



UNIVERSIDADE FEDERAL DE ITAJUBÁ – UNIFEI
INSTITUTO CIÊNCIAS PURAS E APLICADAS
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

MATHEUS KIM

PEDRO AUGUSTO GODOY

RELATÓRIO 2 - ALGORITMO E ESTRUTURA DE DADOS 2

ITABIRA

2023

TEMA:

Máquina de busca utilizando arquivo invertido com armazenamento em trie compacta.

IMPLEMENTAÇÃO

O código fornecido é um programa de pesquisa em arquivos de texto. Ele utiliza uma estrutura de dados chamada Trie compacta para indexar as palavras encontradas nos arquivos, permitindo que o usuário realize consultas eficientes.

A implementação começa abrindo um arquivo chamado "robots.txt", que contém os nomes dos arquivos de texto a serem processados. Em seguida, o programa percorre cada linha desse arquivo e abre cada arquivo de texto listado. Para cada arquivo de texto, o programa lê linha por linha e divide as linhas em palavras. Cada palavra é inserida na Trie compacta, juntamente com o número da linha em que ocorre.

A Trie compacta é uma estrutura de dados em árvore que organiza as palavras em uma estrutura hierárquica. Cada nó da Trie representa uma letra do alfabeto. O programa percorre a Trie de acordo com os caracteres de cada palavra, criando novos nós se necessário. O último nó de uma palavra é marcado como o fim da palavra e é atualizado um índice invertido para armazenar o número da linha em que a palavra ocorre.

Para realizar consultas, o programa permite que o usuário digite uma palavra ou frase. Se a consulta contém apenas uma palavra, o programa verifica se a palavra está presente na Trie. Se a palavra for encontrada, o programa lista as linhas em que ela ocorre, destacando a palavra nessas linhas. Se a consulta contém mais de uma palavra, o programa verifica se a frase completa ocorre em alguma linha dos arquivos de texto. Se ocorrer, o programa imprime a linha correspondente.

Em resumo, o código implementa um sistema de pesquisa em arquivos de texto que utiliza uma Trie compacta para indexar as palavras. Ele permite que o usuário faça consultas interativas para encontrar palavras ou frases nos arquivos de texto e destaca as correspondências encontradas.

LINKS:

<https://youtu.be/J4pl5zm8c4w>

https://youtu.be/Q3HrB_9-rN8

DIFICULDADE:

Dificuldade alta, pois são coisas que nunca havíamos pensado em codar em nossa vida acadêmica. Além disso tivemos dificuldade em pensar em como fazer.

REFERÊNCIA BIBLIOGRÁFICA:

ZIVIANI, Nivio. Projeto de Algoritmos com implementação em Pascal e C.
3a Edição. Ed. Cengage Learning, 2011