

Comparação de busca com grande quantidade de dados utilizando Árvores Rubro Negras

Claudio da Silva T. Júnior Pietro N. Neiva

March 2020

Abstract - The search for data is indispensable in any application. Spreadsheets are available on the Brazilian government portal and, with this data, we will look for solutions with the Red Black Tree to optimize the consultation, insertion and removal of records.

Index Terms: RBT (Red Black Tree), Algorithm.

1 Introdução

Os dados disponibilizados no portal do governo brasileiro (dados.gov.br) para acesso ao público são disponibilizados em planilhas, uma forma ineficiente para lidar com tamanha quantidade de informações. De posse de um *dataset* de preços de medicamentos, iremos tentar otimizar esse tratamento de informações com a ajuda da RBT e uma *Structured Query Language*, intermediados por uma aplicação em JAVA com um CRUD básico.

2 Objetivo Geral

Implementação de um sistema de consulta e registro de medicamentos utilizando uma RBT para aperfeiçoar a velocidade de interação com o usuário na Busca, inserção e manipulação dos dados sobre medicamentos.

3 Objetivo Específico

Consultar a base de dados do governo para entender a periodicidade que estes são atualizados por meio de planilhas, converter esses dados para uma *query language*, implementar uma RBT utilizando a linguagem JAVA, implementar um CRUD básico para interação com a RBT com o usuário.

4 Referencial Teórico

- Dev-C++, ambiente de desenvolvimento integrado livre que utiliza os compiladores do projeto GNU para compilar programas para o sistema operacional Microsoft Windows. - C++, linguagem de programação compilada multiparadigma (seu suporte inclui linguagem imperativa, orientada a objetos e genérica) e de uso geral. Desde os anos 1990 é uma das linguagens comerciais mais populares, sendo bastante usada também na academia por seu grande desempenho e base de utilizadores. - Java, linguagem de programação orientada a objetos desenvolvida na década de 90 pela empresa Sun Microsystems. - Eclipse, IDE desenvolvida em Java pela IBM, seguindo o modelo open source de desenvolvimento de software. - MariaDB, SGDB que surgiu como *fork* do MySQL, criado pelo próprio fundador após sua aquisição pela Oracle. - Árvore Rubro Negra, é um tipo de árvore binária de busca balanceada usada na ciência da computação tipicamente para implementar vetores associativos inventada em 1972 por Rudolf Bayer.

5 Desenvolvimento

Utilizando a linguagem C++ com a teoria de árvore binária Rubro-negra. Implementando no Dev C++ e utilizando o referencial teórico, conseguimos implementar o balanceamento da árvore a medida que são feitas novas inserções, entretanto o balanceamento a medida que são excluídos os elementos da árvore tem encontrado dificuldade, especialmente se o balanceamento for feito com a raiz envolvida. Após a implementação em C++, devido às complicações com a comunicação entre o banco de dados, foi decidido a implementação em Java, tendo em vista também a facilidade para a criação de um CRUD básico com interface gráfica. De posse da base de dados em planilha, foi convertida para SQL, com cerca de 29 mil registros de informações de medicamentos.

6 Conclusão

7 Proposta futura