

# Cálculo da complexidade do algoritmo de Huffman

Allan Cordeiro Rocha de Araújo.

<sup>1</sup>Centro de ciência e tecnologia– Universidade Federal de Roraima (UFRR)  
Boa Vista– RR – Brasil

[allanps32008@gmail.com](mailto:allanps32008@gmail.com)

**Resumo.** *Será aqui calculado a complexidade do algoritmo de Huffman, usando um pseudocódigo específico como base para tal.*

## 1. Algoritmo de Huffman

A codificação de Huffman é uma técnica de compressão que usa da probabilidade de ocorrência dos símbolos no conjunto de dados a ser comprimido para determinar códigos de tamanho variável para cada símbolo. Criado por David A. Huffman em 1952.

É criada uma árvore binária completa recursivamente a partir da união de dois símbolos de menor probabilidade que são somados em símbolos auxiliares colocados no conjunto de símbolos.

Basicamente ele é usado para calcular uma árvore de custo mínimo.

Definição formal do árvore de custo mínimo: Dada uma família  $p_1, \dots, p_n$  de pesos e uma partição  $F$  de  $\{1, 2, \dots, n\}$ , encontrar, dentre as árvores cujo conjunto de folhas é  $F$ , uma que tenha custo mínimo em relação a  $p$ .

## 2. Cálculo da complexidade do algoritmo de Huffman

Segue o pseudocódigo:

<b>HUFFMAN</b> ( $n, p_1, \dots, p_n$ )
1 $Q \leftarrow \text{CRIA-FILA-VAZIA}()$
2 para $i$ crescendo de 1 até $n$ faça
3 $z \leftarrow \text{CRIA-CÉLULA}()$
4 $\text{índice}[z] \leftarrow i$
5 $\text{peso}[z] \leftarrow p_i$
6 $\text{esq}[z] \leftarrow \text{dir}[z] \leftarrow \text{NIL}$
7 $\text{INSERE-NA-FILA}(Q, z)$
8 enquanto $Q$ tiver 2 ou mais células faça
9 $x \leftarrow \text{EXTRAÍ-MIN}(Q)$

10  $y \leftarrow \text{EXTRAÍ-MIN}(Q)$

11  $z \leftarrow \text{CRIA-CÉLULA}()$

12  $\text{esq}[z] \leftarrow x$

13  $\text{dir}[z] \leftarrow y$

14  $\text{peso}[z] \leftarrow \text{peso}[x] + \text{peso}[y]$

15  $\text{INSERE-NA-FILA}(Q, z)$

16  $r \leftarrow \text{EXTRAÍ-MIN}(Q)$

17 devolva  $r$

Com  $n$  como tamanho da entrada;

Com as operações  $\text{INSERE-NA-FILA}()$  e  $\text{EXTRAÍ-MIN}()$  com custo  $O(m)$  cada uma;

$m$  é o número de células na fila  $Q$ ;

E sendo  $m \leq n$ ;

As linhas 2 até 7 tem um laço de repetição de  $i=1$  até  $n$ ;

As linhas 8 até 15 tem um laço de repetição de  $i=2$  até  $n$ .

Com isso o cálculo da complexidade de Huffman é:

Handwritten mathematical derivation of Huffman's algorithm complexity:

$$\begin{aligned} \text{Tendo-se } m \leq n \\ \sum_{i=1}^m m + \sum_{i=2}^m m+m+m &= \sum_{i=1}^m m + \sum_{i=2}^m 3m \\ &= m \cdot (m-1+1) + 3m \cdot (m-2+1) \\ &= m \cdot m + 3mm - 3m = 4mm - 3m \\ \text{Assim } O(m \cdot m), \text{ porém } m \leq n, \text{ logo } O(n^2) \end{aligned}$$

## **Referências**

“Árvore de Huffman”; IME; Disponível em  
<[https://www.ime.usp.br/~pf/analise\\_de\\_algoritmos/aulas/huffman.html](https://www.ime.usp.br/~pf/analise_de_algoritmos/aulas/huffman.html) >. Acesso em  
12 de julho de 2018.

“Algoritmo de Huffman para compressão de dados”;IME. Disponível em  
<<https://www.ime.usp.br/~pf/estruturas-de-dados/aulas/huffman.html>>. Acesso em 12  
de julho de 2018.