

Comunicação de Dados

Tecnologia em Redes de Computadores
Prof. Macêdo Firmino

Aula 02

Representação das Informações em Binário

“Se não houver frutos, valeu o perfume das flores. Se não houver flores, valeu a sombra das folhas. Se não houver folhas, valeu a intenção das sementes” (Henfil)

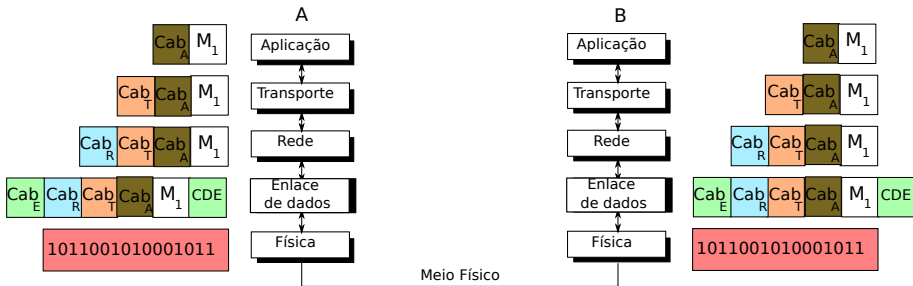


INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
RIO GRANDE DO NORTE

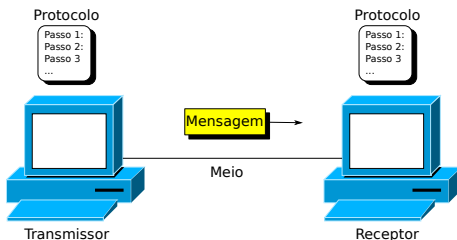
O que Aprenderemos?

Compreender e realizar a conversão de diferentes tipos de informações (números, texto, imagens e vídeos) para o formato binário, entendendo sua importância para a computação e transmissão de dados.





Para que informações como texto, números, imagens, áudio e vídeo sejam armazenadas e transmitidas, é necessário convertê-las em *bits*. Com a representar os dados de maneira adequada, é possível minimizarmos erros de transmissão, aumentar a velocidade de comunicação e a interoperabilidade entre dispositivos de diferentes fabricantes.



Representação dos Dados – Texto

Textos são representados em binário através de um processo chamado codificação de caracteres. Neste, cada caractere (letras, números e símbolos) é mapeado para um valor binário único.

Existem várias codificações de caracteres diferentes, sendo a mais comum a ASCII (*American Standard Code for Information Interchange*) e Unicode.



Representação dos Dados – Texto (ASCII)

Desenvolvida em 1960, a codificação ASCII é composto por 128 caracteres, que incluem letras maiúsculas e minúsculas do alfabeto latino, números de 0 a 9, sinais de pontuação comuns, caracteres especiais (como o espaço em branco, tabulação e nova linha) e alguns caracteres de controle (como o DEL, usado para apagar caracteres). Ele utiliza 7 bits para representar todos os seus símbolos.



Representação dos Dados – Texto (ASCII)

Bin	Oct	Dec	Hex	Sinal
0010 0000	040	32	20	(espaço)
0010 0001	041	33	21	!
0010 0010	042	34	22	"
0010 0011	043	35	23	#
0010 0100	044	36	24	\$
0010 0101	045	37	25	%
0010 0110	046	38	26	&
0010 0111	047	39	27	'
0010 1000	050	40	28	(
0010 1001	051	41	29	)
0010 1010	052	42	2A	*

Bin	Oct	Dec	Hex	Sinal
0100 0000	100	64	40	@
0100 0001	101	65	41	A
0100 0010	102	66	42	B
0100 0011	103	67	43	C
0100 0100	104	68	44	D
0100 0101	105	69	45	E
0100 0110	106	70	46	F
0100 0111	107	71	47	G
0100 1000	110	72	48	H
0100 1001	111	73	49	I
0100 1010	112	74	4A	J

Bin	Oct	Dec	Hex	Sinal
0110 0000	140	96	60	.
0110 0001	141	97	61	a
0110 0010	142	98	62	b
0110 0011	143	99	63	c
0110 0100	144	100	64	d
0110 0101	145	101	65	e
0110 0110	146	102	66	f
0110 0111	147	103	67	g
0110 1000	150	104	68	h
0110 1001	151	105	69	i
0110 1010	152	106	6A	j



Representação dos Dados – Texto (Unicode)

Com a disseminação dos computadores surgiu a necessidade de representar caracteres diferentes daqueles apresentados no ASCII. Para isso, foi criado o por um consórcio de empresas (como IBM, Apple, Microsoft, Nokia, HP e outras) uma nova codificação chamada de Unicode. Esse código utiliza até 32 *bits*. Ele abrange vários idiomas (por exemplo, árabes, grego, hebraico, japonês e latino), símbolos matemáticos, padrões Braile, emojis, entre outros.



Representação dos Dados – Texto (Unicode)

O Unicode pode ser implementado de várias formas, sendo a UTF-8, UTF-16 e UTF-32 as mais comuns. Cada uma dessas formas utiliza um número diferente de bytes para representar os caracteres. A UTF-8 é a mais utilizada na web, pois é eficiente em termos de espaço e é compatível com a codificação ASCII. Ele usa entre 1 a 4 bytes para representar cada caractere. Caracteres comuns são codificados usando apenas 1 byte enquanto que caracteres menos comuns com 4 bytes.



Representação dos Dados – Texto (Unicode)

Periodic Table of the Latin-1

Andrew Plotkin – this image is public domain



Values 20 to 7F (040 to 177)

Values 00 to 1F are non-printable

Space	(	0	8	@	H	P	X	`	h	p	x
040 20	050 28	060 30	070 38	100 40	110 48	120 50	130 58	140 60	150 68	160 70	170 78
!	)	1	9	A	I	Q	Y	a	i	q	y
041 21	051 29	061 31	071 39	101 41	111 49	121 51	131 59	141 61	151 69	161 71	171 79
"	*	2	:	B	J	R	Z	b	j	r	z
042 22	052 2A	062 32	072 3A	102 42	112 4A	122 52	132 5A	142 62	152 6A	162 72	172 7A
#	+	3	;	C	K	S	[	c	k	s	{
043 23	053 2B	063 33	073 3B	103 43	113 4B	123 53	133 5B	143 63	153 6B	163 73	173 7B
\$	,	4	<	D	L	T	\	d	l	t	
044 24	054 2C	064 34	074 3C	104 44	114 4C	124 54	134 5C	144 64	154 6C	164 74	174 7C
%	-	5	=	E	M	U	]	e	m	u	}
045 25	055 2D	065 35	075 3D	105 45	115 4D	125 55	135 5D	145 65	155 6D	165 75	175 7D
&	.	6	>	F	N	V	^	f	n	v	~
046 26	056 2E	066 36	076 3E	106 46	116 4E	126 56	136 5E	146 66	156 6E	166 76	176 7E
'	/	7	?	G	O	W	_	g	o	w	Delete
047 27	057 2F	067 37	077 3F	107 47	117 4F	127 57	137 5F	147 67	157 6F	167 77	177 7F



INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
RIO GRANDE DO NORTE

Representação dos Dados – Imagem

A representação de uma imagem em um computador é feita por meio de *pixels*. Um *pixel* é a menor unidade de uma imagem digital. Os *pixels* em uma imagem são agrupados em linhas e colunas para formar uma grade, também conhecida como matriz de *pixels*. Por exemplo, uma imagem com resolução de 800x600 pixels possui 800 pixels em cada linha (horizontal) e 600 pixels em cada coluna (vertical).



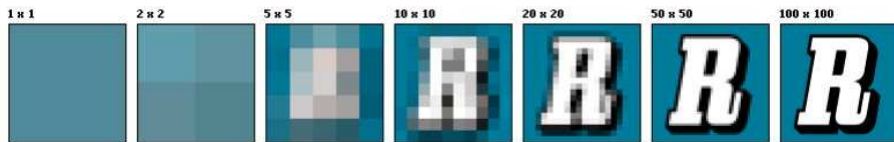
Representação dos Dados – Imagem

A representação de imagem em binário é feita atribuindo um valor numérico a cada *pixel* da imagem. Existem diferentes formatos de arquivo de imagem que utilizam a representação binária, como BMP, PNG, JPEG, entre outros. Cada formato de arquivo de imagem tem sua própria maneira de organizar e armazenar os valores dos *pixels*, bem como métodos de compressão para reduzir o tamanho do arquivo.



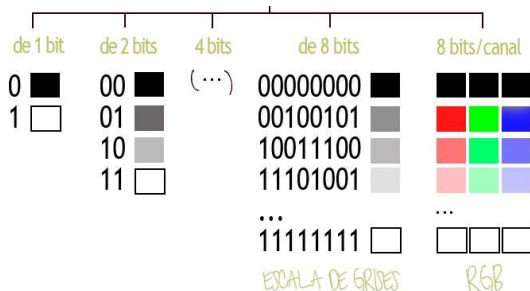
Representação dos Dados – Imagem

A resolução de uma imagem refere-se à quantidade de detalhes que ela contém, geralmente medida em *pixels*. Ele determina a qualidade da imagem. Quanto mais *pixels* uma imagem tiver, maior será sua resolução e, teoricamente, maior será a qualidade da imagem quando exibida em um dispositivo digital, contudo será necessário mais memória para armazená-la.



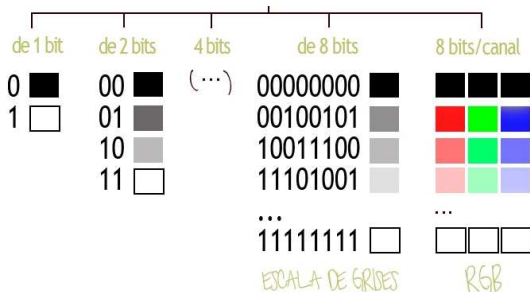
Representação dos Dados – Imagem

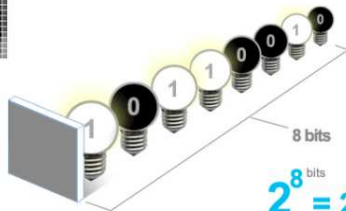
A escala de cinza é uma representação de imagem que consiste em tons de cinza, variando do preto ao branco, sem a presença de cores. Na escala de cinza, o valor 0 representa o preto absoluto, enquanto o valor máximo (geralmente 255 em um sistema de 8 *bits*) representa o branco absoluto. Os tons de cinza intermediários são representados por valores entre 0 e 255.



Representação dos Dados – Imagem

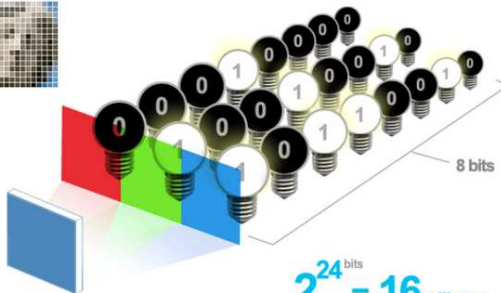
Para representarmos as cores, cada pixel terá uma combinação de *bits*, composto por três valores numéricos que representam a intensidade das cores vermelha, verde e azul (conhecidas como o modelo RGB). Esses valores numéricos são geralmente representados em um intervalo de 0 a 255. Por exemplo, para obter a cor amarela, você pode combinar intensidades máximas de vermelho e verde, enquanto a cor branca é obtida combinando-se intensidades máximas de vermelho, verde e azul.





$$2^8 \text{ bits} = 256 \text{ cores}$$

datos posibles: 0/1



$$2^{24} \text{ bits} = 16 \text{ millones de cores}$$

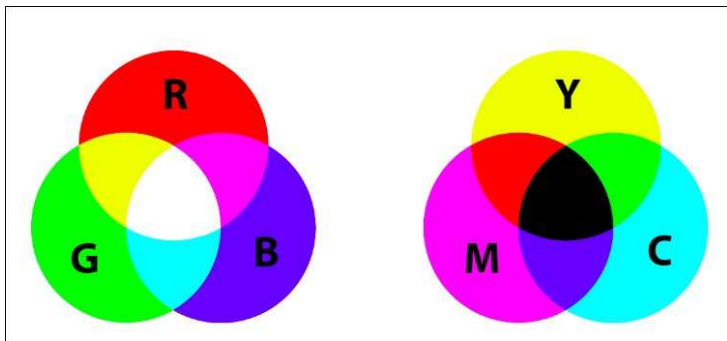
datos posibles: 0/1



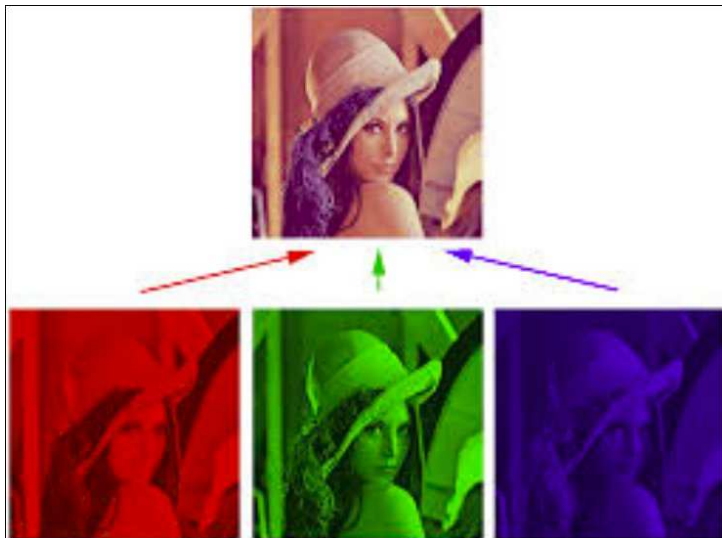
INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
RIO GRANDE DO NORTE

Representação dos Dados – Imagem

Existem vários métodos para a representação de imagens coloridas. Entre elas temos o RGB (usado em monitores) e o YMC (usado em impressoras), onde cada cor é representada por 24 *bits*, dando um alcance de 16.777.216 cores diferentes.



Representação dos dados – Imagem



Representação dos Dados – Vídeo

Para representar vídeo, são utilizadas técnicas semelhantes às de representação de imagens, mas com o acréscimo de uma dimensão temporal. Um vídeo é uma sequência de imagens em movimento, chamadas de *frames* (quadros), que são exibidas em rápida sucessão para criar a ilusão de movimento.



Representação dos Dados – Vídeo

Os vídeos digitais são frequentemente representados usando formatos de arquivo de vídeo, como MP4, AVI, MKV, entre outros. Esses formatos de arquivo armazenam as informações dos *frames* de vídeo, juntamente com informações adicionais, como áudio, legendas e metadados.

Ao reproduzir um vídeo, o dispositivo exibe os *frames* em rápida sucessão, geralmente a uma taxa de 30 *frames* por segundo, criando a ilusão de movimento contínuo. A reprodução de áudio sincronizada com o vídeo também é comum, para fornecer uma experiência audiovisual completa.



Representação dos Dados – Vídeo

Entretanto, o vídeo gera um problema da grande quantidade de memória necessária para armazenar e transmitir. Por exemplo, se utilizarmos uma resolução de 800×600 , colorida, e um padrão de imagem estilo RGB;

- Cada quadro (imagem): $800 \times 600 = 480.000$ *pixels*;
- Considerando 16 milhões de cores possíveis: 24 *bits*, o tamanho de cada imagem: $480.000 \times 24 = 11.520.000 \approx 1,44$ *MBytes*;
- Sendo 30 imagens por segundo: $1,44 \times 30 \approx 43,2$ *MBytes/s*;
- Considerando um filme de 2 horas (7.200 s): $7.200 \times 43,2 \approx 311$ *GBytes*;



Representação dos Dados – Vídeos

A representação binária dos vídeos só é possível utilizando técnicas de compressão de dados. Desta forma, criou-se o conceito de *codec*. Um *codec* é um programa que codifica e comprime os dados referentes a uma imagem para que ela seja transmitida, armazenada e visualizada.



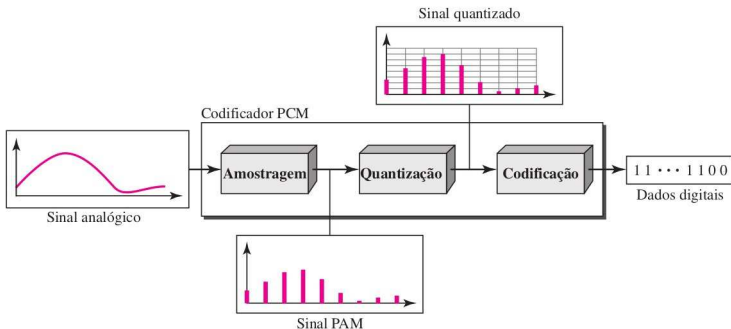
Representação dos dados (Áudio)

O áudio é, por natureza, diferente de texto, números ou imagens, pois ele é analógico. Inicialmente, utilizamos um microfone para transformar sons em um sinal elétrico para posteriormente convertê-lo em um sinal digital.



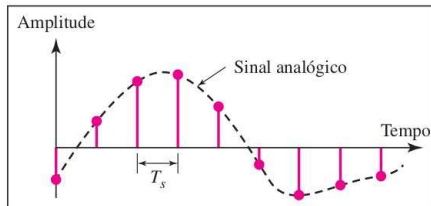
Conversão Analógica-Digital

O sinal analógico possui infinitos valores no tempo e na amplitude. Para poder armazenar e transmitir deverá torná-los finitos. A técnica de conversão mais comum é chamada de PCM (Modulação por Código de Pulso). Ele possui três etapas:



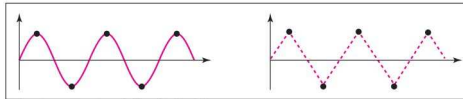
Conversão Analógica-Digital – Amostragem (Tempo)

Inicialmente, torna finito os valores referente ao tempo. O sinal de áudio é amostrado em intervalos de tempo regulares.

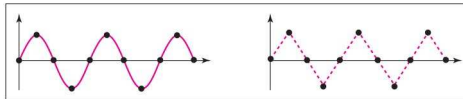


Para reproduzir com a quantidade mínima de perdas, a taxa de amostragem deverá ser o dobro da frequência mais elevada contida no sinal original (Teorema de Nyquist). Por exemplo, o sinal da voz humana tem até 3.400 Hz (variações por segundo) terá que ter pelo menos 6.800 amostras deste sinal por segundo.

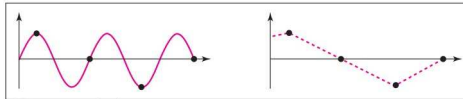
Se tiver menos amostras, não seria possível compreender o audio, se tiver mais o tamanho do arquivo seria gigantesco e sem melhoria significativa na qualidade do áudio.



a. Amostragem com taxa de Nyquist: $f_s = 2f$



b. Sobreamostragem: $f_s = 4f$



c. Subamostragem: $f_s = f$

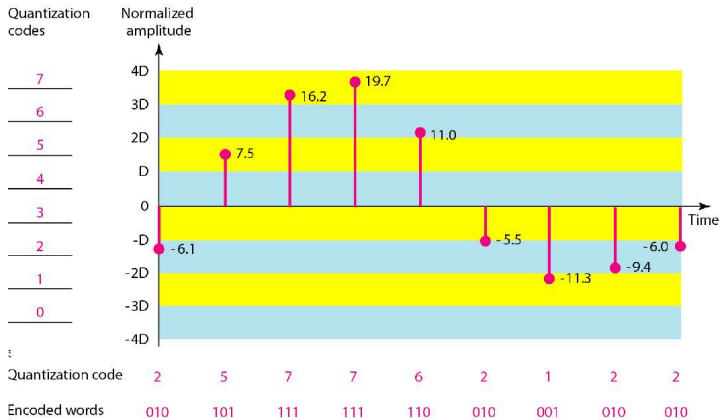
Conversão Analógica-Digital – Quantização e Codificação

O resultado da amostragem ainda é um sinal analógico com infinitas possibilidades de valores de amplitude. Na quantização os valores serão arredondados em um número finito de possibilidades (chamado de níveis). Por exemplo, se estivermos usando uma resolução de 16 bits, cada amostra será quantizada para um valor entre 0 e 65.535.

Após cada amostra ter sido quantizada, será atribuído um código binário para cada valor obtido de acordo com o nível atribuído a ele. Este código, os níveis e o tempo de amostragem deverá ser padronizado para que seja possível reproduzir o sinal original.



Conversão Analógica-Digital – Quantização



Dúvidas



INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
RIO GRANDE DO NORTE