

Rappresentazione digitale delle informazioni

- Il sistema binario: bit e Byte
- Codifica del testo
- Il Byte come U.d.M. dell'informazione
 - Multipli del Byte
 - Ordini di grandezza
- Codifica delle immagini
- Codifica dei suoni
- Codifica delle immagini in movimento (filmati)

• Il sistema binario: bit e Byte

- Le informazioni che possono essere memorizzate sono espresse in forma binaria poiché i dispositivi sono in grado di rappresentare due stati: 0 e 1 (BIT=Binary digiT).
 - Un bit è una entità che può assumere due valori, che si possono indicare con 0 e 1.
 - In generale, n bit consentono di specificare uno stato scelto tra 2^n stati.
- Per rappresentare una lettera (maiuscola e minuscola), un numero e un qualsiasi simbolo sono necessari più di un bit, per esempio un BYTE=8 bit.
- Quando una qualsiasi informazione è rappresentata da bit, si dice che è espressa in formato DIGITALE.

Codifica del testo

- Lunghezza fissa (Ascii, EBCDIC, etc.): ad ogni simbolo corrisponde un byte: si possono rappresentare fino a 2^8 caratteri (256).

ASCII = American Standard Code for International Interchange

- Standard Unicode usa 2 byte per la codifica di ogni carattere potendone rappresentare così 2^{16} (65536), quindi tutti i caratteri delle diverse lingue.

Esempi di codifiche Ascii & Unicode

ASCII/8859-1 Text

A	0100 0001
S	0101 0011
C	0100 0011
I	0100 1001
I	0100 1001
/	0010 1111
8	0011 1000
8	0011 1000
5	0011 0101
9	0011 1001
-	0010 1101
l	0011 0001
	0010 0000
t	0111 0100
e	0110 0101
x	0111 1000
t	0111 0100

Unicode Text

A	0000 0000 0100 0001
S	0000 0000 0101 0011
C	0000 0000 0100 0011
I	0000 0000 0100 1001
I	0000 0000 0100 1001
	0000 0000 0010 0000
天	0101 1001 0010 1001
地	0101 0111 0011 0000
	0000 0000 0010 0000
س	0000 0110 0011 0011
س	0000 0110 0100 0100
س	0000 0110 0011 0111
س	0000 0110 0100 0101
	0000 0000 0010 0000
α	0000 0011 1011 0001
≠	0010 0010 0111 0000
γ	0000 0011 1011 0011

Multipli del Byte

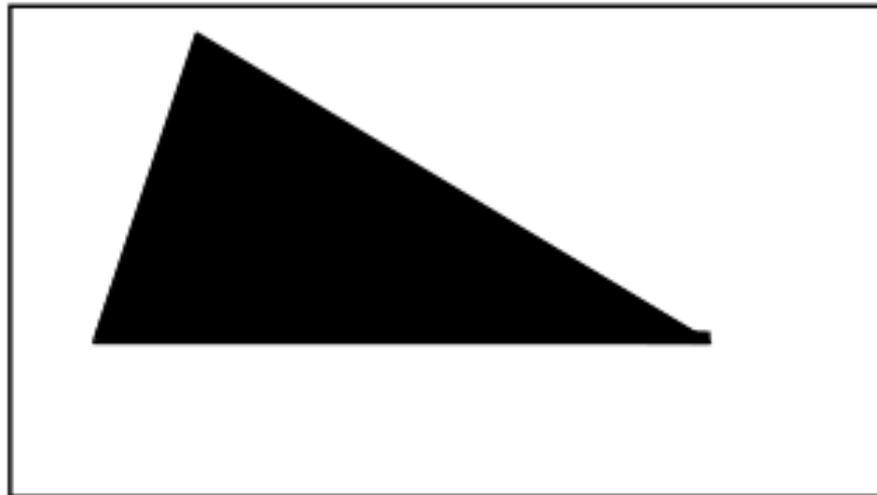
- L'unità di misura della quantità di informazione è il Byte.
- I multipli del byte rappresentano potenze di 2 e sono:
 - KiloByte: 1 KB = 1024 Byte = 2^{10} Byte (K=Kilo $\approx 10^3$)
 - MegaByte: 1 MB = 1024 KB = 2^{20} Byte (M=Mega $\approx 10^6$)
 - GigaByte: 1 GB = 1024 MB = 2^{30} Byte (G=Giga $\approx 10^9$)
 - TeraByte: 1 TB = 1024 GB = 2^{40} Byte (T=Tera $\approx 10^{12}$)

Ordini di grandezza

Una frase:	qualche decina di Byte
Una pagina:	alcuni KB
Il testo di una tesi:	diverse decine di KB
Il testo di un libro:	qualche MB
Una tesi con immagini:	alcuni MB
Un file :	da pochi Byte a decine o centinaia di MB
Una directory (o cartella , è un contenitore di file):	può contenere innumerevoli file

Codifica delle immagini (raster)

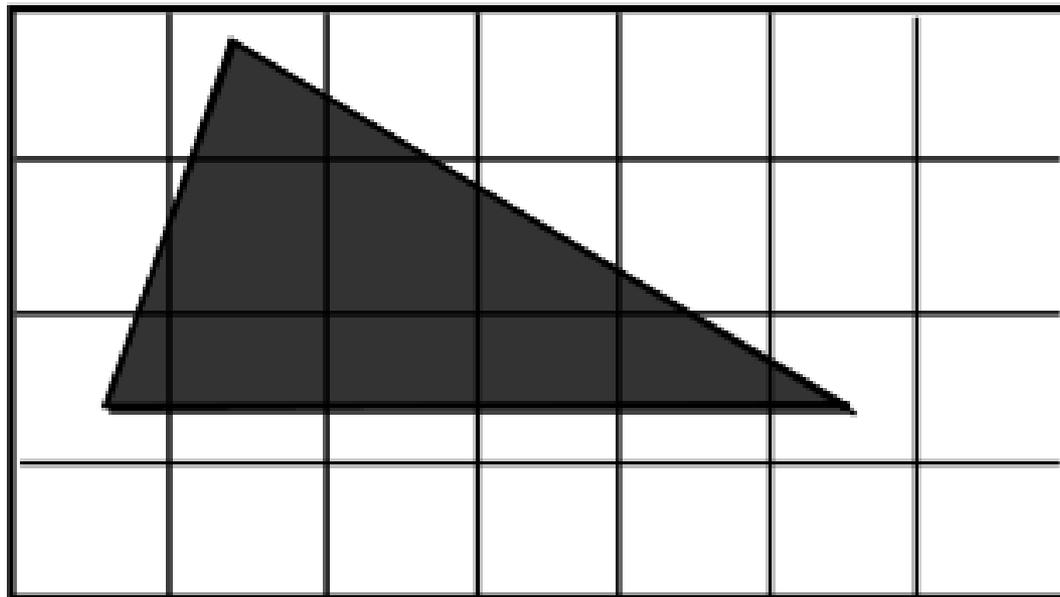
Consideriamo inizialmente per semplicità una figura contenente due soli colori: bianco e nero.



Come si rappresenta digitalmente?

Codifica delle immagini (raster)

- Si costruisce una griglia costituita da tanti quadrati:



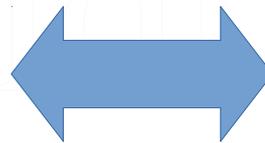
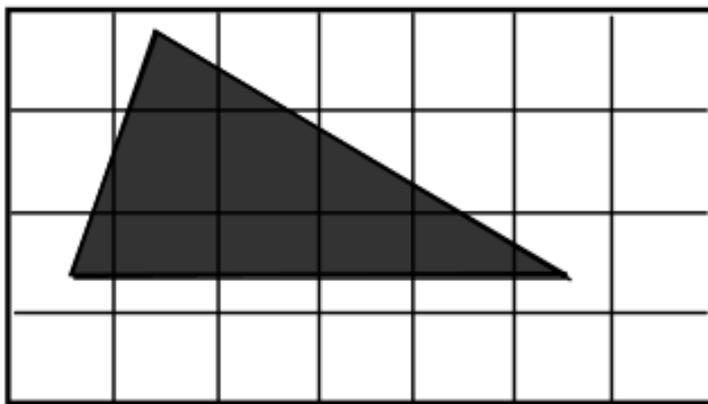
Codifica delle immagini (raster)

- Si assegna convenzionalmente un valore di un bit (per esempio, 1) al colore bianco e l'altro al colore nero (es. 0):

Bianco \longleftrightarrow 1

Nero \longleftrightarrow 0

- Si associa il colore di ciascun quadratino al valore del bit corrispondente, ottenendo una mappa di bit (bitmap)



0	1	0	0	0	0	0
0	1	1	0	0	0	0
0	1	1	1	1	0	0
0	0	0	0	0	0	0

(ai quadratini che non sono completamente bianchi o completamente neri si associa il bit corrispondente al colore prevalente)

Codifica delle immagini (raster)

- Si definisce una convenzione attraverso la quale la bitmap si può rappresentare come una serie di bit consecutivi (per esempio, riscrivendo le righe dall'alto al basso):

0	1	0	0	0	0	0
0	1	1	0	0	0	0
0	1	1	1	1	0	0
0	0	0	0	0	0	0

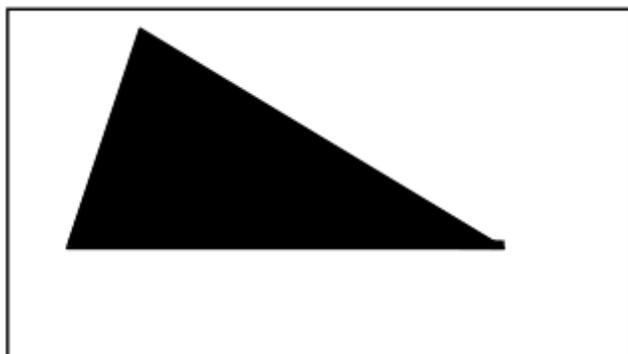


0100000011000001111000000000

Codifica delle immagini (raster)

- Con questa convenzione la rappresentazione della figura sarà data dalla stringa binaria

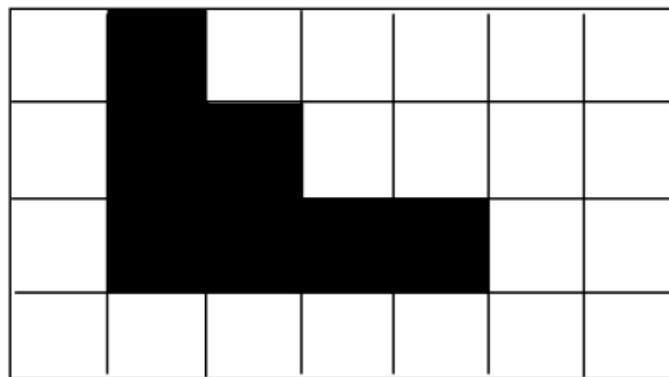
0100000011000001111000000000:



0100000011000001111000000000

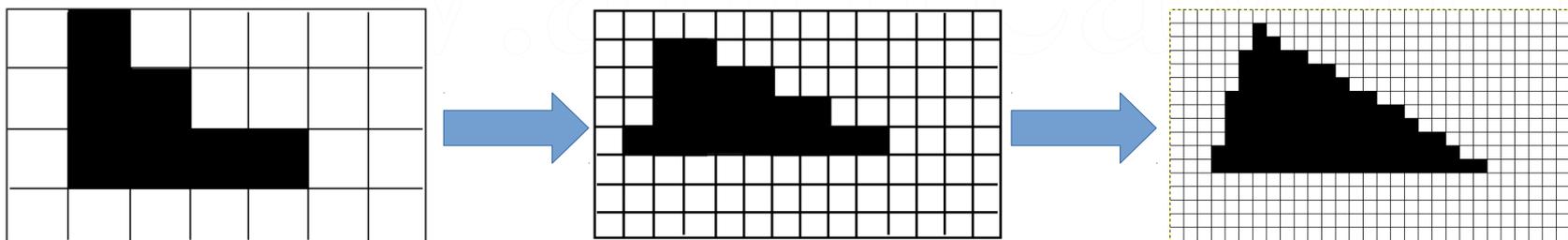
Codifica delle immagini (raster)

- Non sempre il contorno della figura coincide con le linee della griglia. Quella che si ottiene nella codifica è un'approssimazione della figura originaria. Se riconvertiamo la stringa **0100000011000001111000000000** in immagine otteniamo:



Codifica delle immagini (raster)

- Per migliorare la precisione della rappresentazione, e di conseguenza poter rappresentare una immagine di maggiore qualità, basta ridurre la dimensione dei quadratini, e quindi aumentarne il numero:



Codifica delle immagini (raster)

Per rappresentare un'immagine con più di 2 colori si associa ad ogni pixel una diversa combinazione di più bit: per esempio con 2 bit si possono rappresentare 4 colori

00	←→	nero
01	←→	bianco
10	←→	verde
11	←→	rosso

Codifica delle immagini (raster)

In generale, con n bit si ottengono 2^n combinazioni differenti, quindi associando n bit ad ogni pixel si possono rappresentare 2^n diversi colori:

1 bit per pixel	←→	2 colori
2 bit per pixel	←→	4 colori
3 bit per pixel	←→	8 colori
4 bit per pixel	←→	16 colori
8 bit per pixel	←→	256 colori
16 bit per pixel	←→	65536 colori
24, 32 bit per pixel	←→	Milioni di colori

Il numero di bit per ogni pixel è chiamato profondità di colore

Codifica delle immagini (raster)

Esercizio

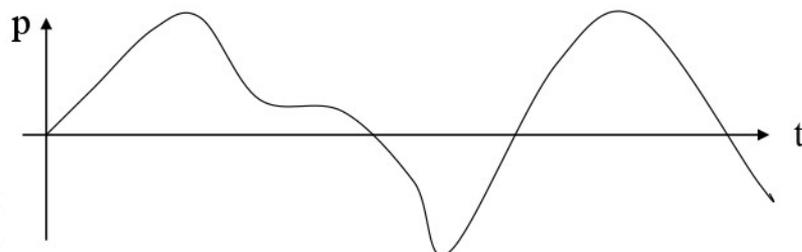
Calcolare l'occupazione di memoria di un'immagine a 65k colori, di dimensioni 4000x3000 pixel

- Calcoliamo il numero di pixel: $4000 \times 3000 = 12M$
- 65k colori \longrightarrow ad ogni pixel sono associati 16 bit, quindi 2 Byte
- Il numero di Byte che rappresentano l'immagine è quindi:

$$12 M \text{ pixel} \times \frac{2 \text{ Byte}}{\text{pixel}} = 24 MB$$

Codifica dei suoni

Un suono è rappresentabile come un'onda, che descrive la variazione della pressione dell'aria nel tempo:



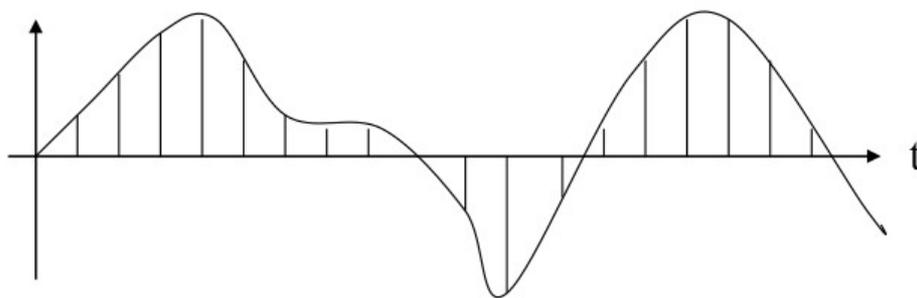
Per rappresentare digitalmente un suono sono necessarie:

- Discretizzazione nel tempo (Campionamento: scelta degli istanti in cui considerare il valore del segnale)
- Discretizzazione dell'ampiezza (Quantizzazione: codifica di ciascun campione con un certo numero di bit)

Codifica dei suoni

Campionamento

Si effettua il campionamento dell'onda (cioè se ne misura il valore a intervalli di tempo regolari)



Il tempo T che trascorre tra due campionamenti successivi è detto periodo di campionamento, e si misura in secondi

Il numero di campioni raccolti per ogni secondo definisce la frequenza di campionamento f , che si misura in Hertz (Hz)

$$f=1/T$$

Codifica dei suoni

Campionamento

- Quanto più frequentemente viene campionato il valore di intensità dell'onda, tanto più precisa sarà la sua rappresentazione: un campionamento più fitto (ovvero con una frequenza di campionamento maggiore) consente di rappresentare i segnali analogici con maggiore fedeltà
- Per segnali audio di tipo vocale (ad es. telefono), la frequenza di campionamento è tipicamente di 8 kHz
- Per segnali audio musicali (ad es. CD audio), la frequenza di campionamento è tipicamente di 44.1 kHz

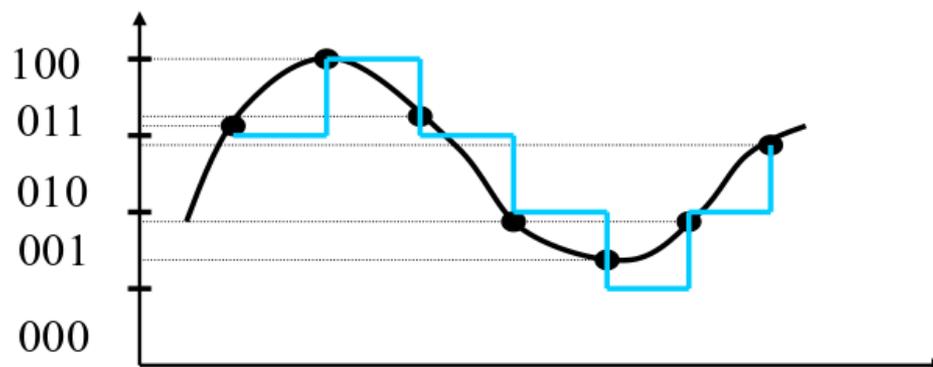
Codifica dei suoni

Quantizzazione

Occorre codificare in formato digitale ciascuno dei campioni estratti

dall'onda: per poter essere rappresentato da un calcolatore, il valore dell'ampiezza deve essere espresso tramite un numero finito di bit

- Si sceglie quanti bit utilizzare per rappresentare ogni campione (in figura, 3 bit)
- Si suddivide l'intervallo dei valori ammissibili in 2^n intervalli, dove n è il numero di bit per campione.
- Si rappresenta il valore di ciascun campione con la sua corrispondente codifica binaria (nell'esempio in figura, il segnale risulta rappresentato dalla stringa **011100011010001010...**)



Codifica dei suoni

Quantizzazione

- La rappresentazione è tanto più precisa quanto maggiore è il numero di bit utilizzati per codificare l'informazione estratta in fase di campionamento
- La differenza tra il valore reale di un campione e il corrispondente livello quantizzato prende il nome di errore di quantizzazione

Codifica dei suoni

Esercizio

Qual è l'occupazione di memoria di un file musicale stereo di 1 minuto, campionato a 16 bit alla frequenza di 44.1 kHz?

- Calcoliamo il numero dei campioni N:

$$2 \times 44.1 \times 10^3 \text{ Hz} \times 60 \text{ s} = 5292000$$

- Per ottenere il numero di bit usati per rappresentare i campioni, si moltiplica N per il numero di bit necessari a rappresentare ciascun campione:

$$5292000 \text{ campioni} \times 16 \text{ bit/campione} = 84672000 \text{ bit} =$$

- $= 10584000 \text{ byte} \approx 1 \text{ MB}$

Rappresentazione digitale delle informazioni

Osservazione

Per calcolare lo spazio occupato da un file di testo, un'immagine, un file audio... , in generale:

- Si trova il numero di unità elementari che costituiscono il file (il numero di caratteri per il testo, il numero di pixel per l'immagine, il numero di campioni per il file audio...)
- si trova lo spazio occupato da ogni unità elementare costituente il file (un carattere per il testo, un pixel per l'immagine, un campione per il file audio), ovvero il numero di bit necessari a rappresentarla
- Si moltiplicano queste due quantità

Rappresentazione delle immagini in movimento

- Memorizzazione mediante sequenze di fotogrammi statici
- La qualità della memorizzazione dipende dal numero di fotogrammi al secondo (chiamati anche frame): fps = frame per second
- Normalmente, le frequenze di fotogrammi più comuni sono 25 fps o 30 fps
- Esercizio: qual è l'occupazione di memoria di un filmato lungo 3 minuti, di risoluzione di 200x100, con 24 fps e profondità di colore di 16 bit?

$$3 \times 60 \times 200 \times 100 \times 24 \times 16 \text{ bit} = 1382400000 \text{ bit} \approx 170 \text{ MB}$$