

CÓDIGOS DE HUFFMAN

- Com três bits podemos armazenar oito valores distintos na memória do computador: 000, 001, 010, 011, 100, 101, 110 e 111.
- Com n bits podemos armazenar 2^n valores distintos.
- Se em um texto aparecem seis caracteres diferentes precisaríamos de três bits, pois $2^2=4$ e $2^3=8$.
- Ficariam sobrando duas representações.
- Considere uma coleção de dados com cerca de 50.000 ocorrências dos seguintes caracteres: A, E, I, O, U e !.
- Os caracteres ocorrem com as seguintes frequências percentuais:

| Caractere | A | E | I | O | U | ! |
|-----------|----|---|----|----|---|----|
| % | 52 | 8 | 12 | 11 | 7 | 10 |

- Poderíamos usar a codificação mostrada abaixo para armazenar os caracteres:

| Caractere | A | E | I | O | U | ! |
|-----------|-----|-----|-----|-----|-----|-----|
| Código | 000 | 001 | 010 | 100 | 110 | 101 |

- São utilizados **150.000** bits para armazenar as 50.000 ocorrências dos caracteres.
- A idéia de Huffman foi usar uma quantidade de bits diferentes para armazenar os caracteres. Os caracteres que aparecessem com maior frequência deveriam usar menos bits e os que aparecessem com menor frequência usariam mais bits.

| Caractere | A | E | I | O | U | ! |
|-----------|---|------|-----|-----|------|-----|
| Código | 1 | 0111 | 010 | 001 | 0110 | 000 |

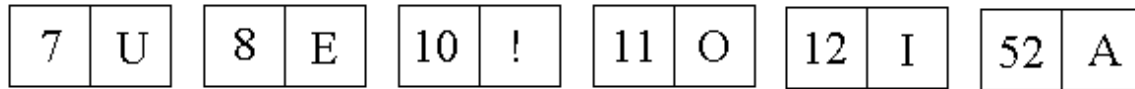
- Para armazenar as 50.000 ocorrências dos caracteres seriam necessários **105.500** bits, conforme demonstrado na tabela abaixo:

| Caractere | A | E | I | O | U | ! |
|-------------------|--------|--------|--------|--------|--------|--------|
| Código | 1 | 0111 | 010 | 001 | 0110 | 000 |
| Nº de bits | 1 | 4 | 3 | 3 | 4 | 3 |
| Nº de ocorrências | 26.000 | 4.000 | 6.000 | 5.500 | 3.500 | 5.000 |
| Nº de bits totais | 26.000 | 16.000 | 18.000 | 16.500 | 14.000 | 15.000 |

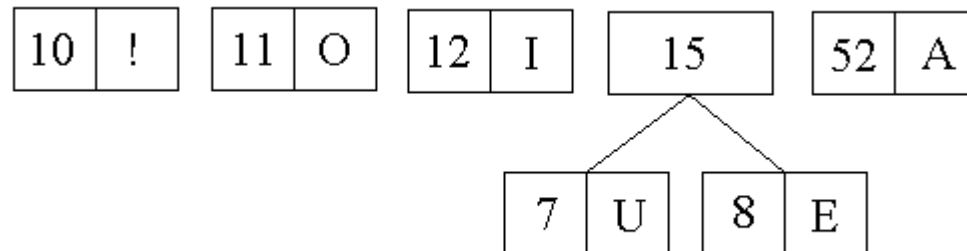
- Economizando aproximadamente 30% de espaço de armazenamento.

CONSTRUÇÃO DA ÁRVORE DE HUFFMAN

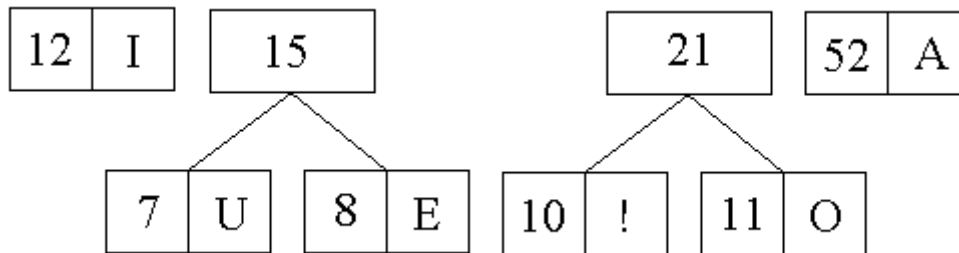
ETAPA 1



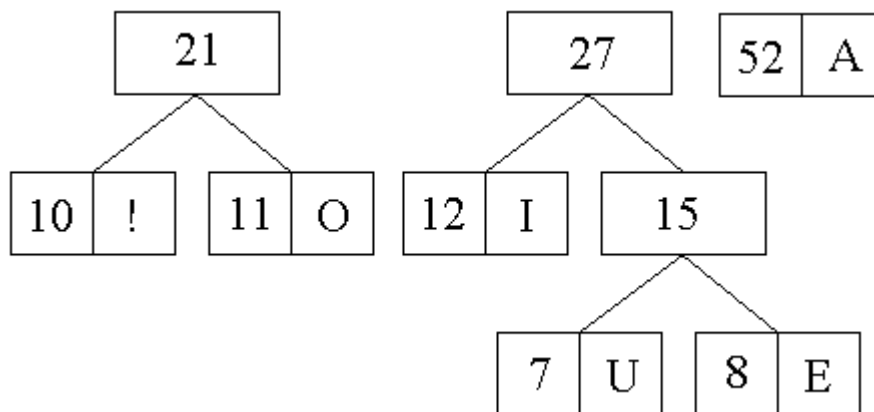
ETAPA 2



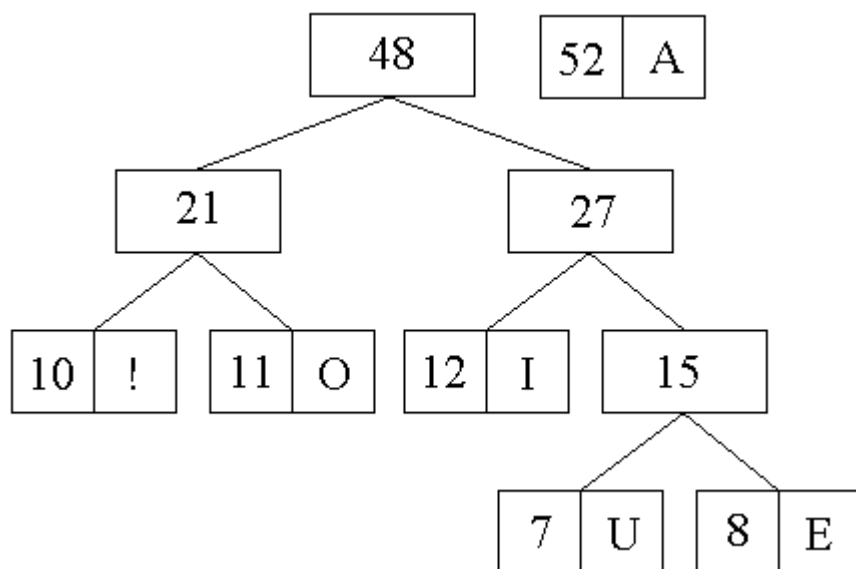
ETAPA 3



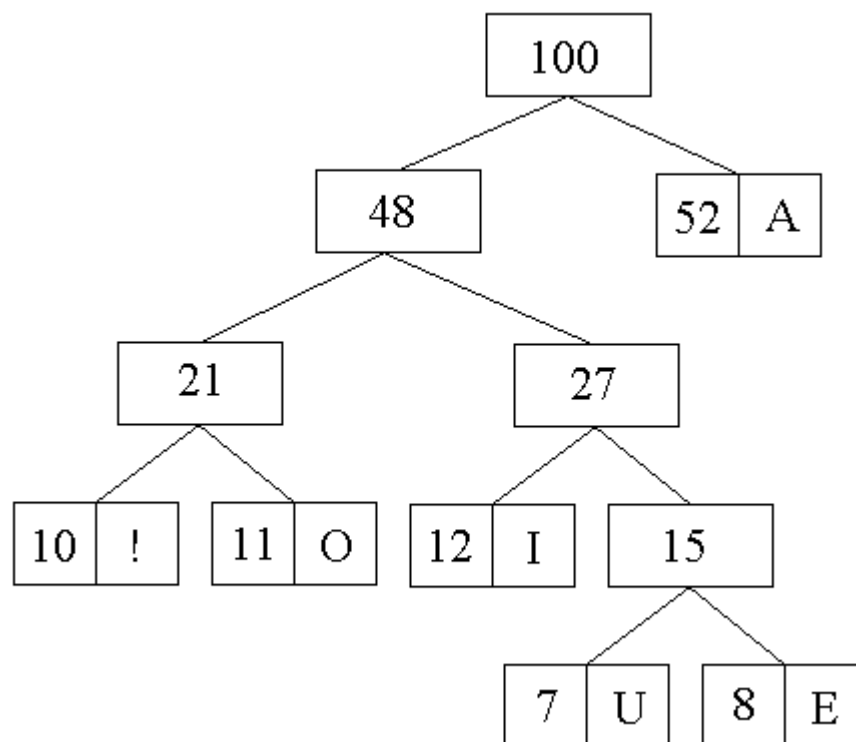
ETAPA 4



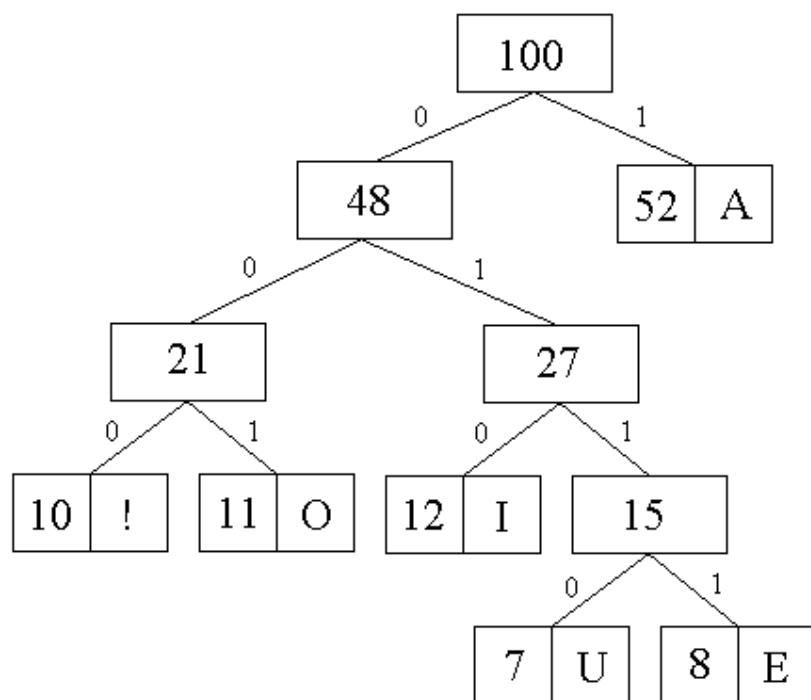
ETAPA 5



ETAPA 6



GERAÇÃO DA TABELA DE CODIFICAÇÃO



| | |
|---|------|
| A | 1 |
| E | 0111 |
| I | 010 |
| O | 001 |
| U | 0110 |
| ! | 000 |

EXERCÍCIOS

1. Com o código de Huffman acima estabelecido:

- Como seria armazenada a seguinte sequência de caracteres
AEIOAAIOU!AI!AEO!
- Decodifique a sequência de bits
001010000101110111011000011010

2. Construa a árvore e a codificação de Huffman para a tabela de caracteres abaixo:

| Caractere | A | B | C | L | S | O |
|-----------|----|---|---|---|---|----|
| % | 35 | 8 | 4 | 5 | 7 | 25 |

LINKS INTERESSANTES

<http://www.cs.auckland.ac.nz/software/AlgAnim/huffman.html>

<http://www.cs.ttu.edu/~eacosta/java/Huffcode/Huffcode.html>