



UNIVERSITÀ DEGLI STUDI DI PALERMO

INFORMATIZZAZIONE DELLA DOCUMENTAZIONE SANITARIA

Anno Accademico 2019/2020

Docente: ing. Salvatore Sorce

Rappresentazione numerica delle informazioni

II parte: Codifica binaria del testo e dei suoni



Rappresentazione di caratteri alfanumerici

- American Standard Code for Information Interchange (ASCII)

- Codice a 8 bit (usati solo i primi 7):

- $B_8 B_7 B_6 B_5 B_4 B_3 B_2 B_1$
 - $B_8 = 0$

ASCII	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1
	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0000	N _U	S _H	S _X	E _X	E _T	E _Q	A _K	B _L	B _S	H _T	L _F	Y _T	F _F	C _R	S ₀	S ₁
0001	D _L	D ₁	D ₂	D ₃	D ₄	N _K	S _Y	E _Σ	C _N	E _M	S _B	E _C	F _S	G _S	R _S	U _S
0010		!	"	#	\$	%	&	'	()	*	+	,	-	.	/
0011	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0100	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0101	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
0110	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
0111	p	q	r	s	t	u	v	w	x	y	z	{		}	~	ᵀ

- Quanti bit per memorizzare la parola "Ciao"?

Rappresentazione di caratteri alfanumerici

- American Standard Code for Information Interchange (ASCII)

- Codice a 8 bit (usati solo i primi 7):

- $B_8 B_7 B_6 B_5 B_4 B_3 B_2 B_1$
- $B_8 = 0$

ASCII	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0000	N _U	S _H	S _X	E _X	E _T	E _Q	A _K	B _L	B _S	H _T	L _F	Y _T	F _F	C _R	S ₀	S ₁
0001	D _L	D ₁	D ₂	D ₃	D ₄	N _K	S _Y	E _Σ	C _N	E _M	S _B	E _C	F _S	G _S	R _S	U _S
0010		!	"	#	\$	%	&	'	()	*	+	,	-	.	/
0011	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0100	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0101	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
0110	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
0111	p	q	r	s	t	u	v	w	x	y	z	{		}	~	ᵖᵀ

- Quanti bit per memorizzare la parola "Ciao"?

C -> 01000011

i -> 01101001

a -> 01100001

o -> 01101111



Rappresentazione di caratteri alfanumerici

- Codificare la stringa "Ciao a tutti." in ASCII:
- 01000011 01101001
01100001 01101111 ("Ciao")
- 00100000 (spazio/blank)
- 01100001 ("a")
- 00100000 (spazio/blank)
- 01110100 01110101
01110100 01110100
01101001 ("tutti")
- 00101110 (".")

ASCII	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1
	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0000	N _U	S _H	S _X	E _X	E _T	E _Q	A _K	B _L	B _S	H _T	L _F	Y _T	F _F	C _R	S _O	S _I
0001	D _L	D ₁	D ₂	D ₃	D ₄	N _K	S _Y	E _Σ	C _N	E _M	S _B	E _C	F _S	G _S	R _S	U _S
0010		!	"	#	\$	%	&	'	()	*	+	,	-	.	/
0011	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0100	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0101	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
0110	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
0111	p	q	r	s	t	u	v	w	x	y	z	{		}	~	ᵀ

Rappresentazione di caratteri alfanumerici

- Codici ASCII estesi per alfabeti nazionali a 8 bit (256 caratteri)
 - Definiti dall'ISO (**I**nternational **S**tandard **O**rganization)
 - Tanti codici per i diversi alfabeti
 - Il nostro è il codice ISO Latin 1

Codice ISO UNICODE per la codifica di tutti i caratteri in una sola tabella

- Codice a 16 bit (65536 caratteri)
- Ancora poco usato
- E' il formato unico dei caratteri in Java

ASCII	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0000	N _U	S _H	S _X	E _X	E _T	E _O	A _K	B _L	B _S	H _T	L _F	Y _T	F _F	C _R	S ₀	S ₁
0001	D _L	D ₁	D ₂	D ₃	D ₄	N _K	S _Y	E _Σ	C _N	E _M	S _B	E _C	F _S	G _S	R _S	U _S
0010		!	"	#	\$	%	&	'	()	*	+	,	-	.	/
0011	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0100	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0101	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
0110	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
0111	p	q	r	s	t	u	v	w	x	y	z	{		}	~	D _T
1000	S ₀	S ₁	S ₂	S ₃	I _N	N _L	S _S	E _S	H _S	H _J	Y _S	P _D	P _V	R _I	S ₂	S ₃
1001	D _C	P ₁	P ₂	S _E	C _C	M _M	S _P	E _P	O _S	O _O	O _A	C _S	S _T	O _S	P _M	A _P
1010	A _o	i	ç	£	¥		§	¨	©	♀	«	¬	-	®	¯	
1011	o	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
1100	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
1101	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
1110	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
1111	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

Rappresentazione di dati multimediali

- I dati multimediali (suoni, immagini, video) sono grandezze **continue**
 - Variano nel tempo e nello spazio senza soluzione di continuità
- Un calcolatore può rappresentare solo informazioni **discrete**
 - Non possono assumere valori infinitamente grandi o infinitamente piccoli
 - Non possono variare in maniera arbitraria tra valori vicini
- Esempio reale: 12647321455.572278990951886933278

Troppo grande

Troppe cifre dopo la virgola

Non c'è spazio



Rappresentazione di dati multimediali

- I dati multimediali vengono acquisiti mediante i due processi di **campionamento** e **quantizzazione**
- Nel campionamento vengono prelevati dei campioni rappresentativi del dato multimediale **ad intervalli regolari di tempo (o spazio)**
- Nella quantizzazione si approssima ogni campione **con il valore digitale più vicino rappresentabile dal calcolatore**