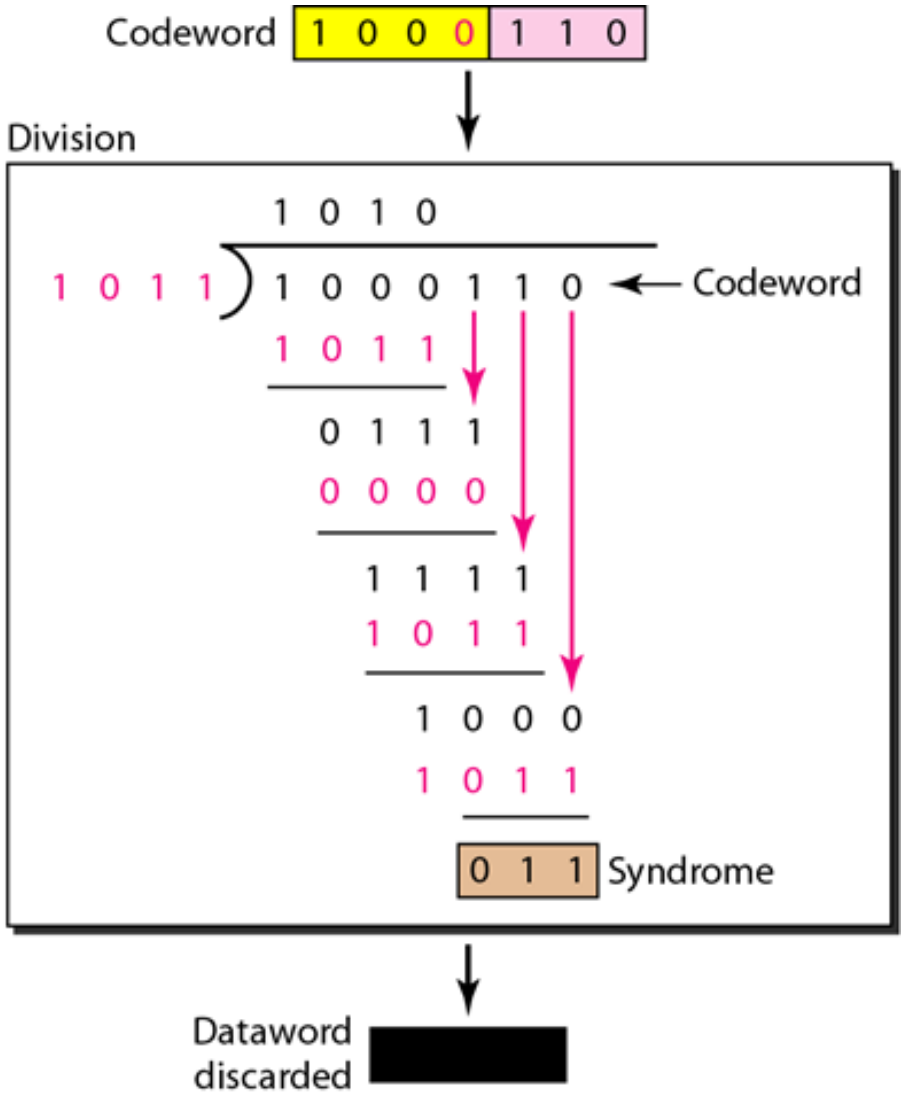
Figure: CRC encoder and decoder

Division in CRC encoder



Division in CRC decoder for two cases





Polynomials

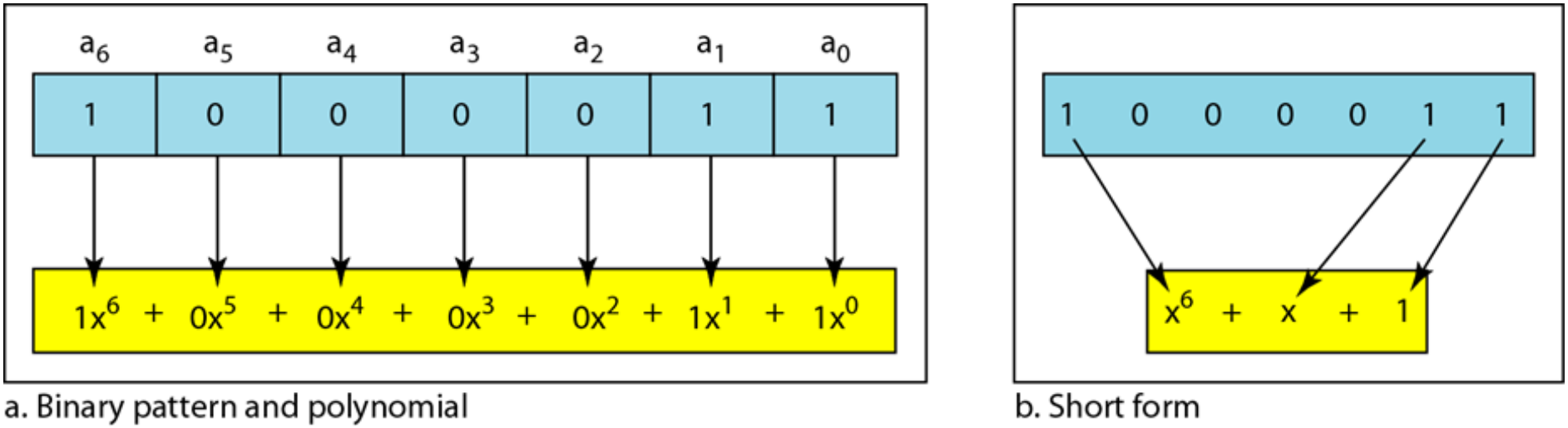
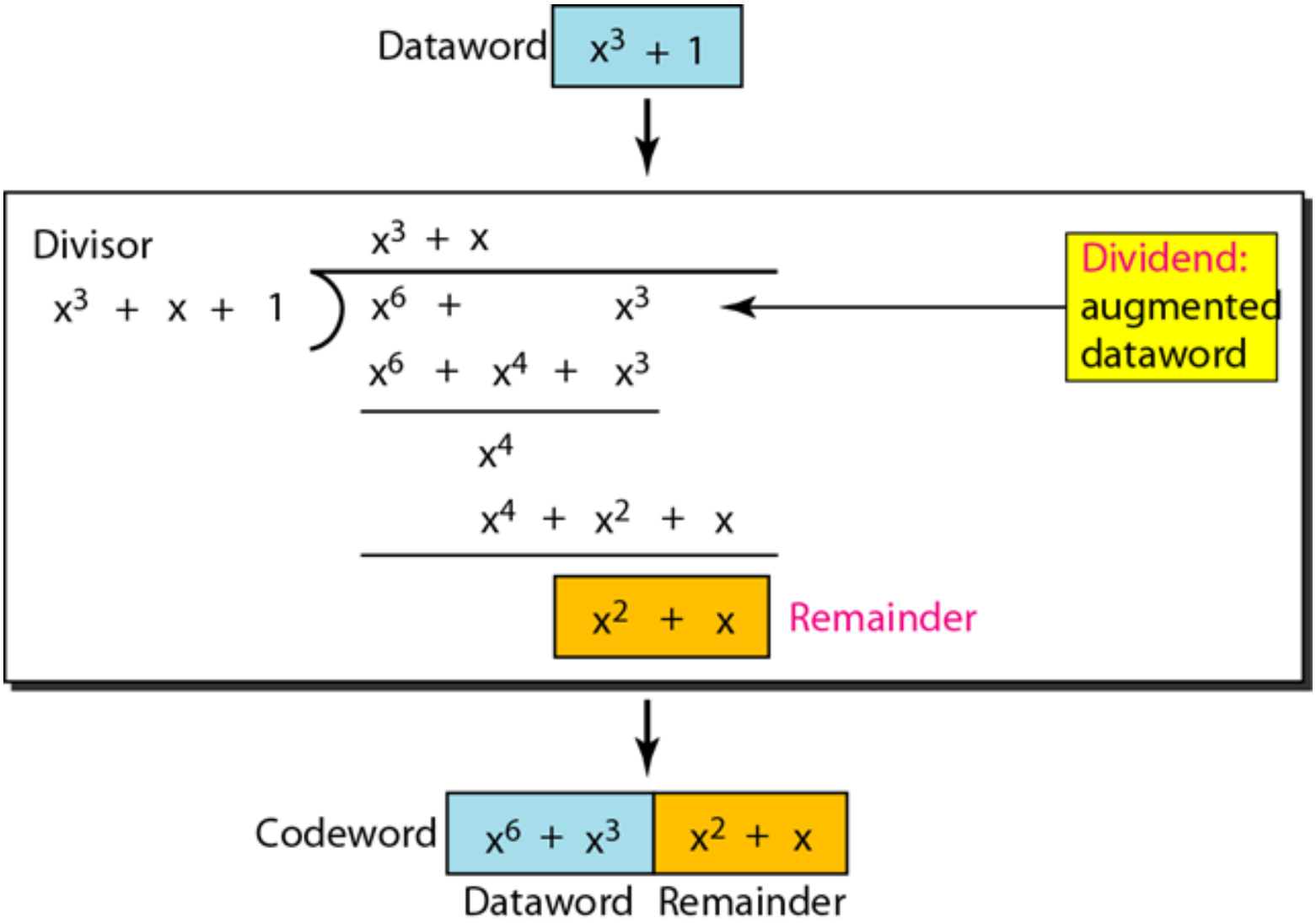


Figure: Polynomial representation

Polynomial Computation

- Adding and Subtracting Polynomials.
- Multiplying or Dividing Terms.
- Multiplying Two polynomials

Cyclic Code Encoder Using Polynomials



Checksum

- It can be applied to a message of any length
- It is used mostly at the network and transport layer rather than the data-link layer.
- Types
  - Traditional Checksum
  - One's Complement addition
  - Other approaches - Fletcher checksum and Adler checksum

Forward Error Correction

- Using Hamming Distance
- Using XOR
- Chunk Interleaving
- Combining Hamming Distance and Interleaving
- Compounding High- and Low-Resolution Packets

References

[1] Behrouz A. Forouzan *Data Communications and Networking*. McGraw-Hill fifth edition.