

Rappresentazione dell'Informazione

Sistema posizionale

- Il valore della cifra dipende anche dalla sua posizione
- Se la base è b ci saranno b cifre: $0, \dots, b - 1$
- Il peso assunto è pari alla potenza della base elevata alla posizione e cresce da dx

Esempi

- Base 10: 10 cifre da 0 a 9

$$101_{10} = 1 * 10^2 + 0 * 10^1 + 1 * 10^0$$

- Base 2: 2 cifre da 0 a 1

$$101_2 = 1 * 2^2 + 0 * 2^1 + 1 * 2^0 (= 5_{10})$$

La base 16

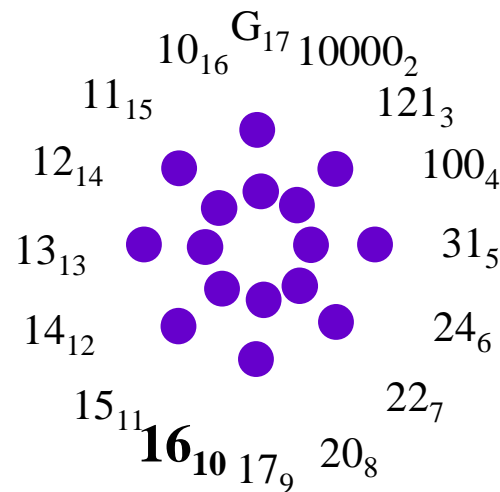
- La base 16 è detta esadecimale
 - ha 16 cifre, quelle da 10 a 15 si scrivono come A, ..., F

$$101_{16} = 1 * 16^2 + 0 * 16^1 + 1 * 16^0 = 257_{10}$$

$$ABC_{16} = 10 * 16^2 + 11 * 16^1 + 12 * 16^0 = 2748_{10}$$

Entità e rappresentazione

- L'entità è il valore “intrinseco” di un numero



Conversione da base 10

- Si divide il numero per la base e si considera il resto intero

e.g. (base 2): 22_{10}

22 / 2	0
11 / 2	1
5 / 2	1
2 / 2	0
1 / 2	1

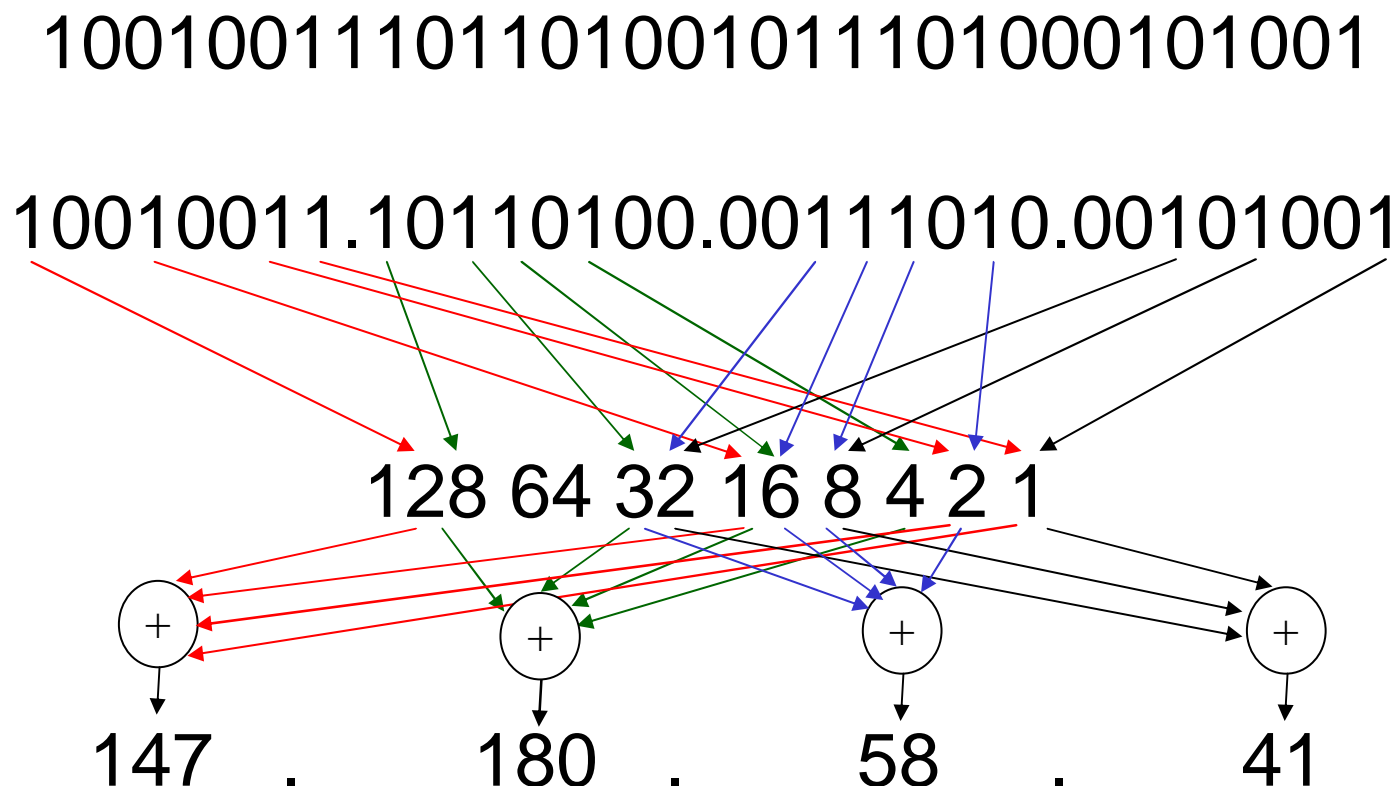
= 10110_2

Codifica di sequenze di bit

- Non c'è spazio per un ulteriore simbolo
 - non si possono codificare i separatori
 - si deve stabilire una lunghezza convenzionale
- Il byte è un compromesso tra
 - espressività: 2^8 valori possibili
 - occupazione di memoria

Applicazioni - Gli indirizzi Internet

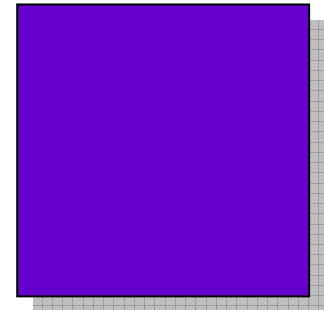
- Da sequenze di *bit* a quaterne di numeri decimali



Applicazioni – i colori in HTML

- Si modificano le componenti RGB (rosso, verde e blu) su base 256

	R	G	B
%	40%	0%	80%
su 256	102	0	204
hex	66	0	CC
risultato	#6600CC		



$$XY = X * 16^1 + Y * 16^0$$

Rappresentare i caratteri

- Il testo in un calcolatore è importante
 - perché è il mezzo principale per la conoscenza
 - per poter comunicare con il calcolatore stesso
- Il contenuto e la forma
 - contenuto: successione di parole su un alfabeto
 - forma: modo con cui il testo viene rappresentato

ASCII e gli altri standard

- L'alfabeto è finito: si può associare un numero ad ogni carattere
 - si parla di codifica (è arbitraria, o meglio, convenzionale)
- Necessità degli *standard*
 - ASCII: 7 o 8 bit (128 o 256 simboli)
 - ISO 8859-1 (per l'Italiano)
 - Unicode UTF: 16 o 32 bit (16 milioni o 4 miliardi di simboli)

Informazione non simbolica

- Non tutto si può rappresentare tramite simboli
- Alcune informazioni dipendono dalle percezioni (soggettive)
 - suoni
 - immagini

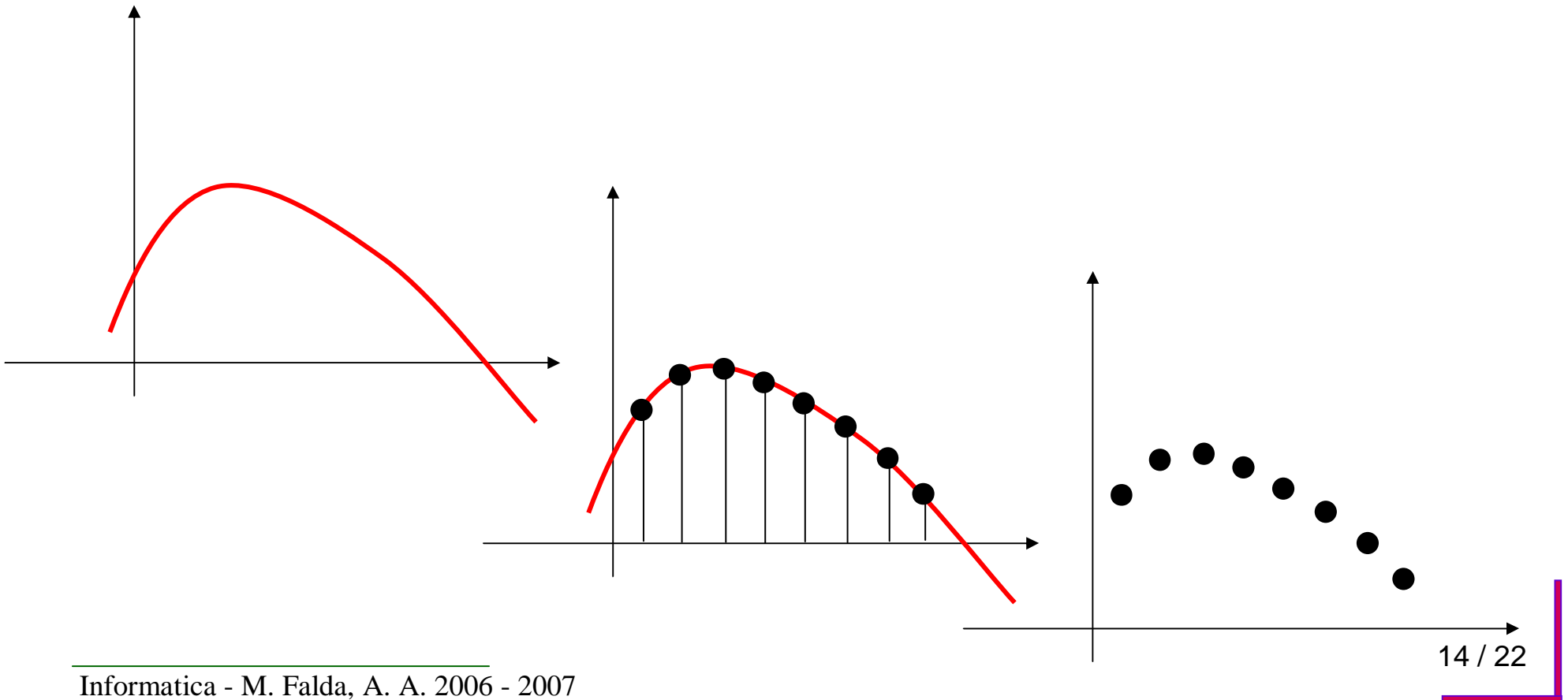
Campionamento e quantizzazione

- Una rappresentazione troppo precisa
 - richiede molto spazio per essere memorizzata
 - può essere inutile

e.g.: la temperatura alle ore 12:13:34:76 era di
27,233245667456754643223 gradi

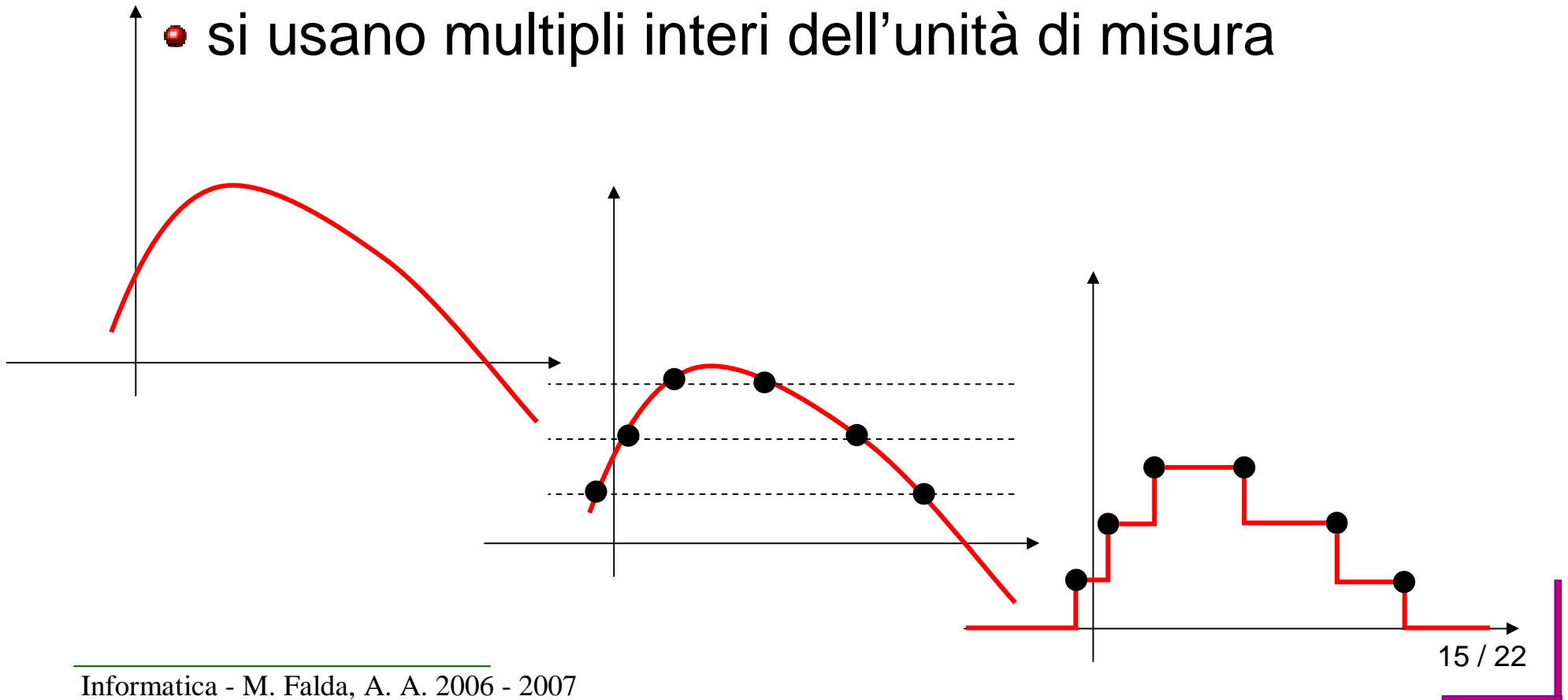
Il campionamento

È il numero di rilevazioni che si effettuano



La quantizzazione

- È l'arrotondamento in base ad una scala
 - si usano multipli interi dell'unità di misura



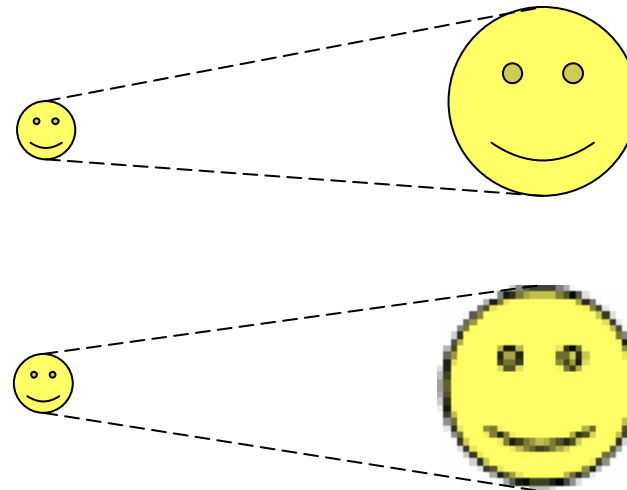
Rappresentazione dei suoni

- Notazione simbolica (le partiture musicali)
 - e.g.: codifica MIDI
- Codifica digitale
 - e.g.: codifica CD

Utilizzo	Campion.	Quantiz.
Telefono	8 KHz	1 byte
CD audio	44,1 KHz	16 byte

Rappresentazione delle immagini

- Notazione simbolica (comandi grafici)
 - e.g.: codifiche PostScript, Metafile, ...
- Codifica digitale
 - e.g.: codifica *bitmap*



Campionamento

1:1

1:4

1:8



Quantizzazione

24 bit

8 bit

1 bit



Compressione

- I dati multimediali hanno dimensioni elevate
- Tipologie
 - senza perdita: i dati sono preservati completamente, ma bassi rapporti di compressione (necessaria per testi, programmi)
 - con perdita: si perde una parte dei dati, ma maggiori rapporti di compressione (immagini, suoni)

Compressione senza perdita

- Tutti gli archiviatori generali: ZIP, RAR
- Formati per le immagini: PNG, TIFF
- Formati per i suoni: WAVE

Compressione con perdita

- Formati per le immagini: JPG

Qualità: 90%

Da 226KB a 24KB

Riduzione: ~ 90%



Qualità: 10%

Da 226KB a 10KB

Riduzione: ~ 95%

- Formati per i suoni: MP3