

Computer science is no more about computers than astronomy is about telescopes.
Edsger Dijkstra

Laboratório – Tabela Hash

Desenvolvendo um sistema de contabilização de ocorrências de palavras em um arquivo texto.

Você desenvolverá um programa que receba como entrada um texto e imprima: uma relação de todas as palavras que ocorrem no texto juntamente com o número de ocorrências de cada palavra, o tamanho da tabela Hash, a quantidade de colisões e o fator de carga (número de elementos dividido pelo tamanho da tabela).

Entrada do programa: um arquivo texto T

Saída: uma lista de palavras e o número de ocorrências e na última linha o tamanho da tabela e quantas colisões ocorreram.

Exemplo de entrada:

Lorem ipsum dolor sit amet, Lorem ipsum dolor sit Lorem.

Exemplo (fictício) de saída:

amet 2
dolor 2
Ipsum 2
Lorem 3
Sit 1
M=10 COLISOES=20 CARGA=1.0

Você deve verificar a eficiência para vários valores de tamanho de tabela M.

Além do código alterado, deverá ser enviado um relatório com o estudo dos tamanho de tabela testados.

Dicas para implementação:

- Armazenamento das palavras lidas e da sua frequência:
 - tabela hash (tratamento de colisão com lista encadeada)
 - usa a própria palavra como chave de busca
- Função para obter a frequência das palavras:
 - dada uma palavra, tente encontrá-la na tabela
 - se não existir, armazene a palavra na tabela com ocorrência 1
 - se existir, incremente o número de ocorrências da palavra
- Imprimir a lista de palavra de ocorrências:
 - Crie uma função obterChave(...) que retorna uma lista com todas as chaves;
 - Percorra a lista e para cada chave
 - Busque a chave na tabela hash
 - Imprima chave e ocorrência
- O código-base da tabela hash com lista encadeada está disponibilizado no AVA.