

Universidade Federal de Viçosa - Campus Florestal
Maio 2023

1º Trabalho Prático de CCF212

Prof. Gláucia Braga e Silva

APLICAÇÃO COM ÁRVORES DIGITAIS

Ana Carolina - 5094
Bárbara Pagnocca - 5061
Jéssica Cristina Carvalho - 4686
Layon Martins - 4220

Sumário

1.	Introdução	3
2.	Metodologia	3
3.	Resultados	4
4.	Considerações finais	6
5.	Referências bibliográficas	6

Introdução

O presente trabalho prático tem como objetivo a implementação de uma máquina de busca dos POCs do curso de Ciência da Computação por meio da criação de um índice invertido, ou seja, um tipo de indexação no qual os termos referenciam os documentos onde estão localizados. Esse método é particularmente apropriado para o problema em questão pois a busca será feita a partir das palavras chaves desejadas, as quais, por meio do código, apresentarão na interface gráfica a lista de documentos, por ordem de relevância. Cada termo diferente é armazenado em uma árvore do tipo PATRICIA, adaptada para guardar palavras, juntamente com o índice invertido. Assim, é possível encontrar muito mais fácil e eficientemente os POCs desejados, baseado apenas nos conceitos base do texto.

A realização de várias partes diferentes foi necessária durante a execução do trabalho, e a princípio cada integrante do grupo trabalharia em uma dessas partes. Porém, à medida que cada uma era executada e conforme as dificuldades foram surgindo, todos os integrantes acabaram trabalhando em praticamente todo o código.

Várias reuniões foram realizadas, de forma presencial e remota, e todo o trabalho foi feito utilizando um repositório no GitHub. As decisões e soluções foram desenvolvidas em conjunto. A estratégia foi implementar primeiro a PATRICIA, e então acrescentar o índice invertido a ela, além de toda a parte de tratamento de arquivos e cálculo de relevância.

Foram criados TADs para a árvore, para o índice invertido, para o tratamento de arquivos e para o cálculo de relevância. Este foi o último a ser criado dentre eles, e só então foi implementada a interface gráfica. Foram utilizados 5 arquivos de texto com o resumo de dissertações em inglês para realizar os testes.

A principal fonte utilizada foi o código da árvore PATRICIA do livro texto da disciplina (Ziviani, 2010).

Metodologia

Vale ressaltar que todas as etapas do trabalho foram explicadas para todos os membros da equipe, a fim de facilitar o entendimento do código como um todo e a resolução de eventuais problemas encontrados.

Desenvolvimento

A estrutura da árvore Patricia utilizada neste trabalho é uma adaptação da implementada no livro Projeto de Algoritmos com Implementações em Pascal e C,

do professor Nivio Ziviani. Nesta, os nós internos guardavam apenas o índice, e os nós externos, os caracteres. Ou seja, era uma árvore que armazenava caracteres. Os nós internos foram adaptados para guardar o índice e um caractere, e os nós externos, para guardar uma string e uma lista de índice invertido. Dessa forma, ela passa a armazenar palavras. Um trabalho anterior também foi utilizado como referência, que foi o do grupo de uma das integrantes deste grupo (consta nas referências).

Resultados

- detalhes técnicos de implementação da PATRICIA para armazenar palavras
- a interface construída
- resultados dos testes

A interface gráfica foi construída por meio da biblioteca GTK em conjunto com a plataforma Glade. Aqui foram desenvolvidas as telas principais para Construir o Índice Invertido, Imprimir e Pesquisar Palavras na Árvore, como mostram os exemplos das Figuras 1 e 2 abaixo.



Figura 1.



Figura 2.

Considerações finais

A principal dificuldade encontrada foi implementar de forma 100% funcional a inserção na PATRICIA. Apesar do estresse causado por esse empecilho, conforme todos tentaram ajudar a resolvê-lo, foi perceptível o lado positivo da situação: o maior entendimento da árvore em questão. Caso não ocorresse este problema, o entendimento geral da funcionalidade da PATRICIA não seria o mesmo.

Além disso, ainda que a lógica faça sentido, ocorreram outras complicações com a linguagem de programação, visto que C é muito peculiar e específico em relação a ponteiros, os quais estavam aos montes neste trabalho. Esses problemas, infelizmente, não foram resolvidos em sua totalidade, mesmo com certo progresso.

Finalmente, é consenso entre o grupo o aprendizado da aplicação das árvores binárias no dia a dia, bem como o funcionamento de mecanismos de autocorreção, busca de palavras, corretores automáticos.

Referências bibliográficas

- N. Ziviani, Projeto de Algoritmos com Implementações em Pascal e C, 3ª ed., Cengage Learning, 2010.
- N. Ziviani. Projeto de Algoritmos com Implementações em Pascal e C. Disponível nos slides da disciplina em <https://ava.ufv.br>.

- Visual Studio Code. Compilador de linguagem C utilizado. Disponível em <https://code.visualstudio.com>.
- Microsoft Live Share. Extensão do Visual Studio Code de edição simultânea do código. Disponível em <https://code.visualstudio.com/learn/collaboration/live-share>.
- GitHub. Plataforma de armazenamento na nuvem e compartilhamento entre integrantes. Disponível em <https://github.com>.
- GTK. Biblioteca de criação de interfaces gráficas. Disponível em <https://www.gtk.org>.

Todos os acessos datam de Maio de 2023.