

Relatório Projeto 3.1 AED 2021/2022

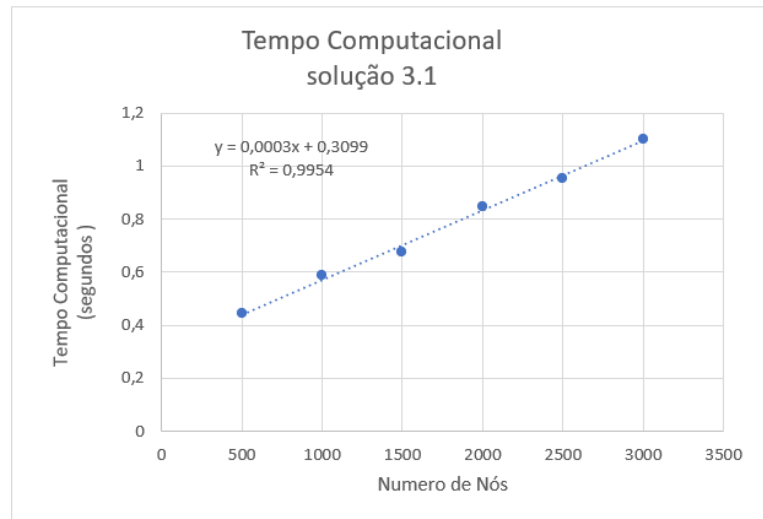
Nome: Tomás Bernardo Martins Dias
PL (inscrição): PL3

NºEstudante: 2020215701
Login no Mooshak: 2020215701

Tabela

Nº Nós	Tempo Computacional					
	500	1000	1500	2000	2500	3000
Tempo	0,0442	0,5859	0,6741	0,8488	0,955	1,1016

Gráfico



A expressão $f(N)$ está de acordo com o esperado? Justifique.

A expressão está de acordo com o esperado, pois pela análise teórica da solução implementada, uma árvore não binária, a sua complexidade seria linear, ou seja, " $O(n)$ ", pois no método recursivo usado para a criação apenas passa uma vez em cada nó, ou seja, percorre os n nós da árvore. Pelo valor do coeficiente da regressão verifica-se que os dados estão muito próximos da regressão obtida havendo três valores que estão na vizinhança da mesma.

O projeto 3.1 pode ser implementado seguindo uma abordagem iterativa e uma recursiva.

Explique sucintamente o essencial das duas implementações em termos de estruturas de dados utilizadas e do cálculo da valorização das categorias e impressão da árvore.

Independentemente da abordagem utilizada a estrutura de dados adequada a implementar é uma árvore não binária.

Seguindo uma abordagem iterativa é necessário salvar sempre o nó, cujos filhos estão a ser adicionados, caso esses filhos também possuam filhos, sendo por isso necessário guardar o pai e o número de filhos do nó. O cálculo da valorização será feito no momento de criação da árvore. Para o print da árvore criamos uma FIFO e antes de começarmos a processar a fila, observamos o número de elementos naquele nível. Criamos um novo nível depois de processados todos os elementos do nível anterior.

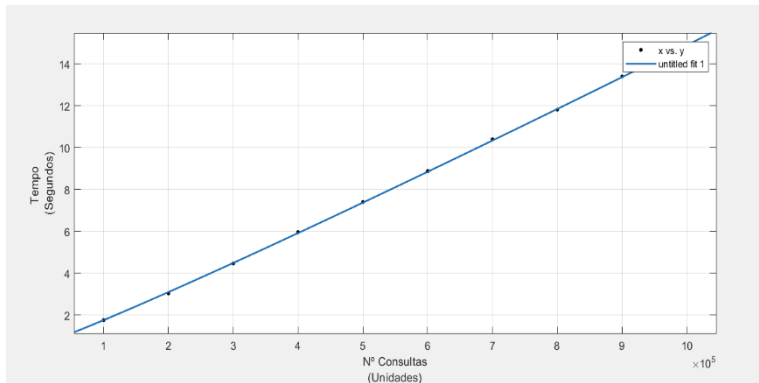
Seguindo uma abordagem recursiva, a criação da árvore e o cálculo da valorização das categorias baseia-se na ideia de criar primeiro o nó, a raiz, verificar se este tem filhos, e adicionar-lhe os filhos e calcular a valorização do nó com base nos mesmos, no passo seguinte é chamada a mesma função onde esses filhos passam a ser a raiz, e assim sucessivamente. O print da árvore tem o mesmo princípio sendo apenas necessário ter uma variável que controle os níveis dos elementos presentes.

Relatório Projeto 3.2 AED 2021/2022

Nome: Tomás Bernardo Martins Dias
PL (inscrição): PL3

Nº Estudante: 2020215701
Login no Mooshak: 2020215701

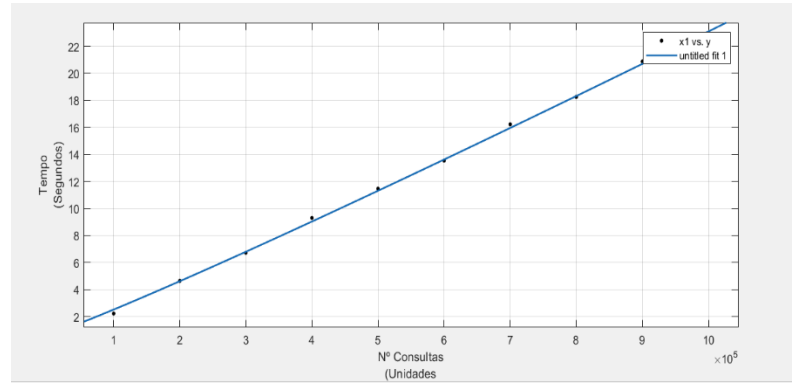
Cenário 1



Nº Consultas	Tempo(s)
1000000	1,755737505
2000000	3,034521394
3000000	4,466276293
4000000	5,964473293
5000000	7,405778792
6000000	8,881395698
7000000	10,4048409
8000000	11,8148419
9000000	13,40827911
10000000	14,8333995

General model:
 $f(x) = a \cdot x \cdot \log(x) + c$
Coefficients (with 95% confidence bounds):
 $a = 1.038e-06$ (1.029e-06, 1.047e-06)
 $c = 0.5629$ (0.4872, 0.6385)
Goodness of fit:
SSE: 0.02009
R-square: 0.9999
Adjusted R-square: 0.9999
RMSE: 0.05011

Cenário 2



Nº Consultas	Tempo(s)
1000000	2,231337202
2000000	4,648415299
3000000	6,729419105
4000000	9,301118598
5000000	11,4626061
6000000	13,54519489
7000000	16,22849231
8000000	18,2609108
9000000	20,8754964
10000000	22,7188898

General model:
 $f(x) = a \cdot x \cdot \log(x) + c$
Coefficients (with 95% confidence bounds):
 $a = 1.624e-06$ (1.582e-06, 1.667e-06)
 $c = 0.6567$ (0.3023, 1.011)
Goodness of fit:
SSE: 0.4408
R-square: 0.999
Adjusted R-square: 0.9989
RMSE: 0.2347

A evolução dos tempos de execução está de acordo com o esperado? Justifique.

Os tempos de execução para a estrutura desenvolvida, uma Splay Tree, estão de acordo com o esperado. Nesta estrutura de dados os nós inseridos/consultados mais recentemente estão mais próximos da raiz da árvore, logo quando 90% dos acessos são feitos a 5% dos artigos os tempos de execução serão mais baixos pois como os artigos estão mais próximos da raiz é necessário percorrer menos nós e fazer menos rotações durante o splay do nó. No cenário 2 como todos os artigos tem sensivelmente os mesmos acessos será necessário fazer mais rotações durante o splay do nó o que conduz a tempos superiores.

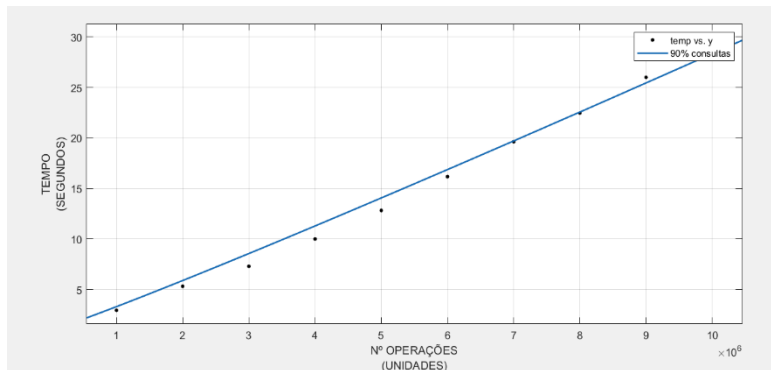
Concluo também que a sua complexidade teórica $O(n \log n)$ também foi comprovada pelos resultados obtidos.

Relatório Projeto 3.3 AED 2021/2022

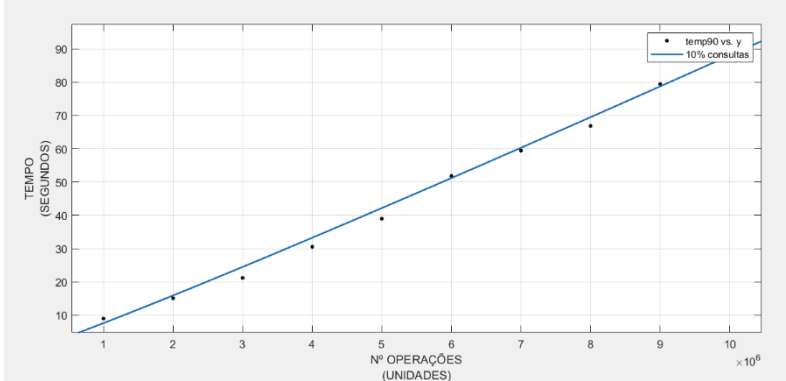
Nome: Tomás Bernardo Martins Dias
PL (inscrição): PL3

Nº Estudante: 2020215701
Login no Mooshak: 2020215701

Cenário 1



Cenário 2



10% INSERÇÕES	
Nº Operações	Tempo(s)
1000000	2,93958
2000000	5,31763
3000000	7,29856
4000000	10,00459
5000000	12,82527
6000000	16,17853
7000000	19,63125
8000000	22,47409
9000000	26,00682
10 000 000	29,96134

General model:
 $f(x) = a \cdot x \cdot \log_2(x) + b$
Coefficients (with 95% confidence bounds):
 $a = 1.178e-07$ (1.067e-07, 1.29e-07)
 $b = 0.9543$ (-0.6233, 2.532)
Goodness of fit:
SSE: 8.628
R-square: 0.9885
Adjusted R-square: 0.9871
RMSE: 1.039

90% INSERÇÕES	
Nº Operações	Tempo(s)
1000000	8,99611
2000000	15,06949
3000000	21,20331
4000000	30,55156
5000000	38,97886
6000000	51,81884
7000000	59,44358
8000000	66,86525
9000000	79,43067
10000000	93,245878

General model:
 $f(x) = a \cdot x \cdot \log_2(x) + c$
Coefficients (with 95% confidence bounds):
 $a = 3.784e-07$ (3.475e-07, 4.093e-07)
 $c = 0.07295$ (-4.304, 4.45)
Goodness of fit:
SSE: 66.43
R-square: 0.9908
Adjusted R-square: 0.9897
RMSE: 2.882

Os tempos de execução estão de acordo com o esperado? Justifique.

Os tempos de execução estão de acordo com o esperado logo, a escolha da estrutura (árvore AVL), foi correta, porque se obteve tempos computacionais mais baixos quando há menos inserções e mais consultas.

Ao nível teórico a complexidade tanto da consulta como da inserção numa AVL com n nós é “ $O(n \log n)$ ” e pela análise dos coeficientes de regressão dos dois cenários, os tempos computacionais obtidos estão muito próximos do esperado, apesar de existirem alguns outliers que podem ser causados pela aleatoriedade dos testes.

Concluído tantos os tempos de execução com a complexidade estão de acordo com o esperado.

Relatório Projeto 3.4 AED 2021/2022

Nome: Tomás Bernardo Martins Dias
PL (inscrição): PL3

Nº Estudante: 2020215701
Login no Mooshak: 2020215701

Estrutura de Dados Principal usada em cada sub-projeto:

PROJ 3.1: Árvore Genérica

PROJ 3.2: Árvore Splay

PROJ 3.3: Árvore AVL

Estruturas de dados usadas	Árvore Genérica	Árvore Splay	Árvore AVL
VANTAGENS GERAIS (max 3)	<ul style="list-style-type: none">• Um nó pode ter vários filhos• Árvore com várias ramificações	<ul style="list-style-type: none">• Nós recentemente inseridos/consultados estão mais próximos da raiz	<ul style="list-style-type: none">• Árvore auto balanceada (altura menor)• Tempos computacionais mais baixos para consultas• Árvore não degenera, devido as operações de balanceamento da árvore
DESvantagens GERAIS (max 3)	<ul style="list-style-type: none">• Tempos computacionais elevados tanto a nível da inserção como consulta• Árvore não equilibrada	<ul style="list-style-type: none">• Necessário fazer splay de todos os nós inseridos/consultados• Árvore pode degenerar e tornar-se numa lista ligada caso os elementos sejam inseridos por ordem crescente	<ul style="list-style-type: none">• Necessário verificar/rebalancear a árvore a cada inserção• Tempos de inserção mais elevados devido as operações de rebalanceamento da árvore
Justificação para a escolha no PROJ 3.1	A escolha foi feita porque “cada nó poderia ter n filhos”, ou seja, teria de ser uma árvore genérica.		
Justificação para a escolha no PROJ 3.2	A escolha foi feita porque “90% dos acessos são feitos a 5%” dos artigos, o que tira proveito de que nas árvores Splay os nós recentemente inseridos/consultados se encontram mais próximos da raiz e será necessário percorrer menos nós para consultar o nó desejado.		
Justificação para a escolha no PROJ 3.3	A escolha foi feita porque realizam se “muito mais consultas ao sistema do que inserções”, ou seja, usando uma árvore AVL tirando partido do facto de ela ser quase-equilibrada e apesar de ser necessário equilibrar a árvore a cada inserção, as consultas são muito mais rápidas tanto para inserir novos cartões como para atualizar cartões.		