

# **Representação da informação**

## **22667 – Organização e Recuperação da Informação**

Jander Moreira

UFSCar – DC

9 de Agosto de 2018



## Símbolos e codificação

## Representações comuns

- Números inteiros

- Números reais

- Valores lógicos

- Cadeias de caracteres

## Sumário

# Representação da informação

## └ Símbolos e codificação



## Símbolos

- Desenhos
- Letras
- Dígitos
- E **bits**

# Representação da informação

## └ Símbolos e codificação

### Representações:

O valor  
*cento e vinte e  
oito*

- ▶ Cento de vinte e oito (texto)
- ▶ 80 (hexadecimal)
- ▶ 200 (octal)
- ▶ 10000000 (binário)
- ▶ 128 (decimal)
- ▶ ρκη (grego antigo)
- ▶ CXXVIII (romano)

Várias representações para *cento e vinte e oito*:

- Cento de vinte e oito (texto)
- 80 (hexadecimal)
- 200 (octal)
- 10000000 (binário)
- 128 (decimal)
- ρκη (grego antigo)
- CXXVIII (romano)

# Representação da informação

## └ Símbolos e codificação

### Código

Conjunto de símbolos e regras para interpretá-los.

### Código

Conjunto de símbolos e regras para interpretá-los.

# Representação da informação

## └ Símbolos e codificação

### Símbolos

0 1 2 3 4 5 6 7 8 9 0 , + -

### Regras

- ▶ Representação posicional
- ▶ Potências de 10
- ▶ Vírgula como separador decimal

Exemplo: números decimais

# Representação da informação

## └ Símbolos e codificação

342,37

$$\begin{aligned} 342,37 &= 3 \times 10^2 \\ &+ 4 \times 10^1 \\ &+ 2 \times 10^0 \\ &+ 3 \times 10^{-1} \\ &+ 7 \times 10^{-2} \end{aligned}$$

Um exemplo de número decimal

**342,37**

$$\begin{aligned} 342,37 &= 3 \times 10^2 \\ &+ 4 \times 10^1 \\ &+ 2 \times 10^0 \\ &+ 3 \times 10^{-1} \\ &+ 7 \times 10^{-2} \end{aligned}$$

# Representação da informação

## └ Símbolos e codificação

### **Símbolos**

I V X L C D M

### **Regras**

- ▶ Valores fixos para os símbolos
- ▶ Composição posicional aditiva ou subtrativa

Exemplo: números romanos



# Representação da informação

## └ Símbolos e codificação

**MCMLXVII**

M = 1000	= 1000
CM = 1000 - 100	= 900
LX = 50 + 10	= 60
VII = 5 + 1 + 1	= 7

MCMLXVII = 1967

Um exemplo de número romano: MCMLXVII

**MCMLXVII**

$$M = 1000 \qquad = 1000$$

$$CM = 1000 - 100 \qquad = 900$$

$$LX = 50 + 10 \qquad = 60$$

$$VII = 5 + 1 + 1 \qquad = 7$$

$$MCMLXVII = 1967$$

# Representação da informação

- └ Representações comuns
  - └ Números inteiros

## Representação binária

Uso dos bits para representar base 2.

## Representação binária

Uso dos bits para representar base 2.

# Representação da informação

- └ Representações comuns
  - └ Números inteiros

$$397 = 110001101_2$$

00000001 10001101

## Representação de valor inteiro com 2 bytes

$$397 = 110001101_2$$

00000001 10001101

# Representação da informação

- └ Representações comuns
  - └ Números inteiros

$$-397 = -110001101_2$$

$$\begin{array}{r} 11111110 \ 01110010 \\ +1 \\ \hline 11111110 \ 01110011 \end{array}$$

## Representação de valor inteiro com 2 bytes

$$-397 = -110001101_2$$

$$\begin{array}{r} 11111110 \ 01110010 \\ +1 \\ \hline 11111110 \ 01110011 \end{array}$$

# Representação da informação

- └ Representações comuns
  - └ Números inteiros

## BCD

Cada dígito é codificado pelo valor binário.

## BCD

Cada dígito é codificado pelo valor binário.

# Representação da informação

- Representações comuns

- Números inteiros

**5049**

5	0	4	9
↓	↓	↓	↓
0101	0000	0100	1001

0101000001001001

**5049**

5	0	4	9
↓	↓	↓	↓
0101	0000	0100	1001

0101000001001001

# Representação da informação

- └ Representações comuns
  - └ Números reais

## Ponto flutuante

Representação binária de um valor com parte decimal.

- ▶  $s$ : sinal
- ▶  $e$ : expoente
- ▶  $m$ : matissa

$$v = s \times 2^e \times m$$

## Ponto flutuante

Representação binária de um valor com parte decimal.

- ▶  $s$ : sinal
- ▶  $e$ : expoente
- ▶  $m$ : matissa

$$v = s \times 2^e \times m$$

# Representação da informação

## └ Representações comuns

### └ Números reais

IEEE 754 (precisão simples)

- ▶  $s$ : sinal (1bit)
- ▶  $e$ : expoente (8 bits)
- ▶  $m$ : matissa (23 bits)

## Padrão IEEE 754.

### IEEE 754 (precisão simples)

- ▶  $s$ : sinal (1bit)
- ▶  $e$ : expoente (8 bits)
- ▶  $m$ : matissa (23 bits)



# Representação da informação

## └ Representações comuns

### └ Números reais

-65,5

11000010100000110000000000000000

1	10000101	000001100000000000000000
↓	133	bits 6 e 7
↓	↓	↓
-1	127 - 133	$1 + 2^{-6} + 2^{-7}$
(s)	(e)	(m)

Exemplo de interpretação do 65,5 no padrão IEEE 754.

-65,5

11000010100000110000000000000000

1	10000101	000001100000000000000000
1	133	bits 6 e 7
↓	↓	↓
-1	127 - 133	$1 + 2^{-6} + 2^{-7}$
(s)	(e)	(m)

# Representação da informação

- └ Representações comuns

- └ Números reais

$$\begin{aligned}v &= s \times 2^e \times m \\&= (-1) \times (2^{133-127}) \times (1 + 2^{-6} + 2^{-7}) \\&= (-1) \times (64) \times (1,0234375) \\&= -65,5\end{aligned}$$

$$\begin{aligned}v &= s \times 2^e \times m \\&= (-1) \times (2^{133-127}) \times (1 + 2^{-6} + 2^{-7}) \\&= (-1) \times (64) \times (1,0234375) \\&= -65,5\end{aligned}$$

# Representação da informação

- └ Representações comuns
  - └ Valores lógicos

Lógico

Verdadeiro ou falso.

**Lógico**

Verdadeiro ou falso.

## Representação da informação

- └ Representações comuns
  - └ Valores lógicos

- ▶ Todos os bits 0: **false**
- ▶ Algum bit diferente de zero: **verdadeiro**

Um ou mais bytes podem ser usados para armazenar um valor lógico.

- ▶ Todos os bits 0: **false**
- ▶ Algum bit diferente de zero: **verdadeiro**

# Representação da informação

- └ Representações comuns
  - └ Valores lógicos

## Mapas de bits

Cada bit representa um valor lógico individual.

## Mapas de bits

Cada bit representa um valor lógico individual.

# Representação da informação

- └ Representações comuns

- └ Cadeias de caracteres

## Com terminador

A cadeia de caracteres é finalizada com um código sentinela.

## Com terminador

A cadeia de caracteres é finalizada com um código sentinela.

# Representação da informação

- └ Representações comuns
  - └ Cadeias de caracteres

"disciplina"

100 105 115 99 105 112 108 105 110 97 0  
64 69 73 63 69 70 6c 69 6e 61 00

**"disciplina"**

100 105 115 99 105 112 108 105 110 97 0  
64 69 73 63 69 70 6c 69 6e 61 00

# Representação da informação

- └ Representações comuns
  - └ Cadeias de caracteres

## **Prefixo de tamanho**

A cadeia de caracteres é prefixada pelo número de bytes que contém.

## **Prefixo de tamanho**

A cadeia de caracteres é prefixada pelo número de bytes que contém.



# Representação da informação

- └ Representações comuns
  - └ Cadeias de caracteres

"disciplina"

10 d i s c i p l i n a  
10 100 105 115 99 105 112 108 105 110 97  
0a 64 69 73 63 69 70 6c 69 6e 61

**"disciplina"**

10 d i s c i p l i n a  
10 100 105 115 99 105 112 108 105 110 97  
0a 64 69 73 63 69 70 6c 69 6e 61

## Representação da informação

- └ Representações comuns
  - └ Cadeias de caracteres



### Onde usar essas codificações:

- Na memória principal
- Na memória secundária
- Na transmissão de dados
- etc.

## Sumário

- Codificação:
  - Conjunto de símbolos
  - Regras de uso dos símbolos
- Computação: representações usuais para dados

# Representação da informação

## └ Créditos

Jander Moreira

<http://www.dc.ufscar.br/~jander>

[jander@dc.ufscar.br](mailto:jander@dc.ufscar.br)

Universidade Federal de São Carlos

<http://www.ufscar.br>

Departamento de Computação

<http://www.dc.ufscar.br>

# Jander Moreira

<http://www.dc.ufscar.br/~jander>

[jander@dc.ufscar.br](mailto:jander@dc.ufscar.br)

# Universidade Federal de São Carlos

<http://www.ufscar.br>

# Departamento de Computação

<http://www.dc.ufscar.br>