

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
EMERSON PATRYCK DA SILVA

RELATÓRIO TÉCNICO EXERCÍCIO DE PROGRAMAÇÃO 3
VERIFICADOR ORTOGRÁFICO

SÃO MATEUS – ES
MAIO DE 2021

SUMÁRIO

OBJETIVO	3
ORTOGRAFIA	4
METODOLOGIA.....	4
RESULTADOS E DISCUSSÕES.....	6
APLICAÇÕES.....	8
REFERÊNCIA.....	9

OBJETIVO

O objetivo desse relatório é apresentar resultados obtidos comparando o desempenho da implementação da árvore balanceada e da implementação da árvore convencional, afim de verificar a eficiência entre elas.

Para realizar o experimento foi desenvolvido o algoritmo em C em que a função dele é buscar palavras sugestões para determinados tipos de texto no qual poderia ou não haver erros ortográficos.

ORTOGRAFIA

Segundo o dicionário Oxford Languages, a ortografia pode ser definida como o *'conjunto de regras estabelecidas pela gramática normativa que ensina a grafia correta das palavras, o uso de sinais gráficos que destacam vogais tônicas, abertas ou fechadas, processos fonológicos como a crase, os sinais de pontuação esclarecedores de funções sintáticas da língua e motivados por tais funções etc.'*

Com os avanços no campo da tecnologia a ortografia passou a ser algo automático no nosso dia-a-dia. Implementado em diversas linguagens, os corretores ortográficos estão presentes na maioria dos celulares, tablets e computadores, sem mencionar em buscadores de sites, browser e outros programas onde as inserções de palavras e textos sejam úteis. O exemplo de aplicação dos corretores ortográficos está na elaboração desse trabalho, que consiste basicamente em inserir uma palavra errada ou estranha no documento e o software retornar sugestões de palavras corretas ou semelhantes.

Como proposto no título desse trabalho, há diferentes formas de implementar o corretor ortográfico e basicamente atentaremos no estudo da implementação utilizando árvores balanceadas e não balanceadas. Objetivando não o estudo do corretor ortográfico, mas a eficiência de cada árvore implementada utilizando o corretor ortográfico como tema desse estudo.

METODOLOGIA

A metodologia aplicada para estudo desse trabalho consiste na implementação da TAD's contendo:

- Árvore balanceada
- Árvore não balanceada
- Fila

E ainda, a aplicação e análise em um texto dividido em três etapas:

- Na primeira etapa será analisado o texto correto, sem alterações.
- Na segunda etapa será analisado o texto com algumas alterações.
- Na terceira etapa será analisado o texto quase incorreto, totalmente alterado.

Primeiro começamos pelas implementações. Foram implementadas as árvores balanceada e não balanceada no qual irão receber o dicionário para análise do texto inserido. Na árvore não balanceada foram implementadas as funções **criarNo**, **liberarNo**, **buscarNo**, **inserirNo**, **removerNo**, **imprimirNo** e **compara**. E na árvore balanceada foram implementadas as funções **criarNoAVL**, **liberarNoAVL**, **buscarNoAVL**, **altura**, **fatorBalanceamento**, **rotacaoDireita**, **rotacaoEsquerda**, **inserirNoAVL**, **removerNoAVL**, **imprimeArvore**, **comparaAVL**. Por questão de espaço, a implementação poderá ser vista nos arquivos enviados junto com esse relatório.

Após implementar as árvores balanceadas e não balanceadas, implementamos a estrutura que irá receber os textos de entrada para comparação e para isso foi escolhida a estrutura de dados fila. A fila foi escolhida por facilidade e simplicidade no momento da implementação. Basicamente foram implementadas as funções referentes a fila, **criarFila**, **criarNoFila**, **filaVazia**, **enqueue**, **dequeue**, **buscarNoFila**, **liberaNoFila** e **liberaFila**.

O objetivo dessas implementações se resume a comparar cada dado contido no nó da fila, inserido pelo usuário, com os dados das árvores e assim retornando os dados semelhantes aos dados inseridos.

Para parte experimental, consideramos os textos abaixo como base. Para análise do comportamento do algoritmo primeiro utilizamos o texto sem alterações. O texto mostrado abaixo é o exemplo utilizado:

```
O empenho em analisar a consolidação das estruturas oferece uma interessante oportunidade para verificação dos índices pretendidos.
O que temos que ter sempre em mente é que o desafiador cenário globalizado desafia a capacidade de equalização das novas proposições.
```

Para diversificar os resultados e observar o comportamento de cada algoritmo, introduzimos algumas alterações no texto. Então temos o seguinte texto para segunda análise:

```
O empenho em analisar a consolidação das estruturas oferece uma interessante oportunidade para verificação dos índices pretendidos.
O que temos que ter sempre em mente é que o desafiador cenário globalizado desafia a capacidade de equalização das novas proposições.
```

Para ampliar diversificação dos resultados e observar o comportamento, alteramos o texto quase por completo. Como mostrado no texto abaixo:

```
O empenho em analisar a consolidação das estruturas oferece uma interessante oportunidade para verificação dos índices pretendidos.
O que temos que ter sempre em mente é que o desafiador cenário globalizado desafia a capacidade de equalização das novas proposições.
```

Cada texto possui 39 palavras e serão analisadas as seguintes características:

1. Tempo de criação de ambas as árvores.
2. Tempo de busca das palavras nas árvores.
3. Quantidade de palavras erradas.
4. Quantidade de palavras sugeridas.
5. Tempo máximo de execução do programa.

Abaixo serão apresentados os resultados desenvolvimento as conclusões obtidas durante a execução do experimento.

RESULTADOS E DISCUSSÕES

Pela relação de proximidade dos resultados teremos abaixo a apresentação de todos os dados onde será feita um apanhado geral.

Para o texto sem nenhuma alteração:

Análise Comportamental - Texto Sem Alterações			
Análises	Árvore Balanceada	Árvore Não Balanceada	Diferença (%)
Tempo de Criação da Arvore (segundos)	0,154706	639,390175	
Tempo Máximo de Execução (segundos)	0,555665	639,556288	
Quantidade de Palavras Erradas	2	2	9475%
Quantidade de Sugestões de Palavra	3	3	
Tempo Médio de Busca (segundos)	0,00403	0,076795	

Para o texto com algumas alterações:

Análise Comportamental - Texto Com Pequenas Alterações			
Análises	Árvore Balanceada	Árvore Não Balanceada	Diferença (%)
Tempo de Criação da Arvore (segundos)	0,15231	662,244713	
Tempo Máximo de Execução (segundos)	1,970379	662,424088	
Quantidade de Palavras Erradas	11	11	9605%
Quantidade de Sugestões de Palavra	111	111	
Tempo Médio de Busca (segundos)	0,003709	0,093933	

E para o texto com muitas alterações:

Análise Comportamental - Texto Com Muitas Alterações			
Análises	Árvore Balanceada	Árvore Não Balanceada	Diferença (%)
Tempo de Criação da Arvore (segundos)	0,149805	519,412729	
Tempo Máximo de Execução (segundos)	2,505453	519,613748	
Quantidade de Palavras Erradas	22	22	9646%
Quantidade de Sugestões de Palavra	667	667	
Tempo Médio de Busca (segundos)	0,003394	0,095899	

O primeiro ponto a ser analisado é o tempo de criação das árvores. Na árvore balanceada o tempo é inferior a 1 segundo em todos os testes, isso acontece quando inserimos os nós na árvore temos o controle da altura. Isso significa quando alocados

os nós na árvore, a maneira que são inseridos faz com que a altura da árvore seja menor, ou seja, a distância entre a raiz e as folhas são relativamente pequenas e isso é feito através do fator balanceamento e também das rotações, que ajustam a árvore para que sempre que possível tenha a altura relativamente pequena que implica na velocidade de leitura da árvore.

Na árvore não balanceada não temos o fator de balanceamento e as rotações, então a altura da árvore se torna suficientemente grande para que quando inserimos uma grande quantidade de dados e realizamos uma busca na árvore o tempo de resposta seja muito alta.

Na árvore não balanceada percebemos que independente das alterações que foram feitas no texto há um tempo consideravelmente constantes no momento da execução do programa, pois sua variação é consideravelmente mínima. Agora observando a árvore balanceada, vemos que o tempo de execução do programa aumentou pelo fator da árvore realizar diversas rotações em seus nós para que a distância entre o nó e a raiz seja mais pequeno possível.

O tempo médio de busca é a média do tempo gasto pelo programa para fazer a busca das palavras do texto na árvore. O gráfico abaixo mostra o tamanho da dispersão entre a árvore balanceada e a árvore não balanceada.

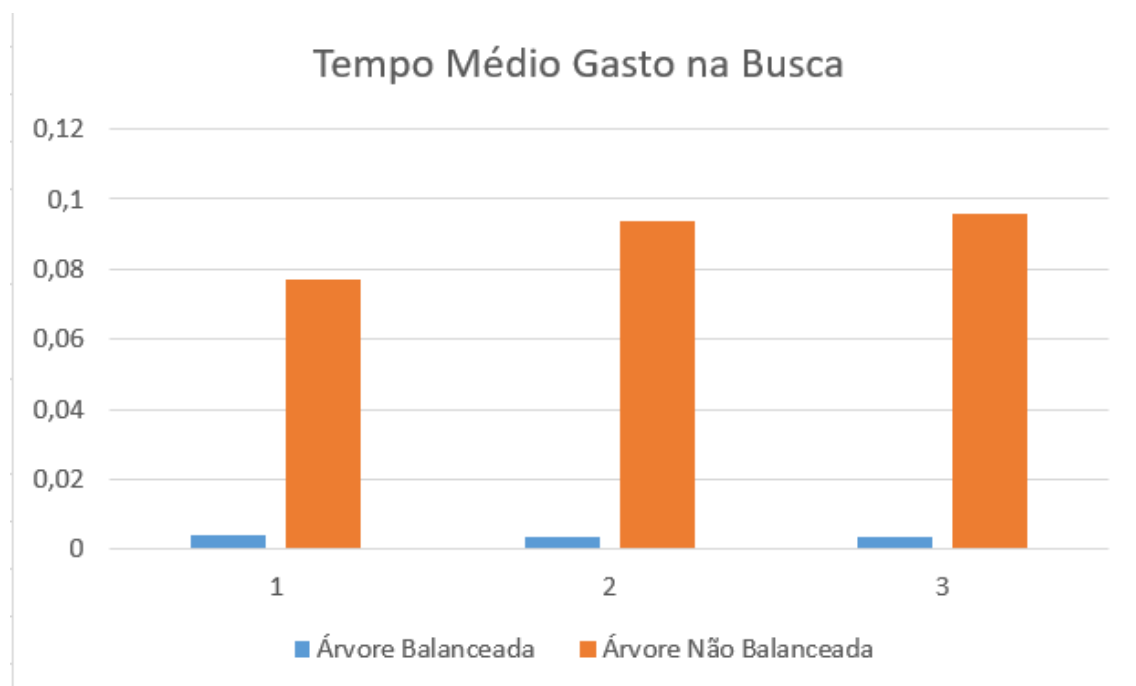


Figura 1 - (1) Texto sem alteração. (2) Texto com poucas alterações. (3) Texto com muitas alterações.

Percebe que o tempo médio de resposta de cada busca é praticamente constante em qualquer situação imposta no texto. E também percebemos que dado o mesmo texto, a diferença de resposta é exorbitante, mostrando que a árvore balanceada é mais eficiente que a árvore não balanceada.

Por último temos a comparação de eficiência entre os valores de tempos médio gasto no momento da busca. O cálculo utilizado mostra que a árvore balanceada é pouco menos de 10000% eficiente que a árvore não balanceada em todos os casos seguindo a formula abaixo.

APLICAÇÕES

Figura 2 - Aplicação do texto sem alterações.

Figura 3 - Aplicação do texto com poucas alterações

Figura 4 - Aplicação do texto com muitas alterações.

REFERÊNCIA

Ma. Luciana Kuchenbecker Araújo. **Ortografia.** Disponível em:<
<https://www.portugues.com.br/gramatica/ortografia.html>> Acesso em:12 de maio de
2021.