

IL CODICE EAN

<i>INTRODUZIONE</i>	2
<i>CODICI RIVELATORI DI ERRORI</i>	3
<i>IL CODICE A BARRE</i>	5
<i>Struttura di un simbolo</i>	5
<i>IL CODICE EAN</i>	7
<i>Codifica dei Caratteri</i>	7
<i>Struttura di un simbolo EAN 13</i>	8
<i>Struttura di un simbolo EAN 8</i>	9
<i>Dimensioni</i>	9
<i>Calcolo del CHECK DIGIT</i>	10
<i>Interpretazione del codice EAN</i>	12
<i>Passaggio da ISBN a EAN</i>	13
<i>GLOSSARIO</i>	14

INTRODUZIONE

In generale, la trasmissione di un messaggio da una sorgente ad un destinatario attraverso un canale di comunicazione comporta un insieme di alterazioni dei segnali impiegati per tale scopo, dovute in parte alle caratteristiche del mezzo trasmissivo ed in parte alle apparecchiature di trasmissione e di ricezione.

In particolare, la presenza di “rumore” nel canale di comunicazione, cioè l’esistenza di disturbi, interferenze casuali o altro, possono provocare degli errori casuali che modificano le informazioni immesse nel canale: è necessario quindi trovare dei metodi che permettano di rilevare tali errori e in alcuni casi poter risalire alla informazione vera. In modo più preciso, per errore si intende una alterazione dei simboli trasmessi.

La teoria dei codici rivelatori e correttori di errori si occupa di risolvere i problemi causati dal rumore presente nel canale: si aggiunge all’informazione che si trasmette un ulteriore “pezzo”, strettamente dipendente dall’informazione stessa, che permette anche in presenza di errori, di ricostruire i dati trasmessi, o almeno di scoprire se i dati ricevuti contengono errori.

Quando si vuole risalire al messaggio trasmesso dall’informazione ricevuta si parla di **Codici correttori di errori**; quando invece si vuole solo stabilire se l’informazione ricevuta è il messaggio trasmesso si parla di **Codici rivelatori di errori**.

CODICI RIVELATORI DI ERRORI

Sia dato un insieme A di simboli che prende il nome di **Alfabeto** (es: le cifre decimali). Si chiama **Codice di lunghezza n** sull'alfabeto A un insieme C di n-ple di simboli A. Gli elementi di C prendono il nome di **parole**.

Un codice rivelatore di errori è tale che cambiando un solo simbolo in una qualsiasi parola, la n-pla che si ottiene non è una parola del codice. In questo modo la generazione di un errore su una parola appartenente al codice produce una parola non appartenente al codice.

Formalmente:

un codice (n,k) con $n > k$ è un codice con parole di lunghezza n di cui k bit di informazione.

Si definisce inoltre **peso di un errore** il numero di bit "corrotti" durante la trasmissione. In sistemi binari si hanno solo due casi di errore:

- Trasmetto 0 ricevo 1
- Trasmetto 1 ricevo 0

Si definisce **distanza di Hamming $d(x,y)$** fra due parole x, y di un codice C il numero di posizioni (bit) per cui differiscono:

$$d(10010,01001) = 4$$

$$d(11010,11001) = 2$$

La distanza minima di un codice è allora:

$$d_{\min} = \min(d(x,y)) \text{ per ogni } x \text{ diverso da } y \text{ appartenenti a } C$$

Un codice a distanza minima d è capace di rivelare errori di peso d-1

Es:

codice 1

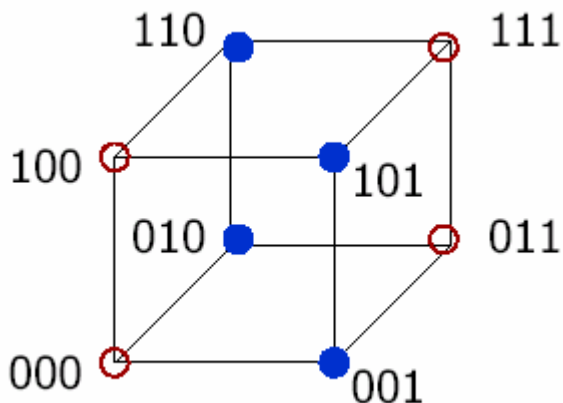
A = 000

B = 100

C = 001

D = 111

$d_{\min} = 1$



codice 2

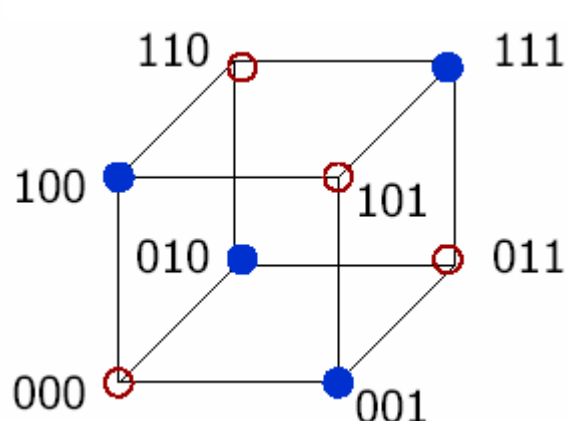
A = 000

B = 011

C = 101

D = 110

$d_{\min} = 2$



Il prototipo dei codici rivelatori di errori è il **Codice di controllo parità**.

In esso ad ogni sequenza di simboli binari, che rappresenta un messaggio, ad esempio una lettera dell'alfabeto, si aggiunge un altro simbolo binario, cioè 0 oppure 1, in modo tale che la somma di tutti i simboli sia un numero pari. Questo ultimo simbolo si chiama **Simbolo di Controllo**. Se in una sequenza entra un solo errore (o un numero dispari di errori), tale errore viene scoperto, perché in questo caso la somma dei simboli è un numero dispari.

Il **Codice di Parità** è un codice a $d_{min} = 2$ e può essere costruito utilizzando le seguenti espressioni:

$$d_1 + d_2 + d_3 + \dots + d_n + p = 0 \text{ parità oppure}$$

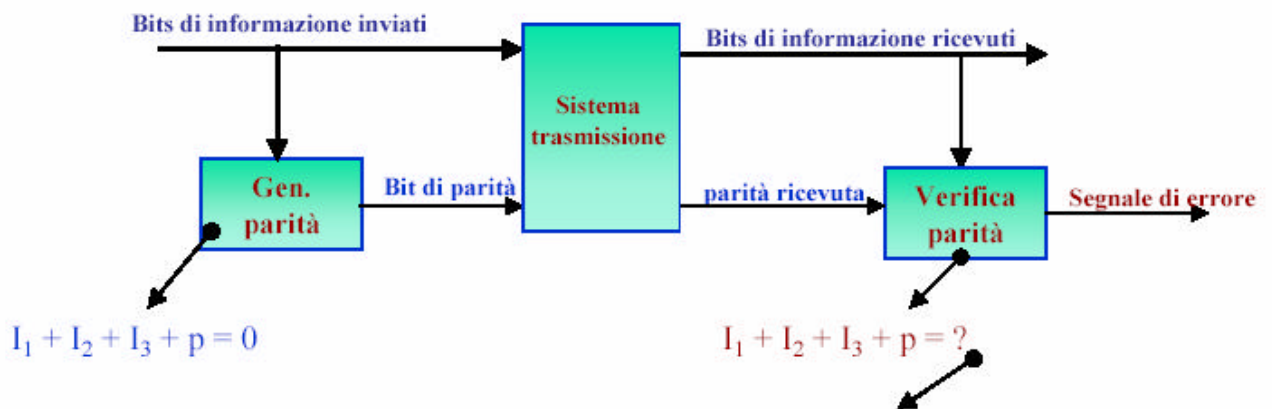
$$d_1 + d_2 + d_3 + \dots + d_n + p = 1 \text{ disparità}$$

dove n è il numero di bit usati per rappresentare in binario gli oggetti (informazione), $+$ è l'operatore di somma modulo 2 e p il bit di "parità" da aggiungere a quelli di informazione per costruire parole de codice (CODIFICA)

Bits di informazione	Parità	Disparità
000	000 0	000 1
001	001 1	001 0
010	010 1	010 0
011	011 0	011 1
100	100 1	100 0
101	101 0	101 1
110	110 0	110 1
111	111 1	111 0

E' un codice di distanza minima pari a 2 che permette di rivelare errori di peso 1

Es: devo trasmettere l'informazione 101



Il generatore di parità calcola il bit di parità $1 + 0 + 1 + p = 0$ cioè $p = 0$ e trasmetto 1010

Il ricevitore riceve 1110, ne verifica la parità $1 + 1 + 1 + 0 = 1 \neq 0$ quindi si è verificato un errore.

Se avessi ricevuto 1111 $\Rightarrow 1 + 1 + 1 + 1 = 0$ **tutto ok?** No, i doppi errori sono sfuggiti al check.

IL CODICE A BARRE

Il codice a barre può essere definito come una simbologia o un alfabeto per la codifica di informazioni in un formato tale da poter essere acquisito automaticamente da opportuni lettori. La codifica si basa su un concetto binario. La diversa logica di codifica, dettata da diverse esigenze applicative, ha portato a diversi tipi di codici a barre. Alcuni di essi sono ampiamente diffusi, altri vengono usati solo in speciali settori, altri ancora solo in determinati paesi.

Di codici a barre ne esistono di diversi tipi, con caratteristiche diverse. Agli inizi sono state proposte diverse tecniche di codifica, con il passare degli anni alcune di queste proposte sono quasi sparite dall'uso pratico, mentre altre si sono evolute; attualmente solo una mezza dozzina di simbologie sono praticamente usate, tra queste alcune sono presenti nella nostra realtà quotidiana, mentre altre sono destinate ad applicazioni speciali.

Ogni simbologia comporta la definizione di un set di caratteri (alfabeto) finito ed univoco, per ciascun carattere ne viene data una rappresentazione come sequenza binaria e da questa, con un opportuno criterio, si passa alla rappresentazione ottica.

1. Codici a due spessori: sono codici i cui elementi (barre e spazi) possono assumere due soli spessori, in rapporto compreso tra 2 e 3.
 - Codice 39 Full ASCII
 - MSI-code
 - Codice BCD Matrix
2. Codici a più di due spessori: sono codici i cui elementi (barre e spazi) possono assumere più di due spessori diversi. In generale ciò richiede una migliore qualità di stampa.
 - Codice EAN
 - Codice UPC
 - Codice 128
3. Codici a due dimensioni: sono codici i cui elementi (barre e spazi) possono assumere più di una dimensione. In generale ciò richiede una migliore qualità di stampa.
 - MaxiCode
 - PDF417
 - SoftStrip

Struttura di un simbolo

Un tipo di codifica è caratterizzata da:

- Rapporto tra spessore degli elementi larghi e spessore degli elementi stretti.
- Spessore dell'elemento stretto. Viene spesso definito "spessore del modulo", e viene usato come unità di misura degli spessori degli elementi larghi.
- Sequenze di barre e spazi per la codifica dei caratteri rappresentabili. I singoli caratteri sono codificati con 4,5,7,8 o 9 elementi.
- Caratteri rappresentabili. Tutti i codici possono rappresentare le cifre decimali. Alcuni possono codificare anche lettere dell'alfabeto e/o caratteri speciali.
- Numero di caratteri rappresentabili con una etichetta. Alcuni codici rappresentano un numero fisso di caratteri.
- Sequenze di barre e spazi per la codifica dei caratteri di start/stop. Alcuni codici prevedono la stessa sequenza per entrambi; altri sequenze distinte.
- Tolleranze per i singoli elementi. Le tolleranze ammesse possono variare da pochi

centesimi a qualche decimo di millimetro.

- **Cifra di controllo.** Per aumentare la sicurezza nella lettura, in aggiunta ai caratteri utili, può essere inserita una cifra di controllo (check digit).

IL CODICE EAN

La funzione della codifica EAN è rappresentare il codice numerico di ogni articolo soggetto alle specifiche EAN (European Article Number) in un formato leggibile automaticamente in entrambe le direzioni. Il codice EAN è interamente compatibile con il codice UPC (Universal Product Code) utilizzato negli USA.

Un simbolo EAN è composto da un codice di identificazione a 13 cifre e da una corrispondente rappresentazione grafica consistente in un simbolo a barre verticali destinato alla lettura ottica.

Gli elementi possono assumere 4 diversi spessori, multipli interi del modulo. Ogni cifra è codificata con 7 moduli. Il carattere di controllo centrale è composto da 5 moduli; quello di start/stop da 3 moduli. I simboli EAN sono disposti "simmetricamente" attorno ad un carattere centrale di controllo, e delimitati agli estremi da caratteri laterali di controllo (start/stop).



Le caratteristiche fondamentali del codice EAN sono:

- codice numerico
- codice continuo
- codice self-checking

Inoltre presenta una alta densità di informazione, ma ha un numero fisso di cifre e gli spazi sono significativi.

Il set di caratteri è composto da:

- 10 cifre
- 1 carattere di start/stop
- 1 carattere di controllo centrale (solo nella versione EAN 13)

Codifica dei Caratteri

Le cifre decimali sono rappresentate attraverso tre diverse codifiche, siglate A, B e C come segue:

Carattere	Codifiche A	Codifica B	Codifica C
0	0001101	0100111	1110010
1	0011001	0110011	1100110
2	0010011	0011011	1101100
3	0111101	0100001	1000010
4	0100011	0011101	1011100
5	0110001	0111001	1001110
6	0101111	0000101	1010000
7	0111011	0010001	1000100
8	0110111	0001001	1001000
9	0001011	0010111	1110100

Nella tabella alla cifra binaria "1" corrisponde un **modulo** scuro ed alla cifra "0" un **modulo** chiaro. Conseguentemente sarà chiamata "**barra**" una qualunque successione continua di moduli scuri, "**spazio**" una qualunque successione continua di moduli chiari. In ogni caso la forma codificata di un carattere include due barre scure.

Le codifiche A e B cominciano sempre con uno spazio e finiscono con una barra; mentre la codifica C, al contrario, comincia sempre con una barra e termina sempre con uno spazio.

La codifica A ha sempre un numero dispari di moduli scuri, mentre le codifiche B e C ne contengono sempre un numero pari. Perciò la codifica A è una codifica a “parità dispari” e le codifiche B e C sono a “parità pari”.

I caratteri ausiliari di controllo sono codificati come segue:

- controllo laterale: 101
- controllo centrale: 01010

Struttura di un simbolo EAN 13

Il simbolo EAN 13 è costituito dalla seguente sequenza, descritta da destra verso sinistra:

- carattere di controllo laterale (start/stop)
- 6 caratteri rappresentati secondo la codifica C (numerati da 1 a 6), che costituiscono la “metà destra” del simbolo;
- carattere di controllo centrale;
- 6 caratteri con codifica A o B (numerati da 7 a 12), che costituiscono la “metà sinistra” del simbolo;
- carattere di controllo laterale (start/stop)

Il tredicesimo carattere è associato ad una data combinazione di codifiche A e B dei caratteri 7 –12 nella metà sinistra del simbolo:

13° Carattere	Codifiche Metà Sinistra					
Carattere	12	11	10	9	8	7
0	A	A	A	A	A	A
1	A	A	B	A	B	B
2	A	A	B	B	A	B
3	A	A	B	B	B	A
4	A	B	A	A	B	B
5	A	B	B	A	A	B
6	A	B	B	B	A	A
7	A	B	A	B	A	B
8	A	B	A	B	B	A
9	A	B	B	A	B	A

Considerando una qualsiasi coppia di caratteri adiacenti, siano entrambi caratteri utili oppure uno dei due sia un carattere ausiliario, l’ultimo modulo del carattere di sinistra è sempre diverso dal primo modulo del carattere di destra. Ciò significa che il confine tra due caratteri può sempre essere distinto visivamente, cosa indispensabile per la decodifica.

Struttura di un simbolo EAN 8

Il simbolo EAN 8 è costituito dalla seguente sequenza, descritta da destra verso sinistra:

- carattere di controllo laterale (start/stop)
- 4 caratteri rappresentati secondo la codifica C (numerati da 1 a 4), che costituiscono la “metà destra” del simbolo;
- carattere di controllo centrale;
- 4 caratteri con codifica A o B (numerati da 5 a 8), che costituiscono la “metà sinistra” del simbolo;
- carattere di controllo laterale (start/stop)

Anche per i simboli EAN 8 vale il discorso fatto per i simboli EAN 13: considerando una qualsiasi coppia di caratteri adiacenti, siano entrambi caratteri utili oppure uno dei due sia un carattere ausiliario, l'ultimo modulo del carattere di sinistra è sempre diverso dal primo modulo del carattere di destra. Ciò significa che il confine tra due caratteri può sempre essere distinto visivamente, cosa indispensabile per la decodifica.

Dimensioni

Lo spessore nominale del modulo è 0.33 mm. Le dimensioni nominali dell'etichetta possono essere moltiplicate per un “fattore di ingrandimento” compreso fra 0.8 e 2. Le barre dei caratteri di controllo centrale e laterali si estendono per 5 moduli al di sotto delle barre che rappresentano i caratteri utili. L'area di margine chiaro circondante l'area stampata corrisponde a:

- 7 moduli alla destra del simbolo;
- 11 moduli (EAN 13) o 7 moduli (EAN 8) alla sinistra del simbolo;
- 1 modulo sopra il simbolo;
- 1 modulo tra l'estremo inferiore delle barre e l'interpretazione OCR posta al di sotto di esse.

La lunghezza del simbolo EAN 13 è quindi di 95 moduli, più 18 moduli di overflow. Quella del simbolo EAN 8 è di 67 moduli, più 14 moduli per le zone di overflow.

Calcolo del CHECK DIGIT

Sia nella versione 13 che in quella 8, il carattere più a destra assume la funzione di check digit (cifra di controllo), ed è calcolata sulla base delle restanti 12 o 7 cifre con l'algoritmo seguente (le posizioni dei caratteri sono numerate da destra verso sinistra):

1. Partendo dalla posizione 2 sommare i valori dei caratteri in posizione pari.
2. Moltiplicare per 3 il risultato dell'operazione 1
3. Partendo dalla posizione 3 sommare i valori dei caratteri in posizione dispari
4. Sommare i risultati delle operazioni 2 e 3
5. Il check digit è il più piccolo numero che sommato al risultato dell'operazione 4 da un numero multiplo di 10

Un altro algoritmo del tutto analogo al precedente per calcolare il check digit è il seguente:

1. Associare un peso ad ogni cifra:
 - peso 3 ai valori dei caratteri in posizione pari partendo dalla posizione 2
 - peso 1 ai valori dei caratteri in posizione dispari partendo dalla posizione 3
2. Moltiplicare ogni cifra per il suo peso
3. Sommare tutti i prodotti
4. Il check digit è il più piccolo numero che sommato al risultato dell'operazione 3 da un numero multiplo di 10

Esempio 1: Calcolare la cifra di controllo

Posizione del carattere	13	12	11	10	9	8	7	6	5	4	3	2	1
Esempio di codice	4	2	7	6	2	2	1	3	5	7	4	6	?

1. $6+7+3+2+6+2 = 26$
2. $26*3=78$
3. $4+5+1+2+7+4 = 23$
4. $78+23 = 101$
5. $110 - 101 = \mathbf{9}$ (cifra di controllo)

Esempio 2: Verificare la correttezza della cifra di controllo

Posizione del carattere	13	12	11	10	9	8	7	6	5	4	3	2	1
Esempio di codice	4	5	5	6	7	8	9	0	3	4	4	6	1

1. $6+4+0+8+6+5 = 29$
2. $29*3 = 87$



3. $4+3+9+7+5+4 = 32$
4. $87+32 = 119$
5. $120 - 119 = \mathbf{1}$ OK

Interpretazione del codice EAN

Vediamo come si associa ad un prodotto la sequenza numerica:

- Le prime due cifre del numero indicano la nazione di produzione

Cifre	Nazione
00-09	USA, Canada
30-37	Francia
40-43	Germania
45	Giappone
50	Gran Bretagna
76	Svizzera
80-83	Italia
87	Olanda
90-91	Austria

- Il gruppo delle 5 cifre seguenti indica la ditta di produzione
- Le 5 cifre successive indicano il tipo di prodotto (che viene completamente individuato, di conseguenza da queste 5 cifre si deduce anche il prezzo)
- L'ultima cifra è quella di controllo

Le prime due cifre sono assegnate internazionalmente, le altre cinque sono stabilite in sede nazionale dall'Istituto preposto alla codifica, mentre il codice prodotto (ultime cinque cifre) è stabilito dal produttore: l'insieme dei numeri individua un, e un solo, prodotto in tutto il circuito commerciale mondiale.

Passaggio da ISBN a EAN

Il codice EAN può essere dotato di un “postambolo” a 2 o a 5 caratteri per essere utilizzato anche per la codifica dei libri e dei periodici.

Nella sua utilizzazione, il codice EAN - collegato all'ISBN - ha la seguente struttura:

Le prime tre cifre sono fisse e identificano internazionalmente il prodotto libro (978).

Le cifre dalla 4a alla 12a corrispondono alle prime nove cifre dell'ISBN.

Esempio 3: **Algoritmo per generare il codice EAN da un codice ISBN**

Partendo da un codice ISBN ad esempio:

88 7075 179 1

si toglie il codice di controllo 1 e si premettono le tre cifre 978 che - come si è detto - identificano internazionalmente il prodotto libro. Si otterrà quindi:

978 88 7075 179

Moltiplicando le cifre in posizione dispari per 1 e quelle in posizione pari per 3 si otterranno le seguenti cifre:

9 21 8 24 8 21 0 21 5 3 7 27

Si divide poi la somma dei prodotti così ottenuta per 10.

Se la divisione è senza resto il numero di controllo è 0. Quando vi fosse un resto, il numero di controllo è il complemento a 10 del resto stesso. Nel nostro caso, poiché la somma dei prodotti è 154, il numero di controllo è 6.

$154 : 10 = 15$ con resto di 4

$10 - 4 = 6$

l'EAN completo sarà dunque: 978 88 7075 179 6

GLOSSARIO

CODICE A BARRE

Successione di elementi scuri (barre) ed elementi chiari (spazi) alternati. Talvolta si intende con "codice" tipo di codifica.

ELEMENTO

Denominazione sia di uno spazio che di una barra.

BARRA

Elemento scuro di un codice

SPAZIO

Elemento chiaro tra due barre.

SIMBOLO DI CODICE A BARRE

Codice a barre completo, costituito da una zona chiara prima del codice, un carattere di start, uno o più caratteri contenenti l'informazione, un carattere di stop ed una ulteriore zona chiara.

CODICE DISCRETO

Codice a barre in cui non tutti gli elementi sono significativi. Alcuni servono solo da "separatori", e sui loro spessori sono ammesse tolleranze relativamente ampie.

CODICE CONTINUO

Sia le barre che gli spazi contengono informazione.

CODICE NUMERICO

Codice a Barre composto solo da numeri decimali.

SPAZIO INTERCARATTERE

Spazio che separa due caratteri adiacenti. In alcuni codici gli spazi all'interno di un carattere contengono informazione, gli spazi intercarattere no.

MODULO

Il più piccolo spessore di un elemento in un codice. Costituisce l'unità di misura degli spessori degli elementi larghi.

CARATTERE

Successione di barre e spazi che codifica una singola cifra decimale o un carattere alfabetico o speciale.

CARATTERE DI

Ogni simbolo comincia con un carattere di start e termina con un carattere di stop. Questo per permettere l'identificazione del tipo di codice e l'individuazione della direzione di lettura.

ZONE DI OVER-FLOW

Gli spazi chiari immediatamente precedente il carattere di start e seguente il carattere di stop.

CODICE DI SELF-CHEKING

Codice strutturato in maniera tale che da un conteggio delle barre e degli spazi letti (bar count) si possa verificare l'esattezza del simbolo.

RAPPORTO DI STAMPA

Nei codici a due spessori, è il rapporto tra lo spessore dell'elemento largo e quello dell'elemento stretto.