

# Buscador de Tweets

Universidade Federal do Rio Grande do Sul

(INF01203) Estruturas de Dados

Prof. Viviane Moreira

---

Arthur Oliveira de Rosso (00318973)

Yasmin Katerine Beer Zebrowski (00277765)

---

## Objetivo

Este trabalho foi elaborado com o objetivo de analisar o número de comparações entre uma árvore AVL (*Adelson-Velskii e Landis*) e uma ABP (*Árvore Binária de Pesquisa*), sendo essas comparações realizadas ao montar as árvores e ao consultar palavras dentro das mesmas. Ademais, também foram comparadas as alturas, o número de nodos e o número de rotações realizadas.

## Árvores AVL

Por definição, árvores AVL são árvores que a cada inserção de nodos realizam rotações conforme o fator de balanceamento. Desse modo, essas árvores têm a característica de possuírem a menor altura possível, com diferença de altura entre as subárvores direita e esquerda de no máximo 1.

## Árvores ABP

Árvores ABP, por outro lado, são muito mais simples, pois elas não se auto regulam. A cada nodo inserido, é verificado se o nodo é maior, igual ou menor que o nodo anterior, e então inserido na árvore conforme essa classificação - menor à esquerda, maior à direita. No caso deste trabalho, caso um nodo seja igual, será inserido na lista de ocorrências e não repetido na árvore.

Além disso, nesse tipo de árvore não existe um fator de balanceamento, e como também não há rotações, a ordem de inserção altera drasticamente a árvore resultante, podendo ter desde uma altura ótima, como podendo ter a altura igual ao número de nodos.

## Resultados Esperados

Nesse sentido, dado um mesmo número de nodos e mesma ordem de inserção, podemos concluir já pelas definições anteriores que uma AVL tenderá a ter menor altura que uma ABP, e conseqüentemente, o número de comparações de consulta e de inserções da AVL também tenderá a ser menor. Ambas encontrarão as palavras consultadas, mas a AVL será mais eficiente conforme o número de nodos aumentar.

## Implementação

Foram utilizados os códigos proporcionados pelos materiais disponíveis pela professora. A implementação da contagem de comparações e da contagem de ocorrências são implementações próprias.

## Resultados do Programa

### Base de 10 Tweets

	AVL	ABP
Indexação		
Nodos	84	
Comparações	1384	1557
Rotações	51	0
Altura	8	12
Consultas		
Comparações	43	51

### Base de 100 Tweets

	AVL	ABP
Indexação		
Nodos	704	

Comparações	27063	29304
Rotações	500	0
Altura	11	22
Consultas		
Comparações	122	153

## Base de 10.000 Tweets

10 palavras pesquisadas

	AVL	ABP
Indexação		
Nodos	15553	
Comparações	3520731	3531541
Rotações	11108	0
Altura	17	34
Consultas		
Comparações	133	152

1000 palavras pesquisadas

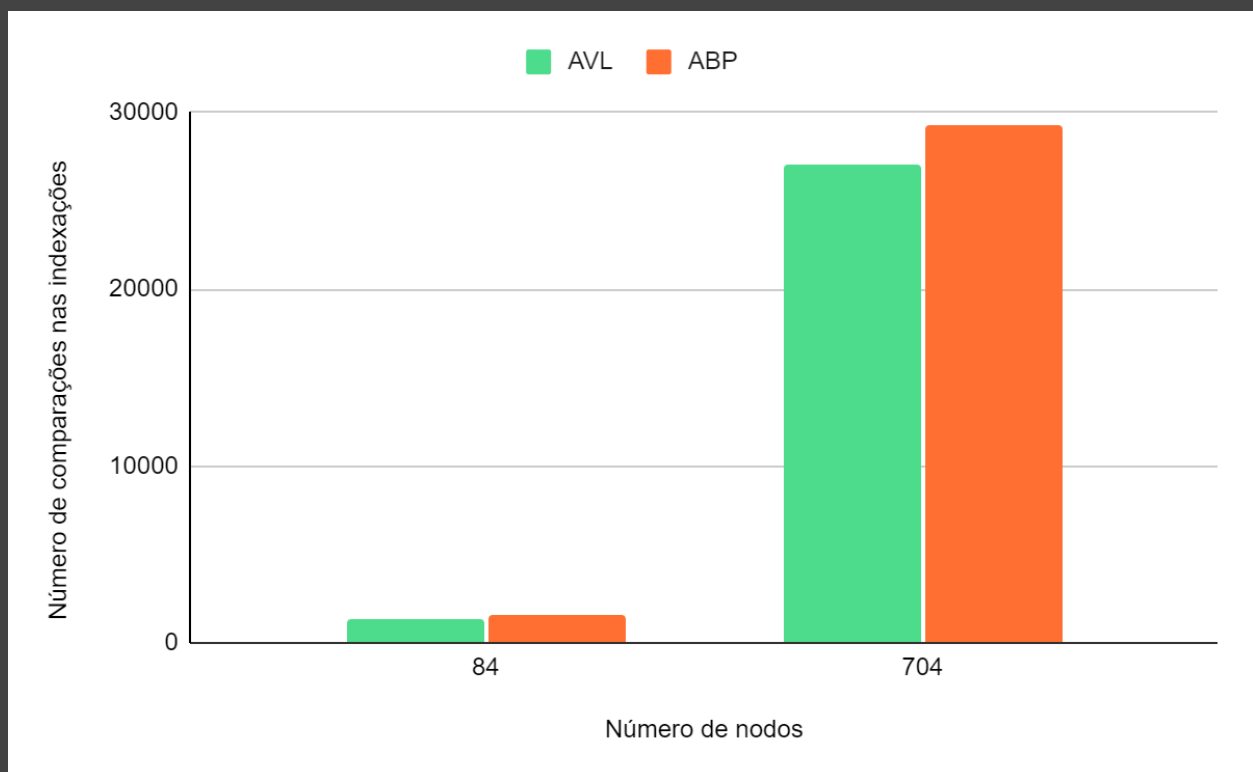
	AVL	ABP
Indexação		
Nodos	15542	
Comparações	3520685	3531412
Rotações	11105	0
Altura	17	34
Consultas		
Comparações	20426	25399

## Base Completa de 900.000 tweets

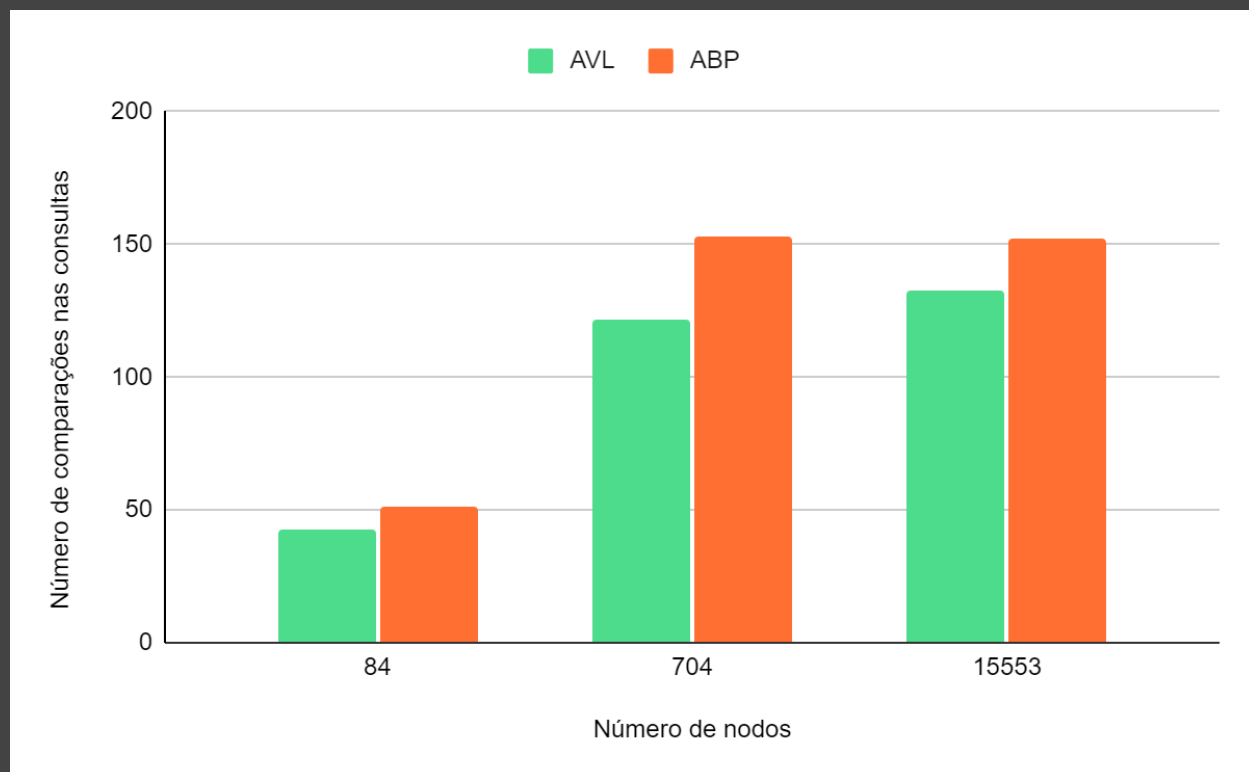
	AVL	ABP
Indexação		
Nodos	174346	
Comparações	332231800	312420772
Rotações	123929	0
Altura	21	45
Consultas		
Comparações	138	152

## Análise dos Dados

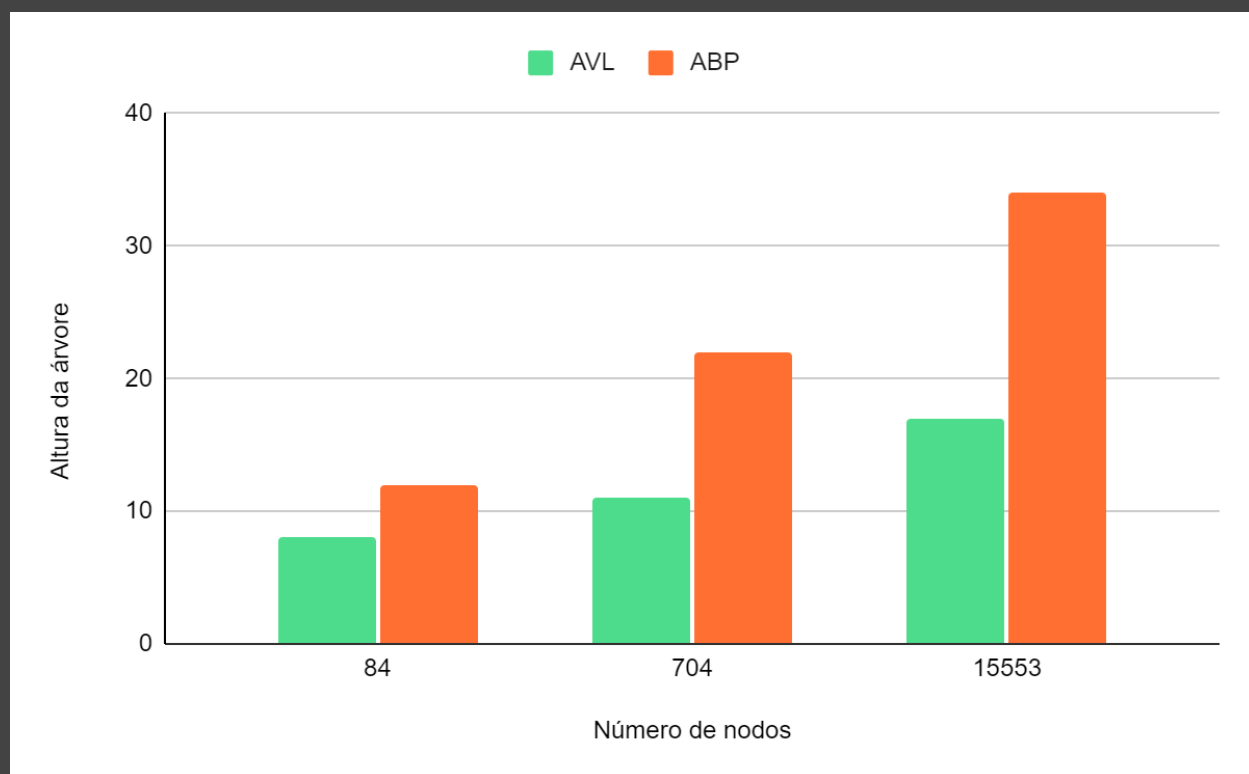
Com isso podemos resumir os resultados em gráficos para melhor visualização dos dados:



**Gráfico 1.** Análise do crescimento de nodos com o número de comparações na indexação. Os dados para a base de 10.000 tweets foram retirados para melhor visualização dos dados anteriores.



**Gráfico 2.** Análise do crescimento de nodos com o número de comparações nas consultas com 10 palavras.



**Gráfico 3.** Análise do crescimento de nodos com a altura.

Podemos então concluir que os resultados esperados foram atingidos. As árvores AVL tiveram suas alturas crescendo em um ritmo lento devido ao balanceamento, enquanto que as árvores ABP cresceram com quase o dobro de tamanho das AVL, com os testes usados.

Já, analisando os números de comparações, também confirmamos que as ABP continuam mostrando maiores resultados. No caso específico da indexação, observamos que as diferenças crescem proporcionalmente. Nas bases de 10, 100 e 10.000 tweets, temos diferenças de 173, 2241 e 10.810 comparações, respectivamente. Além disso, entre as pesquisas de 10 e 1000 palavras da busca de 10.000 tweets, a proporção também se mantém, onde a AVL é mais eficiente que a ABP.

Nas comparações de indexação da base completa, o mesmo não se verificou, mas acreditamos que tenha sido devido à montagem e à execução do programa que levou 5 horas para concluir.

Para as comparações de consultas, podemos observar que houve uma melhoria na ABP entre os totais de nodos 704 e 15553, mas isso pode ser devido tanto à montagem da árvore como nas palavras buscadas. As demais bases seguiram o resultados esperados. Assim, a ABP segue sendo não tão eficiente quanto a AVL.

## Conclusão

Por fim, em ambos os casos podemos afirmar que a árvore AVL é mais eficaz para indexações e consultas, sempre apresentando números menores de comparações em relação à ABP.

## Referências

1. EDELWEISS, N; GALANTE, R. **Estruturas de Dados**. 1. ed. Editora Bookman – Série de Livros Didáticos Informática UFRGS, 2009.
2. MOREIRA, Viviane. **Árvores Binárias de Pesquisa**. 15 slides. Disponível em: [https://moodle.inf.ufrgs.br/pluginfile.php/64754/course/section/10775/abp\\_ERE.pdf](https://moodle.inf.ufrgs.br/pluginfile.php/64754/course/section/10775/abp_ERE.pdf)
3. MOREIRA, Viviane. **Árvores Balanceadas**. 20 slides. Disponível em: <https://moodle.inf.ufrgs.br/pluginfile.php/64754/course/section/10776/INF01203-17-avl.pdf?time=1557776915501>