

Atividade Prática 2 (Individual)
30% da segunda avaliação
Prazo: 31/05/2019

Elabora um programa que apresente um número de vezes da ocorrência de cada palavra em um dado texto lido de arquivo.

Sua entrada deverá ser um arquivo. Seu programa deverá retornar uma lista de palavras, em ordem decrescente do número de ocorrência de cada palavra no arquivo informado.

Você deverá implementar e comparar a solução do problema utilizando:

- 1) Hash (Tratamento de colisão com lista encadeada e hash duplo)
- 2) Árvore AVL
- 3) Árvore Rubro-Negra
- 4) Árvore B

Utilize a própria palavra como chave de busca.

Implemente uma funcionalidade para obter a frequência das palavras. Ou seja, dada uma palavra, tente encontra-la na tabela. Se não existir, armazene-a na tabela. Caso exista, incremente o número de ocorrências da palavra.

Implemente uma funcionalidade para exibir as ocorrências em ordem decrescente.

Observações:

- a) Uma palavra é considerada como uma sequência de letras e dígitos, começando com uma letra. Portanto, ignore sinais de pontuação. Você pode assumir que os textos não terão acentuação.
- b) Palavras com letras em maiúsculas devem ser primeiramente transformadas para minúsculas antes da inserção no índice. Desta forma, a mesma palavra apresentada com letras minúsculas ou maiúsculas não serão diferenciadas.
- d) Palavras constituídas por menos do que C caracteres devem ser desconsideradas.
- e) Uma palavra pode ocorrer múltiplas vezes na mesma linha de um documento, ou mesmo em múltiplas linhas de um mesmo documento.
- f) Você será responsável por criar suas coleções de documentos para testes. Submeta junto com o seu trabalho (no pacote zip) de coleção usada.

ação.

O Tamanho da tabela deverá ser gerenciamento automaticamente pelo seu programa.

Análise o desempenho, em termos do tempo de execução e do consumo de memória, de cada um dos TADs para diferentes cenários. Utilize contadores bem como chamadas do sistema para medições das estatísticas solicitadas. Execute diversos experimentos, variando a entrada e o tamanho M da tabela hash.

4. Entrega

- Código fonte do programa em C/C++ ou Java (bem indentado e comentado).
- Relatório do trabalho
- Upload no SIGAA.

O Relatório deve apresentar:

1. Introdução: descrição do problema a ser resolvido e justificativa apresentado exemplos de aplicações reais
2. Implementação: descrição sobre a implementação do programa. Devem ser detalhadas as estruturas de dados utilizadas (de preferência com diagramas ilustrativos), o funcionamento das principais funções e procedimentos utilizados, o formato de entrada e saída de dados, bem como decisões tomadas relativas aos casos e detalhes de especificação que porventura estejam omissos no enunciado.
3. Testes: apresentação dos testes realizados.
4. Conclusão: comentários gerais sobre o trabalho e as principais dificuldades encontradas em sua implementação
5. Referências: referências utilizadas no desenvolvimento do trabalho.