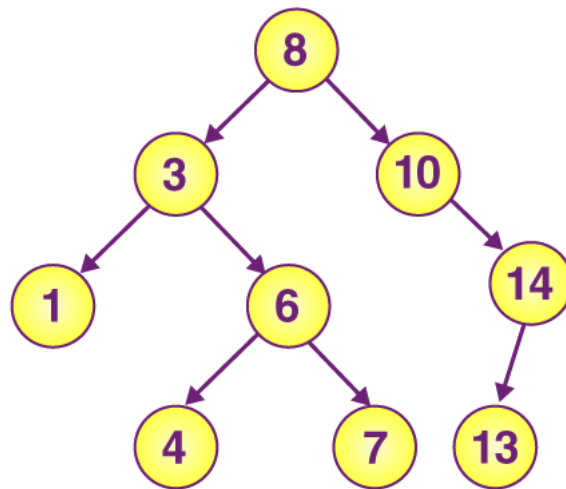


Trabalho Prático: Árvores Binárias de Pesquisa - ABP
Disciplina: FGA0147 - Estruturas de Dados e Algoritmos
Prof. Dr. Nilton Correia da Silva
Faculdade UnB Gama - FGA
Universidade de Brasília - UnB

Aplicação de Árvores Binárias de Pesquisa - ABP para Geração de Relatórios Ordenados.



Gama, Distrito Federal

Sumário

1	Introdução	3
2	Datasets	4
2.1	Grupo A	4
2.2	Grupo B	4
2.3	Grupo C	4
2.4	Grupo D	4
2.5	Grupo F	4
2.6	Grupo G	4
2.7	Grupo H	4
3	Objetivo	5
3.1	Opção Carregar Arquivo de Dados	5
3.2	Opção Emitir Relatório	5
3.3	Opção Sair	5
4	Condições de Contorno da Solução	5
5	Entregáveis e Notas	5
5.1	Carregar Arquivo de Dados - 4 pontos	6
5.2	Emitir Relatório - 3 pontos	6
5.3	Opção Sair - 3 pontos	6



Orientações

Trabalho Prático para composição da média final da disciplina de Estruturas de Dados e Algoritmos. Este trabalho foi elaborado considerando que sua resolução deve ser distribuída para um grupo de alunos.

A solução deverá ser em linguagem C ou C++. No caso do grupo adotar C++, poderá se valer de uma solução orientada a objetos, contudo sua solução não deve usar classes e contêineres já prontos da linguagem C++ - por exemplo, algoritmos de ordenação, busca, objetos vector, list, matrix, etc.



1 Introdução

Árvores Binárias de Pesquisa (ABP) são Estruturas de Dados (ED) flexíveis utilizadas para aplicações em ordenação e buscas. São frequentemente demandadas em ordenações de dados. Aplicações em banco de dados, por vezes, requerem que o algoritmo realize ordenações e buscas em um dataset tendo como chave uma ou mais de suas variáveis - geralmente, as variáveis vêm em forma de colunas do dataset.

Neste trabalho teremos a oportunidade de exercitar ABPs em um caso concreto representado por um dataset que foi preparado para que engenheiros de IA (Inteligência Artificial) e outros profissionais pudessem apresentar soluções para um problema envolvendo o referido dataset. Nosso objetivo aqui, contudo, não será apresentar uma solução final para o problema apresentado, mas sim, apresentar uma solução mais simples que será detalhada nas seções posteriores.

Caso o dataset disponibilizado tenha uma quantidade de registros maior do que os seus recursos computacionais conseguem processar, você pode excluir registros ou colunas (desnecessárias à solução), deixando apenas a quantidade máxima de dados que seus recursos computacionais conseguem processar.



2 Datasets

Abaixo é apresentado os detalhes do dado atribuído a cada grupo: link do dataset e os campos chave para a ordenação do relatório.

2.1 Grupo A

Fonte de dados: Breast Cancer

Campo Chave: Age + Tumor Size

2.2 Grupo B

Fonte de dados: Animal Kingdom Taxonomy

Campo Chave: Kingdom + Phylum + Class

2.3 Grupo C

Fonte de dados: Fire Incidents

Campo Chave: Latitude + Longitude

2.4 Grupo D

Fonte de dados: Cars - Purchase Decision Dataset

Campo Chave: User ID

2.5 Grupo F

Fonte de dados: COVID-19 Numerical Claims Open Research Dataset

Campo Chave: country + publish_time

2.6 Grupo G

Fonte de dados: CO2 Emissions Around the World

Campo Chave: Region + Country Name

2.7 Grupo H

Fonte de dados: Data Science Job Salaries

Campo Chave: salary_in_usd



3 Objetivo

A solução deste trabalho consiste em desenvolver um programa para gerar um relatório ordenado por um campo chave com a utilização de uma Árvore Binária de Pesquisa (ABP). Este campo chave pode ser uma concatenação de um ou mais campos do dataset.

Seu programa deve ter as seguintes opções:

1. Carregar Arquivo de Dados
2. Emitir Relatório
3. Sair

3.1 Opção Carregar Arquivo de Dados

Nesta opção seu programa deve carregar um arquivo de dados. O nome do arquivo deve ser solicitado ao usuário. Neste momento o programa deve gerar uma ABP pelo campo chave. Não carregue todos os campos para a ABP. Os nós devem armazenar apenas o campo chave e a posição de ocorrência (Linha) do arquivo.

3.2 Opção Emitir Relatório

Mostre na tela um relatório ordenado pelo campo chave contendo todos os registros (linhas) com todos seus campos. Caso não haja registros carregados, imprima a mensagem “Arquivo Vazio!!!”.

3.3 Opção Sair

Nesta opção de encerramento de seu programa, atente para fazer a gerência de alocação de memória adequada – não deixe lixo na memória. Você deve aplicar um algoritmo que desaloque todos os nós de sua ABP, bem como desalocar outras ED que tenham sido alocadas dinamicamente.

4 Condições de Contorno da Solução

Sua solução deve atender às seguintes condições:

1. O arquivo de entrada (dataset) deve ser lido por seu programa em sua apresentação original - não deve ser ordenado antes por outro programa.
2. Seu programa deve ser construído utilizando funções e/ou procedimentos ou programação orientada a objetos..

5 Entregáveis e Notas

A avaliação deste trabalho se dará mediante a apresentação do código-fonte contendo os itens abaixo funcionando corretamente e conforme instruções detalhadas acima.



- 5.1 Carregar Arquivo de Dados - 4 pontos
- 5.2 Emitir Relatório - 3 pontos
- 5.3 Opção Sair - 3 pontos

