

18/05/2023

RELATÓRIO: CÓDIGO DE ATIVIDADE COM ÁRVORE BINÁRIA DE BUSCA

1. CÓDIGO DE INSERÇÃO (QUESTÃO 1):

1.1. OBJETIVO E FUNCIONALIDADE DO CÓDIGO:

Esse código visa criar uma árvore binária de busca a partir de vetores de números e realizar inserções. De uma forma mais detalhada, nas seguintes etapas:

- Gera um vetor com números sequenciais de 10 a 100.000;
- Embaralha aleatoriamente o vetor para evitar repetições ou sobreposições dos valores;
- Insere os valores na árvore em ordem crescente de 100 em 100 elementos;
- Calcula a quantidade de comparações feitas durante as inserções.
- Guarda o valor das comparações em um vetor;
- Grava os dados das comparações em um arquivo em duas colunas para posteriormente gerar um gráfico no Gnuplot relacionando o número de inserções com o tamanho do grupo de números inseridos.

1.2. ESTRUTURAS E FUNÇÕES:

Primeiramente, é definida uma estrutura 'arv' que representa um nodo da árvore. A estrutura contém um valor inteiro e dois ponteiros para os filhos esquerdo e direito.

Funções utilizadas no programa:

- **arv_criavazia()** (cria uma árvore vazia);
- **arv_cria()** (cria um novo nodo);
- **arv_insere()** (insere um valor na árvore);

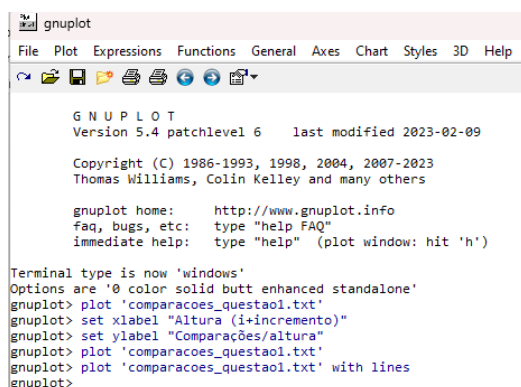
- **contar_comparacoes()** (conta o número de comparações necessárias para encontrar um valor na árvore);
- **liberar_arvore()** (libera a memória ocupada pela árvore);
- **embaralhar_vetor()** (embaralha os elementos de um vetor).

1.3. EXECUÇÃO E ARQUIVO DE SAÍDA:

Na `main()`, é criado um vetor chamado 'numeros' para armazenar os números que serão inseridos na árvore. O vetor é preenchido com valores sequenciais de `TAM_INICIO(10)` a `TAM_FINAL(100.000)`. Em seguida, o vetor é embaralhado usando a função `embaralhar_vetor()`.

Um arquivo chamado "comparações_questao1.txt" é aberto para escrita. Logo após, é alocado dinamicamente um vetor para armazenar o número de comparações para cada tamanho de grupo de números inseridos na árvore. Um loop é iniciado para inserir os valores na árvore em ordem crescente, começando a partir de `TAM_INICIO`. Para cada tamanho de grupo, uma nova árvore é criada e o número de comparações é contado usando a função `contar_comparacoes()`. O valor do número de comparações é armazenado no vetor 'comparacoesVetor'. Posteriormente, o número de comparações e o tamanho do grupo são escritos no arquivo.

1.4. ANÁLISE DO GRÁFICO:



```

gnuplot
File Plot Expressions Functions General Axes Chart Styles 3D Help

G N U P L O T
Version 5.4 patchlevel 6   last modified 2023-02-09

Copyright (C) 1986-1993, 1998, 2004, 2007-2023
Thomas Williams, Colin Kelley and many others

gnuplot home:      http://www.gnuplot.info
faq, bugs, etc:   type "help FAQ"
immediate help:    type "help" (plot window: hit 'h')

Terminal type is now 'windows'
Options are '0 color solid butt enhanced standalone'
gnuplot> plot 'comparacoes_questao1.txt'
gnuplot> set xlabel "Altura (i+incremento)"
gnuplot> set ylabel "Comparações/altura"
gnuplot> plot 'comparacoes_questao1.txt'
gnuplot> plot 'comparacoes_questao1.txt' with lines
gnuplot> _

```

Figura 1 – Script Gnuplot Q1 (Inserção)

Primeiramente são definidos os rótulos dos eixos x ("Altura") e y ("Comparações/altura"), e o título do gráfico. Depois, é plotado o arquivo 'comparacoes_questao1.txt' com as linhas do gráfico, onde a primeira coluna representa o tamanho do grupo e a segunda coluna representa o número de comparações

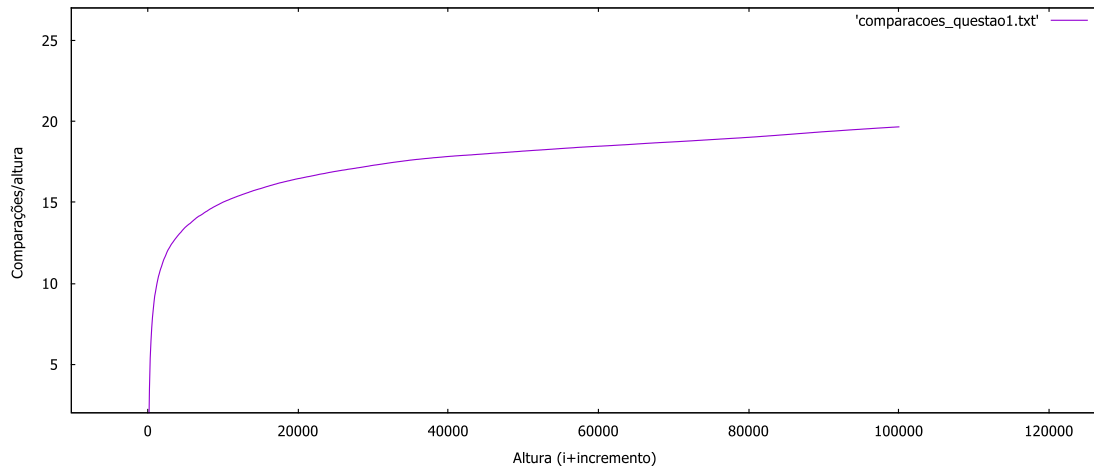


Figura 2 – Gráfico Gnuplot Q1 (Inserção)

O gráfico gerado com a saída do código 1 no Gnuplot mostra a relação entre o número de comparações e o tamanho do grupo de números inseridos na árvore (Altura). Os dados mostram um padrão crescente, mas a taxa de crescimento diminui à medida que o valor de x aumenta. Isso sugere um comportamento logarítmico, no qual o crescimento é mais rápido inicialmente e se torna mais lento à medida que x aumenta.

2. CÓDIGO DE REMOÇÃO (QUESTÃO 2):

2.1.OBJETIVO E FUNCIONALIDADE DO CÓDIGO:

No código da questão 2, o objetivo é remover elementos de uma árvore binária de busca. As principais etapas são:

- Cria uma árvore binária de busca com 100.000 números;
- Realiza a remoção de elementos da árvore em grupos de tamanho 1000;
- Conta o número de comparações feitas durante as remoções;
- Grava os dados das comparações em um arquivo relacionando o tamanho atual da árvore com o número de comparações, para posteriormente gerar um gráfico no Gnuplot.

2.2. ESTRUTURAS E FUNÇÕES:

Assim como no código 1, é definida uma estrutura de dados chamada ``Arv``, que representa um nó da árvore binária de busca. Cada nó contém um valor (``val``) e dois ponteiros para os nós filhos esquerdo (``esq``) e direito (``dir``).

Funções utilizadas no programa:

- **`arv_cria(int v)`** (cria um novo nodo);
- **`arv_insere(Arv *a, int v)`** (insere um valor na árvore);
- **`arv_remove(Arv *a, int v)`** (remove um valor da árvore);
- **`liberar_arvore()`** (libera a memória ocupada pela árvore);

2.3. EXECUÇÃO E ARQUIVO DE SAÍDA:

Na `int main()`, é declarado um vetor ``numeros`` com o tamanho inicial de 100.000, logo após é aberto um arquivo chamado "comparacoes_questao2.txt". Dois vetores são alocados dinamicamente: ``comparacoesVetor``, para armazenar o número de comparações realizadas em cada grupo, e ``tamanhos``, para armazenar o tamanho dos grupos removidos. Em seguida, uma árvore binária de busca vazia é criada e os números do vetor ``numeros`` são inseridos na árvore usando a função `arv_insere`.

A partir desse ponto, o código entra em um loop para realizar a remoção dos números em grupos de 1.000. Para cada grupo, o número de comparações é contabilizado e armazenado, o tamanho atual é registrado e é atualizado subtraindo o tamanho do grupo removido. Esse processo é repetido até que todos os grupos tenham sido removidos. Após as remoções, os resultados são escritos no arquivo de saída, que armazena em duas colunas o tamanho atual da árvore e o número de comparações feitas, para posteriormente gerar um gráfico no Gnuplot.

2.4. ANÁLISE DO GRÁFICO:

```
gnuplot
File Plot Expressions Functions General Axes Chart Styles 3D Help

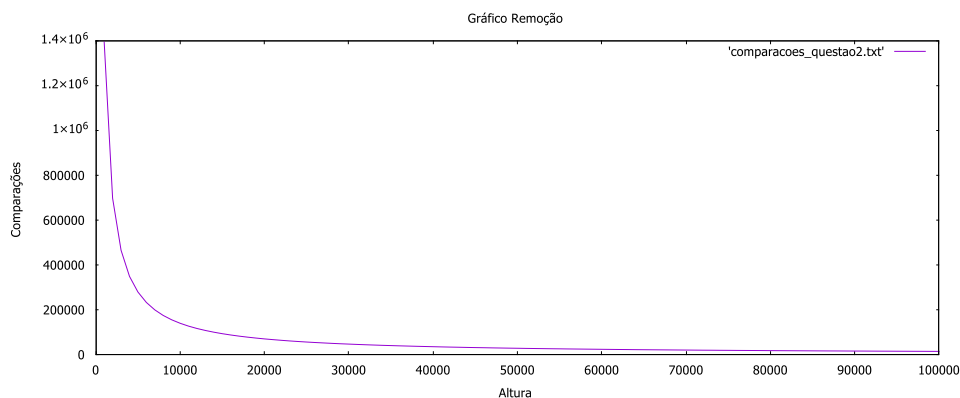
GNU PLOT
Version 5.4 patchlevel 6 last modified 2023-02-09

Copyright (C) 1986-1993, 1998, 2004, 2007-2023
Thomas Williams, Colin Kelley and many others

gnuplot home: http://www.gnuplot.info
faq, bugs, etc: type "help FAQ"
immediate help: type "help" (plot window: hit 'h')

Terminal type is now 'windows'
Options are '0 color solid butt enhanced standalone'
gnuplot> set xlabel "Altura"
gnuplot> set ylabel "Comparações"
gnuplot> set title "Gráfico Remoção"
gnuplot> plot 'comparacoes_questao2.txt' with lines
gnuplot> save 'q2_script.plt'
gnuplot> _
```

Figura 3 – Script Gnuplot Q2 (Remoção)



O gráfico mostra a relação entre o tamanho do grupo processado e o número médio de comparações realizadas por elemento nesses grupos. O eixo x representa o tamanho do grupo (altura), que diminui à medida que nos movemos da esquerda para a direita no gráfico. O eixo y representa o número médio de comparações, que aumenta à medida que nos movemos de baixo para cima no gráfico.

Observamos uma tendência geral de aumento no número médio de comparações por elemento à medida que o tamanho do grupo diminui, um comportamento logarítmico. Inicialmente, o aumento é gradual, mas à medida que o tamanho do grupo diminui ainda mais, o número médio de comparações aumenta de forma mais significativa. Isso sugere que o algoritmo utilizado se torna menos eficiente à medida que o tamanho do grupo diminui.

São definidos os rótulos dos eixos x ("Altura") e y ("Comparações"), e o título do gráfico. Depois, é plotado o arquivo 'comparacoes_questao2.txt' com as linhas do gráfico, onde a primeira coluna representa o tamanho do grupo e a segunda representa o número de comparações.

3. PSEUDO DESAFIO (QUESTÃO 3):

3.1.COMPORTAMENTO DO GRÁFICO:

3.1.1. CÓDIGO DE INSERÇÃO (QUESTÃO 1):

Os dados mostram um padrão crescente, mas a taxa de crescimento diminui à medida que o valor de x aumenta. Isso sugere um comportamento logarítmico, no qual o crescimento é mais rápido inicialmente e se torna mais lento à medida que x aumenta.

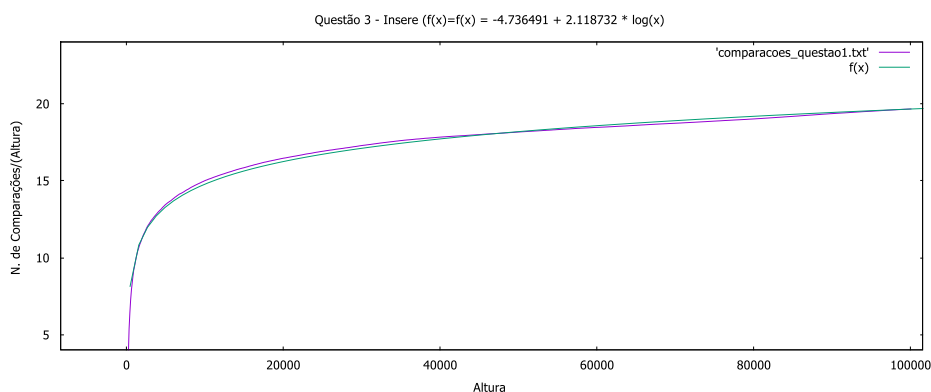


Figura 5 – Gráfico Gnuplot Q3 (Inserção)

3.1.2. CÓDIGO DE REMOÇÃO (QUESTÃO 2):

Os dados mostram um padrão decrescente, mas variação diminui à medida que o valor de x aumenta. Assim como no gráfico anterior, é um comportamento logarítmico, no qual o crescimento é mais rápido inicialmente e se torna mais lento à medida que x aumenta.

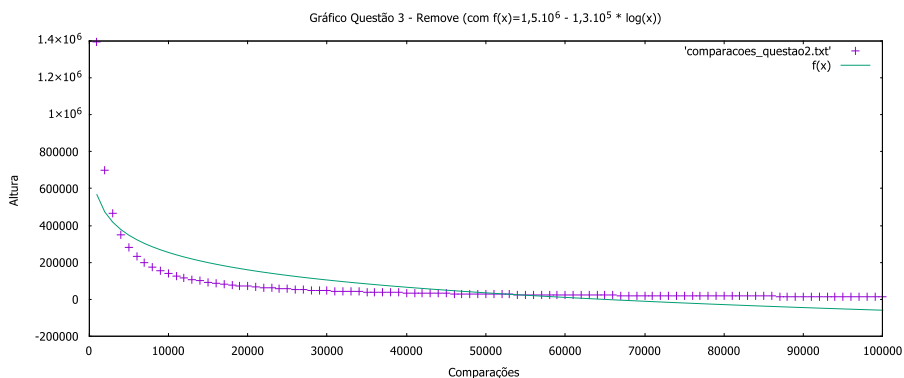


Figura 6 – Gráfico Gnuplot Q3 (Remoção)

3.2.CARACTERÍSTICAS DA FUNÇÃO:

$$f(x) = a + b * \log(x)$$

Equação 1- Função base

3.2.1. CÓDIGO DE INSERÇÃO (QUESTÃO 1):

A função $f(x) = -4,73649 + 2,11873 \log(x)$ pode ser utilizada com os parâmetros obtidos através do ajuste para modelar um comportamento específico. Os parâmetros a e b , $-4,73649$ e $2,11873$, respectivamente, desempenham papéis importantes na forma da curva da função.

O parâmetro a representa o deslocamento vertical da função base. Isso significa que a curva da função é deslocada verticalmente para para baixo em relação ao eixo y, nesse caso, a curva estará posicionada em torno do valor de $-4,7$

Por outro lado, o parâmetro b determina a inclinação da curva, que no código de inserção o valor de b é $2,11873$. A inclinação positiva indica que a curva tem uma tendência a crescer à medida que x aumenta. Quanto maior o valor de b , maior será a inclinação da curva. Nesse caso, a curva terá uma inclinação positiva relativamente acentuada. Vide figura 7 na próxima página para visualizar o resultado completo do ajuste gerado pelo Gnuplot.

3.2.2. CÓDIGO DE REMOÇÃO (QUESTÃO 2):

A função $f(x) = 1,50653e + 06 - 136004 * \log(x)$ pode ser utilizada com os parâmetros obtidos através do ajuste para modelar um comportamento específico. Portanto, ao utilizar a função base com os parâmetros $a = 1,50653 \times 10^6$ e $-1,36004 \times 10^5$, é possível modelar um comportamento em que a curva é deslocada verticalmente para cima em torno do valor de $1,5$ e apresenta uma inclinação negativa relativamente acentuada, o que indica que a curva tem uma tendência a decrescer à medida que x aumenta. Vide figura 8 na próxima página para visualizar o resultado completo do ajuste gerado pelo Gnuplot.

```

gnuplot
File Plot Expressions Functions General Axes Chart Styles 3D Help

faq, bugs, etc: type "help FAQ"
immediate help: type "help" (plot window: hit 'h')

Terminal type is now 'windows'
Options are '0 color solid butt enhanced standalone'
gnuplot> plot 'comparacoes_questao1.txt'
gnuplot> f(x) = a + b * log(x)
gnuplot> fit f(x) 'comparacoes_questao1.txt' via a, b
iter   chisq      delta/lim  lambda  a          b
0 3.7641004068e+04  0.00e+00  7.50e-00  1.000000e+00  1.000000e+00
1 2.8499503993e+02 -1.31e+07  7.50e-01  2.880679e-01  1.644833e+00
2 6.7234867048e+01 -3.24e+05  7.50e-02 -4.427157e+00  2.089574e+00
3 6.6406419862e+01 -1.25e+03  7.50e-03 -4.736288e+00  2.118712e+00
4 6.6406419505e+01 -5.36e-04  7.50e-04 -4.736491e+00  2.118732e+00
iter   chisq      delta/lim  lambda  a          b

After 4 iterations the fit converged.
final sum of squares of residuals : 66.4064
rel. change during last iteration : -5.36207e-09

degrees of freedom (FIT_NDF) : 998
rms of residuals (FIT_STDFIT) = sqrt(WSSR/ndf) : 0.257953
variance of residuals (reduced chisquare) = WSSR/ndf : 0.0665395

Final set of parameters      Asymptotic Standard Error
=====
a = -4.73649                +/- 0.08767      (1.851%)
b = 2.11873                +/- 0.008299    (0.3917%)

correlation matrix of the fit parameters:
a      b
a      1.000
b      -0.996 1.000
gnuplot> plot 'comparacoes_questao1.txt' with points, f(x) with lines
gnuplot> print "Parâmetro a: ", a
Parâmetro a: -4.73649108495259
gnuplot> print "Parâmetro b: ", b
Parâmetro b: 2.11873155605228
gnuplot> plot 'comparacoes_questao1.txt' with lines, f(x) with lines
gnuplot> plot 'comparacoes_questao1.txt' with points, f(x) with lines
gnuplot> set xlabel "Altura"
gnuplot> set ylabel "N. de Comparações/(Altura)
gnuplot> set title "Questão 3 - Insere f(x)=f(x) = -4.736491 + 2.118732 * log(x)
gnuplot>

```

Figura 7 – Script Gnuplot Q3 (Inserção)

```

gnuplot
File Plot Expressions Functions General Axes Chart Styles 3D Help

Version 5.4 patchlevel 6      last modified 2023-02-09

Copyright (C) 1986-1993, 1998, 2004, 2007-2023
Thomas Williams, Colin Kelley and many others

gnuplot home: http://www.gnuplot.info
faq, bugs, etc: type "help FAQ"
immediate help: type "help" (plot window: hit 'h')

Terminal type is now 'windows'
Options are '0 color solid butt enhanced standalone'
gnuplot> plot 'comparacoes_questao2.txt'
gnuplot> f(x) = a + b * log(x)
gnuplot> fit f(x) 'comparacoes_questao2.txt' via a, b
iter   chisq      delta/lim  lambda  a          b
0 3.1803732616e+12  0.00e+00  7.52e+00  1.000000e+00  1.000000e+00
1 2.7605796761e+12 -1.52e+04  7.52e-01  2.036861e+04  3.837647e+03
2 1.3883033508e+12 -9.88e+04  7.52e-02  8.698874e+05 -7.608609e+04
3 1.0798968102e+12 -2.86e+04  7.52e-03  1.501793e+06 -1.355579e+05
4 1.0798797451e+12 -1.58e+00  7.52e-04  1.506528e+06 -1.360036e+05
5 1.0798797451e+12 -8.88e-09  7.52e-05  1.506528e+06 -1.360036e+05
iter   chisq      delta/lim  lambda  a          b

After 5 iterations the fit converged.
final sum of squares of residuals : 1.07988e+12
rel. change during last iteration : -8.885e-14

degrees of freedom (FIT_NDF) : 98
rms of residuals (FIT_STDFIT) = sqrt(WSSR/ndf) : 104972
variance of residuals (reduced chisquare) = WSSR/ndf : 1.10192e+10

Final set of parameters      Asymptotic Standard Error
=====
a = 1.50653e+06             +/- 1.203e+05   (7.988%)
b = -136004                +/- 1.137e+04   (8.359%)

correlation matrix of the fit parameters:
a      b
a      1.000
b      -0.996 1.000
gnuplot> plot 'comparacoes_questao2.txt' with points, f(x) with lines
gnuplot> print "Parâmetro a: ", a
Parâmetro a: 1506528.43012314
gnuplot> print "Parâmetro b: ", b
Parâmetro b: -136003.595960713

```

Figura 8 – Script Gnuplot Q3 (Remoção)

4. CONSIDERAÇÕES FINAIS:

Ambos os códigos utilizam árvores binárias de busca para realizar operações de inserção e remoção de elementos. O código 1 analisa o desempenho da inserção em relação ao número de comparações feitas, enquanto o código 2 analisa o desempenho da remoção.