

Universidade Federal de Ouro Preto
Bacharelado em Ciência da Computação

Codificação de Huffman e Tabela Hash

Trabalho Prático III
Estrutura de Dados I

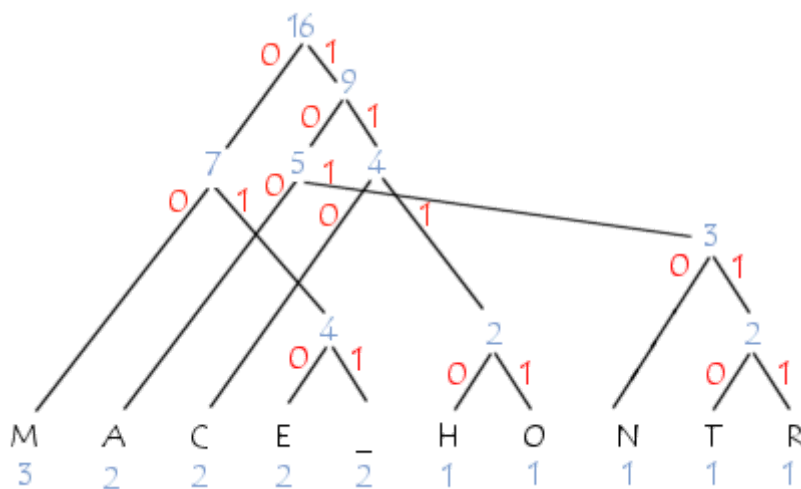
Daniel Monteiro Valério
Fabio Henrique Alves Fernandes

Professora Amanda Nascimento

Ouro Preto, 10 de dezembro de 2019

Questão A

Codificação de Huffman é um algoritmo de compressão e descompressão de arquivos, que usa as probabilidades de ocorrência dos símbolos no conjunto de dados a ser comprimido para determinar códigos de tamanho variável para cada símbolo. É criado com uma árvore binária, onde cada nó é feito recursivamente, a partir da junção dos dois símbolos de menor probabilidade, que são então somados em símbolos auxiliares e estes símbolos auxiliares recolocados no conjunto de símbolos. O processo termina quando todos os símbolos forem unidos em símbolos auxiliares, formando uma árvore binária. A árvore é então percorrida, atribuindo-se valores binários de 1 ou 0 para cada aresta, e os códigos são gerados a partir desse percurso. O código de cada símbolo corresponde à sucessão de códigos ao longo do caminho, que vai desde o caractere até a raiz. Assim, quanto mais profundo o símbolo na árvore, mais a palavra de código será longa.



Exemplo de árvore binária com a nome Mace - Hontr

Questão B

O programa consiste na leitura de um arquivo de texto. pegando cada palavra que o texto possui e, de acordo com a letra inicial, o colocando em uma tabela enumerada de 0 a 25, correspondente às letras do alfabeto. O uso da tabela hash veio para dividir uma letra do alfabeto para cada lista do vetor. Para que não haja colisão, a tabela hash foi feita com vetores de listas encadeadas, para caso tenham várias palavras com a mesma letra inicial não ocorrer interferência e o programa poder inserir a palavra na posição desejada. Ao final da leitura de todas as palavras, o programa faz a impressão de todo o conteúdo da tabela hash, onde se percorre cada lista encadeada e imprime as palavras e a frequência de repetição. Caso uma das listas não possuir nenhuma palavra, a impressão é vazia.