Rappresentazione dell’informazione

"... mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true."

Bertrand Russell

Rappresentare l'informazione

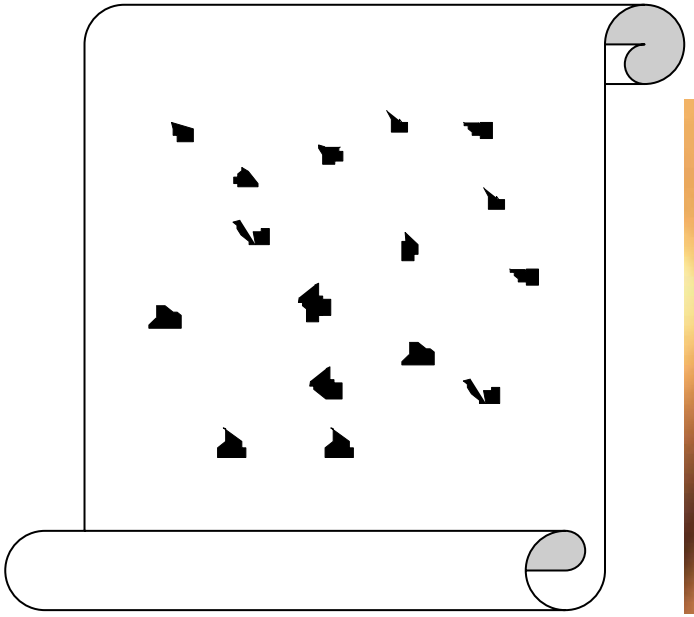
- Per elaborare l'informazione è necessario saperla rappresentare in una forma comprensibile per l'esecutore

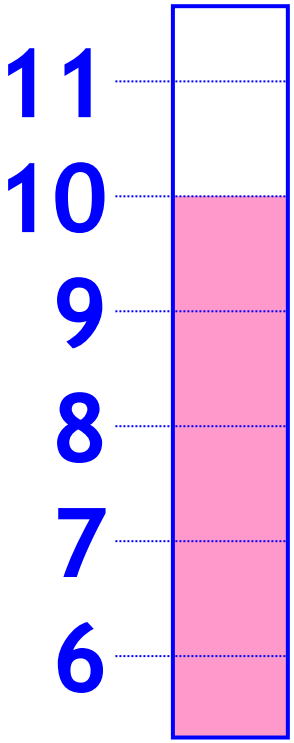
- Bisogna stabilire un ***codice*** che associa a ogni entità di informazione che si desidera rappresentare una configurazione del ***supporto*** su cui l'informazione è trasmessa

Codice e supporto

- Codice
  - E' una regola che a ogni entità di informazione associa una configurazione del supporto su cui l'informazione viene trasmessa
- Supporto
  - Mezzo che può assumere almeno due configurazione diverse

Codice e supporto: esempi





Perché usare un codice?

- Permette l'elaborazione e la memorizzazione di entità non gestibili nella loro forma originale
  - i sistemi di numerazione sono codici
- Permette l'interpretazione dei simboli
  - un dizionario di lingua italiana è un codice
- Aggiunge proprietà ad un sistema di simboli
  - comprimere la lunghezza delle stringhe
  - aumentare l'affidabilità di trasmissione

Che codice usa il calcolatore?

- Il calcolatore è in grado di distinguere solo due tipi di informazione:
  - sta passando corrente / non sta passando corrente

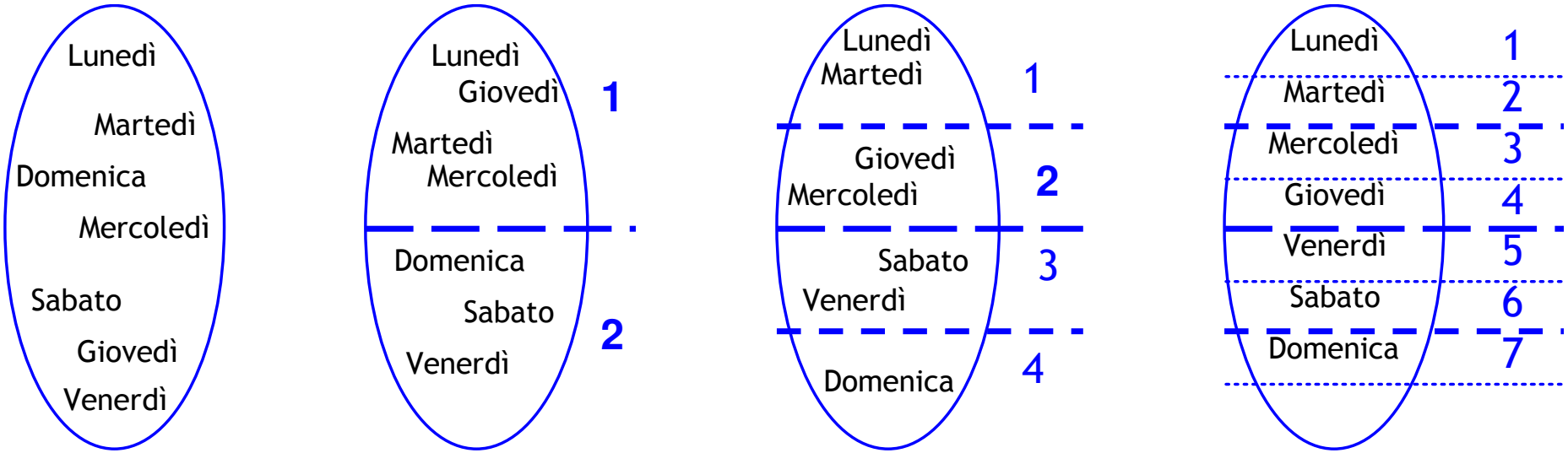
- Il calcolatore usa dispositivi che possono trovarsi in due soli stati
  - acceso / spento

- Quindi gli bastano due numeri (0,1) per rappresentare le uniche due entità di informazione che conosce
  - Il calcolatore usa un ***codice numerico binario (booleano)***

Codifica numerica

- Ad ogni entità di informazione può essere associato un numero (codice numerico)
  - Molte informazioni sono quantitative, e quindi rappresentabili in forma numerica
  - Le informazioni qualitative possono essere comunque associate a numeri tramite un opportuno codice

Codifica numerica: esempio



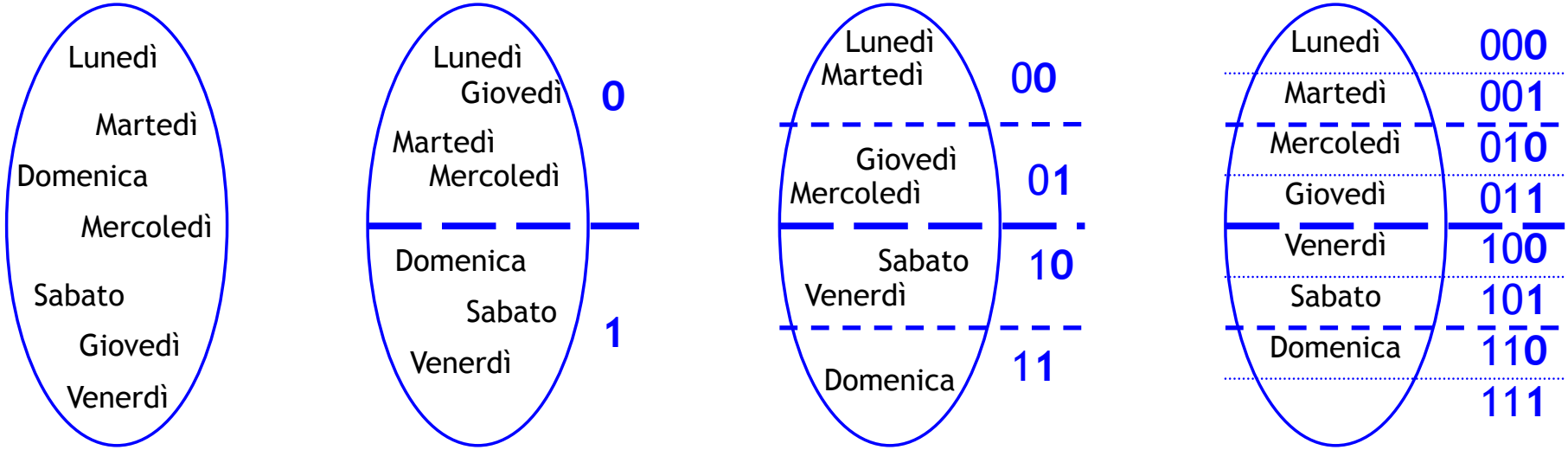
La codifica binaria

- Si usano solo i simboli 0 e 1 (binary digit = **bit**) concatenati in sequenze chiamate stringhe binarie

- Quanti bit mi servono per contare N oggetti:
  - $N \leq 2^k \Rightarrow k \geq \log_2 N \Rightarrow k = \lceil \log_2 N \rceil$ (intero superiore)

- Quanti oggetti posso contare con k bit:
  - 1 bit $\Rightarrow$ 2 stati (0, 1) $\Rightarrow$ 2 oggetti (e.g. Vero/Falso)
  - 2 bit $\Rightarrow$ 4 stati (00, 01, 10, 11) $\Rightarrow$ 4 oggetti
  - 3 bit $\Rightarrow$ 8 stati (000, 001, …, 111) $\Rightarrow$ 8 oggetti
  - …
  - k bit $\Rightarrow$ 2^k stati $\Rightarrow$ 2^k oggetti

Codifica binaria: esempio



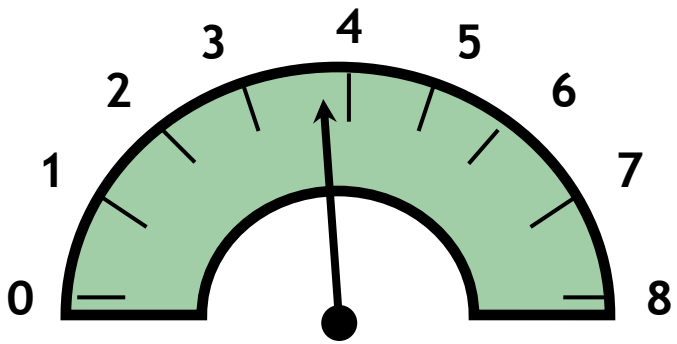
bit, Byte, KiloByte, MegaByte, …

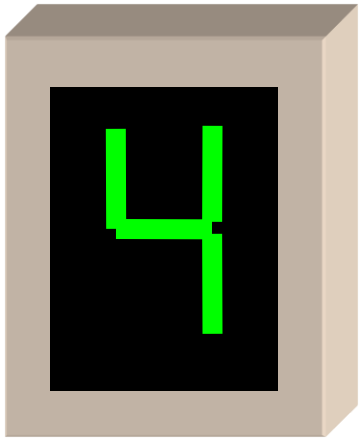
- bit = solo due stati, “0” oppure “1”
- Byte = 8 bit, quindi 2^8 = 256 stati
- KiloByte [KB] = 2^{10} Byte = 1024 Byte ~ 10^3 Byte
- MegaByte [MB] = 2^{20} Byte = 1'048'576 Byte ~ 10^6 Byte
- GigaByte [GB] = 2^{30} Byte ~ 10^9 Byte
- TeraByte [TB] = 2^{40} Byte ~ 10^{12} Byte
- PetaByte [PB] = 2^{50} Byte ~ 10^{15} Byte
- ExaByte [EB] = 2^{60} Byte ~ 10^{18} Byte

Riassumendo…

- Il calcolatore distingue solo 0 e 1
- Concatenando 0 e 1 posso rappresentare qualunque numero decimale
- Ad ogni numero posso associare un'entità di informazione diversa
- Qualunque informazione deve essere ***digitalizzata*** per potere essere rappresentata nel calcolatore

Analogico e digitale





Meta-informazione esplicita nel supporto:
il supporto ha una struttura corrispondente a quella presente tra entità di informazione.

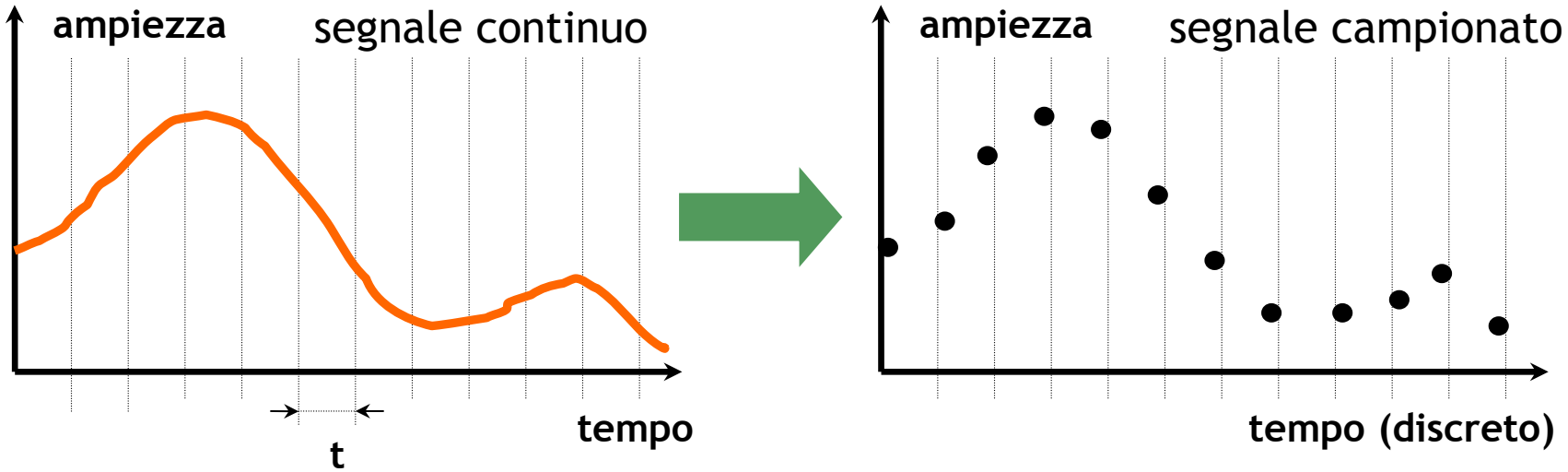
Meta-informazione implicita nella codifica:
al supporto si richiede solo di avere configurazioni molteplici e distinguibili.

Campionamento e quantizzazione

- Gli elaboratori elettronici hanno natura discreta, ovvero ogni grandezza in gioco può essere rappresentata soltanto da un numero finito di elementi

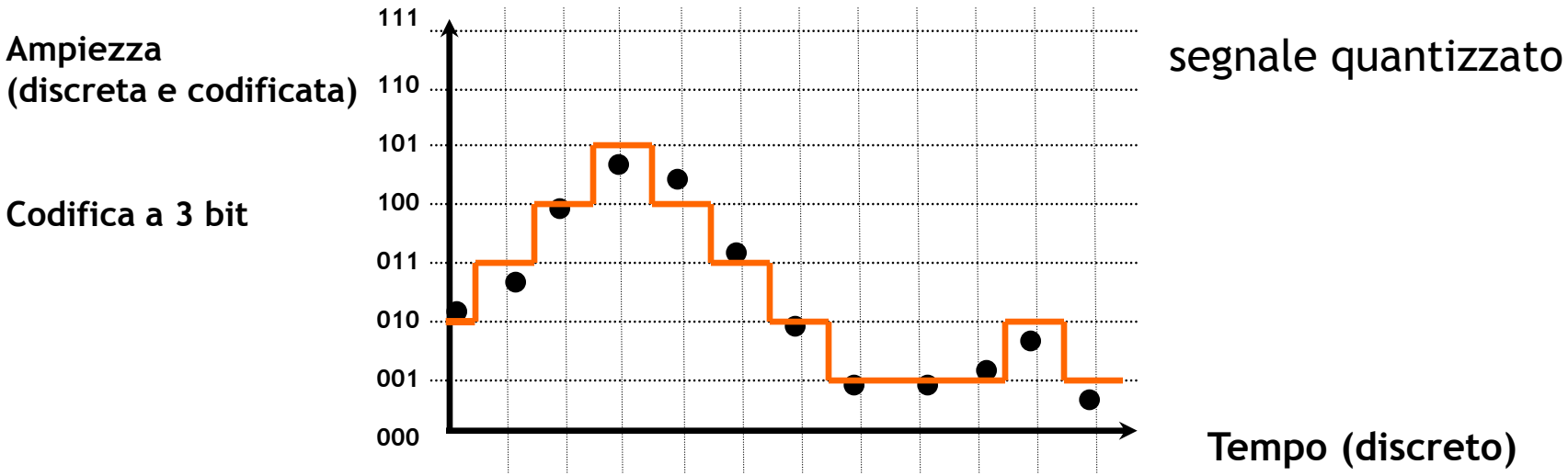
- Per essere elaborati da un calcolatore, segnali intrinsecamente continui quali suoni, immagini, video ecc., devono essere discretizzati (digitalizzati) attraverso operazioni di ***campionamento*** e ***quantizzazione***

Campionamento



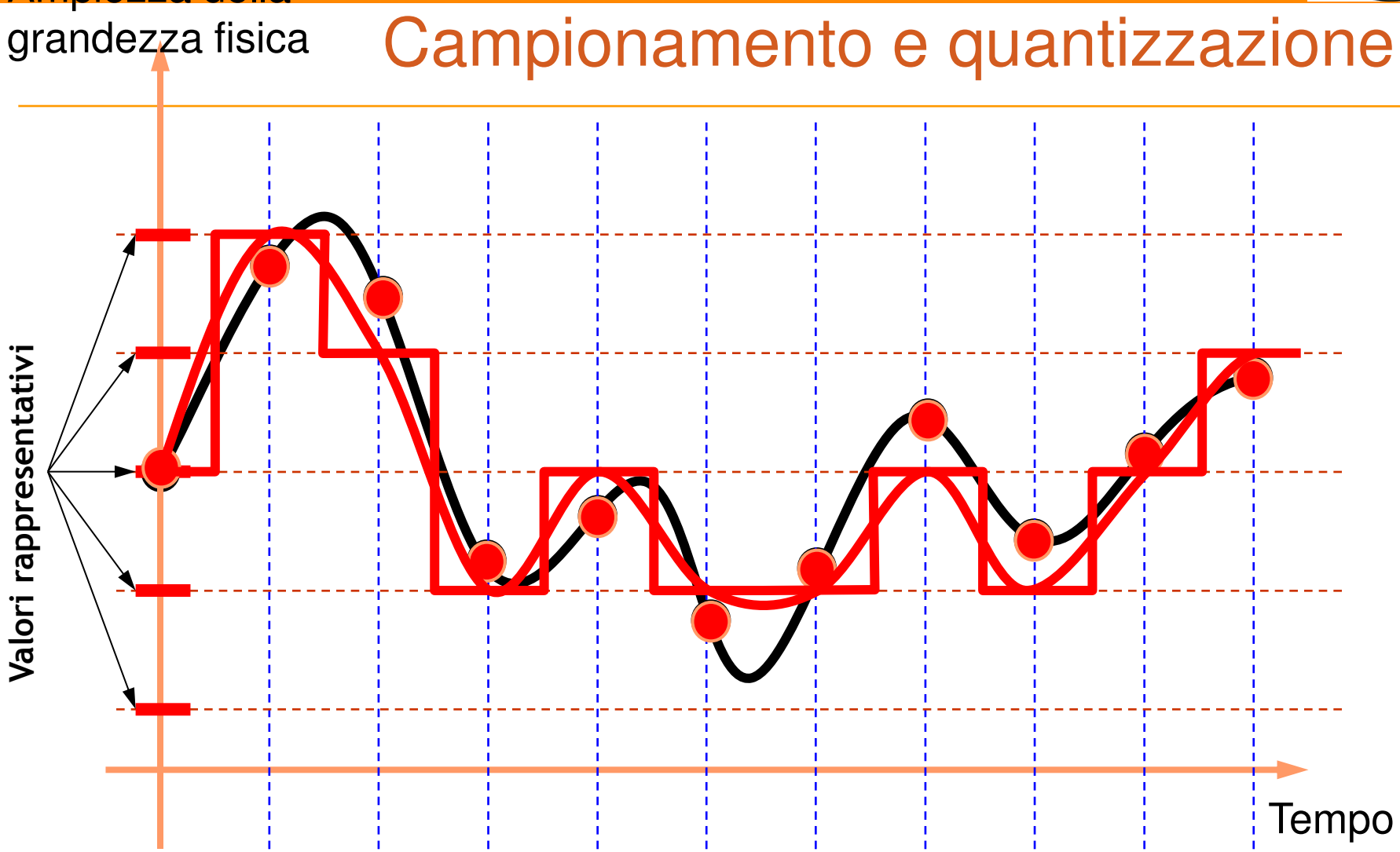
- Il segnale continuo viene campionato ad intervalli di tempo regolari t (t = intervallo di campionamento).
- Il segnale risultante è un insieme finito di punti equidistanti nel tempo. Tuttavia le ampiezze devono essere ancora approssimate ad intervalli discreti, ovvero quantizzate.
- Si noti che campionamento e quantizzazione comportano una perdita di informazione.

Quantizzazione



- La quantizzazione suddivide l'ampiezza in *n* intervalli uguali che vengono poi codificati in binario. Ogni valore di ampiezza del segnale campionato viene approssimato al più vicino valore discreto di ampiezza.
- Più valori (e quindi più bit) si utilizzano per suddividere le ampiezze, più il segnale risultante sarà preciso.

Campionamento e quantizzazione

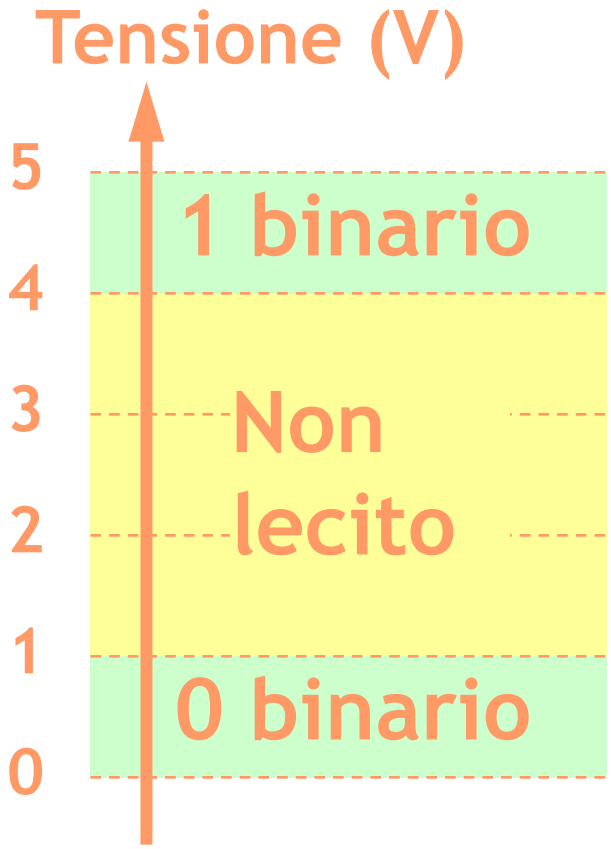


Le immagini digitali

Le immagini digitali non hanno una struttura continua ma sono costituite da un numero finito di componenti monocromatiche (pixel) prodotte dal campionamento dell'immagine reale. I pixel assumono un numero finito di tonalità definite dalla quantizzazione dell'immagine campionata.

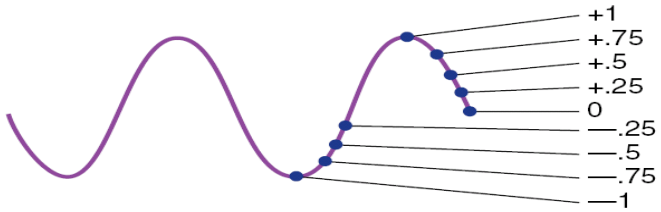
Il successo del digitale

- Rumore: effetto dell'ambiente sul supporto.
- Quanto un supporto è "immune" al rumore?
  - Codifica analogica: ogni configurazione è lecita dal punto di vista informazionale e quindi risulta impossibile distinguere il rumore dal segnale.
  - Codifica digitale: un valore binario è associato a un insieme di configurazioni valide quindi si può
    - riconoscere il rumore che porta in configurazioni non lecite
    - trascurare il rumore che non fa uscire il segnale dall'insieme associato alla stessa configurazione



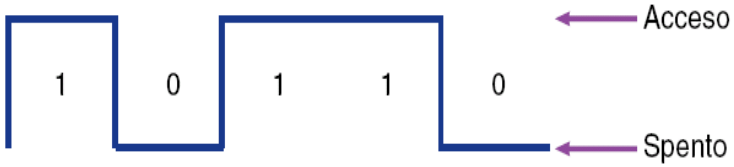
Riassumendo…

- Segnale analogico
  - come un onda che trasporta informazioni, massimi, minimi e tutti i valori intermedi
  - i segnali analogici sono molto sensibili alle interferenze



Segnali analogici

- Segnale digitale
  - assume solo due stati: acceso/spento, sì/no, vero/falso
  - il segnale digitale è più facile da distinguere, quindi risente meno delle interferenze



Segnali digitali

I sistemi di numerazione

- Le informazioni numeriche possono essere elaborate attraverso l'applicazione di operazioni.

- Un sistema di numerazione è una struttura matematica che permette di rappresentare i numeri attraverso dei simboli.

Numerazione non posizionale

- Il significato dei simboli non dipende dalla loro posizione
- E' stabilito in base ad una legge additiva dei valori dei singoli simboli (se posti in ordine crescente)
- Esempio:
  - il sistema di numerazione romano

I = 1

V = 5

X = 10

L = 50

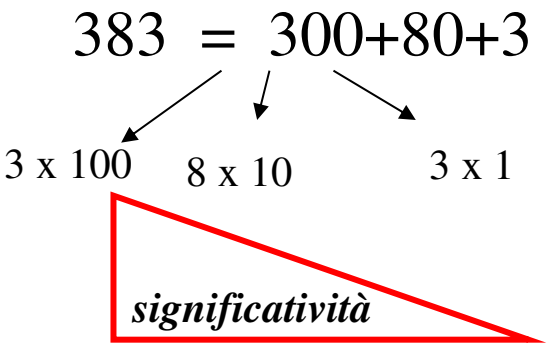
...

LXIV = 50+10 -1+5
= 64

50 10 -1

Numerazione posizionale

- Ai diversi simboli dell'alfabeto (cifre), viene associato un valore crescente in modo lineare da destra verso sinistra
- il significato di un simbolo (il suo valore) dipende ordinatamente dalla sua posizione nella stringa
- Esempio:
  - il sistema di numerazione decimale arabo: 10 simboli (0, 1, 2, ...9)



Sistemi di numerazione non posizionali

- Un simbolo rappresenta un numero.
- Il numero rappresentato da una stringa di simboli si ottiene attraverso regole operazionali applicate ai simboli della stringa
  - Esempio (numeri romani): LXXIV rappresenta 50+10+10-1+5 = 74
- Difficile effettuare operazioni
- Rappresentazione non compatta

Sistemi di numerazione posizionali

- Dato un alfabeto ordinato di b simboli distinti (c_1, c_2, …, c_b) che rappresentano rispettivamente i naturali 0, 1, 2, …, b-1 si rappresenta ogni altro numero x maggiore di b-1 mediante una stringa di simboli dell'alfabeto

- b è la base del sistema di numerazione
  - numero di simboli dell'alfabeto richiesti per rappresentare la serie infinita dei numeri

Valore della posizione

- La posizione di un simbolo all'interno di un numero indica il valore che esso esprime, o più precisamente l'esponente che bisogna dare alla base per ottenere il valore corretto.
- Il valore (o la quantità) di 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9 dipende dalla posizione che ciascuno di essi assume all'interno del numero:
- la prima cifra a destra rappresenta le unità (il coefficiente di 10^0), la seconda le decine (10^1) , la terza le centinaia (10^2) , e così via.

Il numero **3.098.323** è una rappresentazione abbreviata di

$$(3 \times 10^6) + (0 \times 10^5) + (9 \times 10^4) + (8 \times 10^3) + (3 \times 10^2) + (2 \times 10^1) + (3 \times 10^0 = 3 \times 1)$$

Il primo 3 (leggendo da destra a sinistra) rappresenta 3 unità; il secondo 3, sta per 300 unità, o 3 centinaia; infine il terzo 3, per 3 milioni di unità.

Il sistema di numerazione decimale

- È il sistema più conosciuto dall'uomo.
- La base b è pari a 10.
- I simboli utilizzati sono 0,1,2,…,9 dal significato ovvio
  - Esempio: la stringa 2349 rappresenta il numero $2*10^3+3*10^2+4*10^1+9*10^0$.
- I numeri decimali sono facilmente intelligibili.

Il sistema di numerazione binario

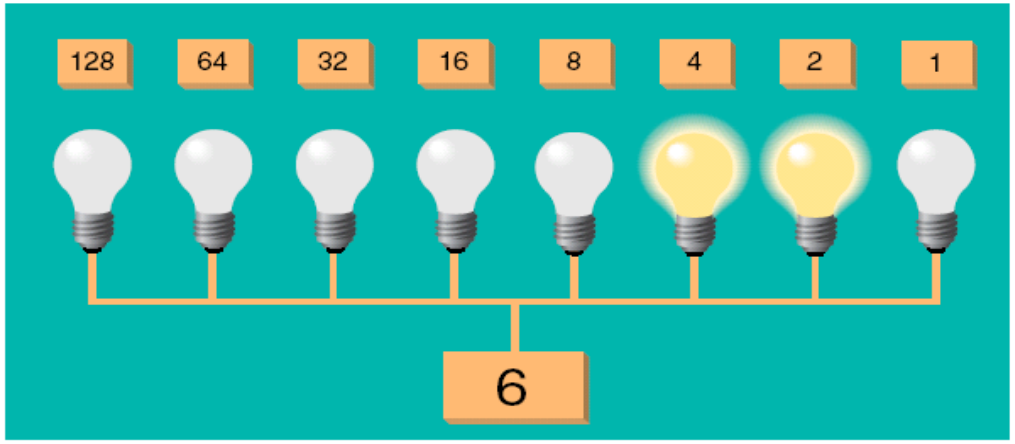
- È il sistema maggiormente utilizzato dai sistemi di elaborazione.
- La base b è pari a 2.
- I simboli utilizzati sono 0 e 1, rappresentanti lo zero e l'uno.
  - Esempio: la stringa binaria 10010 rappresenta il numero $1*2^4+0*2^3+0*2^2+1*2^1+0*2^0$ (=18 in decimale).
- È scarsamente leggibile, specie quando le stringhe sono molto lunghe.

Conversioni

- Da binario a decimale
  - Basta scrivere il numero secondo la notazione posizionale utilizzando già il sistema decimale
- Da decimale a binario
  - Basta dividere ripetutamente il numero decimale per 2, tenere il resto della divisione, dividere il quoziente per 2, tenere il resto della divisone, etc… fino ad arrivare ad avere 0 come quoziente

Da binario a decimale: esempio

- Il numero binario 00000110 corrisponde al valore decimale 6
- $0x2^7+0x2^6+0x2^5+0x2^4+0x2^3+1x2^2+1x2^1+0x2^0 =$
  $= 0+0+0+0+0+4+2+0 = 6$



I bit che danno come somma 6

Da decimale a binario: esempio

(cifra binaria meno significativa)

18 : 2 = 9	resto 0
9 : 2 = 4	resto 1
4 : 2 = 2	resto 0
2 : 2 = 1	resto 0
1 : 2 = 0	resto 1

(cifra binaria più significativa)

10010

137 : 2 =	68	resto 1
68 : 2 =	34	resto 0
34 : 2 =	17	resto 0
17 : 2 =	8	resto 1
8 : 2 =	4	resto 0
4 : 2 =	2	resto 0
2 : 2 =	1	resto 0
1 : 2 =	0	resto 1

10001001

Un esempio di messaggio digitale

- La cavalcata di Paul Revere
  - "Una lanterna se vengono da terra, e due se vengono dal mare"

- Segnale digitale?

- Quanti stati?
  - Lanterne entrambe spente (00)
  - Una lanterna accesa (01 e 10)
  - Entrambe le lanterne accese (11)

”The fundamental cause of the trouble is that in the modern world the stupid are cocksure while the intelligent are full of doubt. ”

Bertrand Russel