Linear Block Codes (2)
Cyclic Codes

Pierangelo Migliorati

DII - University of Brescia

Cyclic codes

Cyclic codes are a subset of the class of linear codes that satisfy the following cyclic shift property: if $\mathbf{C} = [c_{n-1}c_{n-2} \ldots c_1 c_0]$ is a code word of a cyclic code then $[c_{n-2}c_{n-3} \ldots c_0 c_{n-1}]$, obtained by a cyclic shift of the elements of $\mathbf{C}$, is also a code word. That is, all cyclic shifts of $\mathbf{C}$ are code words. As a consequence of the cyclic property, the codes possess a considerable amount of structure which can be exploited in the encoding and decoding operations. A number of efficient encoding and hard-decision decoding algorithms have been devised for cyclic codes that make it possible to implement long block codes with a large number of code words in practical communications systems. A description of specific algorithms is beyond the scope of this book. Our primary objective is to briefly describe a number of characteristics of cyclic codes.

Example

The code with the generator matrix

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

has codewords

$c_1 = 1011100$ $\quad c_2 = 0101110$ $\quad c_3 = 0010111$

$c_1 + c_2 = 1110010$ $\quad c_1 + c_3 = 1001011$ $\quad c_2 + c_3 = 0111001$

$c_1 + c_2 + c_3 = 1100101$

and it is cyclic because the right shifts have the following impacts

$c_1 \to c_2,$ $\quad c_2 \to c_3,$ $\quad c_3 \to c_1 + c_3$

$c_1 + c_2 \to c_2 + c_3,$ $\quad c_1 + c_3 \to c_1 + c_2 + c_3,$ $\quad c_2 + c_3 \to c_1$

$c_1 + c_2 + c_3 \to c_1 + c_2$

parole di lunghezza N → polinomi di grado N-1

$$\bar{a} = (a_1 a_2 \dots a_N) \Rightarrow a(D) = a_1 D^{N-1} + a_2 D^{N-2} + \dots + a_{N-1} D + a_N$$

Le operazioni sui polinomi si intendono in aritmetica modulo 2 ($D^i + D^i = 0, \dots$) e i polinomi possono essere associati a operazioni con registri a scorrimento.

An efficient representation of the code-words is possible using polynomials in GF2 (in D, z, x, …).

Word of length N -> polynomial of degree N-1.

The operations on this polynomial can be effectively implemented using shift registers.

Cyclic shift

N.B. $(a_1\, a_2 \ldots a_N) \Rightarrow a(D)$

SCORRIMENTO CICLICO $\xleftarrow{j}$

$(a_{j+1}, a_{j+2}, \ldots a_N, a_1 \ldots a_j) \Rightarrow a^{(j)}(D)$

$\Longrightarrow D^j a(D) \bmod (D^N + 1)$

$\longrightarrow \quad D^j a(D) = q(D)\,(D^N+1) + a^{(j)}(D)$

quoziente

$a^{(j)}(D)$ è il RESTO della DIVISIONE di $D^j(a(D))$ con (D^N+1)

ES: $\tilde{X} = 0111010$ CODICE H $\overset{(N,K)}{(7,4)}$

$$x(D) = D^5 + D^4 + D^3 + D$$

$$x^{(4)}(D) \rightarrow D^4\, x(D) \bmod (D^7+1)$$

$$
\begin{array}{l|l|l}
D^7+1 & D^9 + D^8 + D^7 + D^5 & D^2 + D + 1 \\
\hline
 & D^9 + \qquad\qquad + D^2 & \downarrow \\
 & \overline{\quad D^8 + D^7 + D^5 + D^2} & \text{QUOZIENTE} \\
 & \quad D^8 + \text{----} - \quad + D & \\
 & \overline{\qquad D^7 + D^5 + D^2 + D} & \\
 & \qquad D^7 + \text{------} \quad + 1 & \\
 & \overline{\qquad\quad D^5 + D^2 + D + 1} & \text{RESTO}
\end{array}
$$

$$
\begin{array}{c|l|l}
D^7+1 & D^9+D^8+D^7+D^5 & D^2+D+1 \\
\hline
 & D^9+ \qquad\qquad +D^2 & \text{QUOZIENTE} \\
\cline{2-2}
 & \quad D^8+D^7+D^5+D^2 & \\
 & \quad D^8+ \cdots \qquad +D & \\
\cline{2-2}
 & \qquad D^7+D^5+D^2+D & \\
 & \qquad D^7+ \cdots \qquad +1 & \\
\cline{2-2}
 & \qquad\quad D^5+D^2+D+1 & \text{RESTO}
\end{array}
$$

Il resto $D^5+D^2+D+1 \rightarrow 0\,1\,00\,111$ si ottiene dalla seq. originale con 4 traslazioni verso sinistra ($011\,1,010$ ← 4 VOLTE)

POLINOMIO GENERATORE

→ DATO un codice ciclico (N, K), ESISTE un unico POLINOMIO di CODICE di grado (N-K) che assume la forma

$$g(D) = D^{N-K} + \dots + 1$$

* TUTTI gli altri polinomi di codice sono multipli di g(D), ed ogni polinomio di grado (N-1) od inferiore che sia divisibile per g(D) deve essere un polinomio di codice.

→ Il polinomio g(D) è DETTO POLINOMIO GENERATORE del codice ciclico

Given a cyclic code (N,K), it exists a unique polynomial of degree (N-K) of the indicated form g(D)=D^{N-k}+ …+1 that is able to generate all the code-words …

All the other polynomials associated to the code-words are multiples of g(D), and all the pol. of degree less or equal to N-1 which are divisible by g(D) are code words.

The polynomial g(D) is said to be the “generator polynomial” of the considered cyclic code.

→ Il polinomio $g(D)$ è DETTO POLINOMIO GENERATORE del CODICE CICLICO

→ Il polinomio generatore di un codice ciclico (N, K) è un DIVISORE di $(D^N + 1)$.

→ ogni divisore di $(D^N + 1)$ di grado $(N - K)$ genera un codice ciclico (N, K)

The polynomial g(D) is said to be the “generator polynomial” of the considered cyclic code.

IMPORTANT FACTS:

1) The generator polynomial of a cyclic code “had to be” a divisor of (D^N + 1).

2) Every divisor of (D^N + 1) of degree N-K generates a cyclic code (N, K).

GENERAZIONE DI CODICI CICLICI

Consideriamo un polinomio $g(D)$ di grado "r".

Indichiamo con $m(D) = m_1 D^{K-1} + \dots + \dots + m_K$ il polinomio corrispondente alla parola di informazione $\bar{m} = (m_1, \dots)$

il polinomio

$$g(D) \cdot m(D) = x(D)$$

di grado $N \leq K+r-1$ può essere assunto come corrispondente ad una parola di codice relativa al blocco $\bar{m}$.

$g(D)$ è il POLINOMIO GENERATORE.

Consider a polynomial g(D) of degree "r".

Indicate with m(D)=m_1 D^{k-1} + … + m_k the polynomial associated to the information word m.

The polynomial x(D)=g(D) m(D)
of degree N <= K+r +1 is a code word x associated to m, where g(D) is the generator polynomial of this code.

The code is linear.

The code is cyclic if g(D) is a divisor of (D^N + 1).

Il codice è SISTEMATICO se uso, come parole di codice, in luogo di $u(D)g(D)$, le sequenze (polinomi)

$$u(D)\,D^{N-K} + \text{resto}\left\{\frac{u(D)\,D^{N-K}}{g(D)}\right\}$$

cioè

$$x(D) = \underbrace{u_1 D^{N-1} + u_2 D^{N-2} + \cdots + u_K D^{N\,\text{[redacted]}\,K}}_{K \text{ bit di informazione}} +$$

$$+ \underbrace{r_1 D^{(N-K-1)} + \cdots}_{} + r_{N-K}$$

$N-K$ bit di parità ottenuti calcolando il resto di $\dfrac{u(D)\,D^{N-K}}{g(D)}$

Usually this code is not systematic.
To obtain the related systematic code we have to use the word obtained using this relation:
……

$$x(D) = \underbrace{u_1 D^{N-1} + u_2 D^{N-2} + \dots + u_K D^{N-K}}_{K \text{ bit di informazione}} +$$

$$+ \underbrace{r_1 D^{(N-K-1)} + \dots + r_{N-K}}_{}$$

Reminder of the division …

N-K bit di parità ottenuti calcolando il resto di $\dfrac{u(D)\,D^{N-K}}{g(D)}$

Parity check bits …

* CIFRE del
* CONTROLLO DI
* PARITÀ
*

$\longrightarrow$ resto $\left(\dfrac{D^{N-K}\; m(D)}{g(D)} \right)$

ES: $H:(7,4)$

$$
G = \begin{array}{cccc|ccc|c}
1 & 0 & 0 & 0 & 1 & 0 & 1 & m_1 \\
0 & 1 & 0 & 0 & 1 & 1 & 1 & m_2 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & m_3 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & m_4 \\
x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 &
\end{array}
$$

$$
G(D) = \left| \begin{array}{cccc|ccc}
D^6 + & & & & D^2 & & +1 \\
& D^5 + & & & D^2 & +D & +1 \\
& & D^4 + & & D^2 & +D & \\
& & & D^3 + & & +D & +1
\end{array} \right|
$$

$$g(D) = D^3 + D + 1$$

$$G(D) = \left| \begin{array}{cc} (D^3+D+1) & g(D) \\ (D^2+1) & g(D) \\ D & g(D) \\ & g(D) \end{array} \right|$$

$$m = 1101 \rightarrow m(D) = D^3 + D^2 + 1$$

$$x(D) = (D^3 + D^2 + 1)(D^3 + D + 1) = D^6 + D^5 + D^4 + D^3 + D^2 + D + 1$$

$\overline{x} \Rightarrow 1111111$ NON È SISTEMATICO

$$m = 1101 \rightarrow x(D) = m(D)\,D^{N-K} + r\left(\frac{m(D)\,D^{N-K}}{g(D)}\right)$$

$x(D) = D^6 + D^5 + D^3 + 1 \rightarrow 1101001 \rightarrow$

La verifica che la parola di codice ricevuta $\bar{y}$ sia priva di errori si può compiere controllando che il polinomio $y(D)$ sia DIVISIBILE per il polinomio generatore $g(D)$.

Possono quindi essere "rilevate" tutte le configurazioni di errore NON DIVISIBILI per $g(D)$

Every code word is divisible by g(D).

Therefore, this code is able to identify all the error configurations associated to polynomials which are NOT divisible by g(D) (i.e., that are NOT code words itself …).

The remainder of the division is related to the syndrome.

Codice di Hamming (7, 4)

$g(D) = D^3 + D + 1$ $\quad$ [N.B.: $D^7 + 1 = g(D)\,h(D) = (D^3+D+1)(D^4+D^2+D+1)$]

$\bar{m}$	$m(D)$	$x(D) = m(D)\,g(D)$	$\bar{x}$
0000			0000000
0001	1	D^3+D+1	0001011
0010	D	D^4+D^2+D	0010110
0011	$D+1$	$D^4+D^3+D^2+1$	0011101
0100	D^2	$D^5+D^3+D^2$	0101100
0101	D^2+1	D^5+D^2+D+1	0100111
0110	D^2+D		
0111	D^2+D+1		
1000			
1001			
1010			
1011			
1100			
1101			
1110			
1111	D^3+D^2+D+1	$D^6+D^5+D^3+1$	1101001

Codice non sistematico.

VERSIONE "SISTEMATICA"

$$x(D) = u(D)\,D^3 + \text{resto}\left\{\frac{u(D)\,D^3}{g(D)}\right\}$$

$\bar{u}$	$x(D)$	$\bar{x}$
0000		0000 | 000
0001	$D^3 + D + 1$	0001 | 011
0010	$D^4 + D^2 + D$	0010 | 110
0011	$D^4 + D^3 + D^2 + 1$	0011 | 101
0100		111
0101		100
0110		001
0111		010
1000		101
1001		110
1010		011
1011		000
$\boxed{1100}$		$\boxed{010}$
1101		001
1110		100
1111	$D^6 + D^5 + D^4 + D^3 + D^2 + D + 1$	1111 | 111

Examples of Important Cyclic Codes: Hamming

- *codici di Hamming*: classe infinita di codici, con coppie di valori di N e K che soddisfano la condizione $N = 2^{N-K} - 1$: (7,4), (15,11), (31,26), (63,57), (127,120), e così via. I corrispondenti polinomi generatori possono essere (ne esiste più d'uno) $D^3 + D + 1$, $D^4 + D + 1$, $D^5 + D^2 + 1$, $D^6 + D + 1$, $D^7 + D^3 + 1$, ... La distanza minima d è però sempre pari a 3, per cui i codici con N grande hanno scarso interesse, se non su canali poco rumorosi. Anche quelli con N piccolo non sono molto interessanti perché troppo semplici; infatti occupano un piccolo numero di dimensioni.

Hamming codes.
Very famous, being the first example of "one error" correcting codes.
Class of many cyclic codes, with N=2^{N-K} -1; (7,4), …
The generator polynomials could be: D^3+D+1, …
The minimum distance dmin is ALWAYS equal to 3 … therefore if N is big the P(E) is not very good (as we will see in more detail later ...).
If N is small, the performance are anyway not so interesting …
They are used as a basic building block to obtain more sophisticated codes …

Examples of Important Cyclic Codes: BCH and RS

12.2.6. BCH and Reed-Solomon Codes

BCH codes, named after the inventors, Bose, Ray-Chaudhuri, and Hocquenghem, are a large class of multiple-error-correcting codes invented around 1960. For any positive integers m and t, there is a t-error-correcting binary BCH code with

$$n = 2^m - 1\ , \qquad k \geq n - mt\ . \tag{12.50}$$

In order to correct t errors, it is clear that the minimum Hamming distance is bounded by

$$d_{H,\min} \geq 2t + 1\ . \tag{12.51}$$

BCH codes are important primarily because practical and efficient decoding techniques have been found [13], and because of the flexibility in the choice of parameters (n and k).

An important class of nonbinary BCH codes are *Reed-Solomon* codes, in which the symbols are blocks of bits. Their importance is again the existence of practical decoding techniques, as well as their ability to correct bursts of errors.

Reed-Solomon Codes

- One of the most error control codes is Reed-Solomon codes.
- These codes were developed by Reed & Solomon in June, 1960.
- The paper I.S. Reed and Gus Solomon, " Polynominal codes over certain finite fields ", Journal of the society for industrial & applied mathematics.
- Reed-Solomon (RS) codes have many applications such as compact disc (CD, VCD, DVD), deep space exploration, HDTV, computer memory, and spread-spectrum systems.
- In the decades, since RS discovery, RS codes are the most frequency used digital error control codes in the world.

Reed-Solomon (RS) code

- An RS code is a cyclic symbol error-correcting code.
- An RS codeword will consist of I information or message symbols, together with P parity or check symbols. The word length is $N=I+P$.
- The symbols in an RS codeword are usually not binary, i.e., each symbol is represent by more than one bit. In fact, a favorite choice is to use 8-bit symbols. This is related to the fact that most computers have word length of 8 bits or multiples of 8 bits.

The parameters of a Reed-Solomon code are the following:

Symbol	m binary digits
Block length n	$=(2^m - 1)$ symbols
	$=m(2^m - 1)$ binary digits
Parity checks $(n-k)$	$= 2t$ symbols
	$= 2mt$ binary digits

These codes are capable of correcting all combinations of t or fewer symbol errors. Alternatively, interpreted as binary codes, they are well suited for correction of bursts of errors (see Section 10.2.10). In fact, one symbol in error means a number of binary digits in error ranging from 1 to m in adjacent positions within the code word. Perhaps the most important application of these codes is in the concatenated coding scheme

Examples ...

Table 13.2-3 Selected cyclic codes

Type	n	k	R_c	d_{min}	$G(p)$
Hamming Codes	7	4	0.57	3	1 011
	15	11	0.73	3	10 011
	31	26	0.84	3	100 101
BCH Codes	15	7	0.46	5	111 010 001
	31	21	0.68	5	11 101 101 001
	63	45	0.71	7	1 111 000 001 011 001 111
Golay Code	23	12	0.52	7	101 011 100 011

Modification to Known Codes

1. Puncturing: delete a parity symbol
   - (n,k) code → (n-1,k) code

2. Shortening: delete a message symbol
   - (n,k) code → (n-1,k-1) code

3. Expurgating: delete some subset of codewords
   - (n,k) code → (n,k-1) code

4. Extending: add an additional parity symbol
   - (n,k) code → (n+1,k) code

Expurgating, Shortening, Extending

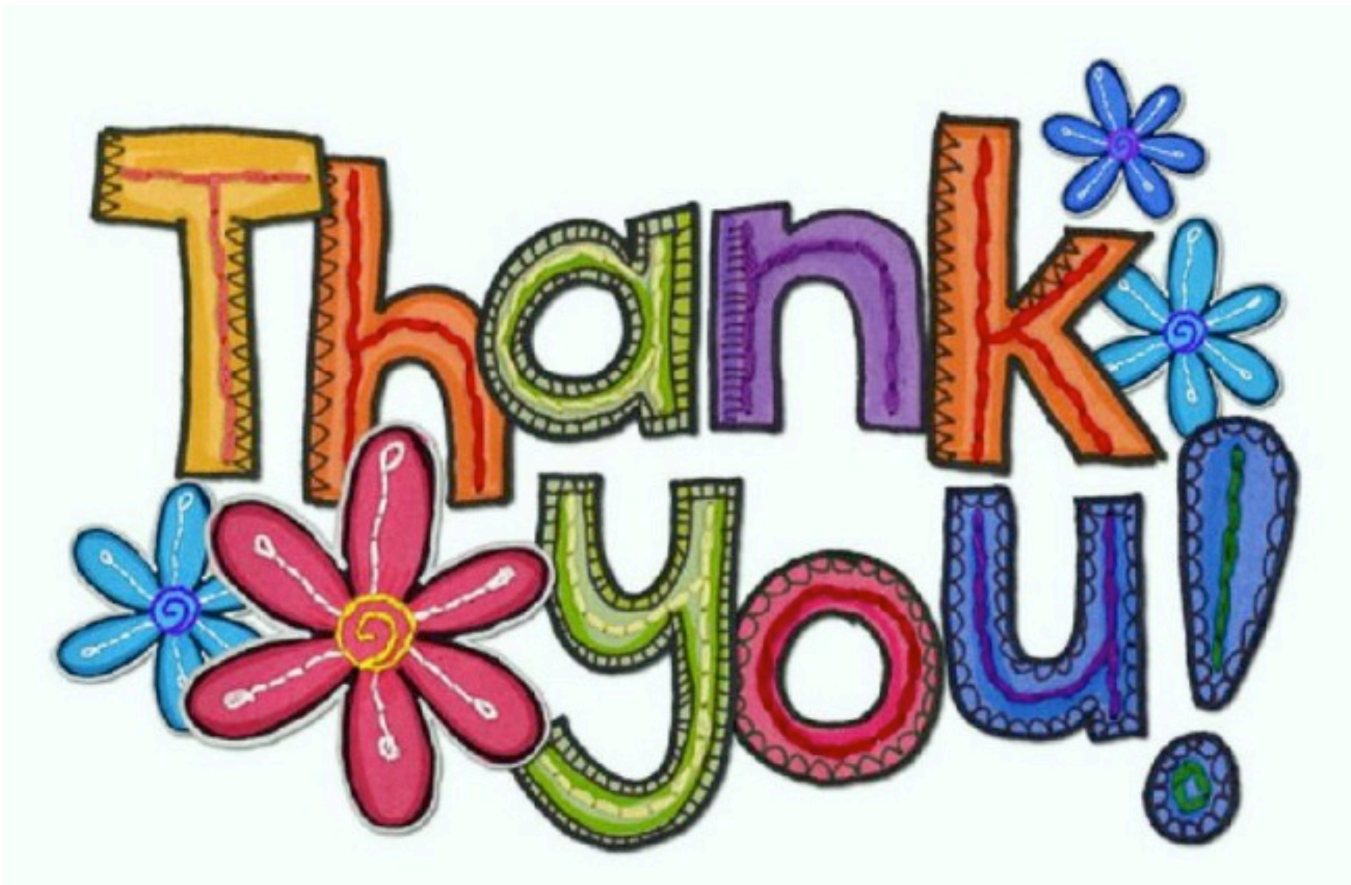
A cyclic code with an odd minimum distance can be expurgated by multiplying the polinomial generator for factor D + 1. ncreasing by one the degree of g (D) is reduced by one K. It is easy to see that all the words in the expurgated code have an even number of ones, and therefore the minimum distance is even, and then increased by one.
The expurgated code is cyclic.

Finally the code can be shortened.
The information bits in the first b positions are reset. Obviously, this data are not transmitted, and a new code is then obtained with K '= K-b and N '= N - b.
The shortened code is not cyclic.

Extended Hamming codes.
Adding to any linear code a general parity check bit, with the same K, we obtain a new code (not cyclic) with an even d_min (at least the same, or greater than, of the starting code).
In case of Hamm. codes d_min becomes therefore 4 … (better than 3 without any big effort).
(N,K) becomes (8,4), …

Shortened cyclic codes

Since the generator polynomial of a cyclic code must be a divisor of $(Z^n + 1)$, it often happens that its possible degree $(n-k)$ does not cover all combinations of n and k that satisfy practical needs. To avoid this difficulty, cyclic codes are sometimes used in a shortened form. To this purpose, the first i information digits are assumed to be always zero and are not transmitted. In this way, a new $(n-i, k-i)$ code is derived whose code words are a subset of the code words of the original code. The code is called *shortened* cyclic code, although it may not be cyclic. The new code has at least the same minimum distance as the code from which it is derived. The encoding and syndrome calculation can be accomplished by the same circuits employed in the original code, since the leading string of zeros does not affect the parity-check computations. Error correction can be accomplished by prefixing to each received vector a string of i zeros, or by modifying accordingly the related circuitry. Therefore, these codes share all the implementation advantages of cyclic codes and are also of practical interest.


Thank you!