

INF01124 - Classificação e Pesquisa de Dados - Exercício 4

Professor João Comba

1 Tabelas Hash usando Endereçamento Fechado

Implemente uma tabela Hash com M entradas para armazenar nomes de uma lista de contatos. O tamanho da tabela M é variável e deve ser informado quando da construção da tabela. Para a implementação da Tabela Hash, realize as seguintes tarefas:

- Escolha uma função de hash para mapear um string para um inteiro entre 0 e $M - 1$;
- Use o método para resolução de conflitos diferentes com endereçamento fechado;
- Implemente a operação de inserção de um elemento na da tabela hash;
- Implemente a operação de busca de um nome na tabela de hash, que retorna -1 caso o nome não esteja na tabela, ou um número $a > 0$ representando o número de acessos necessário para achar o nome;

2 Arquivos de Testes

A implementação da tabela Hash deve ser testada com dois arquivos. O primeiro arquivo, *nomes_10000.txt*, contém 10,000 nomes de contato que devem ser inseridos na tabela Hash. O segundo arquivo, *consultas.txt*, contém listas de 50 nomes que devem ser usadas para acionar consultas na tabela Hash. Ambos arquivos contém em cada linha o nome de um contato no formato :

nome sobrenome

Um exemplo é dado abaixo:

Alfredo Derouen
Janyce Koster
Lesley Trabue
Francesco Lauber
Angel Ridgeway
Buford Montez
Coleen Harley
Viviana Genova
Veronique Zambrana
Naida Meldrum
Artie Smithers
Milton Delgado
Emory Casarez
Debbie Musser
Ines Teal
Jamaal Duty

3 Experimentos

O objetivo deste laboratório é implementar e testar uma tabela Hash com tamanhos variáveis. Para testar sua implementação, iremos usar os arquivos de entrada com valores variáveis de tamanho de tabela: 503, 2503, 5003 e 7507. Para cada um destes tamanhos de tabela, deverá ser realizada um experimento que consiste em inserir todos os 10,000 nomes do arquivo *nomes_10000.txt* na tabela Hash, e depois realizar as consultas descritas no arquivo *consultas.txt*. Uma consulta deve retornar o valor 1, quando o valor é encontrado na tabela Hash e não houver colisões. Quando houverem n colisões, o valor de consultas deve ser $n + 1$. Quando o nome não for encontrado na tabela, devem retornar o número máximo de colisões + 1.

Para cada experimento, gere um arquivo de nome *experimentoM.txt* (onde M possui os valores 503, 2503, 5003 e 7507) contendo o número de consultas realizadas para cada nome, seguido da média e máximo valor de consultas. Os resultados devem ser gerados usando o seguinte formato:

```
NOME01 #CONSULTAS
NOME02 #CONSULTAS
...
NOME50 #CONSULTAS
MEDIA #MÉDIA DAS CONSULTAS
MAXIMO #NRO_MÁXIMO_CONSULTAS
```

4 Entrega

A solução deve ser enviada pelo Moodle dentro de um arquivo .zip, contendo os seguintes arquivos:

- **integrantes.txt**: coloque o nome dos integrantes do grupo (até 2 pessoas) , com um nome por linha
- **readme.txt**: explique a função de hash e a implementação do método de resolução de conflitos;
- **experimento503.txt**: estatísticas do experimento com tamanho de tabela 503
- **experimento2503.txt**: estatísticas do experimento com tamanho de tabela 2503
- **experimento5003.txt**: estatísticas do experimento com tamanho de tabela 5003
- **experimento7507.txt**: estatísticas do experimento com tamanho de tabela 7507
- código fonte correspondente a solução