



Introdução à Informática

Representação digital de dados (Parte II)

Áudio:

Um sinal de áudio é uma representação de uma onda sonora que, por sua vez, é uma somatória de sinais:

- Periódicos: senoidais e cosenoidais
- Quase-periódicos (chirp, variação de fase etc.): ressonâncias, reverberações, ecos etc.
- Não periódicos: ruídos (percussão, sopro, fricção etc.)

Introdução à Informática

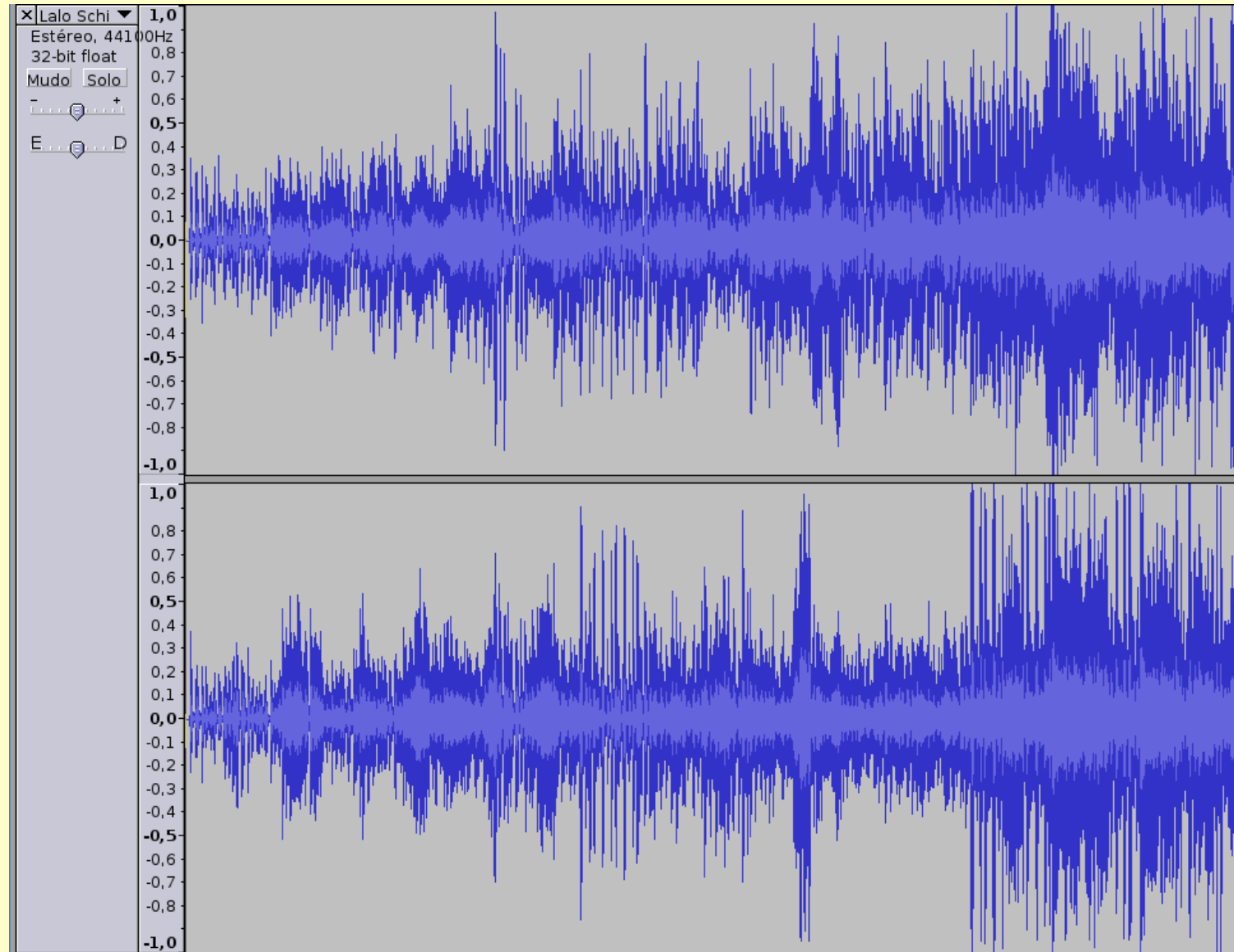
Representação digital de dados (Parte II)

Áudio:

Arquivo de áudio

- 2 canais (estéreo)
- 44.1 kHz
- 32bit float.

(visualizado no
software audacity)





Introdução à Informática

Representação digital de dados (Parte II)

Áudio:

O sinal é amostrado em intervalos regulares de tempo com precisão que pode ser de:

- 8 bits (mono)
- 16 bits (mono ou estéreo)
- 24 bits (mono ou estéreo)
- 32 bits float (mono ou estéreo)

Outras taxas de precisão são possíveis mas mais raramente utilizadas, por exemplo: 12 bits mono MIT/ECG, 32 bits linear (PCM), 16/32 bits fixed point etc..

Introdução à Informática

Representação digital de dados (Parte II)

Áudio:

A taxa de amostragem é registrada em Herz (Hz) mas, na prática é amostras por segundo e a taxa de bits (*bit rate*) em bits/segundo (bps).

Um arquivo de áudio tem um tamanho em Bytes, calculado por:

$$T = S \times \frac{b}{8} \times C \times t$$

Onde,

S = Taxa de amostragem (*sample rate*)

b = resolução em bits/amostra. (*bits per sample*)

C = número de canais de áudio (mono, estéreo, quad)

t = tempo em segundos



Introdução à Informática

Representação digital de dados (Parte II)

Áudio:

Formatos:

- PCM (*Pulse Code Modulation*): RIFF/WAV, AIFF, AU
- Compactados com perda (*lossy compression*): MP3, OGG Vorbis, WMA
- Compactados sem perda (*lossless compression*): FLAC, WavPack, WMA

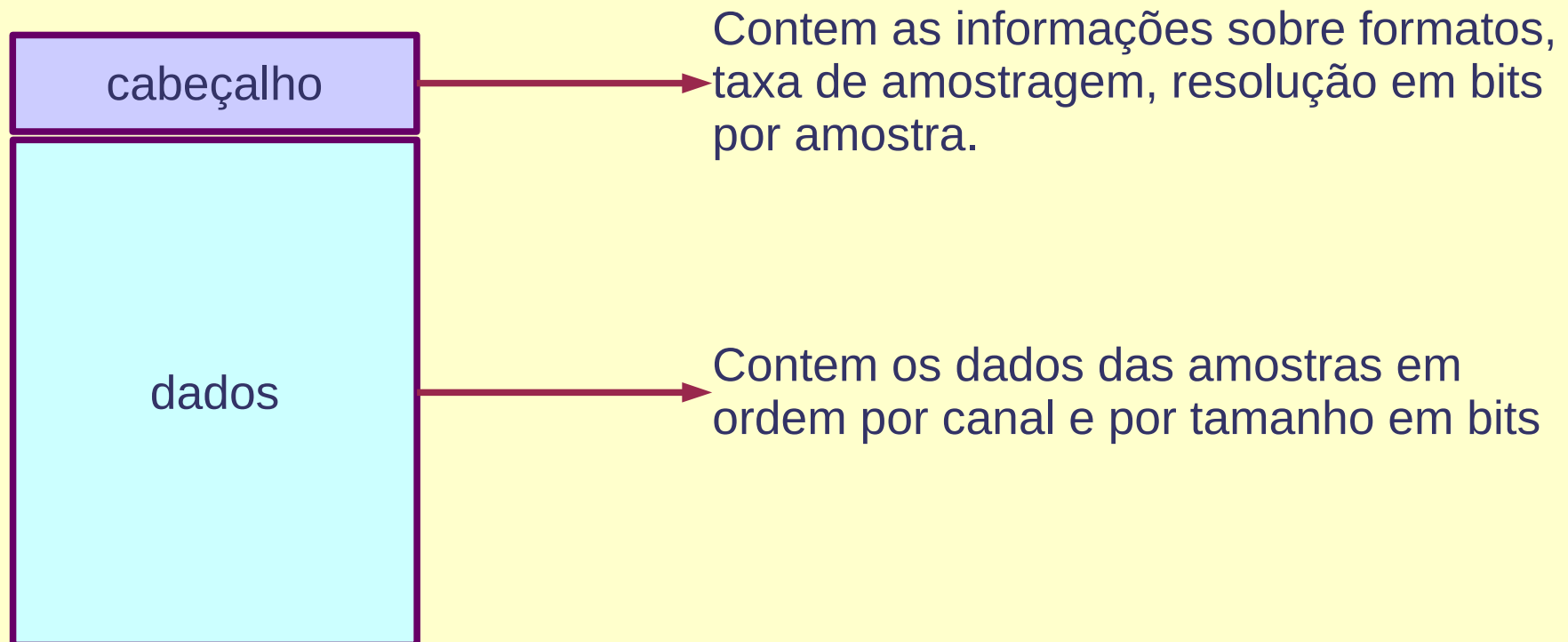
Formato diz respeito ao arquivo que contém os dados de áudio. O processo de codificação e decodificação dos dados de áudio chama-se **CODEC**.

Introdução à Informática

Representação digital de dados (Parte II)

Áudio:

Os arquivos de áudio são, normalmente, formados por um cabeçalho e uma seqüência de dados.

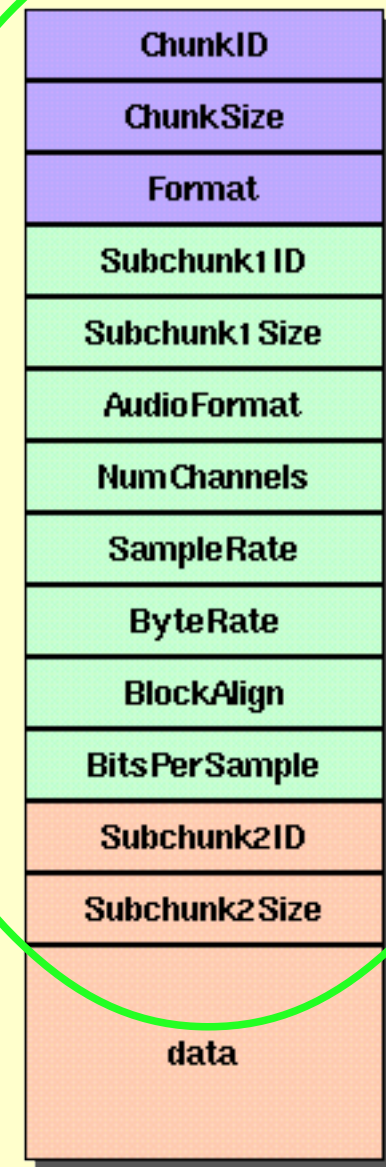
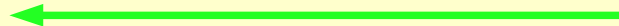


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

44 bytes de
cabeçalho

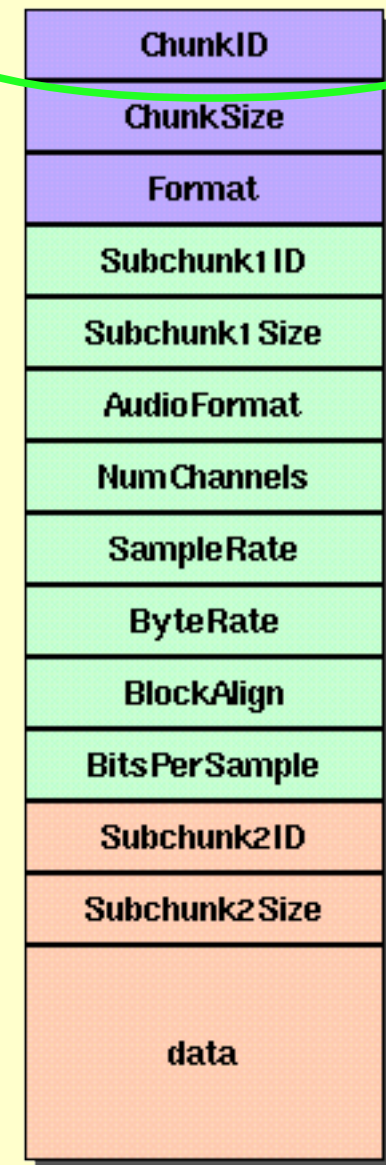


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

ChunkID: 4 bytes (big endian) contendo os caracteres ASCII "RIFF" (0x52494646)



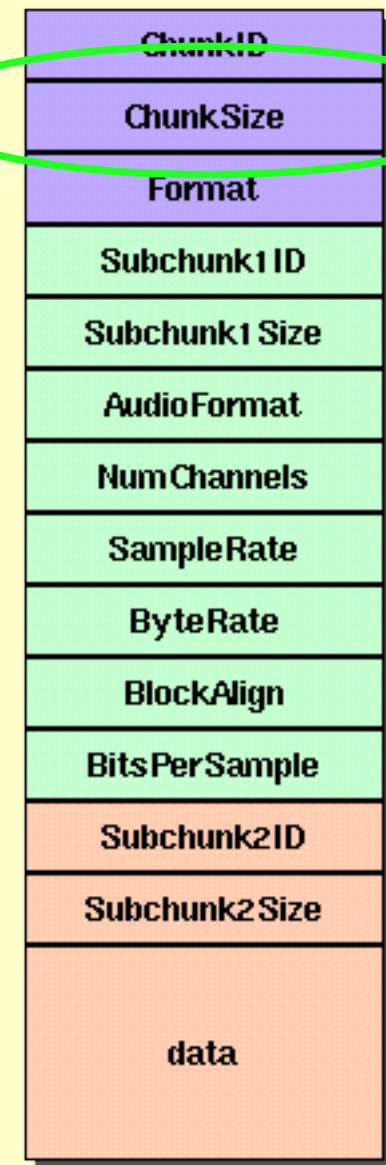
Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

ChunkSize: 4 bytes (little endian) contendo $36 + \text{Subchunk2Size}$, ou seja, o tamanho do arquivo em bytes a partir daqui (o tamanho do arquivo menos 8 bytes).

Por exemplo: Um arquivo com 2 kB de dados de áudio tem 2092 bytes ($2048 + 44$) e o valor do ChunkSize é 2084 ($2092 - 8$). O valor, em hexadecimal neste caso é 0x00000824. Como a notação é little endian deve aparecer: 0x24080000.

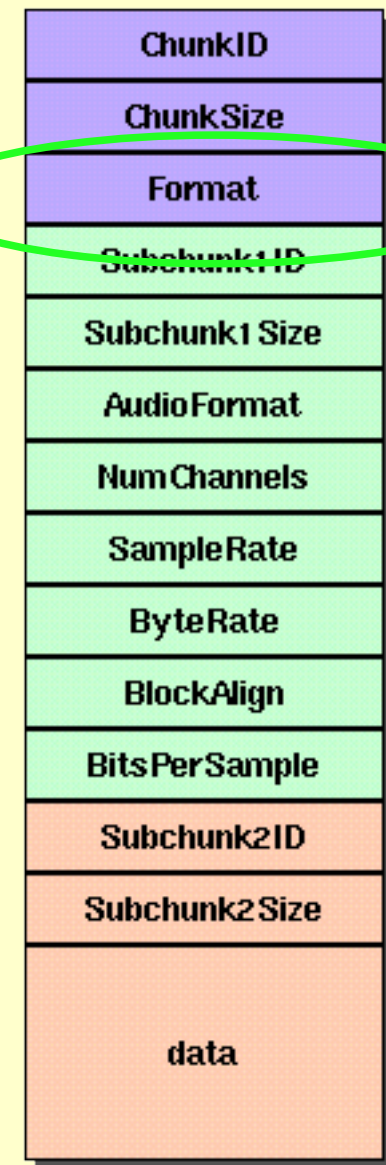


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

Format: 4 bytes (big endian) contendo os caracteres "WAVE" (0x57415645).



Introdução à Informática

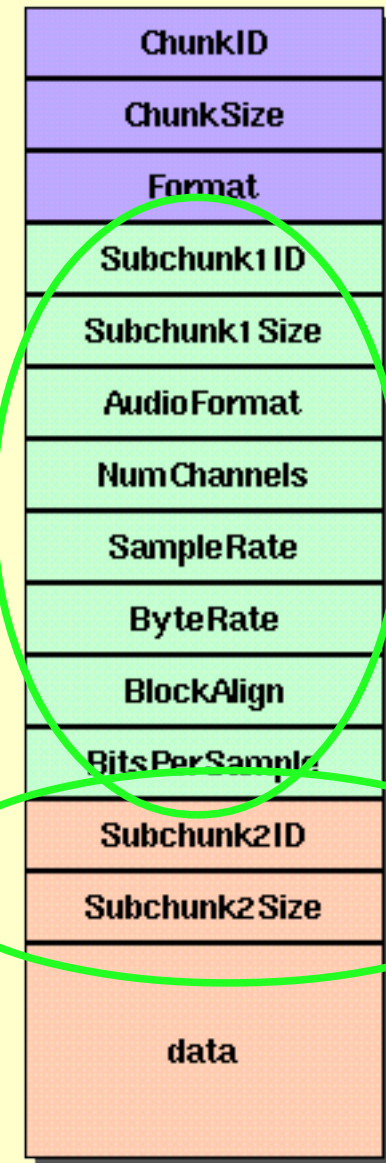
Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

O formato Wave tem dois sub-blocos:

Subchunk “fmt”

Subchunk “data”

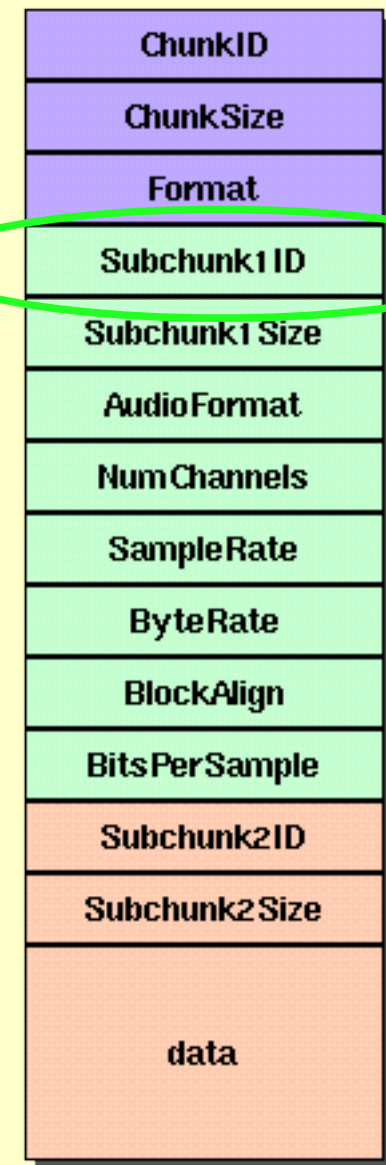


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

Subchunk1ID: 4 bytes (big endian) contendo os caracteres “fmt ” (0x666d7420).

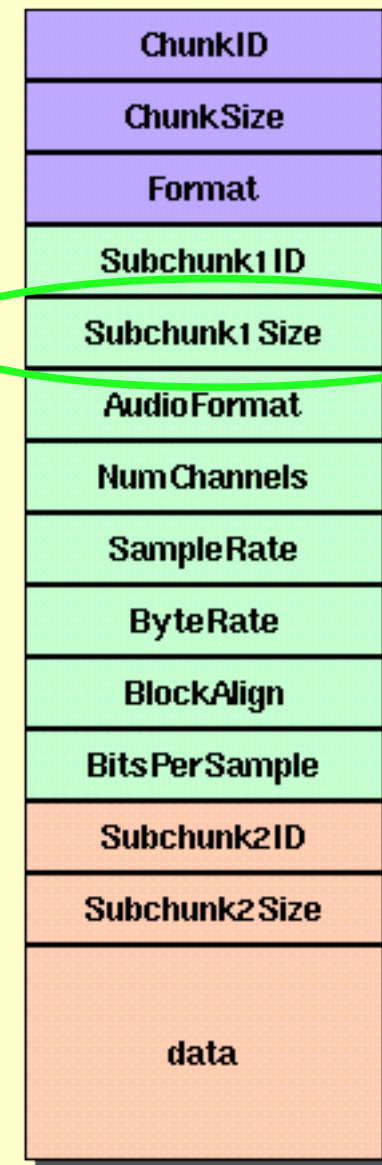


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

Subchunk1Size: 4 bytes (little endian)
contendo o número de bytes deste subchunk,
a partir deste número. Para formato PCM, este
número é 16 (0x10000000 little endian).

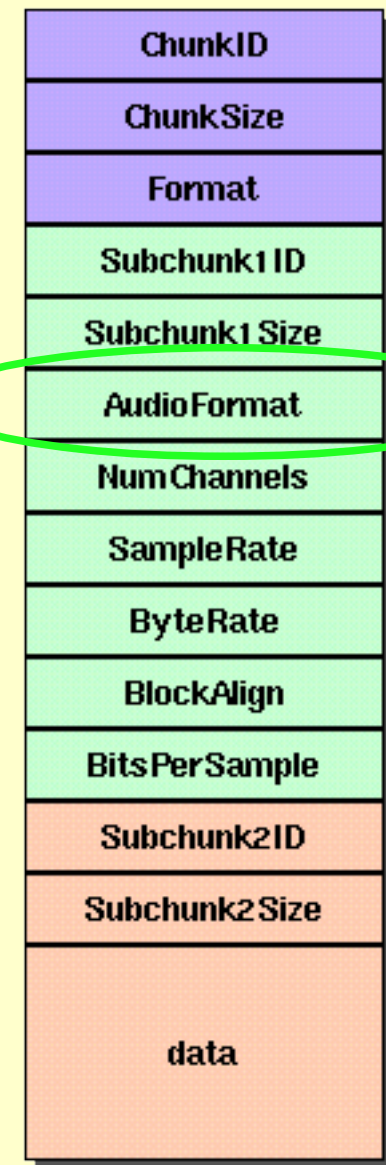


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

AudioFormat: 2 bytes (little endian) contendo “1” para PCM (0x0100). Outros valores indicam algum formato de compressão.

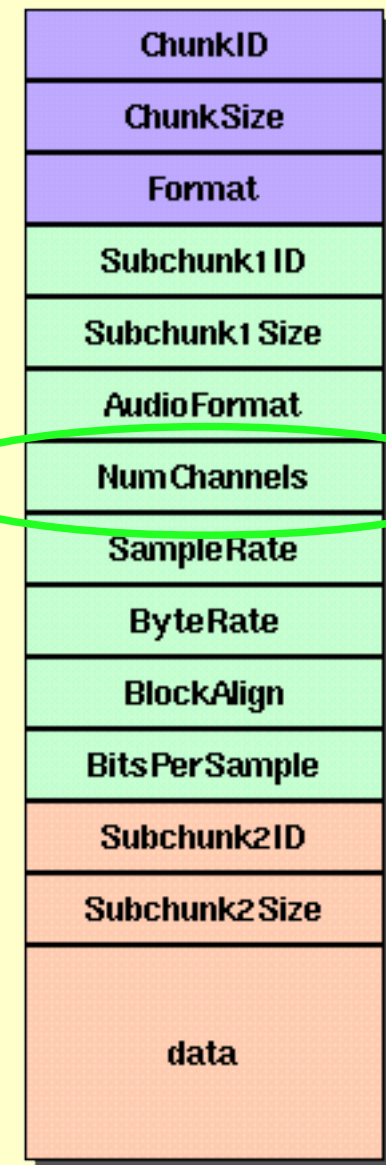


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

NumChannels: 2 bytes (little endian) contendo o número de canais. Por exemplo, um arquivo estéreo por “2” (0x0200).

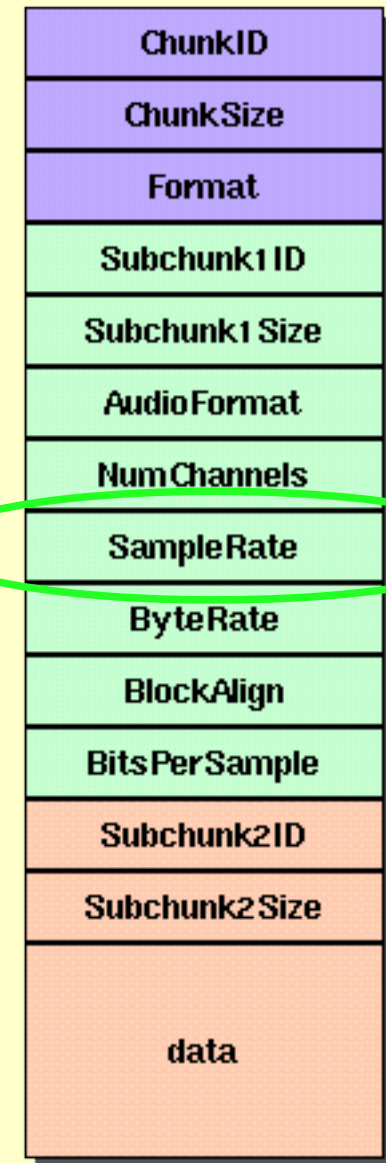


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

SampleRate: 4 bytes (little endian) contendo a taxa de amostragem (8000, 22050, 44100 Hz etc.). Por exemplo, um arquivo com a taxa de amostragem de 22050 = 5622 em hexa (0x22560000 little endian).



Introdução à Informática

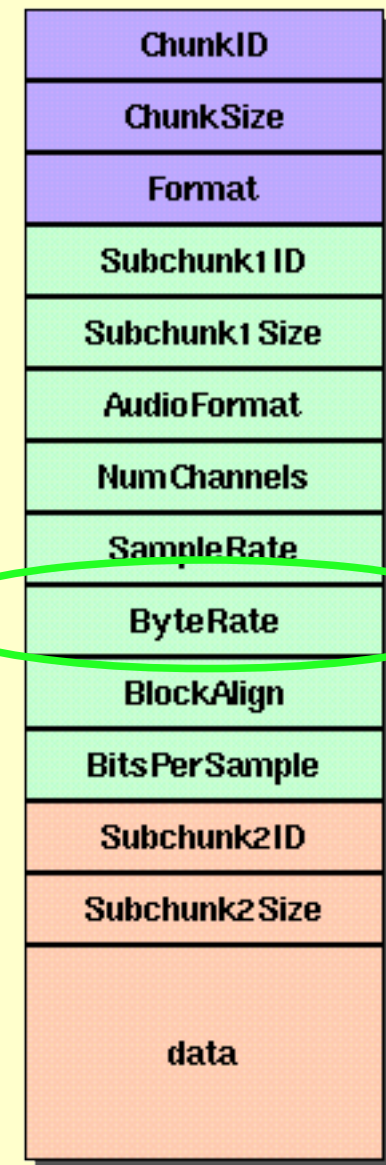
Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

ByteRate: 4 bytes (little endian) contendo a taxa de bytes.

$$ByteRate = S \times C \times \frac{b}{8}$$

Onde S é a taxa de amostragem (sample rate), C é número de canais e b é a resolução em bits (bits por amostra).



Introdução à Informática

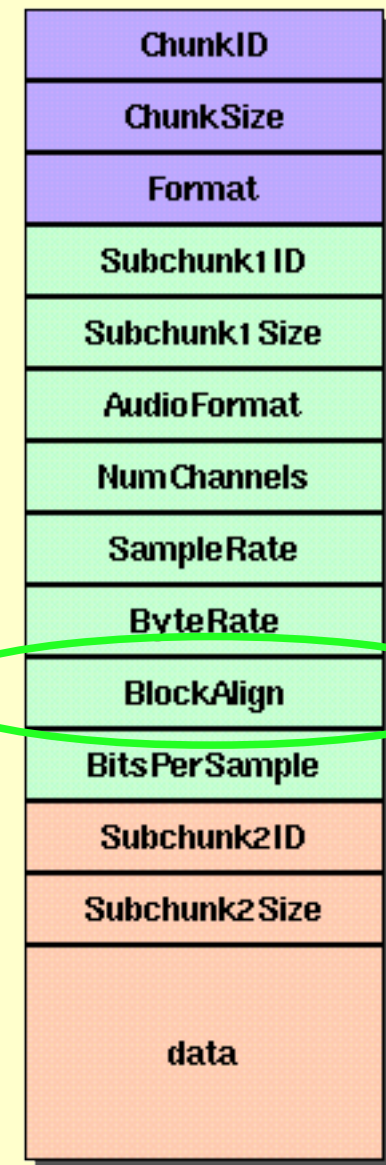
Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

BlockAlign: 2 bytes (little endian) contendo o número de bytes de cada amostra, incluindo os canais.

$$BlockAlign = C \times \frac{b}{8}$$

Onde C é número de canais e b é a resolução em bits (bits por amostra).



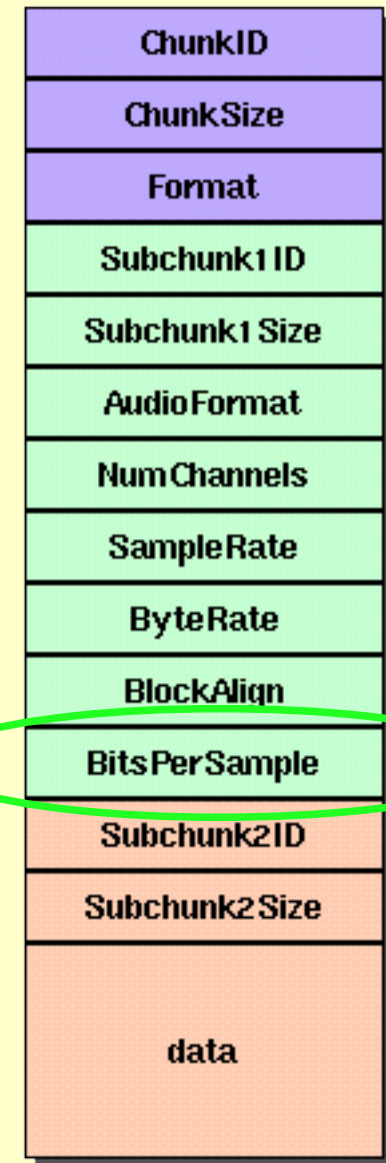
Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

BitsPerSample: 2 bytes (little endian)
contendo o número de bits de cada amostra.

8 bits = 0x0800
16 bits = 0x1000
24 bits = 0x1800
32 bits = 0x2000
etc.

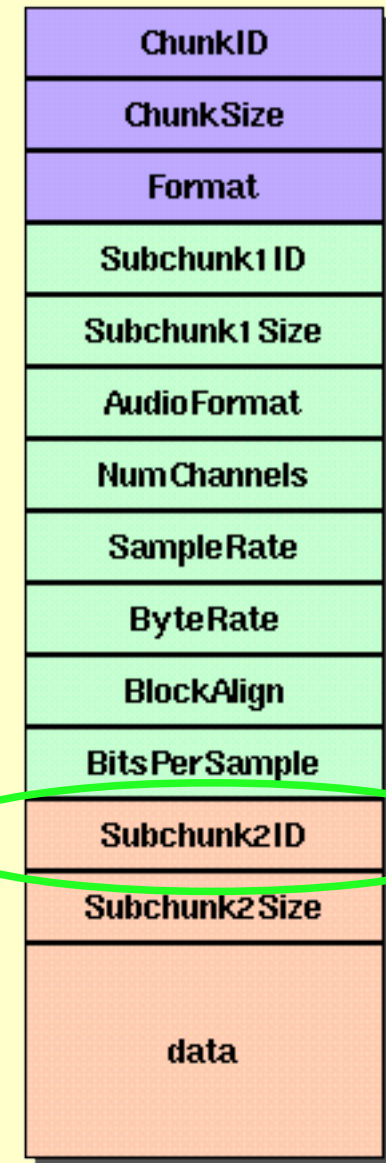


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

Subshunk2ID: 4 bytes (big endian) contendo os caracteres “data” (0x64617461).



Introdução à Informática

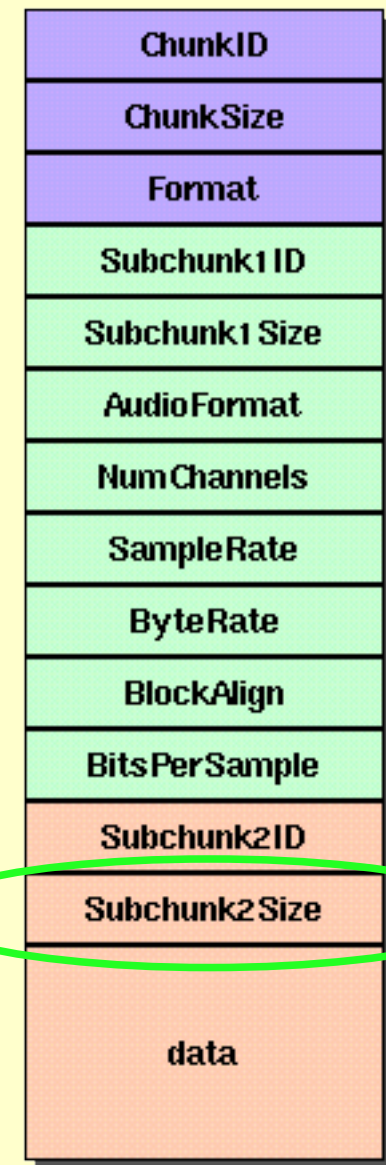
Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

Subchunk2Size: 4 bytes (big endian)
contendo o numero de bytes dos dados de
áudio.

$$Subchunk2Size = N_s \times C \times \frac{b}{8}$$

Onde N_s é o número de amostras, C é número
de canais e b é a resolução em bits (bits por
amostra).

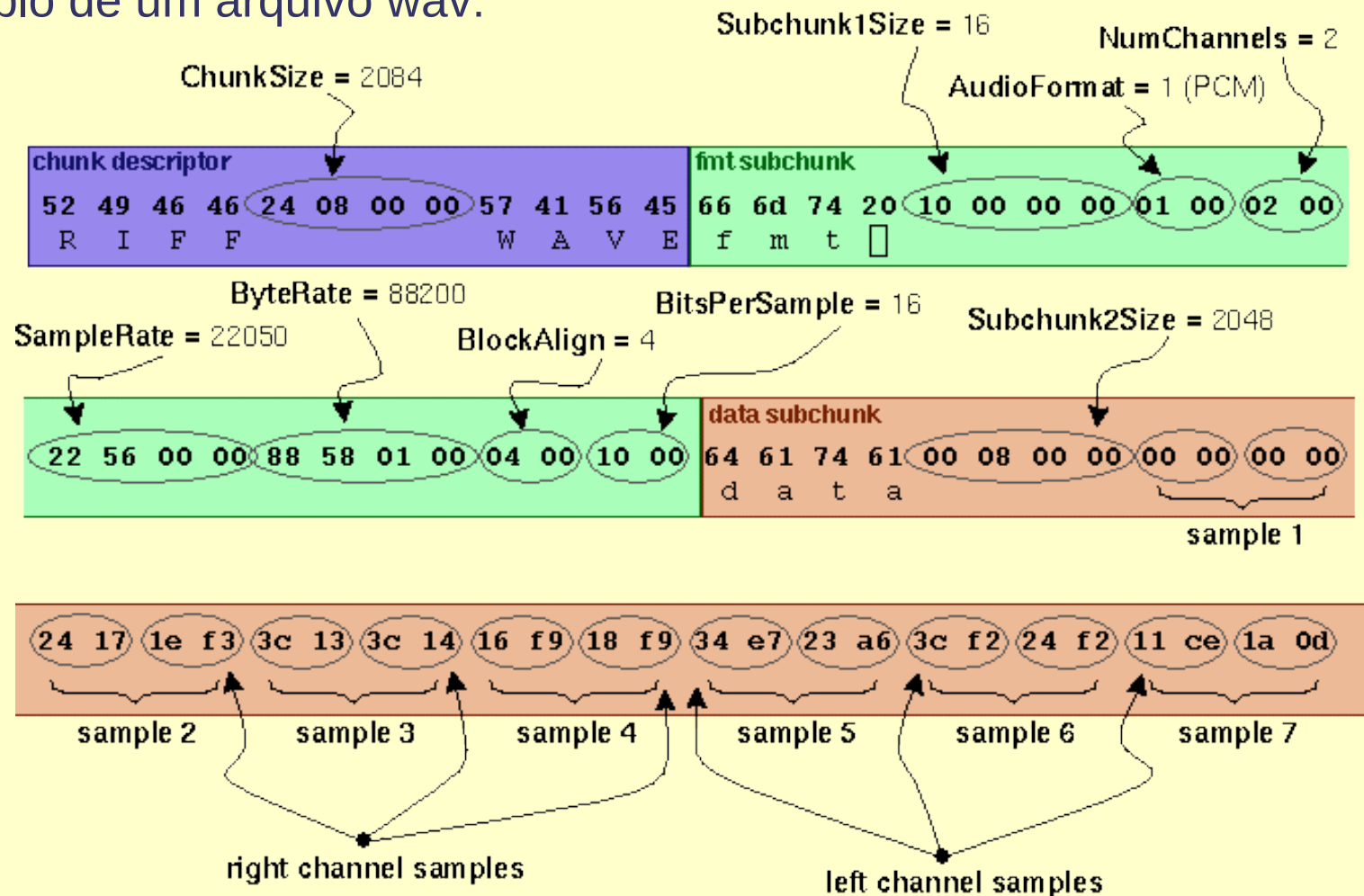


Introdução à Informática

Representação digital de dados (Parte II)

Áudio: cabeçalho RIFF/Wav

Exemplo de um arquivo wav:





Introdução à Informática

Representação digital de dados (Parte II)

Áudio: arquivos AU

Formato AU desenvolvido pela Sun Microsystems

- Cabeçalho de 6 words de 32bits
- Bloco opcional de informações
- Dados no formato big endian

Suporte a vários tipos de codificações de áudio, mais frequentemente utiliza codec μ -law logarítmico, padrão de som para Unices.

Introdução à Informática

Representação digital de dados (Parte II)

Áudio: arquivos AU

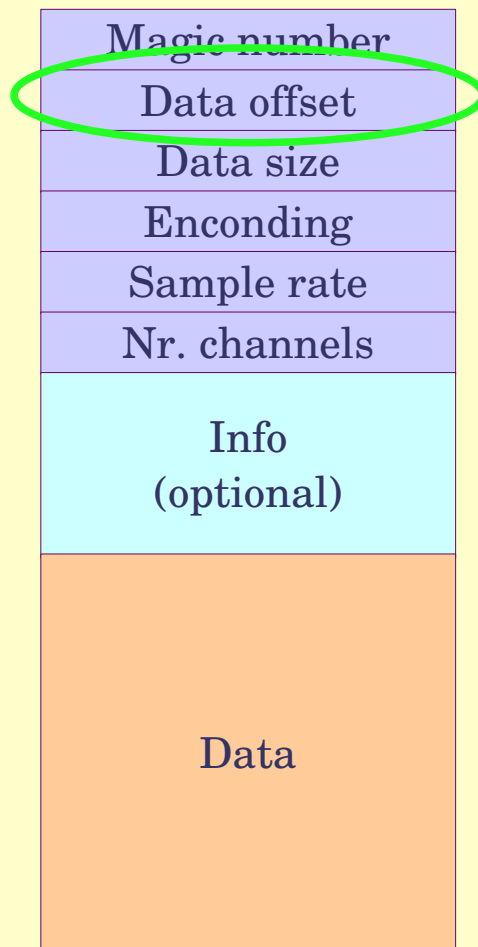


O valor 0x2e736e64 referente aos 4 caracteres “.snd”

Introdução à Informática

Representação digital de dados (Parte II)

Áudio: arquivos AU



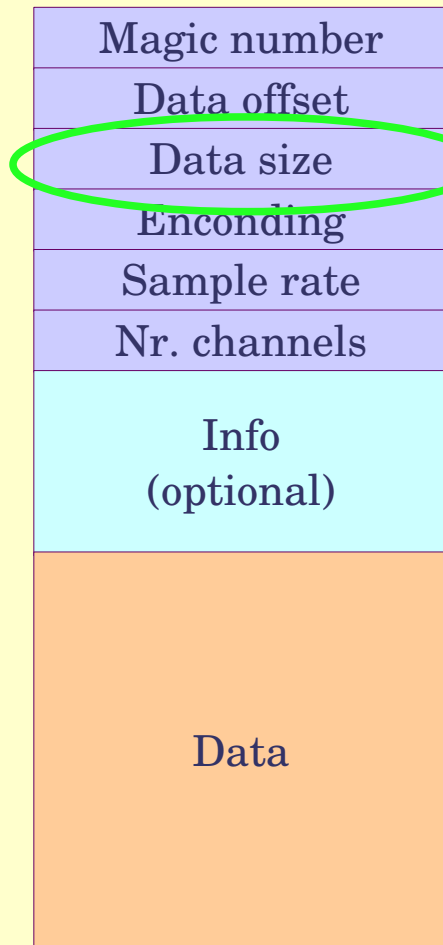
O offset para os dados do início do arquivo até o início dos dados “data”. O valor mínimo é 0x00000018 (24 decimal), já que o cabeçalho tem 6 words de 32 bits (6 x 4 bytes), sem bloco de informação.



Introdução à Informática

Representação digital de dados (Parte II)

Áudio: arquivos AU

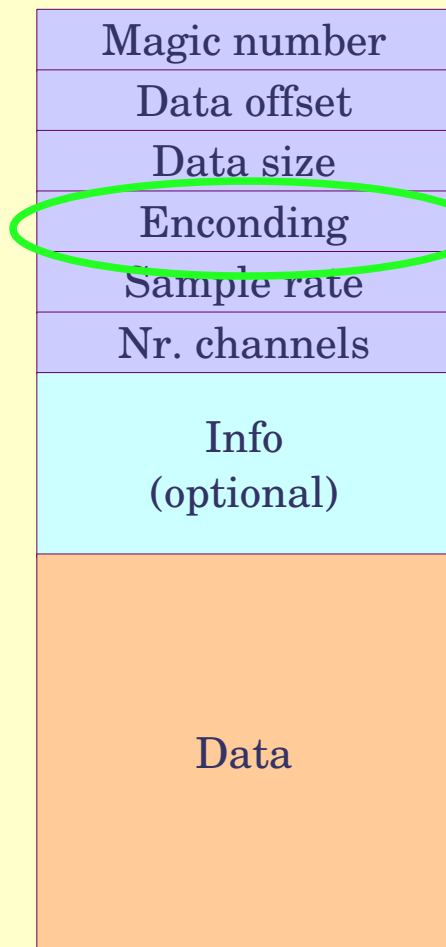


O tamanho do bloco de dados em bytes.
Se for desconhecido o valor deste campo
deve ser 0xffffffff.

Introdução à Informática

Representação digital de dados (Parte II)

Áudio: arquivos AU



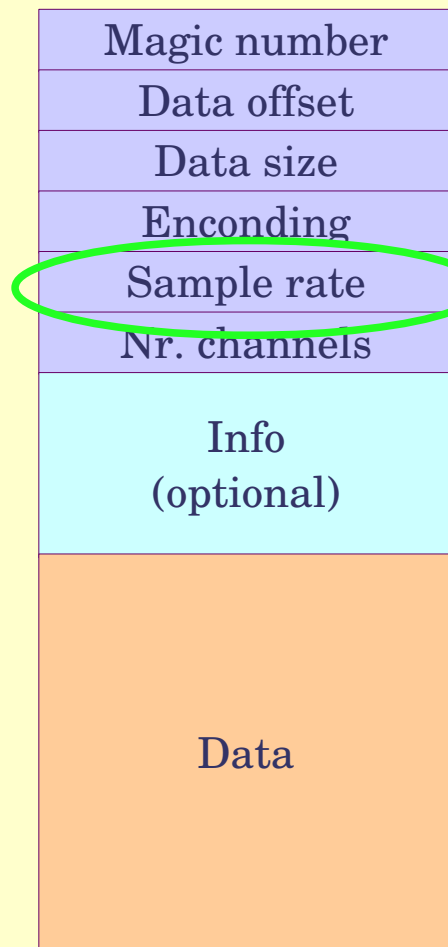
- * 1 = 8-bit G.711 μ -law
- * 2 = 8-bit linear PCM
- * 3 = 16-bit linear PCM
- * 4 = 24-bit linear PCM
- * 5 = 32-bit linear PCM
- * 6 = 32-bit IEEE floating point
- * 7 = 64-bit IEEE floating point
- * 8 = Fragmented sample data
- * 9 = DSP program
- * 10 = 8-bit fixed point
- * 11 = 16-bit fixed point
- * 12 = 24-bit fixed point
- * 13 = 32-bit fixed point
- * 18 = 16-bit linear with emphasis
- * 19 = 16-bit linear compressed
- * 20 = 16-bit linear with emphasis and compression
- * 21 = Music kit DSP commands
- * 23 = 4-bit ISDN u-law compressed
- * 24 = ITU-T G.722 ADPCM
- * 25 = ITU-T G.723 3-bit ADPCM
- * 26 = ITU-T G.723 5-bit ADPCM
- * 27 = 8-bit G.711 A-law



Introdução à Informática

Representação digital de dados (Parte II)

Áudio: arquivos AU

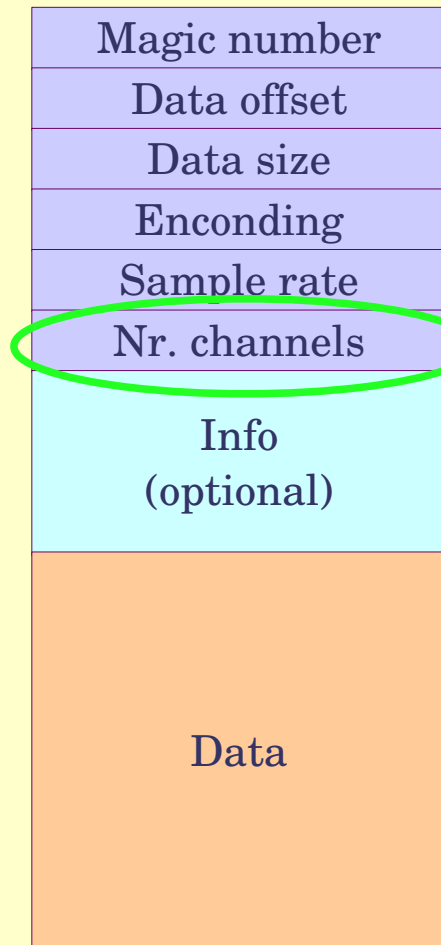


O número de amostras / seg. (p. ex. 8000, 48000, 96000).

Introdução à Informática

Representação digital de dados (Parte II)

Áudio: arquivos AU

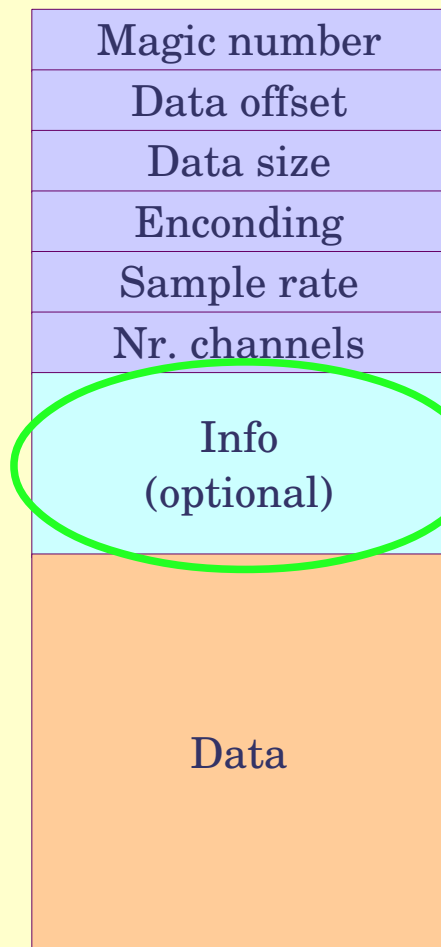


O número de canais de áudio

Introdução à Informática

Representação digital de dados (Parte II)

Áudio: arquivos AU



Outras informações (dados opcionais)
Deve ser múltiplo não zero de 8 bytes
(32 bits) e deve terminar com pelo menos
1 byte de zeros.



Introdução à Informática

Representação digital de dados (Parte II)

Áudio:

OGG audio CODECs

Lossy:

Speex: baixa resolução (8-32 kbits/s/canal)

Vorbis: média a alta resolução (16 a 500 kbits/s/canal)

Lossless:

FLAC: Audio de alta fidelidade.