

FUNDAMENTOS DE PROGRAMAÇÃO

Arquivos

Prof. Bruno Góis Mateus

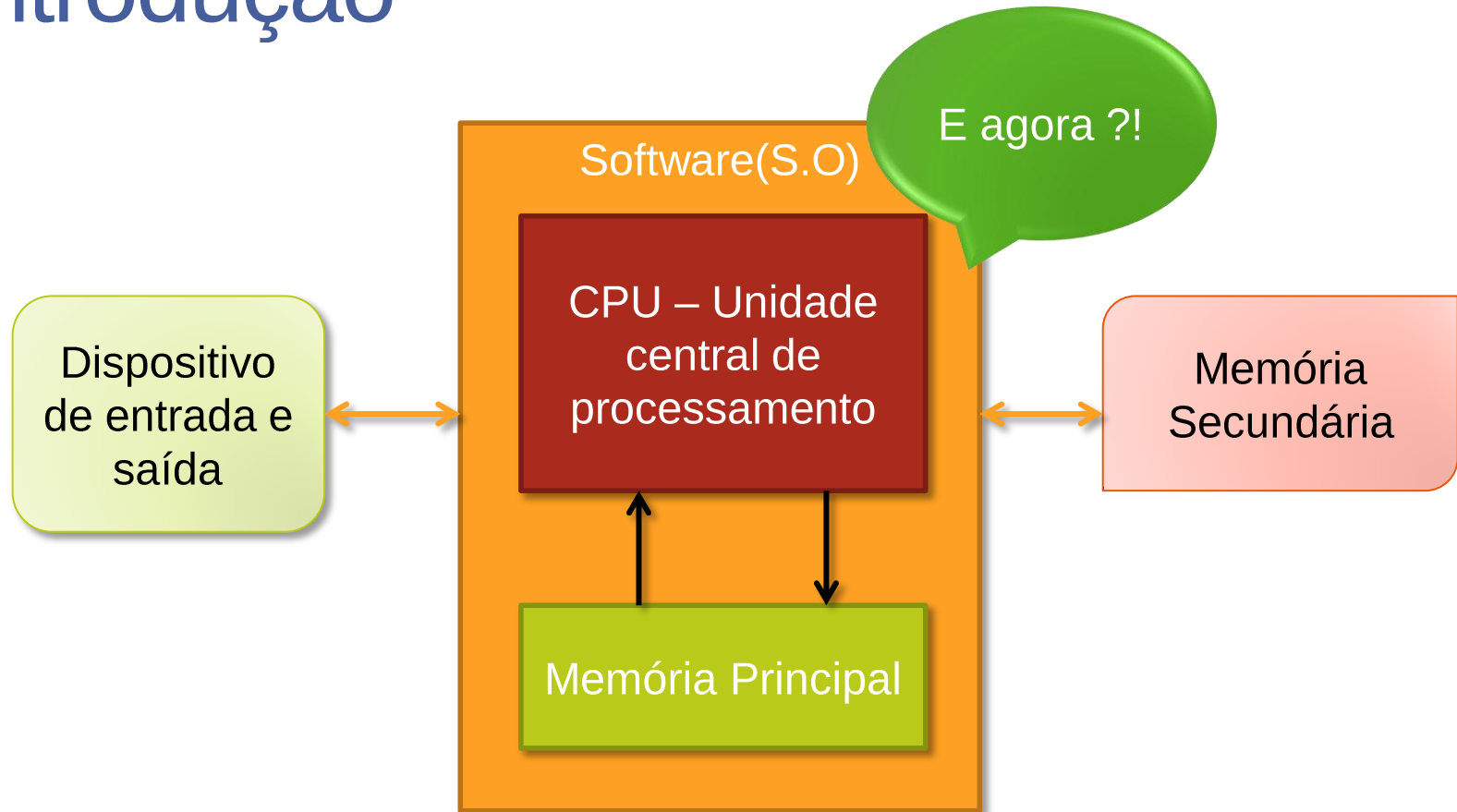


Índice

- Introdução
- Trabalhando com arquivos
- Lendo um arquivo
- Buscando em um arquivo
- Escrevendo arquivos de texto

INTRODUÇÃO

Introdução



Introdução

- Diversos programas utilizam arquivos
 - Editor de imagem
 - Navegadores
 - Processadores de texto
 - Players de música
- Nós iremos trabalhar com arquivos de texto

TRABALHANDO COM ARQUIVOS

Trabalhando com arquivos

- Antes de ler o conteúdo do arquivo, devemos indicar qual arquivo queremos utilizar
- Isso é realizado através da função `open`
 - Retorna um “`file handle`”
 - Variável utilizada para realizar operações em um arquivo
 - Similar ao “`File -> Open`” em um programa qualquer

Trabalhando com arquivos

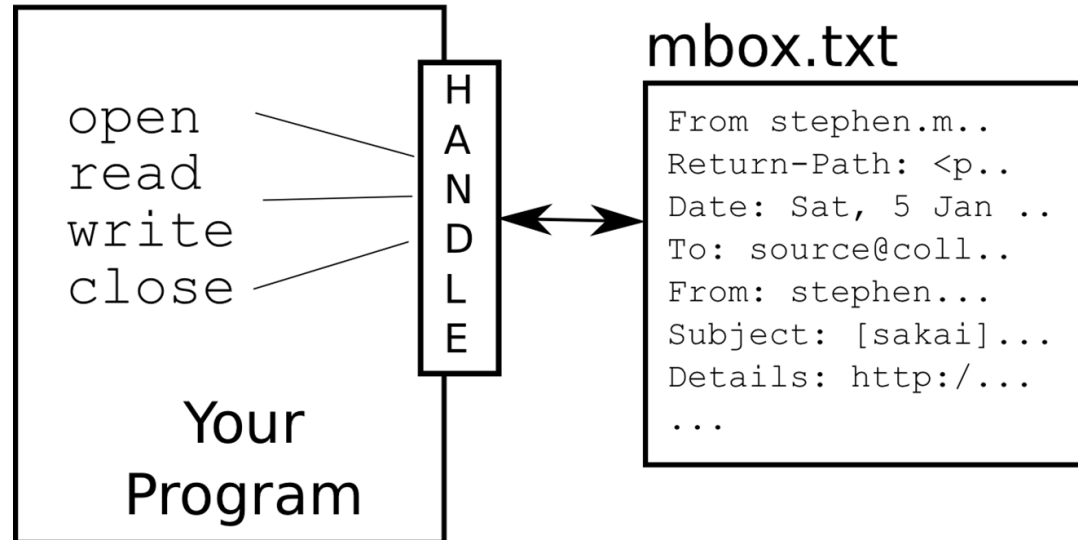
- Usando a função **open**
 - Recebe dois parâmetros
 - String: Nome do arquivo, incluindo o path
 - String: Modo de acesso
 - r : Se você pretende ler o conteúdo do arquivo
 - w : Se você pretende inserir conteúdo no arquivo
 - Retorna um handle para manipular um arquivo

Handle

- Ao abrir um arquivo o seu conteúdo não é lido automaticamente
 - A capacidade do disco rígido é bem maior que a memória
- Handle é uma espécie de conexão entre o programa e o arquivo que vai ser manipulado

Handle

```
>>> fhand = open('mbox.txt')  
>>> print fhand  
<open file 'mbox.txt', mode 'r' at  
0x1005088b0>
```



LENDO UM ARQUIVO

Lendo um arquivo

- Um arquivo pode ser considerado uma sequência de linhas
 - Um arquivo de texto o final de cada linha identificada pelo caractere especial

From stephen.marquard@uct.ac.za Sat Jan 5 09:14:16 2008\n

Return-Path: <postmaster@collab.sakaiproject.org>\n

Date: Sat, 5 Jan 2008 09:12:18 -0500\n

To: source@collab.sakaiproject.org\n

From: stephen.marquard@uct.ac.za\n

Subject: [sakai] svn commit: r39772 - content/branches/\n

\n

Details: <http://source.sakaiproject.org/viewsvn/?view=rev&rev=39772>\n

Lendo um arquivo

- Nós utilizamos um caractere especial para indicar que uma linha terminou
 - `newline`
 - `\n`
 - Apenas um caractere

```
>>> stuff =  
'Hello\nWorld!'  
>>> stuff  
'Hello\nWorld!'  
>>> print stuff  
Hello  
World!  
>>> stuff = 'X\nY'  
>>> print stuff  
X  
Y  
>>> len(stuff)  
3
```

Lendo um arquivo

- Um handle pode ser considerado uma sequência de string
- Podemos usar a instrução for para percorrer um arquivo
 - Conjunto ordenado
 - Lista, String

Contando as linhas em um arquivo

1. Abrir o arquivo como modo de leitura
2. Usar um laço para percorrer cada linha
3. Contar as linhas
4. Imprimir a quantidade de linhas

```
fhand = open('mbox.txt')  
count = 0  
for line in fhand:  
    count = count + 1  
print 'Line Count:', count
```

Lendo o arquivo inteiro

- Podemos armazenar o conteúdo inteiro de um arquivo em uma string
 - Depende do tamanho do arquivo

```
>>> fhand = open('mbox-short.txt')
>>> inp = fhand.read()
>>> print len(inp)94626
>>> print inp[:20]
From stephen.marquar
```


Outras maneira de ler um arquivo

Nome do método	Uso	Explicação
<code>read(n)</code>	<code>handle.read()</code>	Retorna uma string com n caracteres ou o arquivo inteiro quando o n não é informado
<code>readline(n)</code>	<code>handle.readline()</code>	Retorna o texto da próxima linha incluindo o caractere de final de linha. Se o n é informado, apenas n caracteres são retornados.
<code>readlines(n)</code>	<code>handle.readlines()</code>	Retorna uma lista de string, onde cada elemento da lista representa uma linha do arquivo. Se o n é informado apenas n caracteres são retornados. O valor é arredondado para toda linha venha completa.

BUSCANDO EM UM ARQUIVO

Buscando em uma arquivo

- Podemos adicionar alguns condicionais durante a varredura do arquivo para que apenas parte do arquivo seja impresso

```
fhand = open('mbox-short.txt')
for line in fhand:
    if line.startswith('From:') :
        print line
```

Buscando em uma arquivo

- Podemos adicionar alguns condicionais durante a varredura do arquivo para que apenas parte do arquivo seja impresso

```
fhand = open('mbox-short.txt')
for line in fhand:
    if line.startswith('From:') :
        print line
```

Buscando em arquivo

- OPS

- Cada linha possui um caractere de final de linha
- Ao usar o `print` outro caractere de final de linha é inserido

```
From: stephen.marquard@uct.ac.za\n
```

```
\n
```

```
From: louis@media.berkeley.edu\n
```

```
\n
```

```
From: zqian@umich.edu\n
```

```
\n
```

```
From: rjlowe@iupui.edu\n
```

```
\n
```

Buscando em uma arquivo

- Podemos caracteres espaços vazios do final de uma string usando a função `rstrip()`
 - O caractere de final de linha é considerado um espaço vazio

```
fhand = open('mbox-short.txt')
for line in fhand:
    line = line.rstrip()
    if line.startswith('From:') :
        print line
```

```
From: stephen.marquard@uct.ac.za
From: louis@media.berkeley.edu
From: zqian@umich.edu
From: rjlowe@iupui.edu
```

Buscando em uma arquivo

- Podemos pular linha que não são de nosso interesse utilizando a instrução **continue**

```
fhand = open('mbox-short.txt')
for line in fhand:
    line = line.rstrip()
    if not line.startswith('From:') :
        continue
    print line
```

Buscando em uma arquivo

- Podemos verificar se uma linha qualquer possui uma string

- Critério de seleção

```
fhand = open('mbox-short.txt')  
for line in fhand:  
    line = line.rstrip()  
    if not '@uct.ac.za' in line :  
        continue  
    print line
```

From stephen.marquard@uct.ac.za Sat Jan 5 09:14:16 2008

X-Authentication-Warning: set sender to stephen.marquard@uct.ac.za using -f

From: stephen.marquard@uct.ac.za

Author: stephen.marquard@uct.ac.za

From david.horwitz@uct.ac.za Fri Jan 4 07:02:32 2008

X-Authentication-Warning: set sender to david.horwitz@uct.ac.za using

ESCREVENDO ARQUIVOS DE TEXTO

Escrevendo arquivos de texto

- Precisamos abrir o arquivo em modo de **escrita**
 - Utilizamos a função **open**
 - Indicamos o modo de escrita passando a string **'w'** como segundo parâmetro
 - Caso for o interesse ler e modificar o modo informado deve ser **'r+'**, **'w+'** ou **'a+'**

Escrevendo arquivos de texto

First Name	Last Name	Position	Team	Completions	Attempts	Yards	TDs
Colt	McCoy	QB	CLE	135	222	1576	6
Josh	Freeman	QB	TB	291	474	3451	25
Michael	Vick	QB	PHI	233	372	3018	21
Matt	Schaub	QB	HOU	365	574	4370	24

.
.
.
.
.

Escrevendo arquivos de texto

```
infile = open("qbdata.txt", "r")  
outfile = open("qbnames.txt", "w")
```

```
aline = infile.readline()  
while aline:  
    items = aline.split()  
    dataline = items[1] + ',' + items[0]  
    outfile.write(dataline + '\n')  
    aline = infile.readline()
```

```
infile.close()  
outfile.close()
```

McCoy,Colt
Freeman,Josh
Vick,Michael
Schaub,Matt
Rivers,Philip
Hasselbeck,Matt
Clausen,Jimmy
Flacco,Joe
Orton,Kyle
Campbell,Jason
.
.
.

O que vem por aí

- Férias

