

Corso di Informatica

Modulo 1

4-Rappresentazione di informazioni non numeriche

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

1
01/10/2013

Prerequisiti

- Struttura del sistema di elaborazione
- Evoluzione dei sistemi di elaborazione
- Uso elementare della multimedialità

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

2
01/10/2013

Introduzione

Lo scopo di questa Unità è illustrare come attraverso il computer sono rappresentabili le informazioni non numeriche, quali caratteri, suoni e immagini.

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

3
01/10/2013

Il codice ASCII

Il codice di gran lunga più usato all'interno della macchina è il codice **ASCII** (*American Standard Code Information Interchange*), nato nel 1968 per rendere standard la trasmissione di dati tra i diversi sistemi hardware e software.

Utilizza **7 bit** per rappresentare $2^7 (=128_{10})$ caratteri diversi, in particolare:

- lettere maiuscole e minuscole
- caratteri di interpunzione
- cifre e comandi comuni nella periferiche di I/O
- tabulazioni
- a capo
- segnale acustico

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

4
01/10/2013

Codice ASCII e byte

Questo codice fa uso di 7 bit, ma poiché nei calcolatori si usano normalmente celle di 8 bit, spesso si sfrutta il bit più significativo come bit di controllo (detto **bit di parità**) nelle trasmissioni dati.

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

5
01/10/2013

I caratteri di controllo

| Hex | Nome | Significato | Hex | Nome | Significato |
|-----|------|---------------------|-----|------|---------------------------|
| 0 | NUL | Null | 10 | DLE | Data Link Escape |
| 1 | SOH | Start Of Heading | 11 | DC1 | Device Control 1 |
| 2 | STX | Start Of Text | 12 | DC2 | Device Control 2 |
| 3 | ETX | End Of Text | 13 | DC3 | Device Control 3 |
| 4 | EOT | End Of Transmission | 14 | DC4 | Device Control 4 |
| 5 | ENQ | Enquiry | 15 | NAK | Negative Acknowledgement |
| 6 | ACK | ACKnowledgement | 16 | SYN | SYNchronous idle |
| 7 | BEL | BELl | 17 | ETB | End of Transmission Block |
| 8 | BS | BackSpace | 18 | CAN | CANcel |
| 9 | HT | Horizontal Tab | 19 | EM | End of Medium |
| A | LF | Line Feed | 1A | SUB | SUBstitute |
| B | VT | Vertical Tab | 1B | ESC | ESCape |
| C | FF | Form Feed | 1C | FS | File Separator |
| D | CR | Carriage Return | 1D | GS | Group Separator |
| E | SO | Shift Out | 1E | RS | Record Separator |
| F | SI | Shift In | 1F | US | Unit Separator |

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

6
01/10/2013

I caratteri stampabili

| Hex | Char | Hex | Char | Hex | Char | Hex | Char | Hex | Char | Hex | Char |
|-----|---------|-----|------|-----|------|-----|------|-----|------|-----|------|
| 20 | (Space) | 30 | 0 | 40 | @ | 50 | P | 60 | ` | 70 | p |
| 21 | ! | 31 | 1 | 41 | A | 51 | Q | 61 | a | 71 | q |
| 22 | " | 32 | 2 | 42 | B | 52 | R | 62 | b | 72 | r |
| 23 | # | 33 | 3 | 43 | C | 53 | S | 63 | c | 73 | s |
| 24 | \$ | 34 | 4 | 44 | D | 54 | T | 64 | d | 74 | t |
| 25 | % | 35 | 5 | 45 | E | 55 | U | 65 | e | 75 | u |
| 26 | & | 36 | 6 | 46 | F | 56 | V | 66 | f | 76 | v |
| 27 | ' | 37 | 7 | 47 | G | 57 | W | 67 | g | 77 | w |
| 28 | (| 38 | 8 | 48 | H | 58 | X | 68 | h | 78 | x |
| 29 |) | 39 | 9 | 49 | I | 59 | Y | 69 | i | 79 | y |
| 2A | * | 3A | : | 4A | J | 5A | Z | 6A | j | 7A | z |
| 2B | + | 3B | ; | 4B | K | 5B | [| 6B | k | 7B | { |
| 2C | , | 3C | < | 4C | L | 5C | \ | 6C | l | 7C | |
| 2D | - | 3D | = | 4D | M | 5D |] | 6D | m | 7D | } |
| 2E | . | 3E | > | 4E | N | 5E | ^ | 6E | n | 7E | ~ |
| 2F | / | 3F | ? | 4F | O | 5F | _ | 6F | o | 7F | DEL |

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

7
01/10/2013

Il codice ASCII esteso

| | | | | | | | | | | | | | | | |
|-----|---|-----|---|-----|---|-----|---|-----|---|-----|---|-----|---|-----|---|
| 128 | Ç | 144 | É | 161 | í | 177 | ⌘ | 193 | ± | 209 | ⌘ | 225 | β | 241 | ± |
| 129 | ü | 145 | æ | 162 | ó | 178 | ⌘ | 194 | ⌘ | 210 | ⌘ | 226 | Γ | 242 | ≥ |
| 130 | é | 146 | Æ | 163 | ú | 179 | | 195 | ⌘ | 211 | ⌘ | 227 | π | 243 | ≤ |
| 131 | â | 147 | ô | 164 | ñ | 180 | ⌘ | 196 | - | 212 | ⌘ | 228 | Σ | 244 | ∫ |
| 132 | á | 148 | ö | 165 | Ñ | 181 | ⌘ | 197 | + | 213 | ⌘ | 229 | σ | 245 | ∫ |
| 133 | à | 149 | ò | 166 | ª | 182 | ⌘ | 198 | ⌘ | 214 | ⌘ | 230 | μ | 246 | ÷ |
| 134 | ã | 150 | û | 167 | º | 183 | ⌘ | 199 | ⌘ | 215 | ⌘ | 231 | τ | 247 | ≈ |
| 135 | ç | 151 | ù | 168 | ¸ | 184 | ⌘ | 200 | ⌘ | 216 | ⌘ | 232 | Φ | 248 | ° |
| 136 | ê | 152 | - | 169 | - | 185 | ⌘ | 201 | ⌘ | 217 | ⌘ | 233 | ⊙ | 249 | . |
| 137 | ë | 153 | Û | 170 | ¬ | 186 | ⌘ | 202 | ⌘ | 218 | ⌘ | 234 | Ω | 250 | . |
| 138 | è | 154 | Ü | 171 | ½ | 187 | ⌘ | 203 | ⌘ | 219 | ■ | 235 | δ | 251 | √ |
| 139 | ì | 156 | £ | 172 | ¼ | 188 | ⌘ | 204 | ⌘ | 220 | ■ | 236 | ∞ | 252 | - |
| 140 | î | 157 | ¥ | 173 | ¡ | 189 | ⌘ | 205 | = | 221 | ■ | 237 | φ | 253 | ² |
| 141 | ï | 158 | - | 174 | « | 190 | ⌘ | 206 | ⌘ | 222 | ■ | 238 | ε | 254 | ■ |
| 142 | Ä | 159 | f | 175 | » | 191 | ⌘ | 207 | ± | 223 | ■ | 239 | ∩ | 255 | |
| 143 | Å | 160 | á | 176 | ⌘ | 192 | ⌘ | 208 | ⌘ | 224 | α | 240 | ≡ | | |

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

8
01/10/2013

Il codice ASCII esteso

Nella **rappresentazione estesa ASCII** (PC IBM) si usano tutti gli 8 bit per la rappresentazione dei caratteri (non c'è bit di parità), in questo modo si possono rappresentare altri 128 caratteri:

- vocali accentate
- simboli degli alfabeti latini (à, è, é,...)
- simboli matematici e grafici (\pm , \equiv , ...).

Il codice ASCII esteso

Ma rimangono comunque fuori molte lettere di alfabeti internazionali.

Alcuni hanno ridefinito il set in funzione dei propri caratteri, ma così nascono problemi di incompatibilità (uno stesso codice viene interpretato in modo diverso a seconda dell'alfabeto usato dal calcolatore).

Il codice UNICODE

Il consorzio **Unicode** (Apple, HP, IBM, JustSystem, Microsoft, Oracle, SAP, Sun, Sybase, Unisys, etc) è nato negli inizi degli anni '90 con l'obiettivo di

“creare una rappresentazione univoca per «ogni» carattere esistente”

in modo indipendente dalla piattaforma, dal software e dal linguaggio di programmazione usato.

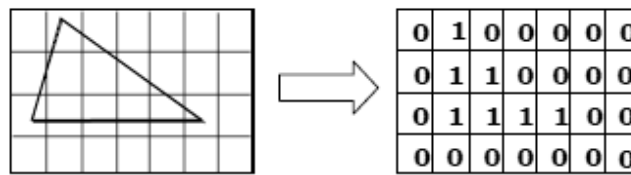
Il lavoro di informatici e linguisti ha portato alla nascita di **Unicode**, che definisce una codifica per i caratteri usati nella maggior parte degli linguaggi attualmente in uso, come gli alfabeti europei, del Middle Eastern right-to-left, e molti alfabeti asiatici.

Il codice UNICODE

Ha il vantaggio di avere la stessa rappresentazione per i caratteri ASCII (perciò può essere usato nel software attuale senza grosse modifiche).

Rappresentazione di immagini

- L'unità minima di rappresentazione sul monitor è il **pixel**
- Le immagini si rappresentano con la tecnica **bitmap** (ogni pixel può essere acceso o spento, supponendo monitor bianco e nero)
- Si nota perdita di informazione (la figura è approssimata)

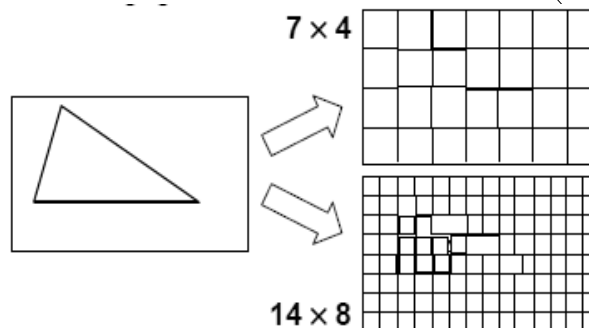


M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

13
01/10/2013

Rappresentazione di immagini

- La **risoluzione** è il numero di pixel della griglia
- Una maggiore risoluzione, dà una figura migliore



M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

14
01/10/2013

Rappresentazione di immagini

- Per il **bianco e nero** è sufficiente un bit (bianco = 1, nero = 0)
- Per la **scala di grigi** è necessario un numero che indica l'intensità (nero = 0, bianco = 255)
- Per il **colore** la codifica è più complessa (si indica con **bpp** = byte per pixel e in genere si hanno 3 byte per pixel)
 - i colori sono codificati con la tecnica **RGB** combinando i **tre colori primari**: rosso (**R**ed), verde (**G**reen) e blu (**B**lue)
 - in particolare, un colore è definito da una terna di numeri che specificano le intensità rispettivamente del rosso, del verde e del blu
 - RGB è utilizzata, tra gli altri, dagli apparecchi TV e dai monitor dei calcolatori

Rappresentazione di immagini

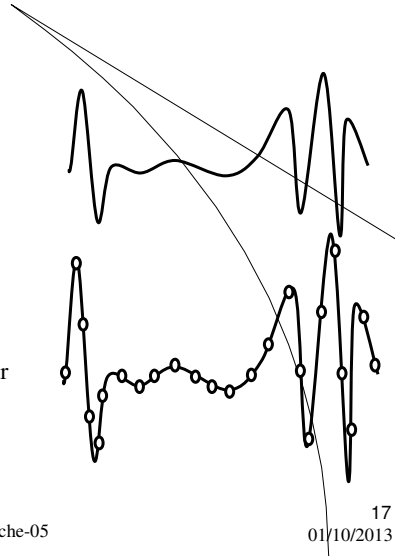
Le immagini sono normalmente memorizzate in forma **compressa**:

- senza perdita di informazione (file .GIF, .PNG)
- con perdita di informazione (.JPG)

Rappresentazione di suoni

Viene realizzata con campionamento:

- Frequenza di campionamento
 - (*sample rate*) quante volte al secondo misuro il suono - espressa in hertz (Hz)
- Risoluzione di campionamento
 - (*sample resolution*) quanti bit occupo per esprimere il suono misurato



M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

17
01/10/2013

Rappresentazione di suoni

File audio non compressi

- WAV (wavetable, tabella delle onde)
- AIFF (audio interchange file format)
- AU (audio)

File audio compressi

- Compressione MP3 - OGG

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

18
01/10/2013

Qualità CD musicale

- Campionamento:
 - frequenza 44.100 Hz
 - risoluzione 16 bit per canale
- Stereo (canale destro e canale sinistro)
- Spazio necessario:
 - un secondo: 176.400 byte (~172 Kb)
 - un minuto: 10.584.000 (~10 Mb)
 - 74 minuti: 783.216.000 (~747 Mb)

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

19
01/10/2013

Argomenti

- Il codice ASCII
- Codice ASCII e byte
- I caratteri di controllo
- I caratteri stampabili
- Il codice ASCII esteso
- Il codice UNICODE
- Rappresentazione di immagini
- Rappresentazione di suoni
- Qualità CD musicale

M. Malatesta 4-Rappresentazione di informazioni non numeriche-05

20
01/10/2013