

INF1010 - ESTRUTURAS DE DADOS AVANÇADAS - 2022.1 - 3WB

Lab3 - Compactador e descompactador de arquivos texto com o algoritmo de Huffman

Nome: Eric Leão

Matrícula: 2110694

Nome: Marina Schuler Martins

Matrícula: 2110075

Relatório do Laboratório

(Enunciado) Gere um relatório indicando como construiu a árvore de prefixos, exiba a frequência dos símbolos em ordem descendente (do mais frequente para o menos frequente) e o percentual de compactação obtido. Exiba o arquivo original e o arquivo descompactado, indicando se a compactação e descompactação funcionou a contento. Indique o que não funcionou, se for o caso.

1. Construção da árvore de prefixos

Construímos a lista de frequência e, posteriormente, a árvore de prefixos utilizando a seguinte estrutura:

```
typedef struct _frequencia {  
    unsigned char ch;  
    int freq;  
    struct _frequencia *prox;  
    struct _frequencia *esq;  
    struct _frequencia *dir;  
}node_freq;
```

Primeiramente, abrimos o arquivo texto contendo o poema pedido (ou qualquer outro texto) para leitura e chamamos a função *node_freq* cria_lst_frequencia(FILE *arq, node_freq *lst, int *tamanho)*, a qual cria a lista de frequência dos caracteres do texto lido em ordem crescente (utilizando-se de algumas funções auxiliares), devolve o tamanho dessa lista e retorna o ponteiro para a cabeça da lista.

Posteriormente, chamamos a função *node_freq* gera_tree_huffman(node_freq* lst)*, passando a lista criada acima como parâmetro. Essa função funciona da seguinte forma:

- Retira os dois elementos menos frequentes, removendo-os da cabeça da lista;
- Cria um novo elemento com a junção dos dois removidos, cuja frequência é a soma das frequências desses dois e seu caracter é simbolizado por um '+'. Com junção entende-se que *novo->esq* aponta para o primeiro removido e *novo->dir* aponta para o segundo removido.
- Por fim, o novo elemento é inserido na lista de frequência mantendo a ordenação correta.

Esse processo é repetido até que reste apenas um elemento na lista de frequência (o mais frequente), que será a raiz da árvore de huffman. Deste modo, a árvore de prefixos é construída.

2. Exibindo lista de frequência

Testando o compactador com um arquivo formado pela letra do poema "Trem de Alagoas" de Ascenso Ferreira, temos a seguinte frequência dos símbolos em ordem descendente:

ch: (603)	ch: - (10)
ch: a (165)	ch: M (6)
ch: o (105)	ch: f (6)
ch: e (98)	ch: V (5)
ch: d (81)	ch: q (4)
ch: (76)	ch: j (4)
ch: n (60)	ch: ◆ (4)
ch: r (59)	ch: D (3)
ch: s (44)	ch: E (2)
ch: t (44)	ch: O (2)
ch: u (42)	ch: S (2)
ch: c (39)	ch: N (2)
ch: . (37)	ch: ◆ (2)
ch: m (34)	ch: P (2)
ch: i (30)	ch: ◆ (2)
ch: p (28)	ch: x (2)
ch: , (22)	ch: T (1)
ch: v (21)	ch: R (1)
ch: l (16)	ch: L (1)
ch: C (16)	ch: G (1)
ch: h (13)	ch: ◆ (1)
ch: ◆ (13)	ch: ◆ (1)
ch: ◆ (13)	ch: z (1)
ch: ◆ (13)	ch: : (1)
ch: A (12)	ch: J (1)
ch: ◆ (12)	ch: ◆ (1)
ch: g (11)	ch: Q (1)
ch: ! (11)	ch: ◆ (1)
ch: b (10)	ch: F (1)

Obs.: Os caracteres acentuados do texto foram impressos como ◆.

3. Percentual de compactação obtido

O tamanho do arquivo texto original com o poema é de 1799 bytes, ou seja, 14392 bits. O tamanho do arquivo binário compactado é de 1213 bytes, ou seja, 9704 bits. Assim, o percentual de compactação é dado por:

$$P = (9704/14392)*100 = 67,43 \%$$

4. Funcionamento da compactação e descompactação

Teste realizado com o arquivo formado pela letra do poema "Trem de Alagoas" de Ascenso Ferreira:

```

root@DESKTOP-NMDRPEU:~/inf1010/lab3# ./teste
Menu:
Para compactar um arquivo texto digite 0;
Para descompactar um arquivo binário digite 1;
Para sair digite -1.
0
Digite o nome do arquivo que deseja compactar: poema.txt
Digite o nome do arquivo compactado: comp.c
Menu:
Para compactar um arquivo texto digite 0;
Para descompactar um arquivo binário digite 1;
Para sair digite -1.
1
Digite o nome do arquivo que deseja descompactar: comp.c
Digite o nome do arquivo descompactado: descomp.c
Menu:
Para compactar um arquivo texto digite 0;
Para descompactar um arquivo binário digite 1;
Para sair digite -1.
-1
root@DESKTOP-NMDRPEU:~/inf1010/lab3# ls
comp.c  descomp.c  lab3_eda_eric_marina.c  poema.txt  teste
root@DESKTOP-NMDRPEU:~/inf1010/lab3# ./teste
Menu:
Para compactar um arquivo texto digite 0;
Para descompactar um arquivo binário digite 1;
Para sair digite -1.
1
Digite o nome do arquivo que deseja descompactar: comp.c
Digite o nome do arquivo descompactado: descomp2.c
Menu:
Para compactar um arquivo texto digite 0;
Para descompactar um arquivo binário digite 1;
Para sair digite -1.
-1
root@DESKTOP-NMDRPEU:~/inf1010/lab3#

```

a. Arquivo original

```
poema.txt ×
```

```
1  TREM DE ALAGOAS
2
3  O sino bate,
4  o condutor apita o apito,
5  solta o trem de ferro um grito,
6  põe-se logo a caminhar...
7
8      | | | | | | | | | | – Vou danado pra Catende,
9      | | | | | | | | | | vou danado pra Catende,
10     | | | | | | | | | | vou danado pra Catende
11     | | | | | | | | | | com vontade de chegar...
12
13  Mergulham mocambos
14  nos mangues molhados ,
15  moleques mulatos,
16  vem vê-lo passar.
17
18     | | | | | | | | | | – Adeus!
19     | | | | | | | | | | – Adeus!
20
21  Mangueiras, coqueiros,
22  cajueiros em flor,
23  cajueiros com frutos
24  já bons de chupar...
```

25
26 – Adeus, morena do cabelo cacheado!

```
28      - Vou danado pra Catende,  
29      vou danado pra Catende,  
30      vou danado pra Catende  
31      com vontade de chegar...
```

33 Na boca da mata
34 há furnas incríveis
35 que em coisas terríveis
36 nos fazem pensar:

38 - Ali mora o Pai-da-Mata!
39 - Ali é a casa das caiporas!

```
41      - Vou danado pra Catende,  
42      vou danado pra Catende,  
43      vou danado pra Catende  
44      com vontade de chegar...
```

46 Meu Deus! Já deixamos
47 a praia tão longe...
48 No entanto avistamos

49 bem perto outro mar...
50
51 Danou-se! Se move,
52 parece uma onda...
53 Que nada! É um partido
54 já bom de cortar...

56 - Vou danado pra Catende,
57 vou danado pra Catende,
58 vou danado pra Catende
59 com vontade de chegar...

61 Cana-caiana
62 cana-roxa
63 cana-fita
64 cada qual a mais bonita,
65 todas boas de chupar...

67 – Adeus, morena do cabelo cacheado!

69 - Ali dorme o Pai-da-Mata!
70 - Ali é a casa das caiporas!

72 – Vou danado pra Catende,
73 vou danado pra Catende,
74 vou danado pra Catende
75 com vontade de chegar...

b. Arquivo compactado

comp.dat x

```
1 :FQJ:zGRLTxP
N S 0 E D j q V f M -
2 b
3 ! g A
4
5
6 h
7 C l v , p i m " . % c ' u * t , s , r ; n <
8 L d Q e b o i a [ \ Z F
9 L - u E T 4 F E z Z , f \ g m ^ ; i K P Y k K o
10 Z _ 8 F & c ~ g R " 8 H - C / 3 o w V U k B H W Y 2 \ 0 k Z N W Z * m ^
+ g V 7 ? g m ^ ; - ! K ^ ; - ! K w S E i f T ) k e { X ! + W M ^ L K + ! m Y
\ M q # U . w B 1 4 # = B t p M P - Ex q D l [ ^ g m @
' C . b l 6 W J - 9 i U z B C e Y z g 2 u ) + Y w p J _ . w B 1 4 # = B
t p M P - Ex q D l [ ^ g m . R 8 & U K # X o a I c ! W
! 3 ; m B , . 6 Z E g o d 1 4 j y 3 . + U D g m * t K o
11 Z _ Ä b h F 7 z - A k " 8 H - C / 3 o z S C p 9 E n C f * M D t 2 % k X
| m E m C ! \ ` a u x s V \ b u e f R W _ z 8
12 % X c ^ F & c ~ g R Z p M P , K P Z " 8 H - C / 3 o r 5 i
```

c. Arquivo descompactado

descomp.txt x

```
1 TREM DE ALAGOAS
2
3 O sino bate,
4 o condutor apita o apito,
5 solta o trem de ferro um grito,
6 põe-se logo a caminhar...
7
8 | | | | | - Vou danado pra Catende,
9 | | | | | vou danado pra Catende,
10 | | | | | vou danado pra Catende
11 | | | | | com vontade de chegar...
12
13 Mergulham mocambos
14 nos mangues molhados ,
15 moleques mulatos,
16 vem vê-lo passar.
17
18 | | | | | - Adeus!
19 | | | | | - Adeus!
20
21 Mangueiras, coqueiros,
22 cajueiros em flor,
23 cajueiros com frutos
24 já bons de chupar...
```

25
26 | | | | | – Adeus, morena do cabelo cacheado!
27
28 | | | | | – Vou danado pra Catende,
29 | | | | | vou danado pra Catende,
30 | | | | | vou danado pra Catende
31 | | | | | com vontade de chegar...
32

33 Na boca da mata

34 há furnas incríveis
35 que em coisas terríveis
36 nos fazem pensar:

37
38 | | | | | – Ali mora o Pai-da-Mata!
39 | | | | | – Ali é a casa das caiporas!
40
41 | | | | | – Vou danado pra Catende,
42 | | | | | vou danado pra Catende,
43 | | | | | vou danado pra Catende
44 | | | | | com vontade de chegar...
45

46 Meu Deus! Já deixamos
47 a praia tão longe...
48 No entanto avistamos

49 bem perto outro mar...

50
51 Danou-se! Se move,
52 parece uma onda...
53 Que nada! É um partido
54 já bom de cortar...
55

56 | | | | | – Vou danado pra Catende,
57 | | | | | vou danado pra Catende,
58 | | | | | vou danado pra Catende
59 | | | | | com vontade de chegar...
60

61 Cana-caiana
62 cana-roxa
63 cana-fita
64 cada qual a mais bonita,
65 todas boas de chupar...
66

67 | | | | | – Adeus, morena do cabelo cacheado!

68
69 | | | | | – Ali dorme o Pai-da-Mata!
70 | | | | | – Ali é a casa das caiporas!

71
72 | | | | | – Vou danado pra Catende,

73 | | | | | vou danado pra Catende,
74 | | | | | vou danado pra Catende
75 | | | | | com vontade de chegar...
76

77 | | | | | Ascenso Ferreira

Compactação e descompactação funcionaram corretamente.