

Relatório de análise: Exercício-Programa

Gabriel Bernardini Schimidt - 12873188
Fernando Webes Soares Brandão - 13672940
Davi Martins - 10337787

Janeiro de 2024

Introdução

Este relatório diz respeito ao estudo e à análise dos resultados obtidos por meio do exercício-programa (EP) da disciplina de **Algoritmos e Estruturas de Dados I**.

Objetivo do Exercício-Programa:

O objetivo do EP baseava-se na implementação de um indexador e buscador de palavras para arquivos de texto (extensão *.txt*). O programa deveria ser capaz de carregar o conteúdo de um arquivo, armazená-lo em memória e indexar suas palavras (isto é, criar uma estrutura que permita verificar a existência ou não de uma palavra **sem que seja necessário verificar o texto linha a linha**).

Desenvolvimento

Desenvolvimento do programa:

Todo o processo de desenvolvimento do programa buscou respeitar as observações citada no arquivo de enunciado, como:

- Ignorar por completo qualquer sinal de pontuação.
- Tratar palavras compostas com o uso do hífen como duas palavras separadas.
- Assumir que a busca é sempre feita considerando palavras inteiras.

- Considerar iguais as versões maiúsculas e minúsculas de um caracteres após a realização da busca.

Todos os códigos utilizados para a realização do programa foram disponibilizados na pasta "ArquivosEp" no arquivo de extensão .zip entregue. Nela, devem constar:

- ep.c
- funcoes.h
- lista.h
- arvore.h

Compilação e Execução:

O passo a passo para a compilação e execução dos códigos fornecidos estão no arquivo "README.txt" presente no .zip entregue.

Experimentos

Abaixo, há duas tabelas com as informações do tempo de execução (em ms) para cada tipo de indexador e para cada tamanho do arquivo de texto dado (em número de linhas). Além disso, criou-se gráficos para melhor visualização dos testes realizados. Os eixos e as legendas estão presentes para auxiliar na compreensão.

Indexação		
Texto	Tempo de Execução (ms)	
nº de linhas	lista	arvore
475	40.24	9.56
1117	181.67	31.59
2449	710.59	128.70
4974	2928.70	615.31

Figure 1: Indexação

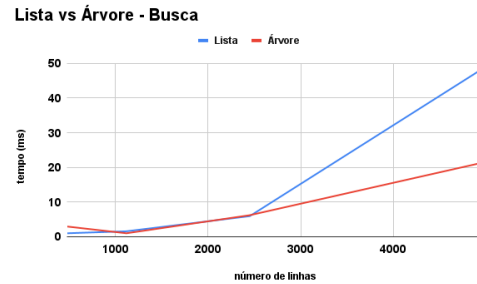


Figure 4: Lista vs Árvore - Busca

Conclusão

Inicialmente, é imprescindível citar que a arquitetura de cada computador utilizado para a execução do programa resultará em diferentes tempos de execução e de busca.

Porém, mesmo assim, conclui-se que devem seguir um mesmo padrão: análises teórica e empírica revelam de forma consistente que ambas as operações propostas, indexação e busca, são mais eficazes em termos de tempo de processamento quando implementadas por meio de uma árvore binária de busca.

Observações:

Por algum motivo não definido e explorado, o comando "fim", executado no terminal quando deseja-se finalizar o processo de busca necessitada de qualquer outra digitação para ser, de fato, concluído.

Além disso, a saída "printada" na tela possui dois campos de "tempo de execução", onde apenas o segundo printa a informação desejada.

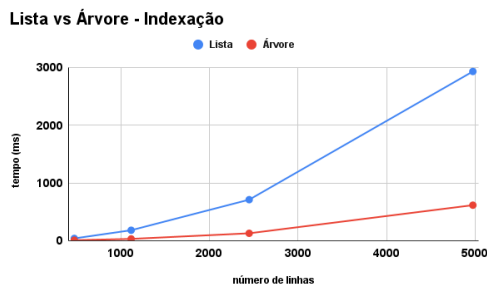


Figure 2: Lista vs Árvore - Indexação

Busca		
Texto	Tempo de Execução (ms)	
nº de linhas	lista	arvore
475	1	2.95
1117	1.57	1.00
2449	5.94	6.21
4974	48.63	21.43

Figure 3: Busca