

Compressão Sem Perdas (Lossless Data Compression)

Refere-se a métodos de compressão de dados aplicados através de algoritmos em que a informação obtida após a descompressão é idêntica à informação original (antes de ser comprimida);

Transferência de texto, arquivos binários, etc.

Codificação Huffman

- Idéia é usar caracteres (símbolos) com número variável de bits
 - Dada uma mensagem, encontrar a melhor (menor) codificação para cada caractere
- Construir a Árvore de Prefixos de Huffman
 - Maior freqüência -> Códigos pequenos
 - Menor freqüência -> Códigos grandes

Características

- Árvore de Prefixos de Huffman
- Folhas = caracteres
- Aresta esquerda = 0
- Aresta direita = 1
- Complexidade $O(M + N \log N)$
 - M é o comprimento da mensagem
 - N é o número de caracteres distintos

Algoritmo Guloso de Huffman

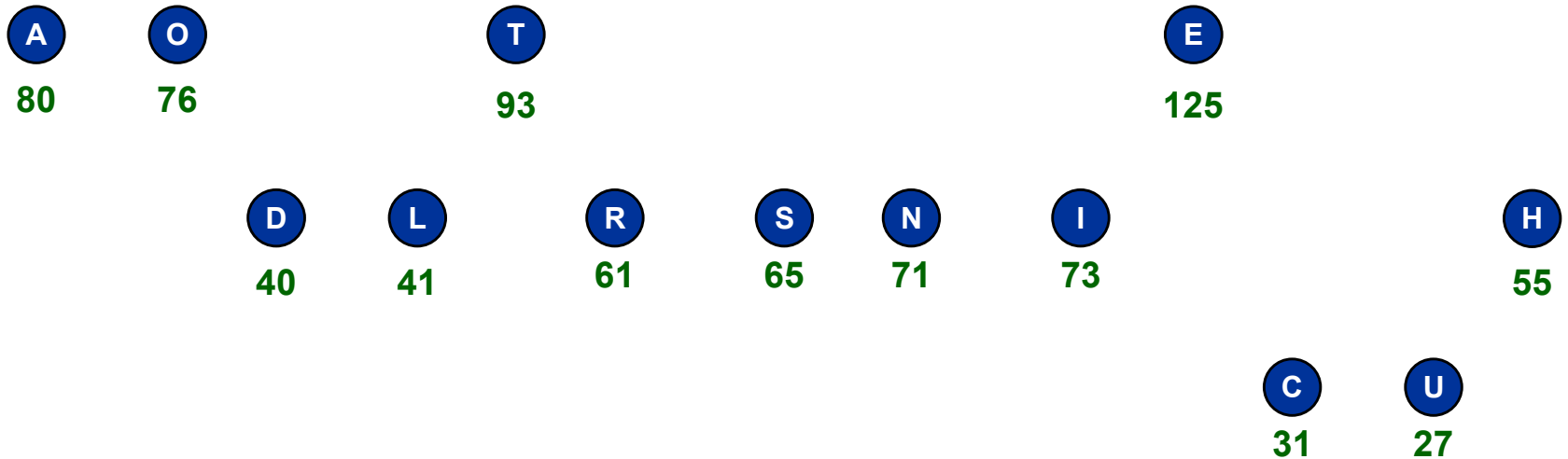
- Tabular frequências dos símbolos
- Montar uma floresta de árvores unitárias com todos os símbolos e suas frequências
- Repetir
 - Localizar os dois elementos com menor frequência F_i e F_j
 - Uní-los numa única árvore com frequência $F_i + F_j$

Exemplo de Codificação de Huffman

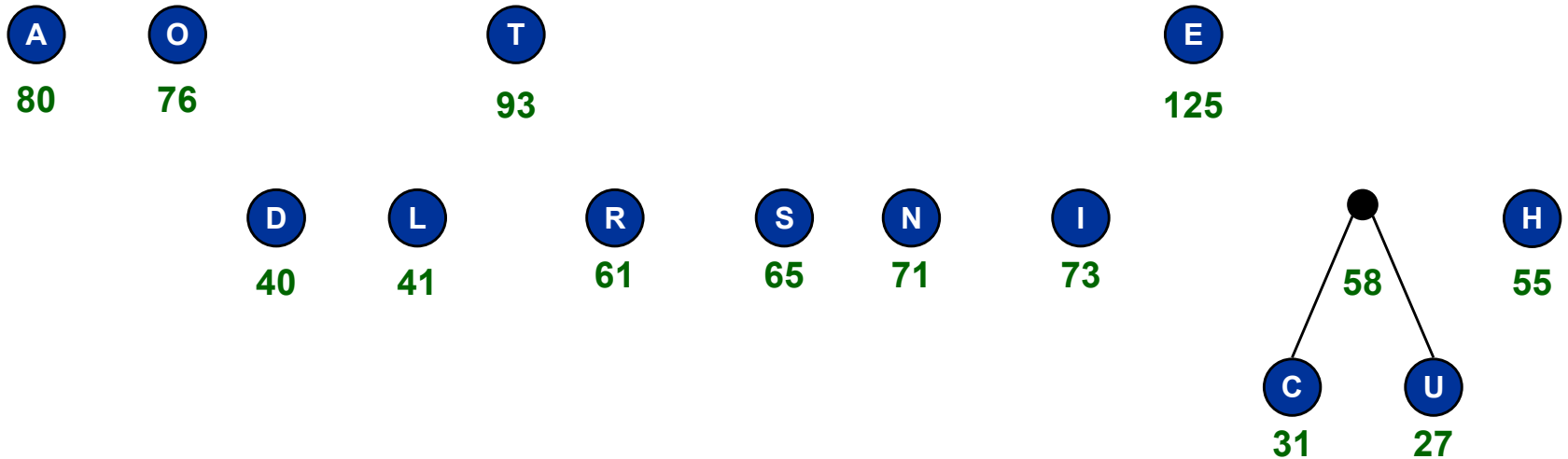
- Freqüências dos caracteres

Char	Freq
E	125
T	93
A	80
O	76
I	72
N	71
S	65
R	61
H	55
L	41
D	40
C	31
U	27

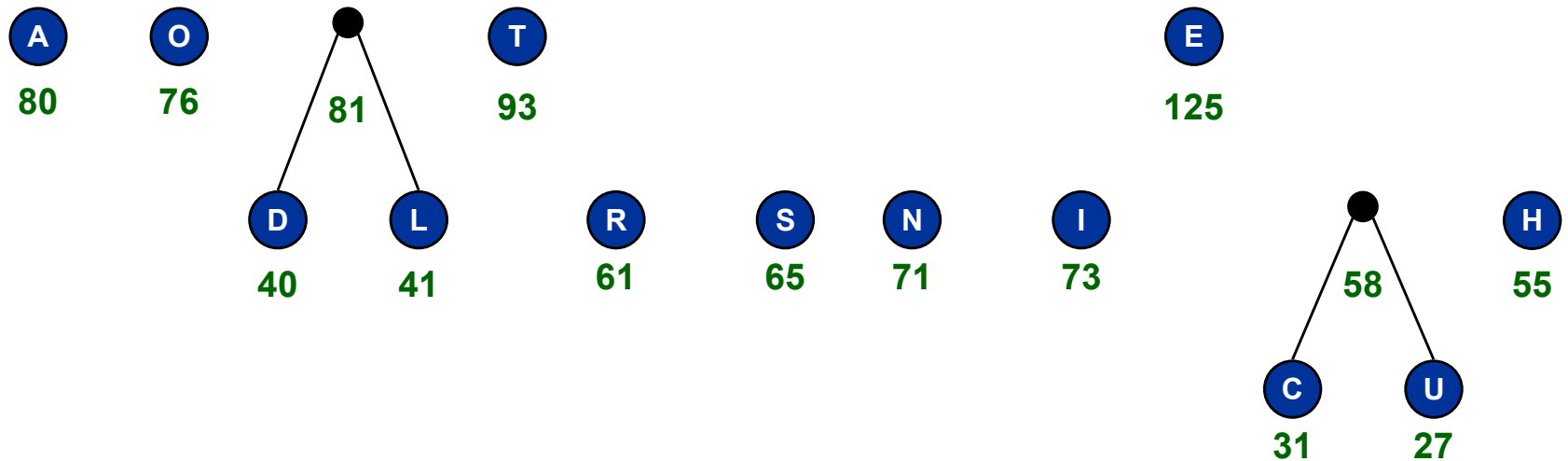
Exemplo de Codificação de Huffman



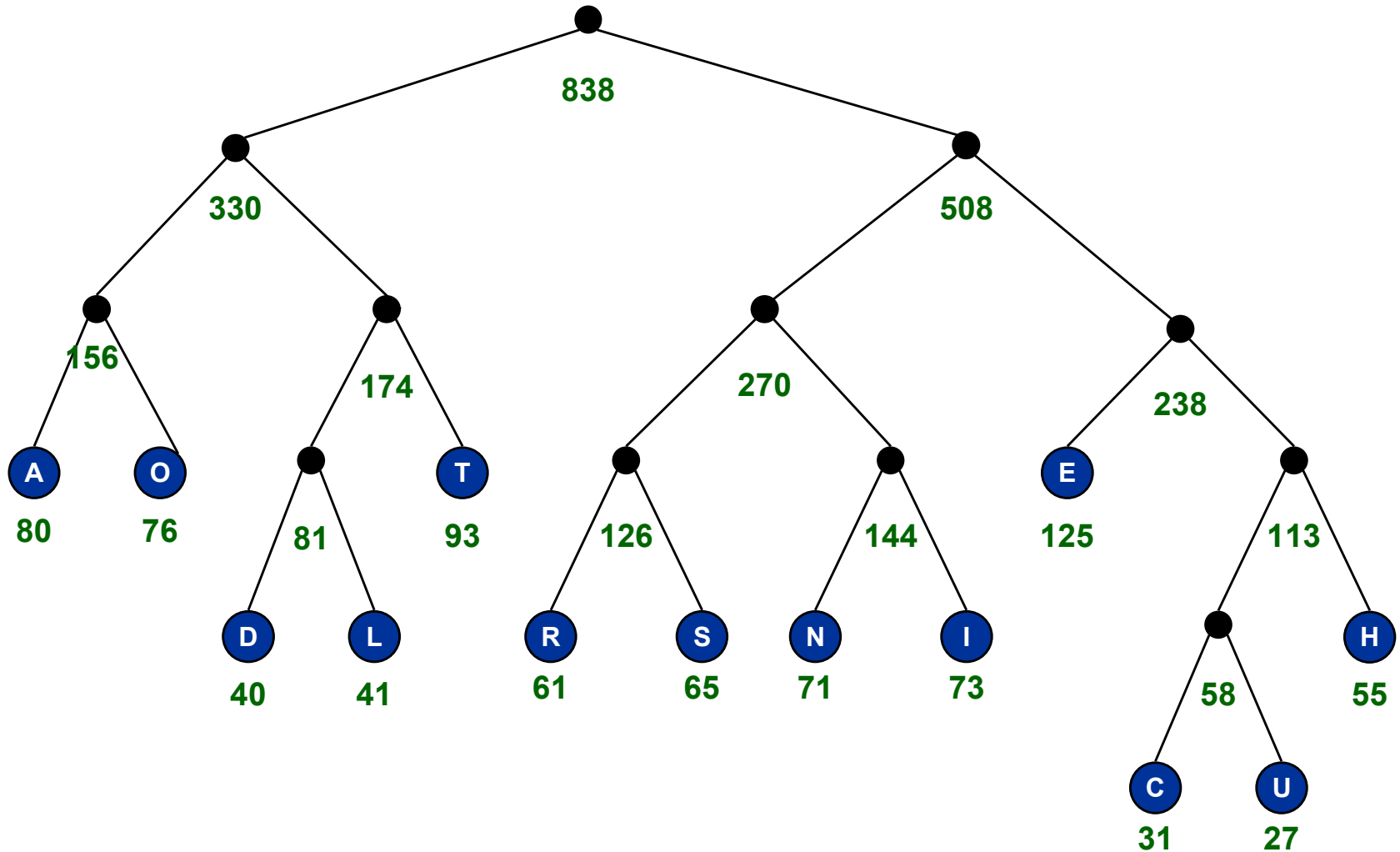
Exemplo de Codificação de Huffman



Exemplo de Codificação de Huffman



Exemplo de Codificação de Huffman



Exemplo de Codificação de Huffman

Char	Freq	Fixo	Huff
E	125	0000	110
T	93	0001	000
A	80	0010	000
O	76	0011	011
I	73	0100	1011
N	71	0101	1010
S	65	0110	1001
R	61	0111	1000
H	55	1000	1111
L	41	1001	0101
D	40	1010	0100
C	31	1011	11100
U	27	1100	11101
	838	4.00	3.62

Compressão de imagens

- O arquivo comprimido deve conter toda a seqüência de códigos junto com a árvore;
- Pode se obter as frequências através do cálculo do histograma da imagem;
- O método não utiliza a relação existente entre os pixels vizinhos abordada em outras técnicas de compressão;

Vantagem

- Redução do código para armazenar caracteres de maior frequência
- A principal vantagem:
 - Processo de descodificação torna-se bastante simples
 - Percorrer a árvore: 0 esquerda, 1 direita;

Dificuldades e Problemas

- Construir a árvore
- Construir frequência dos caracteres ou pixels
- Não é ótimo
 - Número de bits por pixel é inteiro
 - Codificação aritmética: permite número “fracionário” de bits por pixel

Referências Bibliográficas

Algoritmos e Estruturas de Dados. Disponível em: orion.lcg.ufrj.br/algoritmos/Topicos%20Extra.ppt. Acesso em: 17 out. 2008.

Codificação Huffman e Aritmética. Disponível em: www.cic.unb.br/~lamar/te729/Aulas/cod_aula2.pdf. Acesso em: 15 out. 2008.

Processamento e Comunicação Multimédia. Disponível em: www.di.ubi.pt/cursos/mestrados/mei/disciplinas/5052/fichs/Apres_Topico_2.pdf. Acesso em: 17 out. 2008.