



# AEDs3 V1

## 1. Pesquisa Digital

- 1.1. Trie
- 1.2. Árvore Digital Binária
- 1.3. Árvore de Prefixo
- 1.4. Árvore PATRICIA

## 2. Casamento de Padrões

### 2.1. Definições

2.1.1. É uma busca por uma cadeia de símbolos em uma sequência maior

#### 2.1.2. Aplicações

- 2.1.2.1. detecção de plágio
- 2.1.2.2. correção ortográfica
- 2.1.2.3. busca de DNA
- 2.1.2.4. casamento de voz
- 2.1.2.5. ...

### 2.2. Tipos

#### 2.2.1. Casamento Exato

##### 2.2.1.1. Força bruta

- 2.2.1.1.1. Não há fase de pré-processamento
- 2.2.1.1.2. Sempre avança apenas uma posição
- 2.2.1.1.3. Complexidade é  $O(mn)$ ,  $m$  = tam. padrão,  $n$  = tam. texto

##### 2.2.1.2. KMP

- 2.2.1.2.1. Introduz o conceito de autômato de estados e é usado para busca de cadeias.
- 2.2.1.2.2. Quando aparece uma diferença, a palavra tem em si a informação necessária para determinar onde começar a próxima comparação
- 2.2.1.2.3. Função de Falha
  - 2.2.1.2.3.1. Padrão

2.2.1.2.3.1.1. Maior prefixo  $v$  do padrão que é sufixo do texto antes do ponto de erro  $j$

##### 2.2.1.2.3.2. Melhorada

2.2.1.2.3.2.1. Maior prefixo  $v$  do padrão que é sufixo do texto antes do ponto de erro  $j$  e é sucedido por um símbolo diferente do que gerou o erro

2.2.1.2.4. Complexidade é  $O(m + n)$ ,  $m$  = tam. padrão,  $n$  = tam. texto

### 2.2.1.3. Boyer Moore

2.2.1.3.1. Comparações da Direita para a esquerda

2.2.1.3.2. Heurísticas de deslocamento

2.2.1.3.2.1. Caracter rum

2.2.1.3.2.2. Sufixo bom

2.2.1.4. Shift-And Exato

2.2.1.5. Aho Corasick

2.2.2. Casamento aproximado

2.2.2.1. Shift-And Aproximado

## 3. Criptografia

### 3.1. Cifragem

3.1.1. Cifras de substituição

3.1.2. Cifras de transposição

3.1.3. Cifras de fluxo

3.1.4. Cifras de bloco

### 3.2. Tipos

3.2.1. Simétrica

3.2.2. Assimétrica

### 3.3. Algoritmos

3.3.1. DES

3.3.2. 3DES

3.3.3. TDES

3.3.4. AES

## 4. Compressão

### 4.1. Conceitos

4.1.1. Classificação

4.1.1.1. Simetria

4.1.1.1.1. Simétrico

4.1.1.1.2. Assimétrico

4.1.1.2. Perda

#### 4.1.1.2.1. Com

##### 4.1.1.2.1.1. Eliminação de Detalhes

##### 4.1.1.2.1.2. Saída ainda útil

#### 4.1.1.2.2. Sem

##### 4.1.1.2.2.1. Remoção Recuperável de Redundâncias

##### 4.1.1.2.2.2. Dados exatos

#### 4.1.1.3. Adaptabilidade

##### 4.1.1.3.1. Adaptativa

##### 4.1.1.3.2. Semi-Adaptativa

##### 4.1.1.3.3. Não Adaptativa

#### 4.1.2. Métricas

##### 4.1.2.1. Taxa de Compressão

##### 4.1.2.2. Fator de Compressão

##### 4.1.2.3. Ganho de Compressão

##### 4.1.2.4. Percentual de Redução

##### 4.1.2.5. Taxa de Bits

#### 4.2. RLE

##### 4.2.1. Contagem de Sequências longas

##### 4.2.2. Substitui seqüências de caracteres repetidos pelo número de ocorrências seguido do caracter

##### 4.2.3. Sem perda

#### 4.3. Métodos

##### 4.3.1. Estatísticos

##### 4.3.1.1. Entropia

##### 4.3.1.1.1. menor quantidade de bits por símbolos necessária para guardar o conteúdo de informação da fonte

##### 4.3.1.1.2. $P_x$ = Probabilidade do símbolo x

##### 4.3.1.1.3. Entropia $S = -\log_2(P_x)$

##### 4.3.1.2. Estratégia

##### 4.3.1.2.1. Dados na informação original que aparecem com maior frequência são representados por palavras-código menores

4.3.1.2.2. Dados de menor incidência são representados por palavras-código maiores

#### 4.3.1.3. Algoritmos

##### 4.3.1.3.1. Shannon-Fano

4.3.1.3.1.1. Árvore Top-Down

4.3.1.3.1.2. Listas divididas em 2 com somatório de probabilidades semelhantes

4.3.1.3.1.3. Lista de Probabilidades

##### 4.3.1.3.2. Huffman

4.3.1.3.2.1. Lista de Probabilidades

4.3.1.3.2.2. Árvore Bottom-UP

4.3.1.3.2.3. Nodos de menor probabilidade agrupados na árvore

#### 4.3.2. Dicionário

##### 4.3.2.1. Estratégia

4.3.2.1.1. Os símbolos (ou conjunto de símbolos) são substituídos por códigos a partir de um “dicionário”

##### 4.3.2.1.2. Dicionários

4.3.2.1.2.1. Estáticos

4.3.2.1.2.2. Dinâmicos

##### 4.3.2.2. Algoritmos

##### 4.3.2.2.1. LZ77

4.3.2.2.1.1. Parte do conteúdo já lido é usado como dicionário

4.3.2.2.1.2. Utiliza o conceito de “janela deslizante” (sliding window)

4.3.2.2.1.3. Dicionário parcial (janela)

##### 4.3.2.2.2. LZ78

4.3.2.2.2.1. Cria-se um dicionário de prefixos (ou sequências) já encontrados no conteúdo

4.3.2.2.2.2. O dicionário é construído dinamicamente a partir das sequências encontradas e é armazenado em uma tabela na memória primária

4.3.2.2.2.3. Dicionário já usa todo o conteúdo já "visto"

##### 4.3.2.2.3. LZW

4.3.2.2.3.1. Versão melhorada do algoritmo LZ78/LZ79

4.3.2.2.3.2. Os códigos de 0 a 255 representam sequências de tamanho unitário correspondendo ao símbolo de mesmo valor

4.3.2.2.3.3. dicionário modificado que foi inicializado com todos os símbolos do alfabeto

## **5. Memória Secundária**

### 5.1. Estrutura

#### 5.1.1. HD

##### 5.1.1.1. Setores

##### 5.1.1.2. Trilhas

##### 5.1.1.3. Cilindros

##### 5.1.1.4. Cálculos

###### 5.1.1.4.1. Tamanho

###### 5.1.1.4.2. Tempo

###### 5.1.1.4.2.1. Busca

###### 5.1.1.4.2.2. Latência Rotacional

#### 5.1.2. SSD

### 5.2. Tempo das operações

#### 5.2.1. Busca

#### 5.2.2. Transferência de dados

### 5.3. Arquivos e fluxos de entrada e saída

### 5.4. Formas de acesso a arquivos

#### 5.4.1. Sequencial

#### 5.4.2. Aleatório

### 5.5. Arquivo de dados estruturados

#### 5.5.1. Registros

##### 5.5.1.1. Entidades e Atributos

##### 5.5.1.2. Campos

##### 5.5.1.3. Tamanho Fixo

##### 5.5.1.4. Tamanho Variável

###### 5.5.1.4.1. Com Delimitador

###### 5.5.1.4.2. Com totalizador de Bytes

## **6. Arquivos**

## 6.1. Tipos de Arquivos

### 6.1.1. Arquivos Sequenciais

#### 6.1.1.1. CRUD

##### 6.1.1.1.1. Com leitura sequencial

###### 6.1.1.1.1.1. Create

###### 6.1.1.1.1.2. Read

###### 6.1.1.1.1.3. Update

###### 6.1.1.1.1.4. Delete

#### 6.1.1.2. Ordenação Externa

##### 6.1.1.2.1. Intercalação Balanceada

##### 6.1.1.2.2. Segmento de Tamanho Variável

##### 6.1.1.2.3. Seleção por substituição - Heap

##### 6.1.1.2.4. Intercalação Polifásica

### 6.1.2. Arquivos Indexados

#### 6.1.2.1. Tipos de índices

##### 6.1.2.1.1. Direto X Indireto

##### 6.1.2.1.2. Esperso X Denso

##### 6.1.2.1.3. Primário X Secundário

#### 6.1.2.2. CRUD

##### 6.1.2.2.1. Com apoio dos índices

###### 6.1.2.2.1.1. Create

###### 6.1.2.2.1.2. Read

###### 6.1.2.2.1.3. Update

###### 6.1.2.2.1.4. Delete

#### 6.1.2.3. Estruturas de Dados em Arquivos

##### 6.1.2.3.1. Árvore B

###### 6.1.2.3.1.1. B

###### 6.1.2.3.1.1.1. Árvore de Múltiplos Caminhos

###### 6.1.2.3.1.1.2. Ordem = Número de Filhos

6.1.2.3.1.1.3. Menor altura da Árvore

6.1.2.3.1.1.4. Página = Setor ou Multiplo de Setor

6.1.2.3.1.2. B+

6.1.2.3.1.2.1. Todas os registros nas folhas

6.1.2.3.1.2.2. Folhas apontam para a próxima

6.1.2.3.1.2.3. Busca sequencial facilitada

6.1.2.3.1.2.4. Split 1 para 2

6.1.2.3.1.3. B\*

6.1.2.3.1.3.1. Otimizar alocação da página

6.1.2.3.1.3.2. Split 2 para 3

6.1.2.3.2. Como representar registros no arquivo de índices

6.1.2.3.3. Hashing

6.1.2.3.3.1. Função de Dispersão

6.1.2.3.3.2. Tratamento de Colisões

6.1.2.3.3.3. Estático X Dinâmico

6.1.2.3.3.4. Dinâmico

6.1.2.3.3.4.1. Diretório e Buckets

6.1.2.3.3.4.2. Profundidade Global e Local

6.1.2.3.3.4.3. Função hash depende da profundidade Global

6.1.2.3.3.4.4. Dobra de tamanho a cada expansão do diretório

6.1.2.3.4. Listas invertidas

6.1.2.3.4.1. Como estruturar

6.1.2.3.4.2. Vantagens e Aplicações