

Informatica per le discipline umanistiche 2

Rappresentazione dell'informazione

Rappresentazione dell'informazione

Informatica: "studio sistematico degli algoritmi che descrivono e trasformano l'**informazione**: la loro teoria, analisi, progetto, efficienza, realizzazione, applicazione" (ACM)

- Cosa intendiamo per informazione?
- Come viene rappresentata negli elaboratori?
- Come viene rappresentata nel mondo esterno?

Informazione

- **Informazione:** dati *codificati*, su un determinato *supporto* e dotati di un significato
- **Codifica:** modo in cui l'informazione viene rappresentata
- **Decodifica:** modo in cui l'informazione viene interpretata
- Uno stesso dato può essere codificato in modi differenti (es. codifiche equivalenti di un numero)

Informazione

Supporto può portare informazione → può assumere **configurazioni differenti**
Caso elementare: due configurazioni distinguibili
Es. un foglio con una macchia o senza macchia, presenza o assenza di corrente

L'informazione nei sistemi digitali

Quantità di informazione: legata alla *riduzione di incertezza* portata da una certa configurazione

Bit: è l'unità di misura elementare dell'informazione

- permette di distinguere tra due configurazioni differenti: Vero-Falso, due livelli diversi di tensione elettrica, bianco o nero

Informazione

Per codificare e decodificare un messaggio:

- **Principio di composizione:** il messaggio è composto da *parti elementari*
- Nei messaggi di testo le parti sono i singoli simboli (**caratteri**)
- L'insieme dei simboli possibili: **alfabeto**
- Fase successiva: ricomposizione dei simboli nelle parole

Informazione

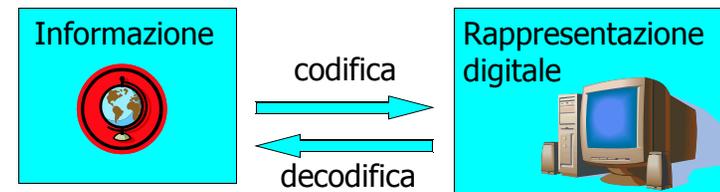
Se vogliamo rappresentare gli elementi di un insieme → **codifica univoca**

Deve permettere:

- codifica
- decodifica

Es. Ad ogni mese associo un nome univoco (o un numero univoco)

L'informazione nei sistemi digitali



Mondo esterno

L'informazione del mondo esterno deve essere codificata

Segnale può assumere solo 2 configurazioni

L'informazione nei sistemi digitali

- L'informazione nei computer è in formato **digitale**
- Informazione rappresentata attraverso due soli valori: i **bit** (*Binary digIT*)
- Le operazioni aritmetiche negli elaboratori sono effettuate con **aritmetica binaria**

L'informazione nei sistemi digitali

Oltre al bit esistono altre unità di misura (ricavate dal bit):

Una successione di 8 bit: **byte**

Con i byte è possibile rappresentare:

```
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 1
0 0 0 0 0 0 1 0
0 0 0 0 0 0 1 1
0 0 0 0 0 1 0 0
...
1 1 1 1 1 1 1 1
```

L'informazione nei sistemi digitali

- **Kilobyte** (KB) $2^{10}=1.024$ byte (circa mille byte)
- **Megabyte** (MB) $2^{20}=1.048.576$ byte (circa un milione di byte)
- **Gigabyte** (GB) $2^{30}=1.073.741.824$ byte (circa un miliardo di byte)
- **Terabyte** (TB) $2^{40}=1.099.511.627.776$ byte (circa mille miliardi di byte)

Transistor

- I computer sono dotati di dispositivi (**transistor**) in grado di assumere due valori: **acceso/spento**
- Ognuno di questi valori associato al valore di un bit

Codifica Binaria

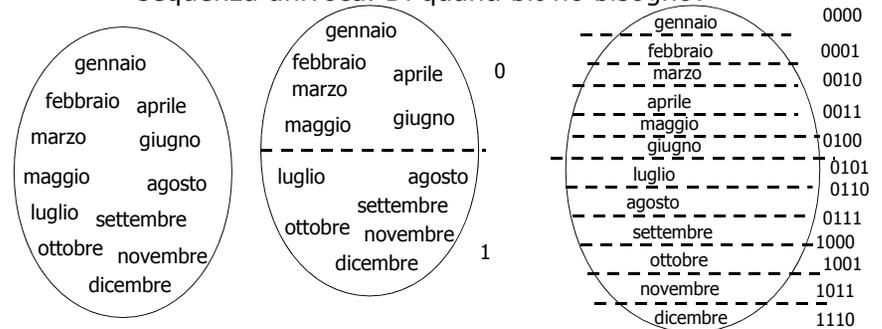
- Supponiamo di avere una successione di k bit: posso rappresentare 2^k successioni diverse.

- $k=2$ 00, 01, 10, 11
- $k=3$ 000, 001, 010, 011,
 100, 101, 110, 111

Se ho k bit posso costruire al più 2^k codifiche univoche

Codifica digitale dei dati

- Supponiamo di dover codificare dati non numerici. Es. i mesi dell'anno
- Codifica dell'insieme: ad ogni elemento associo una sequenza univoca. Di quanti bit ho bisogno?



Se devo rappresentare un insieme di n elementi, scelgo k tale che $2^k \geq n$

L'informazione: caratteri e stringhe

Carattere: simbolo dell'alfabeto

Stringa: successione di caratteri (parola)

Conversione in formato digitale

La rappresentazione dei caratteri

- E' necessario rappresentare con una *codifica binaria* tutti i simboli dell'alfabeto:
 - Lettere maiuscole e minuscole, cifre numeriche e simboli matematici, altri caratteri (es. è, ò, @, ...)
 - Ad ogni carattere → sequenza di bit **univoca**

Codice ASCII

- ▣ Proposta dall'ANSI (*American National Standards Institute*)
- ▣ Codice **ASCII** (*American Standard Code for Information Interchange*): codifica con 7 bit
- ▣ Successivamente Extended ASCII: estensione a 8 bit (1 byte)
 - Il carattere "A" è rappresentato da 0100 0001

Codice ASCII

Carattere	Codice ASCII
...	...
A	0100 0001
B	0100 0010
C	0100 0011
D	0100 0100
E	0100 0101
...	...

Codifica di un testo

"Nel mezzo del cammin di nostra vita"

Traduciamo il testo in codice ASCII:

```
N           e           l           (spazio)           m
01001110 01100101 01101100 00000000 01101101
01100101 01111010 01111010 01101111 00000000
01100100 01100101 01101100 ...
```

Decodifica del codice ASCII

Supponiamo di avere un testo in codice ASCII

```
01001110 01100101 01101100 00000000 01101101 01100101
01111010 01111010 01101111 00000000
```

Per decodificare la sequenza dividiamola in gruppi di 1 byte e traduciamo ogni gruppo

```
01001110 01100101 01101100 00000000 01101101
N           e           l           m
```

Altre codifiche di caratteri

▣ Altre codifiche:

- Codice **Unicode**: 2 byte (16 bit) per rappresentare un carattere.
- Codice Unicode esteso

Rappresentazione dei numeri

I numeri sono codificati in modo particolare

La codifica ASCII:

- molto dispendiosa
- non rispetta le regole aritmetiche

Numeri: **rappresentazione binaria**; cifre utilizzate *0* e *1*

Es. *21* → *10101*

L'informazione: dati numerici

I dati numerici → convertiti in formato *binario*

Quando scriviamo 145, usiamo una codifica decimale:

$$1 * 10^2 + 4 * 10^1 + 5 * 10^0 = 145_{10}$$

Le cifre possono andare da *0* a *9*

Per rappresentare il numero è necessario convertirlo in rappresentazione binaria

$$1 * 2^7 + 0 * 2^6 + 0 * 2^5 + 1 * 2^4 + 0 * 2^3 + 0 * 2^2 + 0 * 2^1 + 1 * 2^0 = 10010001_2$$

Le cifre possono essere solo *0* e *1*

Rappresentazione dei numeri negativi

Per rappresentare i numeri negativi è possibile ricorrere a due tipi di rappresentazione:

- ▣ *complemento a 1*: un bit viene utilizzato per indicare il segno
- ▣ *complemento a 2*: rappresentazione più complicata; con *k* bit si rappresentano i numeri da -2^{k-1} a $2^{k-1}-1$

Informazione analogica e digitale

Informazione vista → classificatoria

Meta-informazione: relazione d'ordine –
metrica

Sistema: meta informazione + informazione

Informazione analogica e digitale

Rappresentazione meta-informazione:

1. rappresentazione **esplicita** →
informazione **analogica**
2. **implicita** nella codifica → informazione
digitale

Rappresentazione esplicita → orologio
analogico

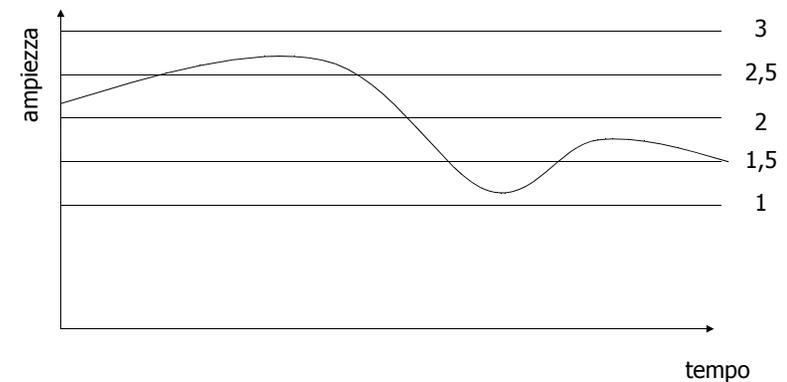
Rappresentazione implicita nella codifica →
orologio digitale

L'informazione nei sistemi analogici

L'informazione analogica è un'informazione
con valori nel **continuo**

Sfumature dei colori, spettro dei suoni: sono
rappresentati da funzioni continue

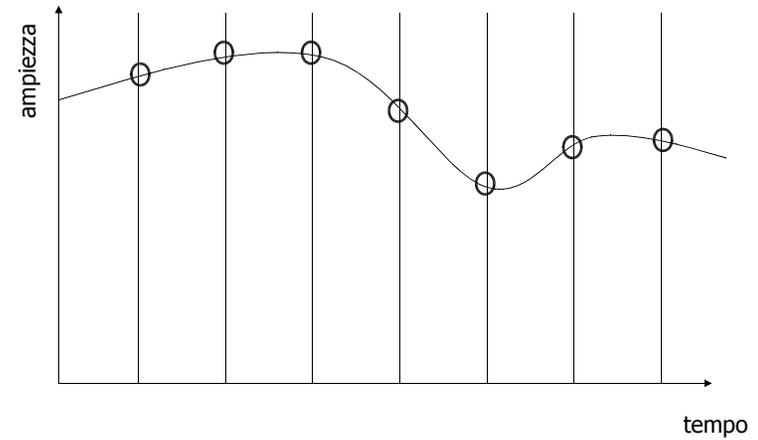
Segnale analogico



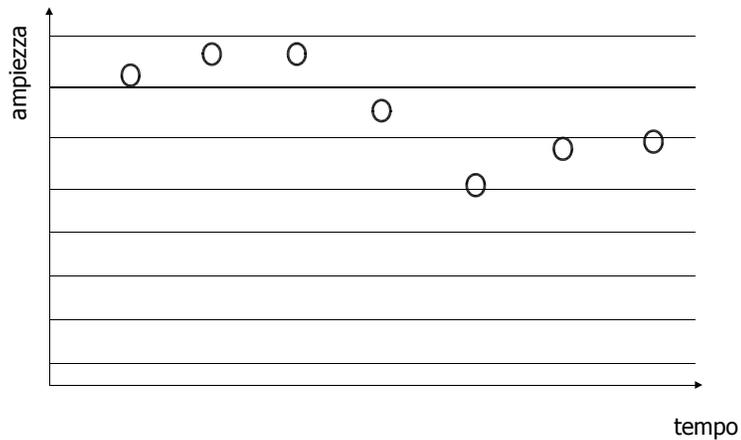
L'informazione nei sistemi analogici

- L'informazione analogica → **discretizzata**: deve assumere un insieme *finito* di valori
- Per trasformazione in informazione discreta:
 - **Campionamento**: si scelgono i valori assunti dal segnale in certi istanti
 - **Quantizzazione**: si divide l'ampiezza del segnale in intervalli discreti

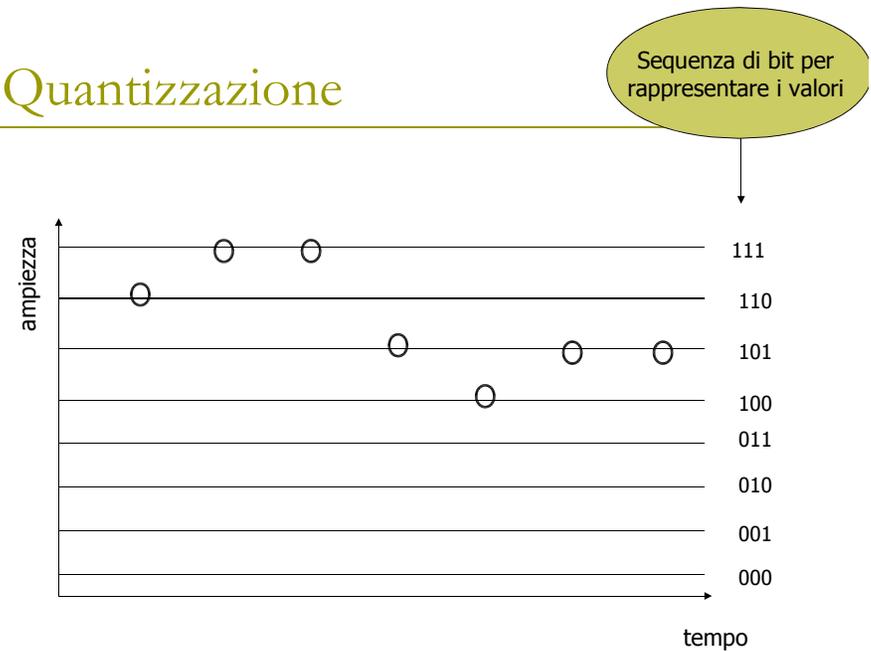
Campionamento



Quantizzazione



Quantizzazione



L'informazione nei sistemi analogici e digitali

- ▣ Aumentando la frequenza di campionamento
- ▣ Aumentando gli intervalli di quantizzazione



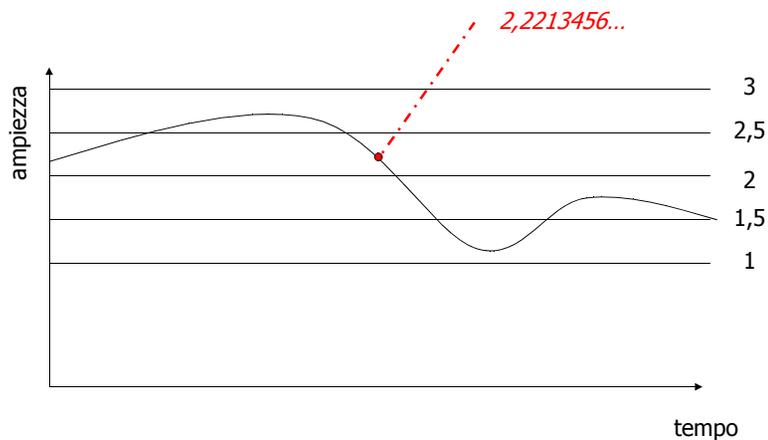
migliora la precisione con cui si rappresenta l'informazione

La lunghezza della sequenza utilizzata deve essere la più **compatta** possibile

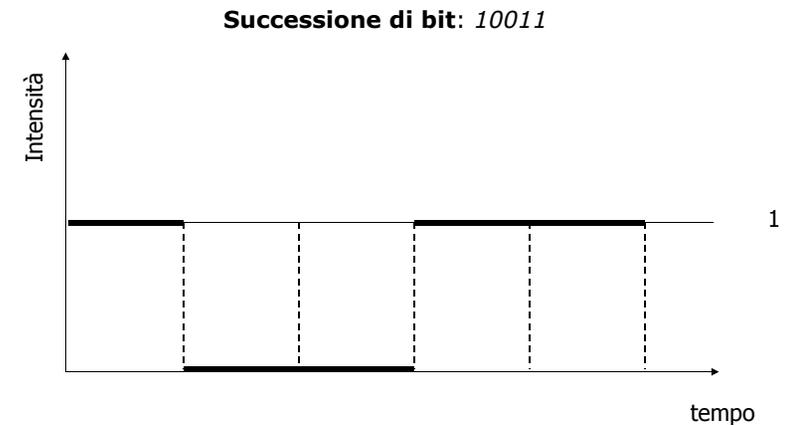
L'informazione nei sistemi analogici e digitali

- ▣ In presenza di rumore, un segnale analogico viene *alterato*: variazione dell'onda → difficile riconoscere l'errore
- ▣ Vantaggi dell'informazione digitale rispetto all'informazione analogica → **tolleranza al rumore, riconoscimento del rumore**

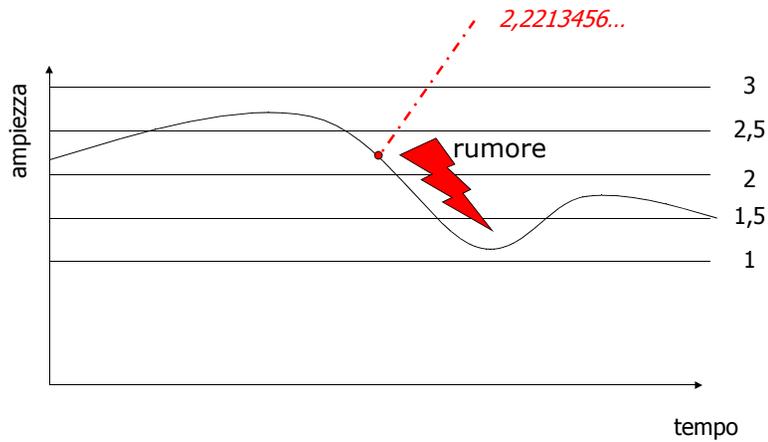
Segnale analogico



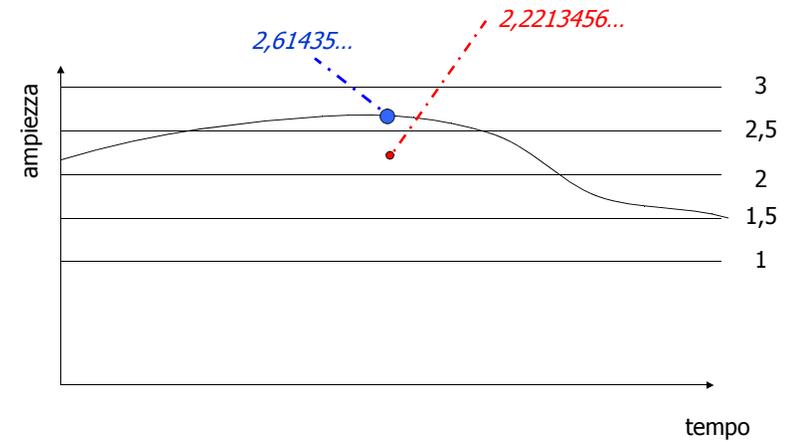
Segnale digitale



Segnale analogico



Segnale analogico



Nei sistemi digitali

Segnale digitale → valori 0 e 1

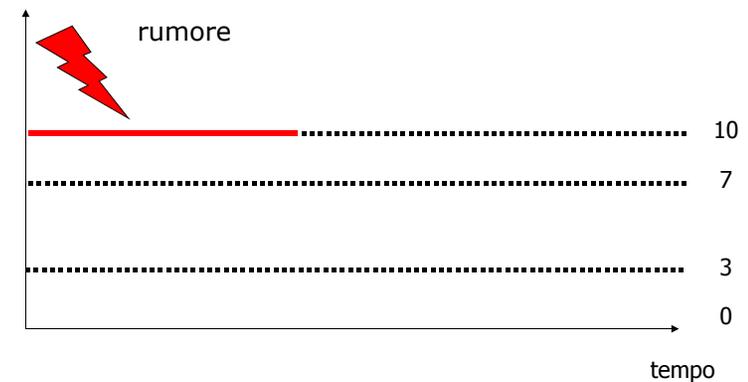
Corrente elettrica → un valore che sta in un intervallo (ad es. da 0 a 10)

Se tale valore è

- ▣ in un intervallo vicino a 10 (es. $[7; 10]$) → tensione elettrica presente → **bit vale 1**
- ▣ in un intervallo vicino a 0 (es. $[0; 3]$) → tensione elettrica assente → **bit vale 0**

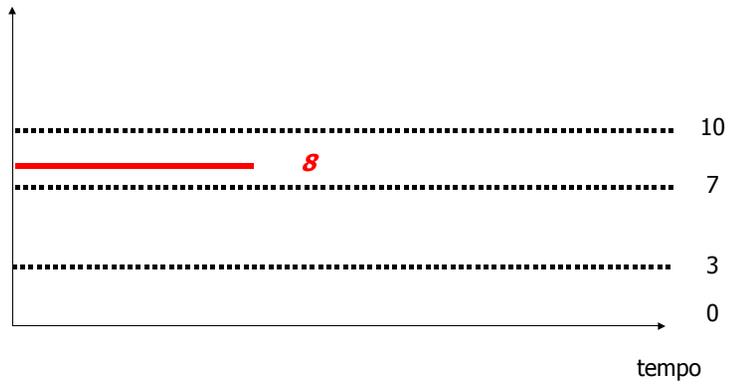
Negli altri casi è interpretato come *errore* → **riconoscimento dell'errore**

Segnale digitale

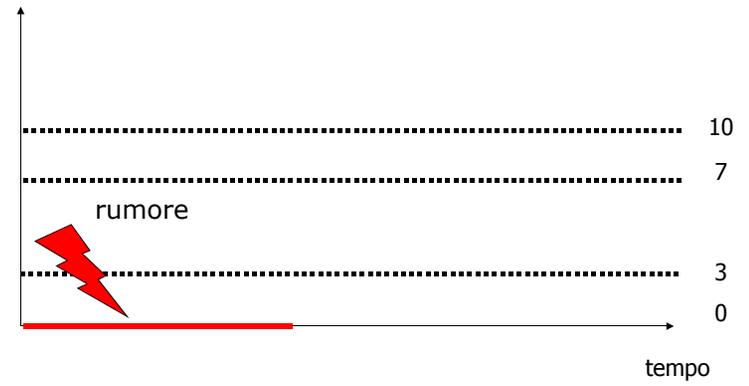


Segnale digitale

Bit di valore 1



Segnale digitale

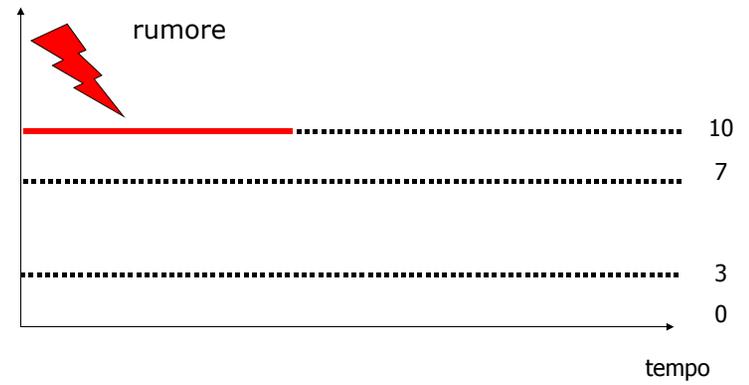


Segnale digitale

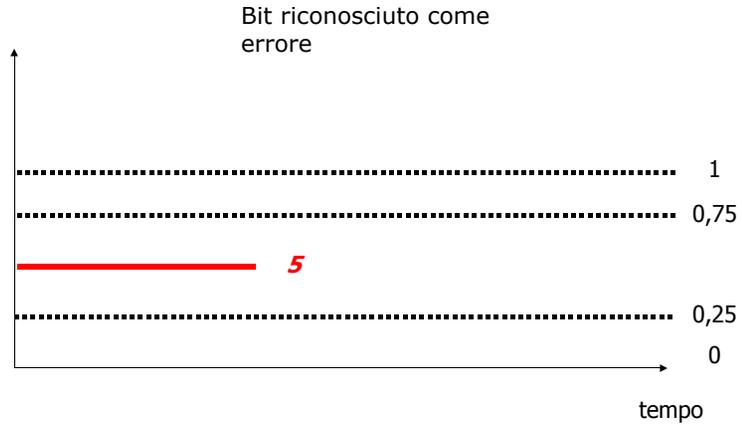
Bit di valore 0



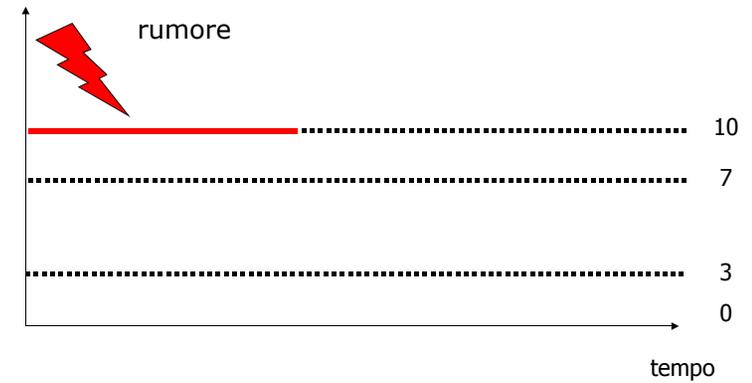
Segnale digitale



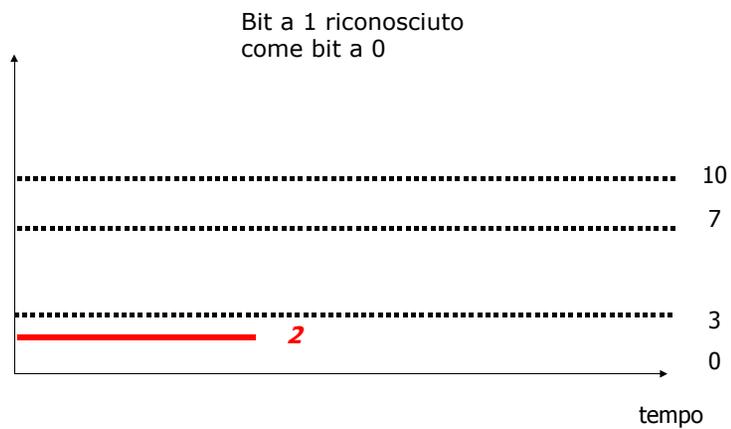
Segnale digitale



Segnale digitale



Segnale digitale



Nei sistemi digitali

Se il rumore non modifica eccessivamente il segnale → **correzione dell'errore**

Posso effettuare un **riconoscimento degli errori**

Errore solo se il rumore è molto forte e modifica in modo consistente il segnale

L'informazione nei sistemi digitali

In alcuni casi si utilizzano codifiche intermedie (codifica esadecimale, i numeri sono visti in base 16)

Rappresentazione dei numeri

Oltre alla rappresentazione binaria:

- ▣ rappresentazione ottale (cifre: $0,1,2,3,4,5,6,7$)
- ▣ rappresentazione esadecimale (cifre: $0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F$)

Rappresentazione dei numeri

Rappresentazione esadecimale utile per la conversione dalla rappresentazione binaria: ogni cifra esadecimale rappresenta 4 bit binari

N. binario: 1000 0011 1010

N. esadecimale: 8 3 A