EDCO4B ESTRUTURAS DE DADOS 2

Aula 08 - Conceitos sobre Estruturas de Arquivos

Prof. Rafael G. Mantovani

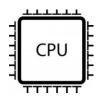


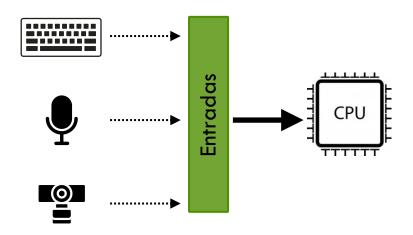
Roteiro

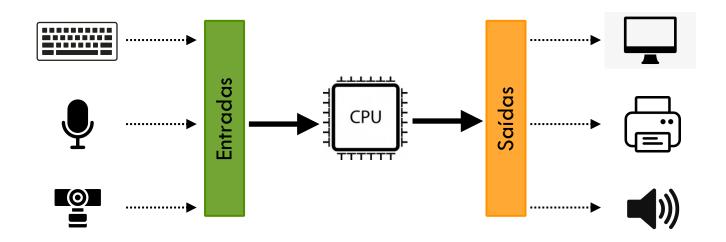
- 1 Introdução
- 2 Estruturas de Campos (*Fields*)
- 3 Estruturas de Registros (*Records*)
- 4 Exercícios
- 5 Revisão
- 6 Referências

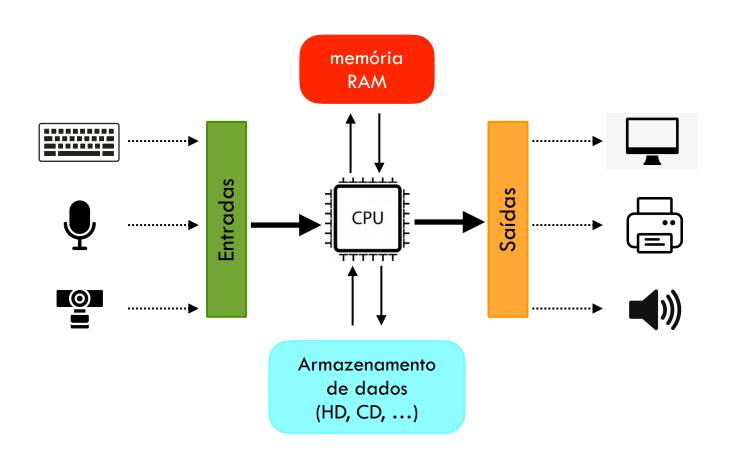
Roteiro

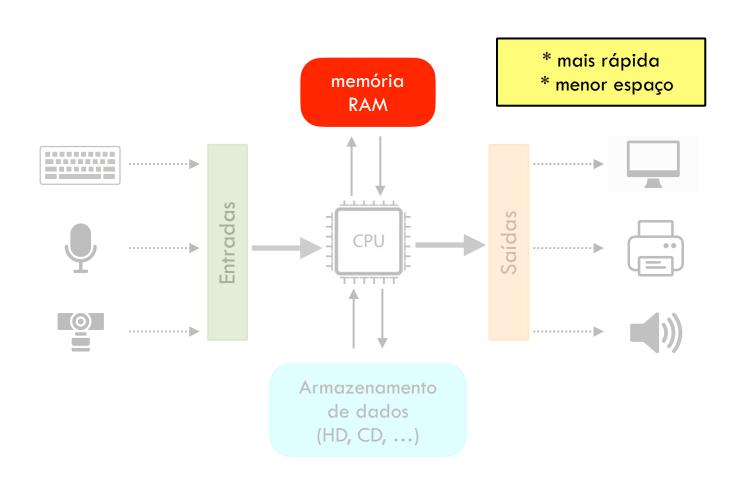
- 1 Introdução
- 2 Estruturas de Campos (*Fields*)
- 3 Estruturas de Registros (*Records*)
- 4 Exercícios
- 5 Revisão
- 6 Referências

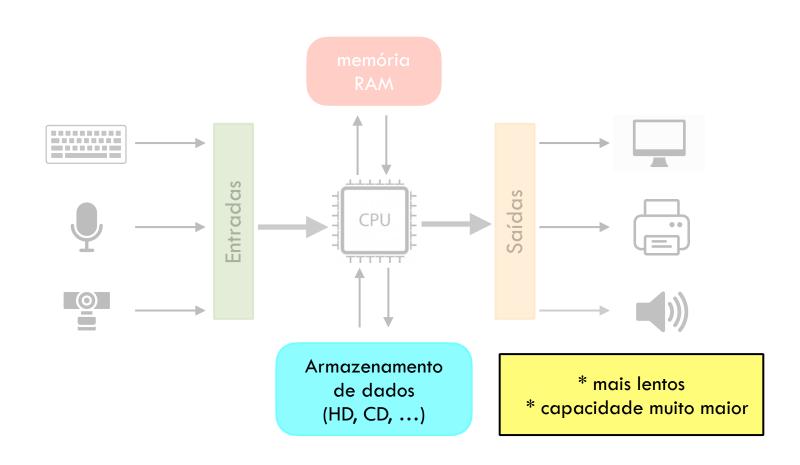


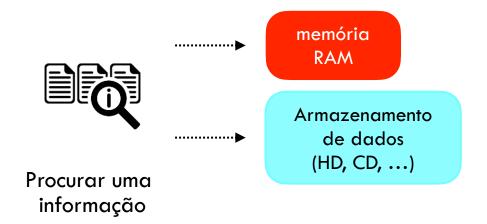


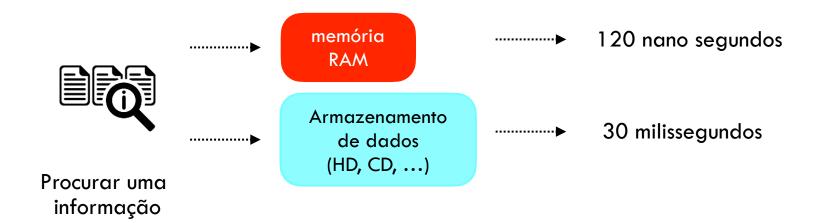




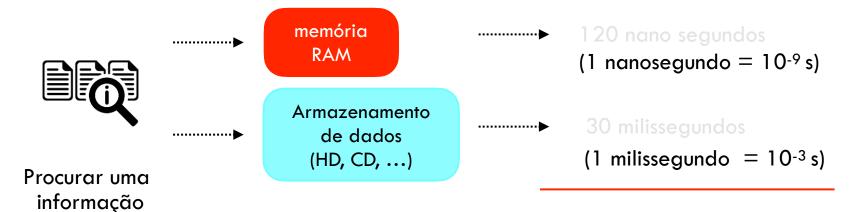








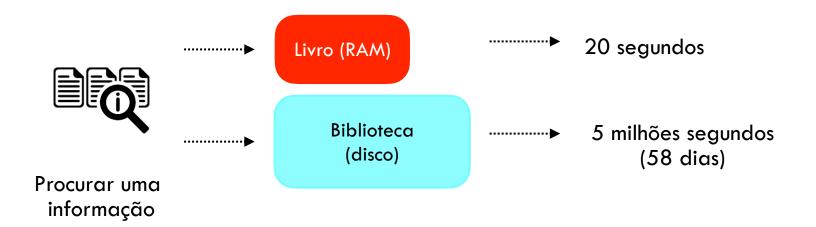




1 milissegundo = 10^6 nanosegundos

1 nano: 1 milhão x mais rápido!

mantendo as devidas proporções de tempo "real" ...



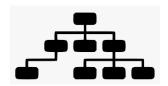
Idealmente:



Idealmente:



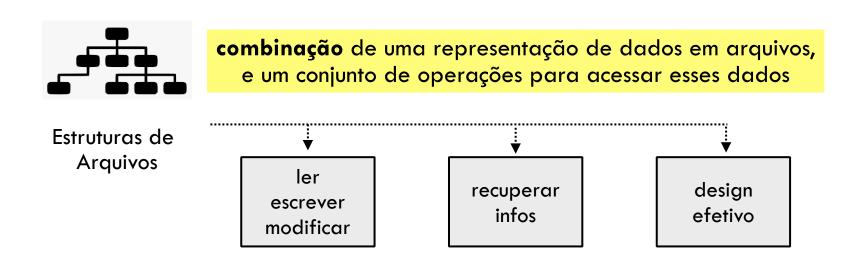
Precisamos de ...:

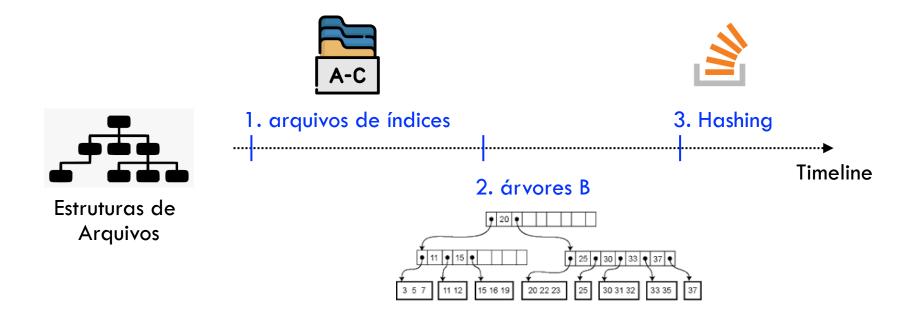


combinação de uma representação de dados em arquivos, e um conjunto de operações para acessar esses dados

Estruturas de Arquivos

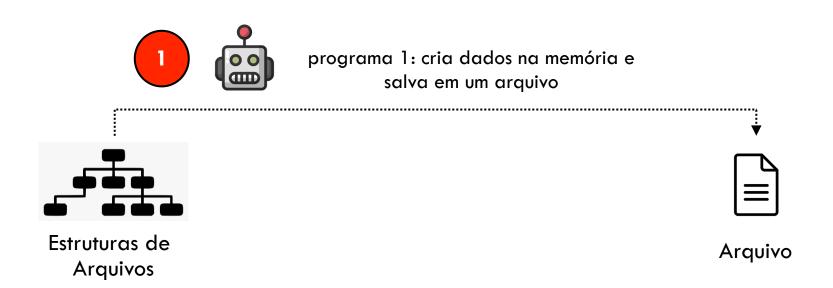
Precisamos de ...:

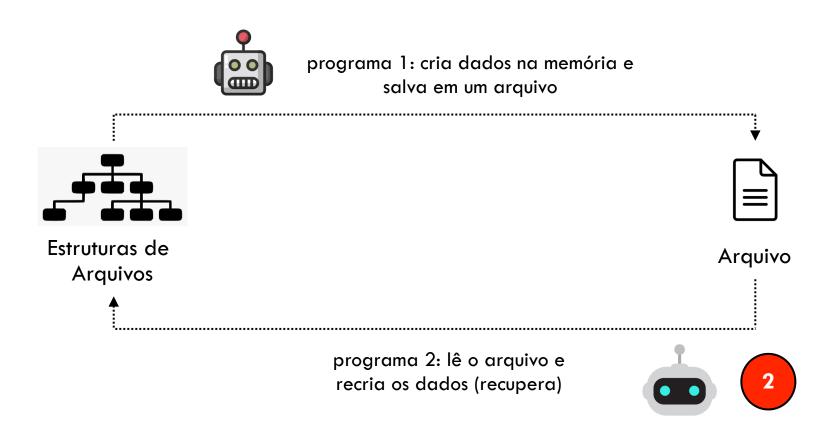


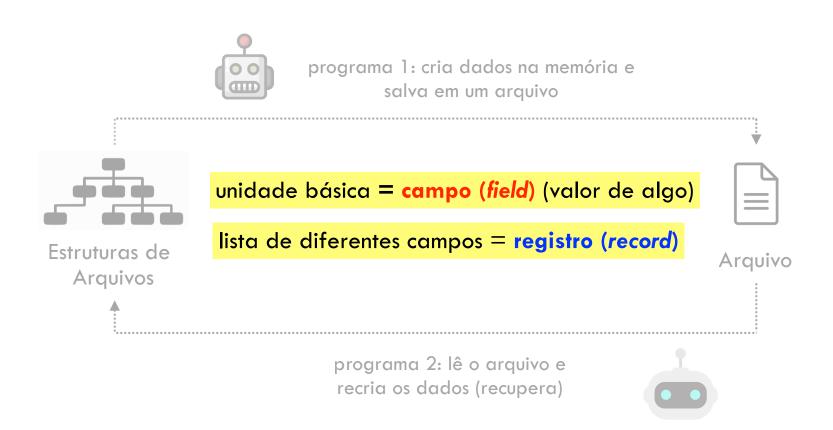


Roteiro

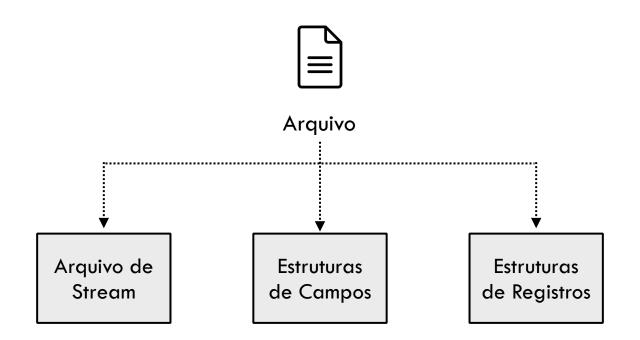
- 1 Introdução
- **2** Estruturas de Campos (*Fields*)
- 3 Estruturas de Registros (Records)
- 4 Exercícios
- 5 Revisão
- 6 Referências



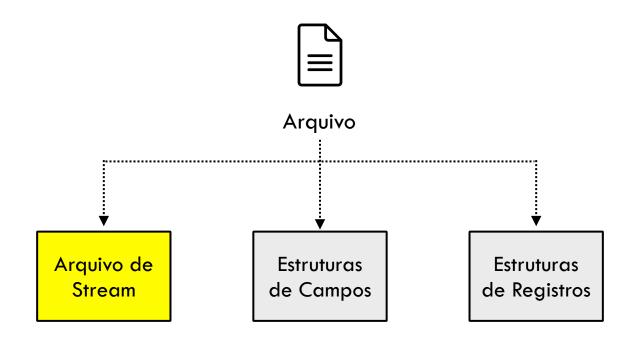




Diferentes formas de representar a organização de arquivos



Diferentes formas de representar a organização de arquivos



Info: coleção de nomes e endereços

Info: coleção de nomes e endereços

Senhor L Rua Incognito X Apucarana PR 86812-460

Pessoa 1

Info: coleção de nomes e endereços

Senhor L Rua Incognito X Apucarana PR 86812-460

Pessoa 1

Tamara
SausageWater
Rua Compiler 277
Londrina PR
86057-970

Pessoa 2

Info: coleção de nomes e endereços

Senhor L Rua Incognito X Apucarana PR 86812-460

Pessoa 1

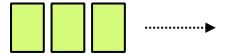
Tamara SausageWater Rua Compiler 277 Londrina PR 86057-970

Pessoa 2

Mauri Chan Rua Controles 69 Apucarana PR 86800-680

Pessoa N

Info: coleção de nomes e endereços



Info: coleção de nomes e endereços

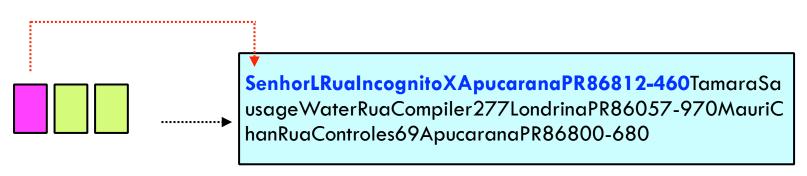


Info: coleção de nomes e endereços

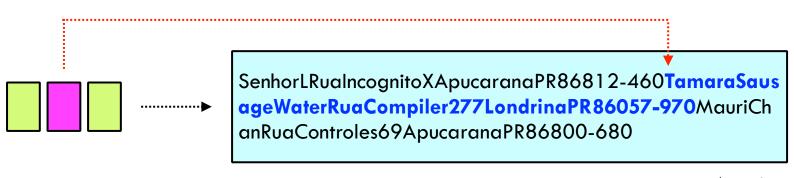


SenhorLRualncognitoXApucaranaPR86812-460TamaraSaus ageWaterRuaCompiler277LondrinaPR86057-970MauriCh anRuaControles69ApucaranaPR86800-680

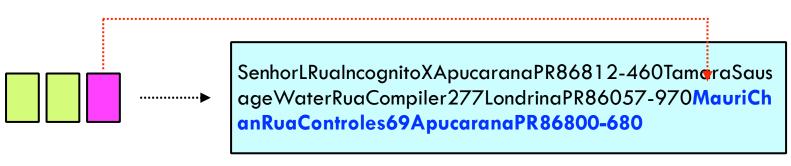
Info: coleção de nomes e endereços



Info: coleção de nomes e endereços



Info: coleção de nomes e endereços



Info: coleção de nomes e endereços



* Problema

Arquivo de Stream

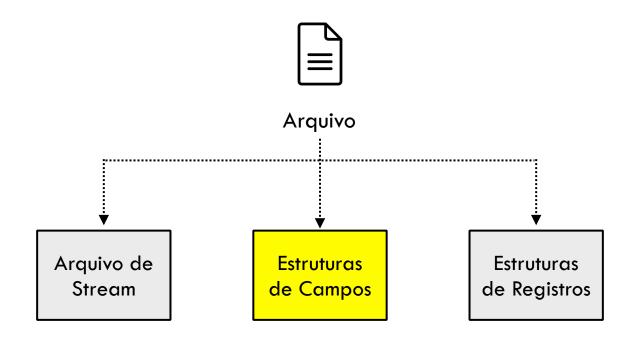
Info: coleção de nomes e endereços



SenhorLRualncognitoApucaranaPR86812-460TamaraSausa geWaterRuaCompilerLondrinaPR86057-970MauriChanRua ControlesApucaranaPR86800-680

- * Problema: perdemos a integridade dos dados
 - Não conseguimos ler novamente na mesma estrutura inicial
- * Precisamos: organizar o arquivo de alguma forma "inteligente"

Diferentes formas de representar a organização de arquivos



Pessoa

Tamara SausageWater Rua Compiler 277 Londrina PR 86057-970

Pessoa

Tamara
SausageWater
Rua Compiler 277
Londrina PR
86057-970

Nome Sobrenome Endereço Cidade Estado CEP

Pessoa

Tamara
SausageWater
Rua Compiler 277
Londrina PR
86057-970

Nome Sobrenome Endereço Cidade Estado CEP

Pessoa

Classe/Objeto (Python)

Tamara
SausageWater
Rua Compiler 277
Londrina PR
86057-970

Nome
Sobrenome
Endereço
Cidade Estado
CEP

class Pessoa:

self.nome
self.sobrenome
self.endereco
self.cidade
self.estado
self.cep



- 1. Forçar os campos a terem tamanhos fixos
- 2. Começar cada campo com um indicador de tamanho
- 3. Usar delimitadores ao fim de cada campo
- 4. Usar expressões keyword=valor



- 1. Forçar os campos a terem tamanhos fixos
- 2. Começar cada campo com um indicador de tamanho
- 3. Usar delimitadores ao fim de cada campo
- 4. Usar expressões keyword=valor

 Funcionamento: tamanhos fixos, podemos prever e recuperar a informação contando os bytes

 Funcionamento: tamanhos fixos, podemos prever e recuperar a informação contando os bytes

Classe/Objeto (Python)

```
class Pessoa:

self.nome
self.sobrenome
self.endereco
self.cidade
self.estado
self.cep
```

 Funcionamento: tamanhos fixos, podemos prever e recuperar a informação contando os bytes

Classe/Objeto (Python)

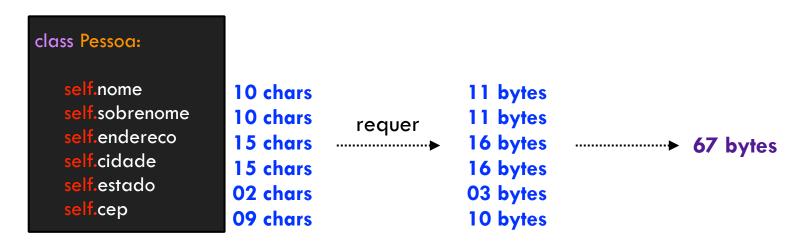
```
class Pessoa:

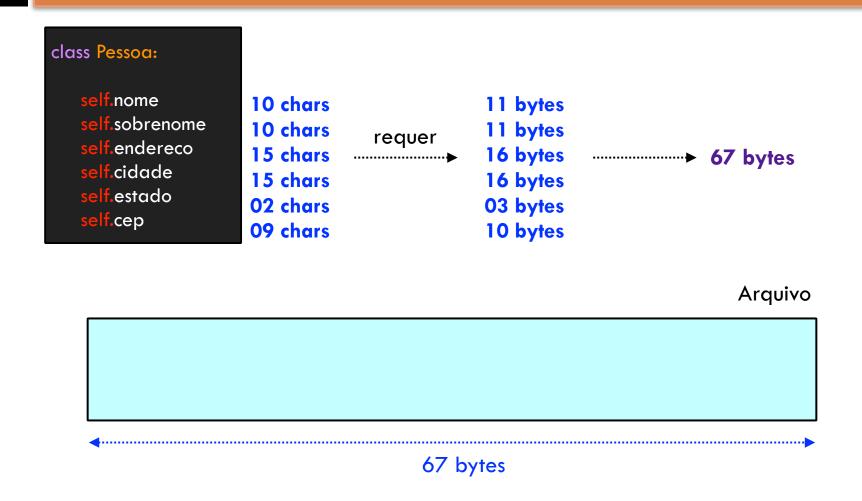
self.nome
self.sobrenome
self.endereco
self.cidade
self.estado
self.cep

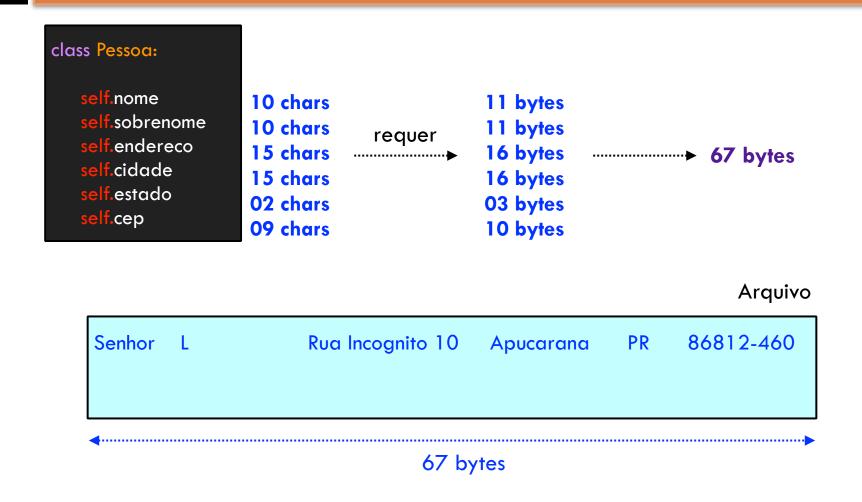
10 chars
10 chars
15 chars
15 chars
02 chars
```

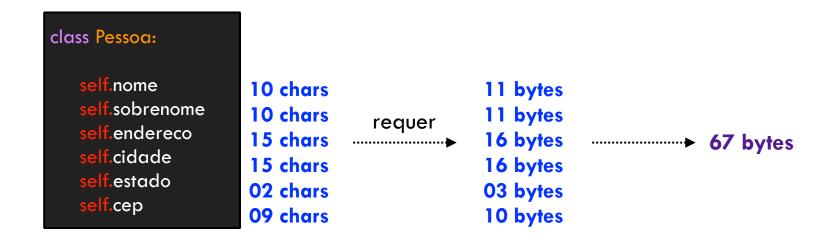
 Funcionamento: tamanhos fixos, podemos prever e recuperar a informação contando os bytes

Classe/Objeto (Python)





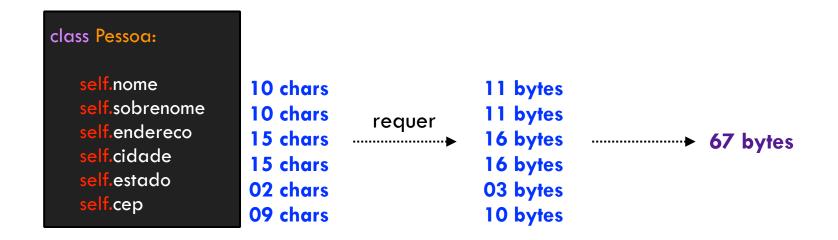




Arquivo



67 bytes



Arquivo



67 bytes

```
10 chars
                                        11 bytes
                 10 chars
                                        11 bytes
                             requer
                 15 chars
                                        16 bytes
                           ....
                                                 ----- 67 bytes
                 15 chars
                                        16 bytes
                 02 chars
                                        03 bytes
                 09 chars
                                        10 bytes
                                                                Arquivo
* Desvantagens:
```

67 bytes

self.nome self.sobrenome self.endereco self.cidade self.estado self.cep 10 chars 10 chars 15 chars 15 chars 15 chars 02 chars 09 chars

```
requer

11 bytes

11 bytes

16 bytes

16 bytes

03 bytes

10 bytes
```

Arquivo

* Desvantagens:

- usar os campos de tamanhos fixos torna o arquivo maior
- inapropriado para dados de grande variabilidade (nomes e endereços)





- 1. Forçar os campos a terem tamanhos fixos
- 2. Começar cada campo com um indicador de tamanho
- 3. Usar delimitadores ao fim de cada campo
- 4. Usar expressões keyword=valor

Funcionamento: manter o tamanho do campo logo a frente dele

Funcionamento: manter o tamanho do campo logo a frente dele

Senhor L Rua Incognito X Apucarana PR 86812-460

Funcionamento: manter o tamanho do campo logo a frente dele

Senhor L Rua Incognito X Apucarana PR 86812-460

Arquivo

Funcionamento: manter o tamanho do campo logo a frente dele

Senhor L Rua Incognito X Apucarana PR 86812-460

Arquivo

Senhor L Rua Incognito X Apucarana PR 86812-460

Funcionamento: manter o tamanho do campo logo a frente dele

Senhor L Rua Incognito X Apucarana PR 86812-460

Arquivo

06Senhor 01L15Rua Incognito X09Apucarana 02PR 0986812-460

Funcionamento: manter o tamanho do campo logo a frente dele

Senhor L Rua Incognito X Apucarana PR 86812-460 Tamara SausageWater Rua Compiler 277 Londrina PR 86057-970

Arquivo

O6SenhorO1L15Rua Incognito XO9ApucaranaO2PRO986812-460

Funcionamento: manter o tamanho do campo logo a frente dele

Senhor L Rua Incognito X Apucarana PR 86812-460 Tamara SausageWater Rua Compiler 277 Londrina PR 86057-970

Arquivo

06Senhor**01L15**Rua Incognito X**09**Apucarana**02**PR**09**86812-460**06**Tamara**13** SausageWater**15**RuaCompiler 277**08**Londrina**02**PR**08**860*5*7-970 ...



- 1. Forçar os campos a terem tamanhos fixos
- 2. Começar cada campo com um indicador de tamanho
- 3. Usar delimitadores ao fim de cada campo
- 4. Usar expressões keyword=valor

 Funcionamento: usar caracteres especiais para separar os campos (vazio, \n, Tab, etc)

□ **Funcionamento**: usar caracteres especiais para separar os campos (vazio, \n, Tab, etc)



 Funcionamento: usar caracteres especiais para separar os campos (vazio, \n, Tab, etc)

Senhor L Rua Incognito X Apucarana PR 86812-460

Arquivo

Senhor | L | Rua Incognito X | Apucarana | PR | 86812-460 |

 □ Funcionamento: usar caracteres especiais para separar os campos (vazio, \n, Tab, etc)

Senhor L Rua Incognito X Apucarana PR 86812-460 Tamara
SausageWater
Rua Compiler 277
Londrina PR
86057-970

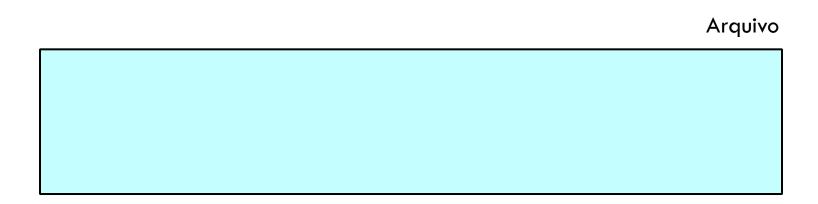
Arquivo

Senhor | L | Rua Incognito X | Apucarana | PR | 86812-460 | Tamara | SausageWater | RuaCompiler 277 | Londrina | PR | 86057-970 ...



- 1. Forçar os campos a terem tamanhos fixos
- 2. Começar cada campo com um indicador de tamanho
- 3. Usar delimitadores ao fim de cada campo
- 4. Usar expressões keyword=valor

Funcionamento: usar expressão explícita de chave e valor



Funcionamento: usar expressão explícita de chave e valor

Senhor L Rua Incognito X Apucarana PR 86812-460

Arquivo

nome=Senhor | sobrenome=L | endereco=Rua Incognito X | cidade=Apucarana | estado=PR | cep=86812-460 |

Funcionamento: usar expressão explícita de chave e valor

Senhor L Rua Incognito X Apucarana PR 86812-460 Tamara
SausageWater
Rua Compiler 277
Londrina PR
86057-970

Arquivo

```
nome=Senhor | sobrenome=L | endereco=Rua Incognito X | cidade=Apucarana | estado=PR | cep=86812-460 | nome=Tamara | sobrenome=SausageWater | endereco=Rua Compiler 277 | cidade=Londrina | estado=PR | cep=86057-970 | ...
```

Funcionamento: usar expressão explícita de chave e valor

* Vantagens: Desvantagem:

Método 4: Expressão chave-valor

Funcionamento: usar expressão explícita de chave e valor

Senhor L Rua Incognito X Apucarana PR 86812-460 Tamara SausageWater Rua Compiler 277 Londrina PR 86057-970

* Vantagens:

- o próprio campo apresenta informação do dado
- bom para manipular eventuais valores ausentes

estad nome

* Desvantagem:

- gasta muito espaço de dados, por causa das chaves

Arquivo arana |

```
it(object):
self.hit_points = kwan
self.current_hit_points
self.level = kwargs.gut
attack(self, enemy: """
Attack enemy unit. Ret
                    Vamos Codificar!
```

Exercícios

 1) Implemente funções de leitura/escrita de Pessoa(s) usando todos os métodos descritos.

```
def escritaTamanhoFixo(arquivo, Pessoa)
def leituraTamanhoFixo(arquivo)

def escritaTamanhoCampos(arquivo, Pessoa)
def leituraTamanhosCampos(arquivo)

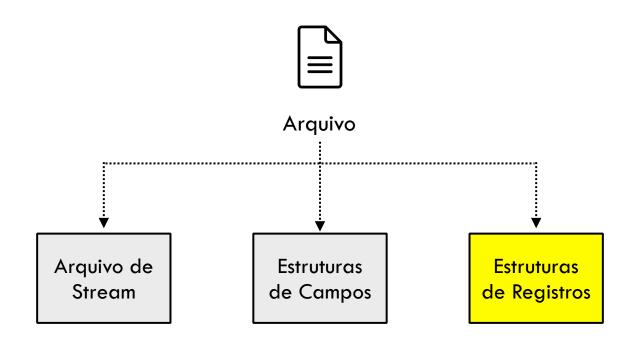
def escritaDelimitador(arquivo, Pessoa)
def leituraDelimitador(arquivo)

def escritaKeywords(arquivo, Pessoa)
def leituraKeywords(arquivo)
```

Roteiro

- 1 Introdução
- 2 Estruturas de Campos (*Fields*)
- 3 Estruturas de Registros (Records)
- 4 Exercícios
- 5 Revisão
- 6 Referências

Diferentes formas de representar a organização de arquivos





Conjunto de campos que representam um conceito/objeto

Registro

Escrever um registro Salvar o estado de um objeto

Ler um registro Recuperar o estado de um objeto

Objeto reside na ... Memória

Registro reside em ... Arquivos



- 1. Registros de tamanhos fixos
- 2. Registros com uma quantidade de campos
- Registro
- 3. Começar cada registro com a quantidade de bytes
- 4. Usar segundo arquivo mantendo o endereço inicial de cada registro
- 5. Usar delimitadores ao final de cada registro



Registro

- 1. Registros de tamanhos fixos
- 2. Registros com uma quantidade de campos
- 3. Começar cada registro com a quantidade de bytes
- 4. Usar segundo arquivo mantendo o endereço inicial de cada registro
- 5. Usar delimitadores ao final de cada registro

- Todos registros tem o mesmo tamanho de bytes
- Um dos métodos mais usados
- Registro de tamanho fixo não implica em um número fixo de campos no registro

- Todos registros tem o mesmo tamanho de bytes
- Um dos métodos mais usados
- Registro de tamanho fixo não implica em um número fixo de campos no registro
 - a) registros com tamanhos fixos com campos de tamanhos fixo

Apucarana PR 86821-460 Senhor L Rua Incógnito X Registro 1

- Todos registros tem o mesmo tamanho de bytes
- Um dos métodos mais usados
- Registro de tamanho fixo não implica em um número fixo de campos no registro
 - a) registros com tamanhos fixos com campos de tamanhos fixo

Senhor	L	Rua Incógnito X	Apucarana	PR 86821-460	Registro
Tamara	SausageWater	Rua Compiler 27	77 Londrina	PR 86057-970	Registro

- Todos registros tem o mesmo tamanho de bytes
- Um dos métodos mais usados
- Registro de tamanho fixo não implica em um número fixo de campos no registro
 - a) registros com tamanhos fixos com campos de tamanhos fixo

```
SenhorLRua IncógnitoApucaranaPR86821-460TamaraSausageWaterRua Compiler 277LondrinaPR86057-970
```

b) registros com tamanhos fixos com campos tamanhos variados

- Todos registros tem o mesmo tamanho de bytes
- Um dos métodos mais usados
- Registro de tamanho fixo não implica em um número fixo de campos no registro
 - a) registros com tamanhos fixos com campos de tamanhos fixo

```
SenhorLRua IncógnitoApucaranaPR86821-460TamaraSausageWaterRua Compiler 277LondrinaPR86057-970
```

b) registros com tamanhos fixos com campos tamanhos variados

```
Senhor | L | Rua Incógnito X | Apucarana | PR | 86821-460 |
Tamara | Sausage Water | Rua Compiler 277 | Londrina | PR | 86057-970 |
```

- Todos registros tem o mesmo tamanho de bytes
- Um dos métodos mais usados
- Registro de tamanho fixo não implica em um número fixo de campos no registro
 - a) registros com tamanhos fixos com campos de tamanhos fixo

```
SenhorLRua IncógnitoApucaranaPR86821-460TamaraSausageWaterRua Compiler 277LondrinaPR86057-970
```

b) registros com tamanhos fixos com campos tamanhos variados

```
Senhor | L | Rua Incógnito X | Apucarana | PR | 86821-460 | 
Tamara | SausageWater | Rua Compiler 277 | Londrina | PR | 86057-970 |
```

Dados não usados



Registro

- 1. Registros de tamanhos fixos
- 2. Registros com uma quantidade de campos
- 3. Começar cada registro com a quantidade de bytes
- 4. Usar segundo arquivo mantendo o endereço inicial de cada registro
- 5. Usar delimitadores ao final de cada registro

Método 2: número fixo de campos

Registros vão conter um número fixo de campos

Método 2: número fixo de campos

Registros vão conter um número fixo de campos

Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 |

Arquivo

6 campos compõem um registro

Método 2: número fixo de campos

Registros vão conter um número fixo de campos

Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 | Tamara | SausageWater | Rua Compiler 277 | Londrina | PR | 86057-970



- 1. Registros de tamanhos fixos
- 2. Registros com uma quantidade de campos
- Registro
- 3. Começar cada registro com a quantidade de bytes
- 4. Usar segundo arquivo mantendo o endereço inicial de cada registro
- 5. Usar delimitadores ao final de cada registro

Método 3: indicador de tamanho

Cada registro começa com um indicativo de tamanho em bytes

48Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 |

Método 3: indicador de tamanho

Cada registro começa com um indicativo de tamanho em bytes

48Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 | 58Tamara | SausageWater | Rua Compiler 277 | Londrina | PR | 86057-970 |



- 1. Registros de tamanhos fixos
- 2. Registros com uma quantidade de campos
- Registro
- 3. Começar cada registro com a quantidade de bytes
- 4. Usar segundo arquivo mantendo o endereço inicial de cada registro
- 5. Usar delimitadores ao final de cada registro

 Usar um arquivo adicional para guardar o endereço inicial de cada registro

 Usar um arquivo adicional para guardar o endereço inicial de cada registro

Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 | Tamara | Sausage Water | Rua Compiler 277 | Londrina | PR | 86057-970 |

 Usar um arquivo adicional para guardar o endereço inicial de cada registro

Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 | Tamara | Sausage Water | Rua Compiler 277 | Londrina | PR | 86057-970 |

Arquivo

Arquivo de indices

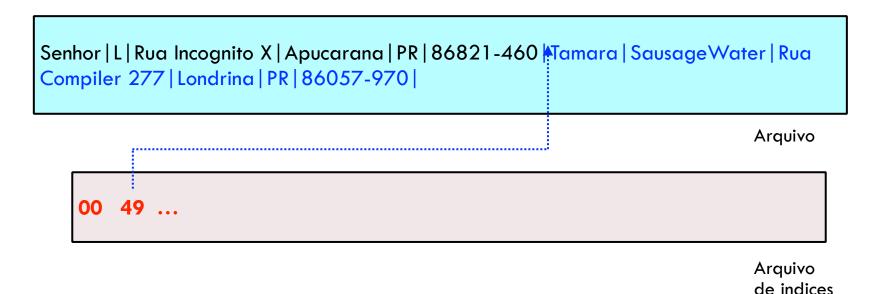
Usar um arquivo adicional para guardar o endereço inicial de cada registro

Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 | Tamara | Sausage Water | Rua | Compiler 277 | Londrina | PR | 86057-970 |

Arquivo

Arquivo de indices

Usar um arquivo adicional para guardar o endereço inicial de cada registro





1. Registros de tamanhos fixos

2. Registros com uma quantidade de campos

Registro

3. Começar cada registro com a quantidade de bytes

4. Usar segundo arquivo mantendo o endereço inicial de cada registro

5. Usar delimitadores ao final de cada registro

Método 5: delimitadores

 No nível de registros, usar um caracter ao fim de cada registro 	•

Método 5: delimitadores

No nível de registros, usar um caracter ao fim de cada registro

Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 | #

Método 5: delimitadores

No nível de registros, usar um caracter ao fim de cada registro

Senhor | L | Rua Incognito X | Apucarana | PR | 86821-460 | #Tamara | Sausage Water | Rua Compiler 277 | Londrina | PR | 86057-970 | #

Roteiro

- 1 Introdução
- 2 Estruturas de Campos (*Fields*)
- 3 Estruturas de Registros (Records)
- 4 Exercícios
- 5 Revisão
- 6 Referências

Precisamos de um Conjunto de Dados:

Precisamos de um Conjunto de Dados:



Dataset: Crunchyroll Animes

```
it(object):
self.hit_points = kman
self.current_hit_point
self.level = kwargs.gat
attack(self, enemy: "limit
Attack enemy unit. Ret
                    Vamos Codificar!
```

Exercícios

 2) Implemente funções de escrita e leitura de registros para Animes usando os seguintes métodos:

```
# Guarde os objetos em texto, simulando uma escrita de bytes

# método 1: tamanhos fixos de registros

# método 2: quantidade definida de campos

# método 3: quantidade de bytes/tamanho antes de cada registro/campo

# método 4: arquivo de índices

# método 5: usando delimitadores
```

Roteiro

- 1 Introdução
- 2 Estruturas de Campos (*Fields*)
- 3 Estruturas de Registros (Records)
- 4 Exercícios
- 5 Revisão
- 6 Referências

Revisão

- O menor nível de organização de um arquivo é um stream de bytes
- Campos (fields): pedaços fundamentais de informação
- Campos são agrupados em registros (objeto)
- Reconhecer campos/registros requer impor uma estrutura de organização nos arquivos

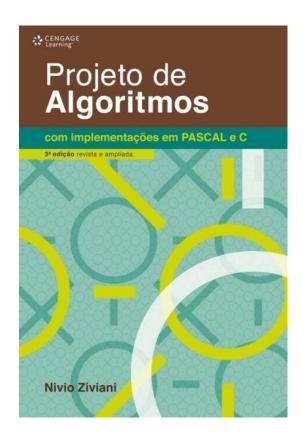
Revisão

- Estruturas: existem algumas opções de organizar informação
 - Fixar tamanho dos campos/registros
 - Começar campos/registros com indicadores de tamanho
 - Usar delimitadores para dividir os campos/registros

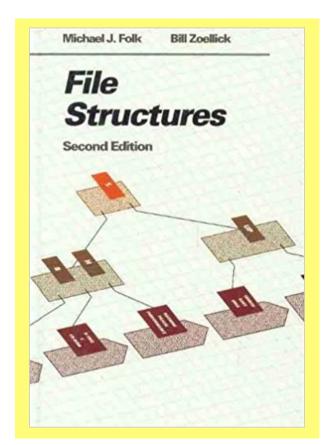
Roteiro

- 1 Introdução
- 2 Estruturas de Campos (*Fields*)
- 3 Estruturas de Registros (Records)
- 4 Exercícios
- 5 Revisão
- 6 Referências

Referências sugeridas

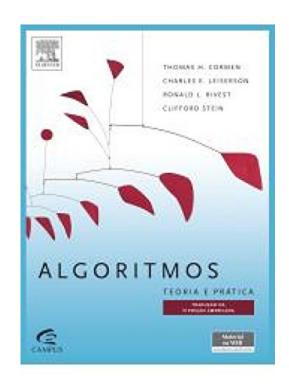


[Ziviani, 2010]



[Folk & Zoellick, 1992]

Referências sugeridas



[Cormen et al, 2018]



[Drozdek, 2017]

Perguntas?

Prof. Rafael G. Mantovani

rafaelmantovani@utfpr.edu.br