Análise de Complexidade de Tempo do Método Quick Sort

Eduardo Costa de Paiva eduardocspv@gmail.com Frederico Franco Calhau fredericoffc@gmail.com Gabriel Augusto Marson gabrielmarson@live.com

Faculdade de Computação Universidade Federal de Uberlândia

18 de dezembro de 2015

Lista de Figuras

2.1	Complexidade de tempo do método Quick Sort (Vetor Aleatório)	11
2.2	Complexidade de tempo do método Quick Sort com mínimos quadrados (Ve-	
	tor Aleatório)	12
2.3	Complexidade de tempo do método Quick Sort (Vetor Ordenado Crescente)	12
2.4	Complexidade de tempo do método Quick Sort com mínimos quadrados (Ve-	
	tor Ordenado Crescente)	13
2.5	Complexidade de tempo do método Quick Sort (Vetor Ordenado Decrescente)	13
2.6	Complexidade de tempo do método Quick Sort com mínimos quadrados (Ve-	
	tor Ordenado Decrescente)	14
2.7	Complexidade de tempo do método Quick Sort (Vetor Parcialmente Ordenado	
	Crescente)	14
2.8	Complexidade de tempo do método Quick Sort com mínimos quadrados (Ve-	
	tor Parcialmente Ordenado Crescente)	15
2.9	Complexidade de tempo do método Quick Sort (Vetor Parcialmente Ordenado	
	Decrescente)	15
2.10	Complexidade de tempo do método Quick Sort com mínimos quadrados (Ve-	
	tor Parcialmente Ordenado Decrescente)	16

Lista de Tabelas

3.1	Vetor Aleatorio	17
3.2	Vetor Ordenado Crescente	18
3.3	Vetor Ordenado Decrescente	18
3.4	Vetor Parcialmente Ordenado Crescente	19
3.5	Vetor Parcialmente Ordenado Decrescente	19
3.6	Dados para Análise de Memória	19

Lista de Listagens

1.1	QuickSort.py	7
	testeGeneric.py	
1.3	monitor.py	9
A.1	testdriver.py	22

Sumário

Li	ista de Figuras	
Li	sta de Tabelas	3
1	Introdução 1.1 Diretório 1.2 Códigos de programas	6 6 7
2	Gráficos	11
3	Tabelas	17
4	Análise	20
5	Citações e referências bibliográficas	21
\mathbf{A}	pêndice	22
A	Códigos extensos	22
	A.1 testdriver.py	22

Introdução

Este documento foi feito com o intuito de exibir uma análise do algoritmo Quick Sort com relação a tempo. Além disso, será feita uma comparação da curva de tempo do que se espera do algoritmo, ou seja, O(nlgn) com melhor caso e $\theta(n^2)$ no pior caso.

1.1 Diretório

Dada a seguinte organização das pastas, utilizamos o arquivo testdriver.py, executando, uma função conveniente por vez. Para mais informações vá até ao apêndice.

OBS.: É necessário instalar o programa tree pelo terminal. Isso pode ser feito da seguinte maneira.

```
> sudo apt-get install tree
```

A seguir é mostrada a organização das pastas sendo que os diretórios significativas para o projeto são Codigos e Relatorio além do raíz:

```
tree --charset=ASCII -d
|-- Codigos
  |-- Bubble
   | `-- __pycache__
  |-- Bucket
      `-- <u>__</u>pycache__
   |-- Counting
       `-- __pycache__
  |-- Heap
       `-- __pycache__
 |-- Insertion
  | `-- __pycache__
   |-- Quick
      `-- <u>pycache</u>
   |-- Quick
   | `-- __pycache__
    `-- Selection
    `-- __pycache__
|-- Other
```

```
|-- Plot
|-- __pycache__
`-- relatorio
    |-- imagens
        |-- Bubble
        |-- Bucket
       |-- Counting
        |-- Heap
        |-- Insertion
        |-- Quick
        |-- Quick
        |-- Radix
        `-- Selection
    |-- Relatorio_Bubble
    |-- Relatorio_Bucket
    |-- Relatorio_Counting
    |-- Relatorio Heap
    |-- Relatorio_Insertion
    |-- Relatorio_Quick
    |-- Relatorio_Selection
    `-- Resultados
        |-- Bubble
        |-- Bucket
        |-- Counting
        |-- Heap
        |-- Insertion
        |-- Quick
        |-- Quick
        `-- Selection
47 directories
```

1.2 Códigos de programas

Seguem os códigos utilizados na análise de tempo do algoritmo Quick Sort.

1. QuickSort.py: Disponível na Listagem 1.1.

Listagem 1.1: QuickSort.py

```
1 #@profile
2 def Particiona(A,p,r):
    x = A[r] #pivo é o último elemento
    i = p-1
    for j in range(p,r):
5
      if A[j] <= x:
6
       i = i + 1
7
        aux = A[i]
        A[i] = A[j]
9
       A[j] = aux
10
    temp = A[i+1]
11
    A[i+1] = A[r]
12
13
   A[r] = temp
   return i+1
14
15
```

```
16 #@profile
17 def quickSort(A,p,r):
    if(p < r):
      q = Particiona(A, p, r)
19
      quickSort (A, p, (q-1))
      quickSort(A,(q+1),r)
21
    return A
22
23
24 @profile
25 def quick(A):
    quickSort (A, 0, (len(A)-1))
28
29
31 #A = [i for i in range(10)]
32 #quick (A)
33 #quickSort (A, 0, 6)
34 #print(A)
```

2. testeGeneric.py Disponível na Listagem 1.2

Listagem 1.2: testeGeneric.py

```
1 ##adicionei - Serve para importar arquivos em outro diretório
2 ### A CADA NOVO MÉTODO MUDAR O IMPORT, A CHAMADA DA FUNÇÃO E O SYS.
     PATH
4 import sys
5 sys.path.append('/home/gmarson/Git/AnaliseDeAlgoritmos/Trabalho_Final
     /Codigos/Radix')
6 sys.path.append('/home/gmarson/Git/AnaliseDeAlgoritmos/Trabalho_Final
7
8 sys.setrecursionlimit(200000)
10 from monitor import *
11 from memoria import *
13 from RadixSort import *
14 import argparse
16 parser = argparse.ArgumentParser()
17 parser.add_argument("n", type=int, help="número de elementos no vetor
      de teste")
18 args = parser.parse_args()
20 v = criavet(args.n)
21 radix(v)
22
25 ## A EXECUÇÃO DESSE ARQUIVO EH ASSIM
26 ## NA LINHA DE COMANDO VC MANDA O NOME DO ARQUIVO E O TAMANHO DO
     ELEMNTO DO vetor
27 ##EXEMPLO testeBubble.py 10
28 ##ele gera um vetor aleatório (criavet) e manda pro bubble_sort
```

3. monitor.py Disponível na Listagem 1.3

Listagem 1.3: monitor.py

```
1 # Para instalar o Python 3 no Ubuntu 14 ou 15
3 # sudo apt-get install python3 python3-numpy python3-matplotlib
      ipython3 python3-psutil
4 #
6 from math import *
7 import gc
8 import random
9 import numpy as np
10
11
12 from tempo import *
14 # Vetores de teste
15 def troca(m, v, n): ## seleciona o nível de embaralhamento do vetor
16
      m = trunc(m)
      mi = (n-m)//2
17
      mf = (n+m)//2
18
      for num in range(mi, mf):
19
          i = np.random.randint(mi, mf)
20
           j = np.random.randint(mi, mf)
21
          #print("i= ", i, " j= ", j)
22
          t = v[i]
23
24
          v[i] = v[j]
25
          v[j] = t
      return v
26
27
28
29 def criavet(n, grau=0, inf=0, sup=0.999999999):
      passo = (sup - inf)/n
30
      if grau < 0.0:
31
          v = np.arange(sup, inf, -passo)
          if grau <= -1.0:
33
               return v
34
35
          else:
               return troca (-grau*n, v, n)
36
      elif grau > 0.0:
37
          v = np.arange(inf, sup, passo)
38
          if grau >= 1.0:
39
               return v
40
          else:
41
               return troca(grau*n, v, n)
42
43
      else:
          #return np.random.randint(inf, sup, size=n)
44
          return [random.random() for i in range(n)] # for bucket sort
45
46
47
48
49 #print(criavet(20))
50
51 #Tipo
                                                  grau
52 #aleatorio
                                                   0
53 #ordenado_crescente
                                                   1
                                                  -1
54 #ordenado_decrescente
55 #parcialmente_ordenado_crescente
                                                  0.5
56 #parcialmente_ordenado_decrescente
                                               -0.5
57
```

```
58
59 def executa(fn, v):
60      gc.disable()
61      with Tempo(True) as tempo:
62      fn(v)
63      gc.enable()
```

 $4.\ test$ driver.py Referenciado no apêndice A.

Gráficos

Seguem os Gráficos utilizadas no processo de análise do método Quick Sort:

- 1. Para um vetor aleatório
 - (a) Complexidade de tempo do método Quick Sort disponível na lista de imagens 2.1.

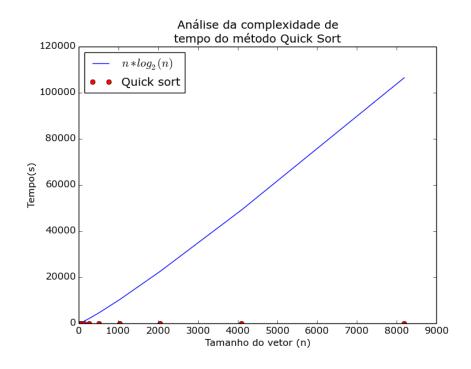


Figura 2.1: Complexidade de tempo do método Quick Sort (Vetor Aleatório)

- (b) Complexidade de tempo do método Quick Sort com mínimos quadrados disponível na lista de imagens 2.2.
- 2. Complexidade de tempo do método Quick Sort disponível na lista de imagens 2.3.
- 3. Complexidade de tempo do método Quick Sort com mínimos quadrados disponível na lista de imagens 2.4.

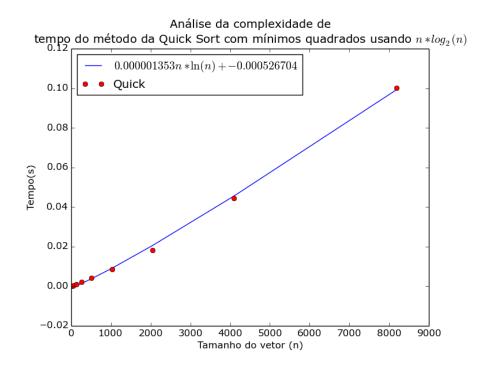


Figura 2.2: Complexidade de tempo do método Quick Sort com mínimos quadrados (Vetor Aleatório)

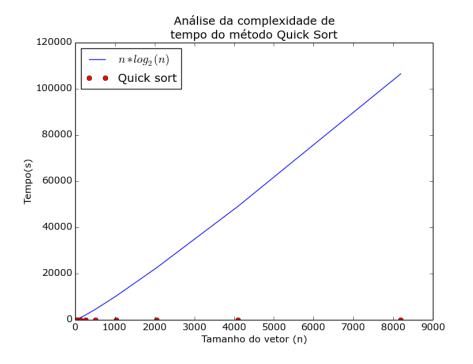


Figura 2.3: Complexidade de tempo do método Quick Sort (Vetor Ordenado Crescente)

- 1. Complexidade de tempo do método Quick Sort disponível na lista de imagens 2.5.
- 2. Complexidade de tempo do método Quick Sort com mínimos quadrados disponível na lista de imagens 2.6.
- 1. Complexidade de tempo do método Quick Sort disponível na lista de imagens 2.7.

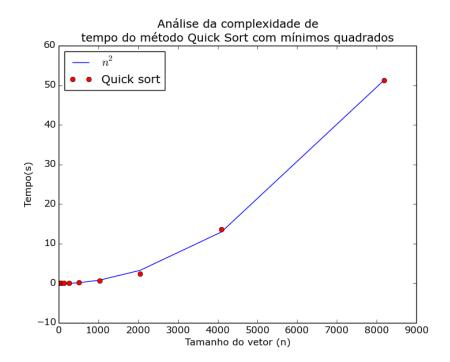


Figura 2.4: Complexidade de tempo do método Quick Sort com mínimos quadrados (Vetor Ordenado Crescente)

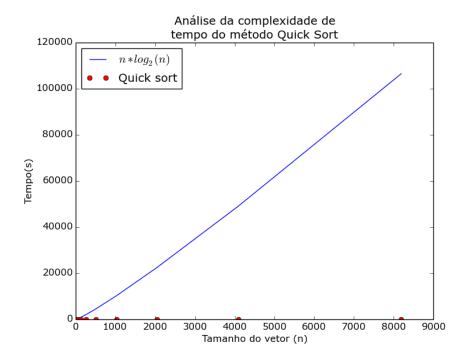


Figura 2.5: Complexidade de tempo do método Quick Sort (Vetor Ordenado Decrescente)

- 2. Complexidade de tempo do método Quick Sort com mínimos quadrados disponível na lista de imagens 2.8.
- 1. Complexidade de tempo do método Quick Sort disponível na lista de imagens 2.9.
- 2. Complexidade de tempo do método Quick Sort com mínimos quadrados disponível na lista de imagens 2.10.

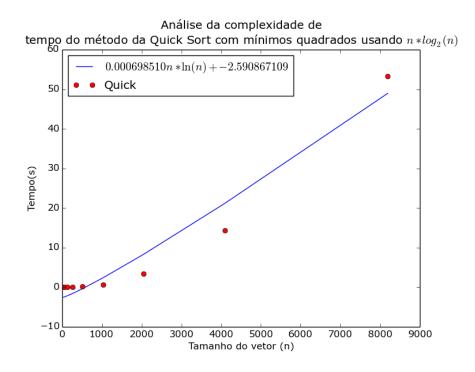


Figura 2.6: Complexidade de tempo do método Quick Sort com mínimos quadrados (Vetor Ordenado Decrescente)

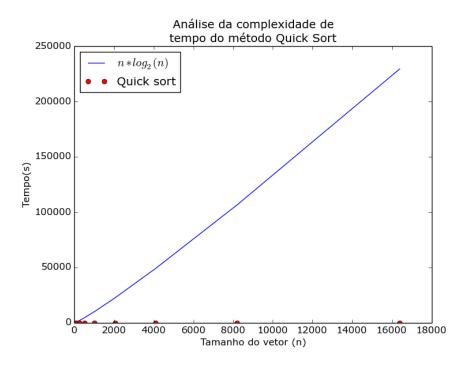


Figura 2.7: Complexidade de tempo do método Quick Sort (Vetor Parcialmente Ordenado Crescente)

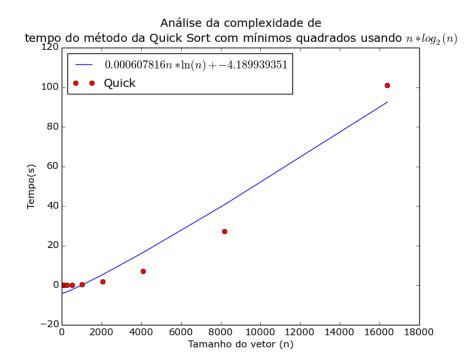


Figura 2.8: Complexidade de tempo do método Quick Sort com mínimos quadrados (Vetor Parcialmente Ordenado Crescente)

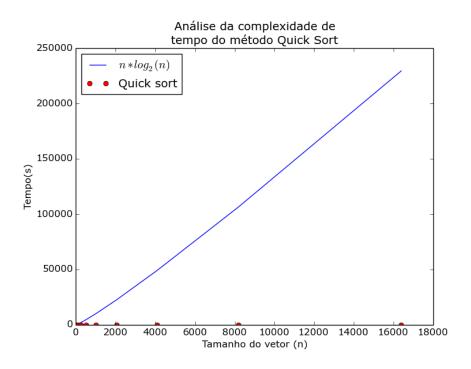


Figura 2.9: Complexidade de tempo do método Quick Sort (Vetor Parcialmente Ordenado Decrescente)

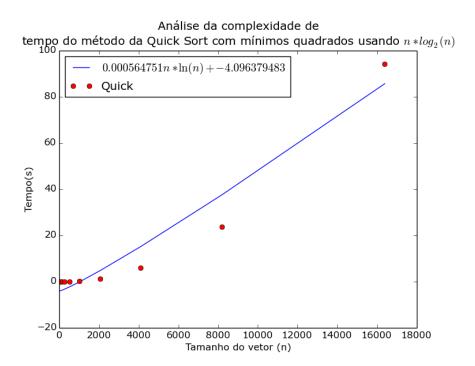


Figura 2.10: Complexidade de tempo do método Quick Sort com mínimos quadrados (Vetor Parcialmente Ordenado Decrescente)

Tabelas

Seguem as tabelas utilizadas para a análise do método Quick Sort.

Tabela 3.1: Vetor Aleatorio

Tamanho do Vetor	Comparações	Tempo(s)
32	95	0.002739
64	219	0.006053
128	475	0.013715
256	1203	0.033465
512	2334	0.067288
1024	5539	0.162378
2048	11118	0.331024
4096	26759	0.837335
8192	51853	1.817348
16384	110392	4.521457
32768	237663	12.608580
65536	491980	39.949010

Tabela 3.2: Vetor Ordenado Crescente

Tamanho do Vetor	Comparações	Tempo(s)
32	94	0.002563
64	239	0.006553
128	476	0.013587
256	1039	0.030059
512	2293	0.066084
1024	5084	0.147069
2048	11228	0.331113
4096	24201	0.766142
8192	51017	1.773804
16384	108998	4.481584
32768	251194	12.579250
65536	497903	39.168110

Tabela 3.3: Vetor Ordenado Decrescente

Tamanho do Vetor	Comparações	$\overline{\text{Tempo}(s)}$
32	95	0.002664
64	231	0.006120
128	500	0.013601
256	1112	0.031692
512	2386	0.069012
1024	5132	0.149191
2048	10950	0.327610
4096	23895	0.753161
8192	52772	1.847495
16384	109071	4.501062
32768	242107	12.890980
65536	497543	39.163110

Tabela 3.4: Vetor Parcialmente Ordenado Crescente

Tamanho do Vetor	Comparações	Tempo(s)
32	92	0.002681
64	188	0.006145
128	485	0.013650
256	1086	0.031333
512	2314	0.067024
1024	5056	0.150097
2048	11063	0.329353
4096	24663	0.797015
8192	52275	1.806541
16384	110549	4.538953
32768	234740	12.476440
65536	514762	40.043390

 Tabela 3.5: Vetor Parcialmente Ordenado Decrescente

Tamanho do Vetor	Comparações	$\overline{\text{Tempo}(s)}$
32	87	0.002646
64	221	0.006141
128	597	0.016766
256	1058	0.030065
512	2316	0.066540
1024	5330	0.157471
2048	10864	0.329800
4096	23856	0.752486
8192	54738	1.895006
16384	114088	4.755551
32768	245564	12.874460
65536	496850	39.231150

Tabela 3.6: Dados para Análise de Memória

Tamanho do Vetor	Memória (MiB)
32	24.102000
64	24.199000
128	23.121000
256	24.078000
512	24.242000
1024	24.215000
2048	24.10864
4096	24.230000
8192	24.168000
16384	24.738000

Análise

O algoritmo Quick Sort é um algoritmo de divisão e conquista que utiliza a função particiona, $\theta(n)$, que tem como objetivo ordenar as partições. O algoritmo quicksort possui o melhor caso em $\theta(nlgn)$ e pior caso em $\theta(n^2)$. O pior caso ocorre quando o vetor já está ordenado. Podemos observar que todas as curvas de todos os gráficos, exceto os de complexidade de tempo sem a interpolação dos mínimos quadrados (Gráficos 2.1,2.3,2.5,2.7,2.9), apresentaram uma correspondência forte com a curva da função $F(x) = x^2$, o que nos permite concluir que, dada a complexidade de tempo do algoritmo Quick Sort por G(x) então F(x) = c * G(x) sendo que c é uma constante maior que zero e $x > x_0$. Portanto, o Quick Sort é $\theta(nlgn)$.

Citações e referências bibliográficas

[1] Algoritmos: Teoria e Prática. Thomas H. Cormen Today

Apêndice A

Códigos extensos

A.1 testdriver.py

Listagem A.1: testdriver.py

```
1 # coding = utf-8
2 import subprocess
3 import numpy as np
4 import matplotlib.pyplot as plt
5 import sys , shutil
8 ##PRA CADA NOVO METODO TEM QUE MUDAR
9 #Sys.path()
11 ## PARA CADA VETOR NOVO OU NOVO METODO TEM QUE MUDAR
12 #Para o executa_teste a chamada das funções e o shutil.move()
13 #para os plots
                        a chamada das funções e o savefig
15 sys.path.append('/home/gmarson/Git/AnaliseDeAlgoritmos/Trabalho_Final/
     Codigos/Bucket') ## adicionei o código de ordenação
16 sys.path.append('/home/gmarson/Git/AnaliseDeAlgoritmos/Trabalho_Final/
     relatorio/Resultados/Bucket') ## adicionei o resultado do executa_teste
17
18
 def executa_teste(arqteste, arqsaida, nlin, intervalo):
      """Executa uma sequência de testes contidos em arqteste, com:
         arqsaida: nome do arquivo de saída, ex: tBolha.dat
21
         nlin: número da linha no arquivo gerado pelo line_profiler contendo
22
               os dados de interesse. Ex: 14
23
         intervalo: tamanhos dos vetores: Ex: 2 ** np.arange(5,10)
25
      f = open(arqsaida, mode='w', encoding='utf-8')
26
      f.write('# n
                         tempo(s)\n')
27
28
      for n in intervalo:
29
          cmd = ' '.join(["kernprof -l -v", "testeGeneric.py", str(n)])
30
          str_saida = subprocess.check_output(cmd, shell=True).decode('utf-8
          linhas = str_saida.split('\n')
32
          #for i in linhas:
```

```
print(i)
34
          #print (linhas)
35
          unidade_tempo = float(linhas[1].split()[2])
36
          tempo_total = float(linhas[3].split()[2])
37
          #lcomp = linhas[nlin].split()
39
          #print ("unidade tempo: ",unidade_tempo )
40
          #print("lcomp: ",lcomp)
41
42
          #print("tempo total", tempo_total)
43
          #num_comps = int(lcomp[1])
44
          str_res = '{:>8} {:13.6f}'.format(n, tempo_total)
45
          print (str_res)
46
          f.write(str_res + '\n')
47
      f.close()
48
      #shutil.move("tBucket_vetor_parcialmente_ordenado_decrescente.dat", "/
          home/gmarson/Git/AnaliseDeAlgoritmos/Trabalho_Final/relatorio/
          Resultados/Bucket/tBucket_vetor_parcialmente_ordenado_decrescente.
          dat")
51 executa_teste("testeGeneric.py", "
     tBucket_vetor_parcialmente_ordenado_decrescente.dat", 46, 2 ** np.
     arange(5,15))
  def executa_teste_memoria(arqteste, arqsaida, nlin, intervalo):
53
      """Executa uma sequência de testes contidos em arqteste, com:
54
         arqsaida: nome do arquivo de saída, ex: tBolha.dat
55
         nlin: número da linha no arquivo gerado pelo line_profiler contendo
56
                os dados de interesse. Ex: 14
57
         intervalo: tamanhos dos vetores: Ex: 2 ** np.arange(5,10)
58
      .....
59
      f = open(arqsaida, mode='w', encoding='utf-8')
60
      f.write('#
                      n
                          comparações
                                             tempo(s)\n')
61
62
      for n in intervalo:
63
          cmd = ' '.join(["kernprof -l -v ", "testeGeneric.py", str(n)])
64
65
          str_saida = subprocess.check_output(cmd, shell=True).decode('utf-8)
66
              ')
67
          linhas = str_saida.split('\n')
68
          for i in linhas:
69
              print(i)
70
71
          print ("Linhas:", linhas[1])
72
73
          unidade_tempo = float(linhas[1].split()[2])
74
75
76
77
          str_res = '{:>8} {:>13} {:13.6f}'.format(n, n, n)
78
          print (str_res)
          f.write(str_res + '\n')
79
      f.close()
80
      #shutil.move("tSelection_memoria.dat", "/home/gmarson/Git/
81
          AnaliseDeAlgoritmos/Trabalho_Final/relatorio/Resultados/Selection/
          tSelection_memoria.dat")
82
83 #executa_teste_memoria("testeGeneric.py", "tSelection_memoria.dat", 14, 2
     ** np.arange(5,15))
```

```
84
85 def plota_teste1(argsaida):
       n, c, t = np.loadtxt(arqsaida, unpack=True)
86
       #print("n: ",n,"\nc: ",c,"\nt: ",t)
87
       #n eh o tamanho da entrada , c eh o tanto de comparações e t eh o
          tempo gasto
       plt.plot(n, n ** 2, label='$n^2$') ## custo esperado bubble Sort
89
       plt.plot(n, c, 'ro', label='selection sort')
90
91
       # Posiciona a legenda
92
       plt.legend(loc='upper left')
93
94
       # Posiciona o título
95
       plt.title('Análise de comparações do método da seleção')
96
97
       # Rotula os eixos
       plt.xlabel('Tamanho do vetor (n)')
99
       plt.ylabel('Número de comparações')
100
101
       plt.savefig('relatorio/imagens/Selection/
102
          selection_plot_1_ordenado_descresente.png')
       plt.show()
103
104
105
106
107 def plota_teste2(arqsaida):
       n, t = np.loadtxt(arqsaida, unpack=True)
108
       plt.plot(n, n , label='$n$')
109
       plt.plot(n, t, 'ro', label='bucket sort')
110
111
       # Posiciona a legenda
112
       plt.legend(loc='upper left')
113
114
       # Posiciona o título
115
       plt.title('Análise da complexidade de \ntempo do método Bucket Sort')
116
117
       # Rotula os eixos
118
       plt.xlabel('Tamanho do vetor (n)')
119
       plt.ylabel('Tempo(s)')
120
121
       plt.savefig('relatorio/imagens/Bucket/
122
          bucket_plot_2_parcialmente_ordenado_decrescente.png')
       plt.show()
123
124
125
126 def plota_teste3(arqsaida):
       n, t = np.loadtxt(arqsaida, unpack=True)
127
128
       # Calcula os coeficientes de um ajuste a um polinômio de grau 2 usando
129
       # o método dos mínimos quadrados
130
131
       coefs = np.polyfit(n, t, 2)
       p = np.poly1d(coefs)
132
133
       plt.plot(n, p(n), label='$n$')
134
       plt.plot(n, t, 'ro', label='bucket sort')
135
136
       # Posiciona a legenda
137
       plt.legend(loc='upper left')
138
```

```
# Posiciona o título
       plt.title('Análise da complexidade de \ntempo do método Bucket Sort
141
          com mínimos quadrados')
142
       # Rotula os eixos
143
       plt.xlabel('Tamanho do vetor (n)')
144
       plt.ylabel('Tempo(s)')
145
146
147
       plt.savefig('relatorio/imagens/Bucket/
          bucket_plot_3_parcialmente_ordenado_decrescente.png')
       plt.show()
148
149
150 #plota_teste1("/home/gmarson/Git/AnaliseDeAlgoritmos/Trabalho_Final/
      relatorio/Resultados/Selection/tSelection_vetor_ordenado_descresente.
      dat")
151 plota_teste2("/home/gmarson/Git/AnaliseDeAlgoritmos/Trabalho_Final/
      relatorio/Resultados/Bucket/
      tBucket_vetor_parcialmente_ordenado_decrescente.dat")
152 plota_teste3("/home/gmarson/Git/AnaliseDeAlgoritmos/Trabalho_Final/
      relatorio/Resultados/Bucket/
      tBucket_vetor_parcialmente_ordenado_decrescente.dat")
153
154
155
  def plota_teste4(arqsaida):
       n, c, t = np.loadtxt(argsaida, unpack=True)
156
157
       # Calcula os coeficientes de um ajuste a um polinômio de grau 2 usando
158
       # o método dos mínimos quadrados
159
160
       coefs = np.polyfit(n, c, 2)
       p = np.poly1d(coefs)
161
162
       plt.plot(n, p(n), label='$n^2$')
163
       plt.plot(n, c, 'ro', label='bubble sort')
164
165
       # Posiciona a legenda
166
       plt.legend(loc='upper left')
167
168
       # Posiciona o título
169
       plt.title('Análise da complexidade de \ntempo do método da bolha')
170
171
       # Rotula os eixos
172
       plt.xlabel('Tamanho do vetor (n)')
173
       plt.ylabel('Número de comparações')
174
175
       plt.savefig('bubble4.png')
176
       plt.show()
177
179 def plota_teste5(arqsaida):
       n, c, t = np.loadtxt(arqsaida, unpack=True)
180
181
182
       # Calcula os coeficientes de um ajuste a um polinômio de grau 2 usando
       # o método dos mínimos quadrados
183
       coefs = np.polyfit(n, c, 2)
184
       p = np.poly1d(coefs)
185
186
       # set_yscale('log')
187
       # set_yscale('log')
188
       plt.semilogy(n, p(n), label='n^2')
189
       plt.semilogy(n, c, 'ro', label='bubble sort')
190
```

```
191
       # Posiciona a legenda
192
       plt.legend(loc='upper left')
193
194
       # Posiciona o título
195
       plt.title('Análise da complexidade de \ntempo do método da bolha')
196
197
       # Rotula os eixos
198
       plt.xlabel('Tamanho do vetor (n)')
199
200
       plt.ylabel('Número de comparações')
201
       plt.savefig('bubble5.png')
202
       plt.show()
203
```