

Guia 1

Proposta de Projeto - Grupo 6

#Equipe

Nicolas Kayamori Huschak - 124110694
Pedro Brito de Melo Trovão - 124110728
Rafael Antonio de Lucena Serey - 124110730
Matteus Marinho Cantisani - 124110576
Lucas Araujo Medeiros - 124110594
Henrique Ponce de Leon Pessoa - 124111032

#Resumo

Tendo em vista a implementação descrita na disciplina, o projeto tem como objetivo comparar o desempenho de diferentes variações de estruturas de *HashMap* na linguagem Java, explorando estratégias distintas de tratamento de colisões e variação do fator de carga.

As estruturas analisadas serão:

1. Endereçamento Aberto com Sondagem Linear
2. Endereçamento Aberto com Sondagem Quadrática
3. Encadeamento Fechado com *LinkedList*
4. Encadeamento Fechado com *ArrayList*

Neste experimento, as estruturas de dados serão implementadas em Java. As implementações dos *HashMap* serão reaproveitadas do repositório Open Source [thealgorithms](#). Além da comparação entre essas implementações, serão avaliadas diferentes operações fundamentais (como *put*, *get* e *remove*) sob diferentes fatores de carga e conjuntos de entradas. Os resultados serão apresentados em forma de tabelas e gráficos, de modo a facilitar a análise comparativa e permitir uma compreensão clara do impacto de cada variação.

#Motivação

No estudo de Estruturas de Dados, compreender o impacto das diferentes estratégias de resolução de colisões é essencial para projetar algoritmos que conciliam rapidez e uso eficiente de memória. Nesse contexto, estruturas amplamente utilizadas, como *HashMap*, podem apresentar desempenhos bastante distintos dependendo da forma como as colisões são tratadas. Estratégias como o encadeamento fechado, que pode ser otimizado com diferentes variações de listas, e o endereçamento aberto, que admite distintas técnicas de sondagem, influenciam diretamente no custo das operações fundamentais, como a inserção de elementos. Assim, este projeto tem como motivação analisar experimentalmente como essas variações se comportam sob diferentes condições de carga e tipos de entrada, de modo a identificar quais configurações se mostram mais adequadas em cada cenário. A partir dessa avaliação, pretende-se evidenciar o papel das escolhas de implementação no desempenho prático de estruturas de dados e oferecer subsídios para decisões mais embasadas em aplicações reais.

#Objetivo

O objetivo do projeto é analisar o desempenho na execução de operações fundamentais em quatro variações de *HashMap* (endereçamento aberto com sondagem linear, endereçamento aberto com sondagem quadrática, encadeamento fechado com *LinkedList* e encadeamento fechado com *ArrayList*) sob o efeito de diferentes fatores de carga. A proposta busca comparar experimentalmente essas abordagens de tratamento de colisões, avaliando como cada variação responde a diferentes níveis de carga e padrões de entrada. Esse sistema de análise permitirá identificar vantagens e limitações de cada técnica, demonstrando em quais cenários cada implementação apresenta maior eficiência e melhor desempenho.

#Metodologia

Esse experimento será conduzido em cinco etapas principais, correspondentes ao ciclo completo de preparação, execução e análise. O objetivo é avaliar a performance de

diferentes estratégias de resolução de colisões em operações de inserção e busca em *HashMap*. Para isso, serão utilizadas quatro implementações distintas: endereçamento aberto com sondagem linear, endereçamento aberto com sondagem quadrática, encadeamento fechado com *LinkedList* e encadeamento fechado com *ArrayList*. Cada etapa foi definida de forma a garantir comparabilidade entre as implementações, possibilitando identificar em quais cenários cada estrutura apresenta maior eficiência.

Configuração de etapa do plano de experimento:

Configuração das estruturas

Inicialmente, serão obtidos *HashMaps* do repositório Open Source [thealgorithms](https://github.com/steve-nouri/thealgorithms), que serão devidamente alterados, para garantir que cada versão do *HashMap* esteja disponível para análise e possível modificação.

Ambientação e configuração

Após a recuperação, será realizada a adaptação das implementações para permitir a variação controlada do fator de carga em cada estrutura, assegurando comparabilidade entre os diferentes métodos de tratamento de colisões.

Geração das entradas

As entradas utilizadas nos experimentos serão geradas de maneira randômica, assegurando diversidade de casos de colisão. Essa abordagem garante que cada implementação seja testada sob condições variadas, simulando cenários próximos a aplicações reais.

Execução dos testes com benchmarks

Para a execução dos testes será utilizado um ambiente controlado, no qual serão aplicados benchmarks em Java, utilizando o *JMH (Java Microbenchmark Harness)*, uma ferramenta de análise de desempenho projetada para medir e analisar o desempenho de código Java. Os testes permitirão a coleta de métricas relacionadas ao custo de memória e à eficiência das operações *put* e *get*. Cada cenário será repetido múltiplas vezes para reduzir a influência de outliers.

Coleta e análise dos resultados

Após a execução, os dados serão processados para permitir a comparação entre os cenários. Os dados obtidos durante a execução dos benchmarks serão registrados em

arquivos de resultados. Em seguida, serão gerados gráficos comparativos que representarão o desempenho de cada variação de *HashMap* em relação às métricas coletadas. Por fim, os resultados serão analisados de forma crítica, buscando identificar padrões de comportamento e compreender em quais cenários cada técnica apresenta maior eficiência.

#Entrega

A entrega será feita via um repositório do GitHub, contendo os códigos utilizados nos experimentos, os dados coletados para embasar as conclusões e os gráficos gerados a partir das análises realizadas. Assim, os resultados obtidos neste projeto fornecerão uma visão comparativa entre estratégias de sondagem e encadeamento, permitindo evidenciar os cenários em que cada implementação apresenta vantagens.

O repositório reunirá tabelas e gráficos construídos a partir dos benchmarks aplicados, facilitando a visualização dos padrões de desempenho encontrados. Dessa forma, os dados produzidos poderão ser utilizados não apenas para sustentar as conclusões deste trabalho, mas também como material de apoio para o aprendizado de outros estudantes, servindo como uma ponte entre o estudo teórico de estruturas de dados e sua aplicação prática em contextos reais.