

# **UNIVERSIDADE DE SÃO PAULO - USP**

Instituto de Ciências Matemáticas e de Computação - ICMC

Bacharelado em Ciências de Computação

Disciplina Organização de Arquivos

Profa. Dra. Cristina Dutra de Aguiar Ciferri

André Moreira Souza - 9778985

Carlos André Martins Neves - 8955195

Guilherme Amorim Menegali - 8531289

Vitor Trevelin Xavier da Silva - 9791285

TRABALHO PRÁTICO - PARTE 2

SÃO CARLOS

2 de Julho de 2017

# SUMÁRIO

1.	Estrutura e Importação de Dados	2
2.	Detalhes Sobre as Funcionalidades	2
2.1.	Importar o Arquivo de Dados	2
2.2.	Gerar Arquivo de Índices	3
2.3.	Remover Registro	3
2.4.	Inserir Registro	3
2.5.	Visualização de estatísticas sobre os arquivos de índice	4
2.6.	Visualização de estatísticas sobre as listas de registros removidos	5
3.	Capturas de Tela da Interface com o Usuário	5
4.	Bateria de Testes	7

## 1. Estrutura e Importação de Dados

Após uma análise de como os dados são estruturados, foi implementado nos arquivos “*reg\_delim.c*” e “*reg\_delim.h*” funções e estruturas que aplicam os conceitos de organização de arquivos para registros de tamanho variável utilizando delimitadores entre registros e indicadores de tamanho para os campos de tamanho variável.

A seguir, apresenta-se uma tabela que representa um registro do arquivo de saída:

Tabela 1: Registro genérico

Campo (Tipo de dados)	Tamanho do campo
Documento (char *)	(19 bytes) // Tamanho fixo
DataHoraAtualiza (char *)	(19 bytes) // Tamanho fixo
DataHoraCadastro (char *)	(19 bytes) // Tamanho fixo
Ticket (int)	(4 bytes) // Tamanho fixo
Dominio (char *)	Tamanho variável
Nome (char *)	Tamanho variável
Cidade (char *)	Tamanho variável
UF (char *)	Tamanho variável

## 2. Detalhes Sobre as Funcionalidades

### 2.1. Importar o Arquivo de Dados

Para a leitura do arquivo de entrada, os campos foram lidos individualmente, aproveitando-se dos delimitadores ‘;’ e ‘\n’ entre registros. Desta maneira amenizados o grande problema do uso de delimitadores, já que a escolha do caractere delimitador não pode estar no domínio de caracteres utilizados (no caso, é problemático se o usuário utilizar o caractere ‘;’ na hora da inserção de um registro).

Em seguida os dados são alocados na memória, escritos nos três arquivos de saída: “1”, “2” e “3”, conforme a especificação e, por fim, liberados.

## **2.2. Gerar Arquivo de Índices**

Com os três arquivos de saída prontos, o usuário pode solicitar esta função para a criação dos três arquivos de índice (um para cada arquivo de saída).

Inicialmente os arquivos são iguais, pois todos são índices primários no campo ticket e os arquivos de saída são iguais, entretanto, quando o usuário começa a fazer inserções e remoções, as diferenças começam a surgir devido às distintas formas de tratamento da fragmentação interna.

## **2.3. Remover Registro**

Depois que o usuário informa o ticket do registro a ser removido, é feita uma busca em cada arquivo de índices e em sequência o registro é removido de cada arquivo de saída. Por fim os arquivos de índices são atualizados.

A remoção lógica é aplicada, em que o registro removido é marcado com uma flag sinalizando a remoção e a possível reaproveitamento desse espaço em uma futura inserção. Para isso lê-se o topo da pilha dos registros removidos no cabeçalho do arquivo. Atualiza-se o topo para apontar para o registro removido. E no registro removido escreve-se a flag que sinaliza remoção, o tamanho do registro e o offset do próximo registro da pilha de registros removidos (antigo topo da pilha).

## **2.4. Inserir Registro**

Para todos arquivos a inserção tenta solucionar a fragmentação interna reaproveitando os espaços de registros removidos. Para todos arquivos, armazena-se uma lista de registros removidos. O topo dessa lista é armazenado no cabeçalho do arquivo. A inserção, nos três arquivos, primeiro percorre essa lista em busca de um espaço para inserir o registro novo. A diferença está, como é pedido na especificação

do trabalho, na estratégia de alocação utilizada para decidir qual espaço utilizar.

Para o arquivo de saída “1”, a estratégia de alocação *first-fit* foi utilizada. Ao percorrer a lista de registros removidos, caso um espaço com tamanho suficiente é encontrado, ele é utilizado imediatamente e a inserção acaba. Caso contrário, o registro é inserido no fim do arquivo de dados.

Para o arquivo de saída “2”, a estratégia de alocação *best-fit* foi utilizada. Percorre-se a lista de registros removidos inteiramente e verifica-se qual é o espaço que, se utilizado, minimiza a fragmentação secundária. Se tal espaço existe, ele é utilizado. Caso contrário, o registro é inserido no fim do arquivo de dados.

Para o arquivo de saída “3”, a estratégia de alocação *worst-fit* foi utilizada. Percorre-se a lista de registros removidos inteiramente e verifica-se qual é o espaço que, se utilizado, maximiza a fragmentação secundária. Se tal espaço existe, ele é utilizado. Caso contrário, o registro é inserido no fim do arquivo de dados.

## **2.5. Visualização de estatísticas sobre os arquivos de índice**

Para visualizar estatísticas sobre os arquivos de índice, criamos uma opção que o usuário pode selecionar e, ou navegar os registros dos índices um por um, ou gerar uma tabela ASCII com as informações pedidas.

Na navegação dos registros um por um, como os três índices sempre terão o mesmo número de registros, os registros de mesmo número de cada um dos arquivos são mostrados lado a lado, um por vez. O usuário tem as opções de ir para o próximo, anterior, um número específico ou voltar ao menu (como pode ser observado na Seção 3, Figura 4).

Na tabela, os três índices aparecem lado a lado (como pode ser observado na Seção 3, Figura 5). Para cada índice, é mostrada a quantidade total de registros armazenada no arquivo de índice e, em seguida, todos os registros do índice com seus campos ticket e offset.

## 2.6. Visualização de estatísticas sobre as listas de registros removidos

Para a visualização, optamos por montar uma tabela com os dados relativos a lista de registros removidos de cada arquivo, mostrando o byte offset e o tamanho do registro removido, além do byte offset do próximo registro.

## 3. Capturas de Tela da Interface com o Usuário

Figura 1: Menu de opções que o usuário pode optar

```
gcc -o trab main.c reg_delim.c -Wall -std=c99
./trab
1. Importar arquivo de dados.
2. Gerar arquivos de indice.
3. Remover registro.
4. Inserir registro.
5. Visualizar estatisticas(arquivo de indice)
6. Visualizar estatisticas(arquivo de dados)
7. Sair
Opcao: █
```

É válido observar que para a correta execução do programa, ao iniciar devemos utilizar as opções 1 e 2 antes de realizar qualquer outra operação.

Figura 2: Exemplo de Inserção

```
Opcao: 4
documento(max: 19): 000.073.957/0001-67
dataHoraCadastro(max: 19): 27/07/1996 12:00:07
dataHoraAtualiza(max: 19): 27/07/1996 12:00:17
ticket(int): 157
dominio: orkut.gov.br
nome: orkut
cidade: Rio de Janeiro
uf: Rio de Janeiro
Registro inserido com sucesso.
```

Figura 3: Exemplo de remoção bem sucedida

```
Opcao: 3
Digite o ticket do registro a ser removido: 157

Registro removido com sucesso.
```

Figura 4: Exemplo de visualização de estatísticas sobre os arquivos de índice, um registro por vez

```
Opcao: 5

Visualizar um por vez? (s/n): s
```

Indice 1				Indice 2				Indice 3			
Entradas: 200				Entradas: 200				Entradas: 200			
Atual: 1				Atual: 1				Atual: 1			
#	Ticket	Offset		Ticket	Offset		Ticket	Offset		Ticket	Offset
	0	798		0	798		0	798		0	798

```

1. Anterior
2. Proximo
3. Ir para
4. Voltar ao menu

Opcao: 
```

Figura 5: Exemplo de visualização de estatísticas sobre os arquivos de índice, todos em uma tabela

```
Opcao: 5
```

TABELA DE INDICES											
Indice 1				Indice 2				Indice 3			
Entradas: 200				Entradas: 200				Entradas: 200			
#	Ticket	Offset		Ticket	Offset		Ticket	Offset		Ticket	Offset
1	0	798		0	798		0	798		0	798
2	1	2076		1	2076		1	2076		1	2076
3	2	8237		2	8237		2	8237		2	8237
4	3	10498		3	10498		3	10498		3	10498
5	5023	16		5023	16		5023	16		5023	16
6	5153	1213		5153	1213		5153	1213		5153	1213
7	5443	2850		5443	2850		5443	2850		5443	2850
8	6787	26100		6787	26100		6787	26100		6787	26100
9	11611	21008		11611	21008		11611	21008		11611	21008
10	12785	2536		12785	2536		12785	2536		12785	2536
11	20961	20711		20961	20711		20961	20711		20961	20711
12	21425	6495		21425	6495		21425	6495		21425	6495

Figura 6: Exemplo de visualização de estatísticas sobre as listas de registros removidos

Opcao: 6

TABELA DE REMOVIDOS						
	Indice 1		Indice 2		Indice 3	
	Removidos: 1		Removidos: 1		Removidos: 1	
	Inicio da Lista: 29681		Inicio da Lista: 29681		Inicio da Lista: 29681	
#	tamanho	offset	tamanho	offset	tamanho	offset
1	( 123,	NULL)	( 123,	NULL)	( 123,	NULL)

Pressione enter para retornar ao menu.

#### 4. Bateria de Testes

Figura 7: Situação inicial da bateria de testes

Opcao: 6

TABELA DE REMOVIDOS									
Indice 1			Indice 2			Indice 3			
Removidos: 0			Removidos: 0			Removidos: 0			
Inicio da Lista: NULL			Inicio da Lista: NULL			Inicio da Lista: NULL			
#	tamanho	offset	tamanho	offset	tamanho	offset			

Pressione enter para retornar ao menu.

1. Importar arquivo de dados.

2. Gerar arquivos de indice.

3. Remover registro.

4. Inserir registro.

5. Visualizar estatisticas(arquivo de indice)

6. Visualizar estatisticas(arquivo de dados)

7. Sair

Opcao: 5

TABELA DE INDICES							
Indice 1		Indice 2		Indice 3			
Entradas: 200		Entradas: 200		Entradas: 200			
#  Ticket	Offset	#  Ticket	Offset	#  Ticket	Offset		
1  0	798	0  798		0  798			
2  1	2076	1  2076		1  2076			
3  2	8237	2  8237		2  8237			
4  3	10498	3  10498		3  10498			



Figura 8: Situação após remover os tickets 2 e 3, e por inserir os tickets 4 e 5

TABELA DE REMOVIDOS									
Indice 1			Indice 2			Indice 3			
Removidos: 1			Removidos: 1			Removidos: 1			
Inicio da Lista: 10498			Inicio da Lista: 8237			Inicio da Lista: 10498			
#	tamanho	offset	tamanho	offset	tamanho	offset	tamanho	offset	
1	( 135,	NULL)	( 149,	NULL)	( 135,	NULL)	( 135,	NULL)	

Pressione enter para retornar ao menu.

1. Importar arquivo de dados.
2. Gerar arquivos de indice.
3. Remover registro.
4. Inserir registro.
5. Visualizar estatisticas(arquivo de indice)
5. Visualizar estatisticas(arquivo de dados)
7. Sair

Opcao: 5

TABELA DE INDICES									
Indice 1			Indice 2			Indice 3			
Entradas: 200			Entradas: 200			Entradas: 200			
#  Ticket	Offset		#  Ticket	Offset		#  Ticket	Offset		
1	0	798	1	0	798	1	0	798	
2	1	2076	2	1	2076	2	1	2076	
3	4	29681	3	4	29681	3	4	29681	
4	5	8237	4	5	10498	4	5	8237	