

# Relatorio do Projeto 1

Vinicius Campos Silva  
16/0041465  
Organizacao de Arquivos - UnB  
vinicius848@gmail.com

Alexandre Coelho Silva  
16/0022878  
Organizacao de Arquivos - UnB  
scoelhoalexandre@gmail.com

## 1. Objetivos

O Projeto 1 tem como objetivo gerar dois arquivos de Indices a partir de dois arquivos de texto armazenados na memória secundária, procedimento de indexação, bem como criar funções que possam manipular esses Indices, como inserir, remover, atualizar os indices já existentes. O desenvolvimento desse projeto visa aprimorar os conhecimentos dos autores acerca do conteúdo de indexação, de como implementa-la, suas funcionalidades e benefícios.

## 2. Introdução

### 2.1. Motivações

O trabalho com indexação é motivado pela procura na melhoria no tempo de acesso à arquivos muito grandes. Imagine um arquivo contendo milhões de registros, seria extremamente demorado para localizar um registro específico, e, buscando resolver esse problema, surgiram as técnicas de indexação.

### 2.2. Arquivo

Um arquivo é um conjunto de estruturas de dados com informações sobre uma determinada Entidade que estão armazenadas na memória secundária do computador. No caso desse Projeto, a Entidade sobre a que estamos trabalhando é a de Estudantes, que possui os seguintes atributos :  
-Matrícula -Nome -OP -Turma -Curso

024312	Leonardo Nunes de Souza	34	G	AB
031272	Andreza Pollyana da Silva	34	G	AB
035830	Rogério Augusto de Castro Neves	34	G	AX
041881	Ana Luisa Costa Miguel	34	G	AX
042290	Bruno Filipe Milare de Macedo	34	G	AX
042780	Daniel Schowantz Silva	34	G	AX
042865	Daniilo Silva Marshall	34	G	AX
043217	Erika Regina Campos de Almeida	34	G	AX
043341	Fabrizio Nallin Violin	34	G	AX
044072	Igor Ribeiro de Assis	34	G	AX
045058	Marcio Alexandre Marar Moreno	34	G	AX
046306	Rodrigo Dias Takase	34	G	AX
046708	Thiago de Moraes Sfredo	34	G	AX
046713	Thiago Fernandes Crepaldi	34	G	AX
046802	Tiago Rezende Campos Falcao	34	G	AX
046944	Vitor Baccetti Garcia	34	G	AX
046946	Vitor Bruno Jacintho de Almeida	34	G	AX
046956	Vitor Gouvea Franchi	34	G	AX

Figura 1. Arquivo de Indices

### 2.3. Chave

Uma chave é uma cadeia de caracteres (string) que utilizamos para identificar uma Entidade com seus atributos específicos.

### 2.4. Indexação

A técnica de indexação utilizada nesse projeto foi de gerar uma chave primária contendo a matrícula de um estudante concatenada com o seu nome, de forma que o tamanho máximo dessa chave fosse de 30 bytes. Com isso, foi possível gerar o arquivo de indices, que possui o seguinte formato:

A vantagem de possuir o arquivo de indices é que a velocidade de busca de um registro é muito maior, pois basta ordenar o mesmo, no caso desse projeto pelo algoritmo Heapsort, e depois fazer uma busca da chave.

### 3. Materiais e Métodos

#### 3.1. Materiais

Para solucionar o Projeto, foi desenvolvido um programa, em linguagem de programação C, que esta dividido nos arquivos "sort.h" e "main.c". Para a compilação foi usado o GCC.

#### 3.2. Métodos

O programa desenvolvido faz um acesso à "lista1.txt" e "lista2.txt" que fazem parte do problema proposto, e para cada uma dessas listas é gerado um arquivo de índices que contém as chaves primárias denominados "indicelista1.ind" e "indicelista2.ind", após gerar esses dois arquivos, ambos são ordenados por meio do algoritmo Heapsort. Com esses dois novos arquivos disponíveis, o programa faz, por meio da técnica de listas invertidas, mais dois arquivos, sendo esses novos "indice-seclista1.ind" e "indice-seclista2.ind", sendo esses os arquivos de índice contendo a chave secundária. Após realizar todos esses procedimentos o programa abre uma interface de Menu, que pode ser vista a seguir:

```
MENU:
Digite qual lista deseja manipular:

[1] : Lista 1;
[2] : Lista 2;
1

Qual operacao deseja realizar na Lista :

[1] : Verificar o arquivo armazenado na memoria :
[2] : Verificar lista de Indices primarios :
[3] : Verificar lista de Indices Secundarios :
[4] : Adicionar um novo elemento na lista :
[5] : Remover um elemento da lista :
[6] : Atualizar um Registro :
[7] : Trocar a lista que deseja manipular :
[0] : Sair do programa.
]
```

Figura 2. Menu do Programa

Também foram desenvolvidas funções que realizam todas as operações mostradas no menu, sendo assim possível adicionar, remover e atualizar registros nos arquivos.

### 4. Resultados

Os resultados obtidos foram os seguintes:  
Uma lista de Chaves contendo as diversas chaves

do arquivo original associadas a distância, em bytes, do Início desse arquivo até o primeiro byte da chave correspondente. Essa lista pode ser vista a seguir:

024312	Leonardo Nunes de Souza	34	G	AB
031272	Andreza Pollyana da Silva	34	G	AB
035830	Rogério Augusto de Castro Neves	34	G	AX
041881	Ana Luisa Costa Miguel	34	G	AX
042290	Bruno Filipe Milare de Macedo	34	G	AX
042780	Daniel Schowantz Silva	34	G	AX
042865	Daniilo Silva Marshall	34	G	AX
043217	Erika Regina Campos de Almeida	34	G	AX
043341	Fabricio Nallin Violin	34	G	AX
044072	Igor Ribeiro de Assis	34	G	AX
045058	Marcio Alexandre Marar Moreno	34	G	AX
046306	Rodrigo Dias Takase	34	G	AX
046708	Thiago de Moraes Sfredo	34	G	AX
046713	Thiago Fernandes Crepaldi	34	G	AX
046802	Tiago Rezende Campos Falcao	34	G	AX
046944	Vitor Baccetti Garcia	34	G	AX
046946	Vitor Bruno Jacintho de Almeida	34	G	AX
046956	Vitor Gouvea Franchi	34	G	AX

Figura 3. Arquivo de Índices

Já a outra lista gerada contém combinações de turmas e op's dos estudantes e logo em seguida todas as chaves primárias que possuem essa mesma turma e op. Essa lista pode ser vista a seguir:

```
34 AB: 024312Leonardo Nunes de Souza 031272Andreza Pollyana da Silv
34 AX: 035830Rogério Augusto de Castr 041881Ana Luisa Costa Miguel 042290Bruno Filipe Milare de M
042780Daniel Schowantz Silva 042865Daniilo Silva Marshall 043217Erika Regina Campos de A
043341Fabricio Nallin Violin 044072Igor Ribeiro de Assis 045058Marcio Alexandre Marar M
046306Rodrigo Dias Takase 046708Thiago de Moraes Sfredo 046713Thiago Fernandes Crepald
046802Tiago Rezende Campos Fal 046944Vitor Baccetti Garcia 046946Vitor Bruno Jacintho de
046956Vitor Gouvea Franchi
```

Figura 4. Arquivo de Índices secundario

### 5. Discussão e Conclusões

Após a conclusão do projeto foi possível observar que a construção de um arquivo de índices não é tão trivial quanto parece, mas, apesar disso, o arquivo de índices é extremamente útil para a localização de uma chave no arquivo principal, visto que é muito fácil ordena-lo uma vez que o algoritmo Heapsort esteja pronto e o arquivo já exista. Vale notar também que a criação do arquivo de índices de chaves secundárias é complicado, e que a técnica de listas invertidas não é de fácil implementação, mas uma vez implementada

permite uma rápida busca de elementos que possuam uma mesma chave secundária.

## **6. References**

[1] Slides apresentados em sala [2]  
<https://en.wikipedia.org/wiki/Heapsort>