

# **Mídia Texto**

**Profa. Débora Christina Muchaluat Saade**

**debora@midia.com.uff.br**

# Compressão de Texto

- ✓ **Texto não tolera erros**
- ✓ **Compactação = Compressão sem perdas**
- ✓ **Classificação das técnicas de compactação:**
  - *Técnicas que usam caracteres únicos como base X*  
*Técnicas que usam strings de caracteres como base*
  - *Codificação estática X Codificação dinâmica*

# Compressão de Texto

- ✓ **Técnicas que usam caracteres únicos como base**
  - ***Codificação por carreira***
    - Só usa se repetir no mínimo 4 vezes
      - *(character, símbolo (!), número de vezes)*
    - entrada: ABCCCCCCCCDEFGGG
    - saída: ABC!4DEFGGG
  - ***Codificação de Shannon-Fano***
  - ***Codificação aritmética***
  - ***Codificação de Huffman***

# Compressão de Texto

- ✓ Técnicas que usam strings de caracteres como base
  - *Codificação por carreira*
    - pode-se substituir sequências maiores que um
    - requer que o tamanho da sequência seja codificado ou pode-se usar um caracter especial de fim
      - *(símbolo (!), número de vezes, seqüência, delimitador de fim (\$))*
    - entrada: UFYUGDUFHUFHUFHUFHUFHBFD
    - saída: UFYUGD!5UFH\$BFD
    - Se símbolos especiais aparecerem nos dados (*character stuffing*)
      - *entrada: U!HIIIIID*
      - *saída: U!!HI!1D*
  - *Codificação de Lempel-Ziv-Welch*

# Compressão de Texto

## ✓ Codificação estática

- *Tabelas de códigos são conhecidas a priori (padronizadas)*
- *Tabelas de códigos são enviadas junto com a sequência codificada*
- *Exemplos:*
  - Codificação de Shannon-Fano
  - Codificação aritmética
  - Codificação de Huffman

## ✓ Codificação dinâmica

- *Tabelas de códigos são calculadas dinamicamente no momento da decodificação*
- *Exemplos:*
  - Codificação de Lempel-Ziv-Welch
  - Codificação de Huffman dinâmica

# Codificação de Huffman Dinâmica

- ✓ **A árvore de Huffman é construída dinamicamente na codificação e na decodificação**
- ✓ **Não é necessário saber a frequência dos símbolos a priori**
- ✓ **Se o caracter não estiver presente na árvore ainda (da primeira vez que aparece), ele não é compactado**
- ✓ **Usa o conceito de uma folha vazia que representa o local onde o próximo caracter será inserido na árvore**

# Codificação de Huffman Dinâmica

