# Algoritmos de Busca

Prof. Luiz Gustavo Almeida Martins

Frequentemente precisamos realizar a busca por um elemento em um determinado conjunto de dados

Frequentemente precisamos realizar a busca por um elemento em um determinado conjunto de dados

Recuperação de informações a partir de uma grande massa de dados previamente armazenada

Frequentemente precisamos realizar a busca por um elemento em um determinado conjunto de dados

Recuperação de informações a partir de uma grande massa de dados previamente armazenada

A informação é organizada em registros

Frequentemente precisamos realizar a busca por um elemento em um determinado conjunto de dados

Recuperação de informações a partir de uma grande massa de dados previamente armazenada

A informação é organizada em **registros**Cada registro possui uma **chave** (*key*) usada na pesquisa

Frequentemente precisamos realizar a busca por um elemento em um determinado conjunto de dados

Recuperação de informações a partir de uma grande massa de dados previamente armazenada

A informação é organizada em registros

Cada registro possui uma **chave** (*key*) usada na pesquisa

Um conjunto de registros é chamado de **tabela** ou **arquivo** 

## **Tipo Abstrato de Dados (TAD)**

É importante tratar os algoritmos de pesquisa como um TAD:

Conjunto de operações associado a uma estrutura de dados

Proporciona uma independência de implementação para as operações

## **Tipo Abstrato de Dados (TAD)**

É importante tratar os algoritmos de pesquisa como um TAD:

Conjunto de operações associado a uma estrutura de dados

Proporciona uma independência de implementação para as operações

#### **Operações mais comuns:**

Inicializar a estrutura de dados

Pesquisar um ou mais registros com uma determinada chave

Inserir um novo registro

Retirar um registro específico

Ordenar um arquivo de acordo com a chave

Mesclar dois arquivos para formar um arquivo maior

Etc.

## Pesquisa em um dicionário

Dicionário é o nome comumente utilizado para descrever uma estrutura de dados para pesquisa

## Pesquisa em um dicionário

Dicionário é o nome comumente utilizado para descrever uma estrutura de dados para pesquisa

Analogia com um dicionário da lingua portuguesa:

**chave** = palavras

**Demais campos** = entradas associadas com as palavras (pronúncia, definição, sinônimos, etc.)

## Pesquisa em um dicionário

Dicionário é o nome comumente utilizado para descrever uma estrutura de dados para pesquisa

Analogia com um dicionário da lingua portuguesa:

**chave** = palavras

**Demais campos** = entradas associadas com as palavras (pronúncia, definição, sinônimos, etc.)

Geralmente possui as seguintes operações:

Inicializa

Insere

Remove

Pesquisa



A escolha do método mais adequado para uma determinada aplicação depende:

A escolha do método mais adequado para uma determinada aplicação depende:

Quantidade de dados envolvidos

A escolha do método mais adequado para uma determinada aplicação depende:

Quantidade de dados envolvidos

Frequência de inserções e remoções no arquivo

A escolha do método mais adequado para uma determinada aplicação depende:

Quantidade de dados envolvidos

Frequência de inserções e remoções no arquivo

Se o conteúdo do arquivo é estável (não muda muito com o tempo), é importante minimizar o tempo de pesquisa através de sua estruturação

A escolha do método mais adequado para uma determinada aplicação depende:

Quantidade de dados envolvidos

Frequência de inserções e remoções no arquivo

Se o conteúdo do arquivo é estável (não muda muito com o tempo), é importante minimizar o tempo de pesquisa através de sua estruturação

O tempo necessário para essa estruturação é compensado pelo ganho na eficiência das futuras pesquisas

Diferentes estratégias podem ser empregadas:

#### Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial

#### Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial Busca binária

#### Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial Busca binária

Árvore binária de busca

#### Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial Busca binária

Árvore binária de busca Árvores balanceadas

## Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial Busca binária

Árvore binária de busca Árvores balanceadas

Tabelas de dispersão (hash table)

## Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial Busca binária

Dados organizados em estruturas lineares (ex: vetores)

Árvore binária de busca Árvores balanceadas

Tabelas de dispersão (hash table)

## Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial Busca binária

Árvore binária de busca Árvores balanceadas

Tabelas de dispersão (hash table)

Dados organizados em estruturas lineares (ex: vetores)

Dados organizados em árvores

### Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial Busca binária

Árvore binária de busca Árvores balanceadas

Tabelas de dispersão (hash table)

Dados organizados em estruturas lineares (ex: vetores)

Dados organizados em árvores

Dados organizados em tabelas

#### Diferentes estratégias podem ser empregadas:

Busca linear ou sequencial Busca binária

Árvore binária de busca Árvores balanceadas

Tabelas de dispersão (hash table)

Dados organizados em estruturas lineares (ex: vetores)

Dados organizados em árvores

Dados organizados em tabelas

Escolha afeta na forma de implementação do TAD

# Formas de implementação

## Implementações Lineares

Um registro é formado por:

Um campo chave

Um ou mais campos com informações adicionais

```
Um registro é formado por:
Um campo chave
Um ou mais campos com informações adicionais
```

#### **Exemplo:**

```
struct registro {
    int chave;
    char nome[100];
    // Poderíamos ter outros campos em nosso registro
};
```

typedef struct registro Registro;

```
Um registro é formado por:
  Um campo chave
  Um ou mais campos com informações adicionais
Exemplo:
struct registro {
    int chave;
    char nome[100];
    // Poderíamos ter outros campos em nosso registro
```

Um arquivo contém um conjunto de registros

Um arquivo contém um conjunto de registros

### **Exemplo:**

```
#define MAX 10

struct arquivo {
    Registro itens[MAX];
    int tamanho;
};
```

Um arquivo contém um conjunto de registros

# #define MAX 10 struct arquivo { Registro itens[MAX]; int tamanho; }; typedef struct arquivo Arq;

# Operações básicas

#### Inicializar a estrutura de dados:

Vetor de itens é alocado estaticamente ao se criar um arquivo

## Operações básicas

#### Inicializar a estrutura de dados:

Vetor de itens é alocado estaticamente ao se criar um arquivo Campo tamanho precisa ser inicializado com ZERO

# Operações básicas

#### Inicializar a estrutura de dados:

Vetor de itens é alocado estaticamente ao se criar um arquivo Campo tamanho precisa ser inicializado com ZERO

#### Implementação em C:

```
Arq * inicia_arq() {
  Arg *arg = (Arg *) malloc(sizeof(Arg));
  if (arg != NULL)
     arq->tamanho = 0;
  return arq;
```

#### Inserir um novo registro no arquivo:

O local mais adequado depende do critério de ordenação

#### Inserir um novo registro no arquivo:

O local mais adequado depende do critério de ordenação

Sucesso depende da disponibilidade de espaço

#### Inserir um novo registro no arquivo:

O local mais adequado depende do critério de ordenação Sucesso depende da disponibilidade de espaço

#### Implementação em C:

```
int insere_registro(Arq *arq, Registro Elem) {
   if( arq->tamanho == MAX ) // ou usar a função arq_cheio() se existir
       return 0;

   arq->items[arq->tamanho] = Elem; // Coloca o registro na última posição
   arq->tamanho++; // Incrementa o número de registros
   return 1;
}
```

#### Remover um registro do arquivo:

Envolve a pesquisa do registro

#### Remover um registro do arquivo:

Envolve a pesquisa do registro

Sucesso depende da existência do registro desejado

#### Remover um registro do arquivo:

Envolve a pesquisa do registro

Sucesso depende da existência do registro desejado

#### Implementação em C:

## Pesquisar um registro no arquivo:

Envolve localizar um registros que possua a chave especificada

## Pesquisar um registro no arquivo:

Envolve localizar um registros que possua a chave especificada

Sucesso também depende da existência do registro

## Pesquisar um registro no arquivo:

Envolve localizar um registros que possua a chave especificada

Sucesso também depende da existência do registro

Retorna a posição do registro ou um valor inválido quando o registro não é encontrado

## Método de pesquisa

## Busca linear ou sequencial

Consiste em pesquisar sequencialmente, a partir do 1º registro, até encontrar a chave procurada ou terminar o arquivo

É a forma mais simples de pesquisa

Consiste em pesquisar sequencialmente, a partir do 1º registro, até encontrar a chave procurada ou terminar o arquivo

É a forma mais simples de pesquisa

#### Pesquisa sequencial em um arquivo não ordenado:

```
int busca_seq(Arq *arq, int chave) {
   if (arq->tamanho == 0)
     return -1;
   int x = 0;
   while (x < arq->tamanho && arq->itens[x].chave != chave)
     x++;
   if (x == arq->tamanho)
     return -1; // Registro não foi encontrado
   return x; // Retorna posição do registro no arquivo
}
```

Algoritmo ineficiente para um número elevado de dados

Algoritmo ineficiente para um número elevado de dados

## Análise do algoritmo:

No pior caso (quando a chave não está no arquivo), serão necessárias verificar TODOS os registros: O(N)

Algoritmo ineficiente para um número elevado de dados

## Análise do algoritmo:

No pior caso (quando a chave não está no arquivo), serão necessárias verificar TODOS os registros: O(N)

No caso médio, considerando que todas as entradas possuem a mesma probabilidade de ocorrência, temos N/2 comparações (comportamento esperado)

#### Considerações:

Busca **sempre encontra o registro** procurado **p**<sub>i</sub> é a probabilidade do *i*-ésimo registro ser o procurado A recuperação do *i*-ésimo registro envolve *i* comparações

#### Considerações:

Busca **sempre encontra o registro** procurado **p**<sub>i</sub> é a probabilidade do *i*-ésimo registro ser o procurado A recuperação do *i*-ésimo registro envolve *i* comparações

$$f(n) = 1 \times p_1 + 2 \times p_2 + 3 \times p_3 + \cdots + n \times p_n$$

Cálculo de f(n) depende do conhecimento sobre a distribuição das probabilidades  $p_i$ 

#### Considerações:

Busca **sempre encontra o registro** procurado **p**<sub>i</sub> é a probabilidade do *i*-ésimo registro ser o procurado A recuperação do *i*-ésimo registro envolve *i* comparações

$$f(n) = 1 \times p_1 + 2 \times p_2 + 3 \times p_3 + \cdots + n \times p_n$$

Cálculo de f(n) depende do conhecimento sobre a distribuição das probabilidades  $p_i$ 

Se todas as entradas são igualmente prováveis ( $p_i = 1/n$ ):

$$f(n) = (1 + 2 + 3 + \dots + n) \times 1/n$$

#### Considerações:

Busca **sempre encontra o registro** procurado **p**<sub>i</sub> é a probabilidade do *i*-ésimo registro ser o procurado A recuperação do *i*-ésimo registro envolve *i* comparações

$$f(n) = 1 \times p_1 + 2 \times p_2 + 3 \times p_3 + \cdots + n \times p_n$$

Cálculo de f(n) depende do conhecimento sobre a distribuição das probabilidades  $p_i$ 

Se todas as entradas são igualmente prováveis ( $p_i = 1/n$ ):

$$f(n) = (1 + 2 + 3 + \dots + n) \times 1/n = (n+n2)/2 \times 1/n$$

#### **Considerações:**

Busca **sempre encontra o registro** procurado **p**<sub>i</sub> é a probabilidade do *i*-ésimo registro ser o procurado A recuperação do *i*-ésimo registro envolve *i* comparações

$$f(n) = 1 \times p_1 + 2 \times p_2 + 3 \times p_3 + \cdots + n \times p_n$$

Cálculo de f(n) depende do conhecimento sobre a distribuição das probabilidades **p**<sub>i</sub>

Se todas as entradas são igualmente prováveis ( $p_i = 1/n$ ):

$$f(n) = (1 + 2 + 3 + \dots + n) \times 1/n = (n+n2)/2 \times 1/n = (n+1)/2$$

Suponha um vetor com 7 posições:



Quantas comparações tenho que fazer para encontrar o número 6?

Suponha um vetor com 7 posições:



Quantas comparações tenho que fazer para encontrar o número 6?

5 comparações

Considere que o número 6 poderia estar em outro local do vetor, com mesma probabilidade:



Quantas comparações tenho que fazer para encontrar o número 6?

Considere que o número 6 poderia estar em outro local do vetor, com mesma probabilidade:



Quantas comparações tenho que fazer para encontrar o número 6?

2 comparações

Em cada uma precisamos de diferentes números de comparações para achar o número 6

Veja algumas opções de vetor:

Na média, quantas comparações temos que fazer para encontrar o número 6?

2	6	5	1	7	3	4
5	2	6	3	4	7	1
5	1	2	3	7	4	6
1	4	2	3	5	6	7
6	7	3	5	4	2	1
7	2	6	3	4	5	1
5	4	1	7	3	6	2
6	2	5	3	4	1	7
1	4	2	5	7	6	3
6	4	2	5	7	3	1
6	7	1	4	3	2	5
5	2	3	6	1	4	7
4	5	3	2	6	7	1
2 5 5 1 6 7 5 6 1 6 5 4 5 7	6 2 1 4 7 2 4 2 4 7 2 5 4 1	5 6 2 2 3 6 1 5 2 2 1 3 6 6 6	1 3 3 5 3 5 4 6 2 7 2	7 4 7 5 4 4 3 4 7 7 3 1 6 3 3	3 7 4 6 2 5 6 1 6 3 2 4 7 1 5	4 1 6 7 1 1 2 7 3 1 5 7 1 2 4
7	1	6	2	3	5	4

Veja algumas opções de vetor:	2	6	5	1	7	3	4
	5	2	6	3	4	7	1
	5	1	2	3	7	4	6
Em cada uma precisamos de	1	4	2	3	5	6	7
diferentes números de	6	7	3	5	4	2	1
comparações para achar o	7	2	6	3	4	5	1
número 6	5	4	1	7	3	6	2
	6	2	5	3	4	1	7
	1	4	2	5	7	6	3
Na média, quantas	6	4	2	5	7	3	1
comparações temos que fazer	6	7	1	4	3	2	5
	5	2	3	6	1	4	7
para encontrar o número 6?	4	5	3	2	6	7	1
	5	4	6	7	3	1	2
	7	1	6	2	3	5	4
	===	====	====	====	====	====	

Quantas vezes o 6 aparece em cada posição? 4 1 4 1 1 3 1

Veja algumas opções de vetor:	2 5	6 2	5 6	1 3	7 4	3 7	4 1
	5	1	2	3	7	4	6
Em cada uma precisamos de	1	4	2	3	5	6	7
diferentes números de	6	7	3	5	4	2	1
comparações para achar o	7	2	6	3	4	5	1
número 6	5	4	1	7	3	6	2
	6	2	5	3	4	1	7
	1	4	2	5	7	6	3
Na média, quantas	6	4	2	5	7	3	1
•	6	7	1	4	3	2	5
comparações temos que fazer	5	2	3	6	1	4	7
para encontrar o número 6?	4	5	3	2	6	7	1
	5	4	6	7	3	1	2
	7	1	6	2	3	5	4
O	===	====	====	====	===	====	===
Quantas vezes o 6 aparece em cada posição?	4	1	4	1	1	3	1
Quantas comparações ocorre por posição?	1	2	3	4	5	6	7

#### Número de comparações:

4x na primeira coluna

$$\Rightarrow$$
 4x1 = 4

1x na segunda

$$= 1x2 = 2$$

4x na terceira

$$\Rightarrow$$
 4x3 = 12

1x na quarta

$$= 1x4 = 4$$

1x na quinta

$$\Rightarrow$$
 1x5 = 5

3x na sexta

$$> 3x6 = 18$$

1x na sétima

$$= 1x7 = 7$$

2.	6	5	1	7	3	4
2 5	2	6	3	4	7	1
5	1	2	3	7	4	6
1	4	2	3	5	6	7
6	7	2 3	5	4	2	1
7	2	6	3	4	5	1
5	2 4	1	3 7	3	6	2
6	2	5		4	1	7
	2 4 4	2	3 5 5 4 6	7	6	3
1 6	4	2	5	7		1
6	7	1	4	3	3 2 4	5
5	7 2 5 4	3	6	1	4	7
4	5	3	2	6	7	1
5 4 5	4	6	2 7	3	1	2
7	1	6	2	3	5	4
===	====			====		===
4	1	4	1	1	3	1
1	2	3	4	5	6	7

#### Número de comparações:

4x na primeira coluna

$$\Rightarrow$$
 4x1 = 4

1x na segunda

$$\Rightarrow$$
 1x2 = 2

4x na terceira

$$\Rightarrow$$
 4x3 = 12

1x na quarta

$$\Rightarrow$$
 1x4 = 4

1x na quinta

$$\Rightarrow$$
 1x5 = 5

3x na sexta

$$> 3x6 = 18$$

1x na sétima

$$\Rightarrow$$
 1x7 = 7

Total: 4+2+12+4+5+18+7 = 52

2	6	5	1	7	3	4
5	2	6	3	4	7	1
5	1	2	3	7	4	6
1	4	2	3	5	6	7
6	7	3	5	4	2	1
7	2	6	3	4	5	1
7 5	4	1	7	3	6	2 7
6	2	5	3	4	1	7
1	2 4	2	5	7	6	3
6	4		5	7	3	1
6	7	2 1	4	3	2	5
5	2	3	6	1	4	7
6 5 4 5 7	5	3	2	6	7	1
5	4	6	7	3	1	2
7	1	6	2	3	5	4
===	====			====		===
4	1	4	1	1	3	1
4	2	3	4	5	6	7

Número de comparações: 52

## Número de vetores:

15

2	6	5	1	7	3	4
5	2	6	3	4	7	1
5	1	2	3	7	4	6
1	4	2	3	5	6	7
6	7	3	5	4	2	1
7	2	6	3	4	5	1
5	4	1	7	3	6	2
6	2	5	3	4	1	7
1	4	2	5	7	6	3
6	4	2	5	7	3	1
6	7	1	4	3	2	5
5	2	3	6	1	4	7
4	5	3		6	7	1
5	4	6	2 7	3	1	2
7	1	6	2	3	5	4
4	1	4	1	1	3	1
4	2	3	4	5	3	7

# Número de comparações: 52

# Número de vetores:

15

Média de comparações: 52/15 = 3,46

2	6	5	1	7	3	4
5	2	6	3	4	7	1
5	1	2	3	7	4	6
1	4	2		5	4 6	7
6	7	3	3 5 3 7	4		1
6 7	2	6	3	4	2 5	1
5	4	1	7	3	6	2
6	2	5		4	1	7
1	4	2	3 5 5	7	6	3 1
1 6	7 2 4 2 4 4 7	2		7	3	
6	7	1	4	3	3 2	5
5	2	3	6	1	4	7
4	5	3	2	6	7	1
6 5 4 5	4	6	2 7	3	1	2
7 ===	1	6	2	3	5	4
	1	====		1	====	1
4	1	4	1	1	3	1 7
1	2	3	4	5	6	7

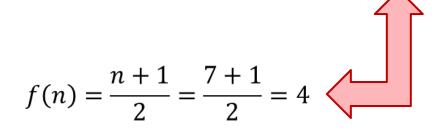
# Número de comparações:

52

#### Número de vetores:

15

Média de comparações: 52/15 = 3,46



2	6	5	1	7	3	4
5	2	6	3	4	7	1
5	1	2	3	7	4	6
1	4	2	3	5	6	7
6	7	3	5			1
7	2	6	3	4 4	2 5	1
7 5	4	1	7	3	6	1 2 7 3 1
6	2	5	3	4	1	7
1	4	2	5	7	6	3
6	4	2	5	7 7	3	1
6	7	1	4	3	2	5
		3	6	1	4	7
4	2 5	3	2	6	7	1
5 4 5	4	6	7	3	1	7 1 2
	1	6	2	3	5	4
7 ===						===
4	1	4	1	1	3	1

## Novo experimento:

Geração aleatória de 10.000 vetores

## Novo experimento:

Geração aleatória de 10.000 vetores

#### Qtde de ocorrências do 6 por coluna:

Coluna 1: 1429

Coluna 2: 1452

Coluna 3: 1422

Coluna 4: 1439

Coluna 5: 1466

Coluna 6: 1405

Coluna 7: 1387

### Novo experimento:

Geração aleatória de 10.000 vetores

#### Número de comparações:

```
Coluna 1: 1429 \times 1 = 1429
```

Coluna 2:  $1452 \times 2 = 2904$ 

Coluna 3:  $1422 \times 3 = 4266$ 

Coluna 4:  $1439 \times 4 = 5756$ 

Coluna 5:  $1466 \times 5 = 7330$ 

Coluna 6:  $1405 \times 6 = 8430$ 

Coluna 7:  $1387 \times 7 = 9709$ 

=====

39.824 comparações

## **Novo experimento:**

39.824 comparações

**10.000 vetores** 

# Análise empírica do caso médio

## Novo experimento:

39.824 comparações

**10.000 vetores** 

Média de comparações: 39.824/10.000 = 3,9824 ≈ 4

# Análise empírica do caso médio

## Novo experimento:

39.824 comparações

**10.000 vetores** 

#### Conclusão:

No caso médio (esperado), o número médio de comparações é (n+1)/2

Considerando uma distribuição uniforme dos números no vetor

Média de comparações: 39.824/10.000 = 3,9824 ≈ 4

$$f(n) = \frac{n+1}{2} = \frac{7+1}{2} = 4$$

## Busca sequencial com sentinela

Consiste em incluir um registro extra com a chave desejada no final do arquivo

#### Mudança na estrutura:

Quantidade de itens precisa prever um registro a mais

```
struct arquivo {
    Registro itens[MAX+1];
    int tamanho;
};
```

#### Mudança na operação:

Elimina o teste lógico de final de arquivo

## Busca sequencial com sentinela

#### Implementação em C:

```
int busca_seq(Arq *arq, int chave) {
  if (arq->tamanho == 0)
     return -1;
  arq->itens[arq->tamanho].chave = chave; // Inclui sentinela
  int x = 0:
  while (arq->itens[x].chave != chave)
     X++;
  if (x == arq->tamanho)
     return -1; // Registro não foi encontrado
  return x; // Retorna posição do registro no arquivo
```

## Busca sequencial com sentinela

Não existe grande mudança na complexidade do algoritmo

#### Pesquisa com sucesso:

Melhor caso: C(n) = 1

Pior caso: C(n) = n

Caso médio: C(n) = (n+1)/2

#### Pesquisa sem sucesso:

$$C(n) = n + 1$$

Redução na quantidade dos testes lógicas

A ordenação do vetor de pesquisa torna o algoritmo um pouco mais eficiente

Não precisa percorrer todo o vetor para determinar a inexistência da chave

A ordenação do vetor de pesquisa torna o algoritmo um pouco mais eficiente

Não precisa percorrer todo o vetor para determinar a inexistência da chave

Evita percorrimento para valores "fora do intervalo"

A ordenação do vetor de pesquisa torna o algoritmo um pouco mais eficiente

Não precisa percorrer todo o vetor para determinar a inexistência da chave

Evita percorrimento para valores "fora do intervalo"

#### **Desvantagem:**

Custo para a ordenação do vetor

#### Implementação em C:

```
int busca_seqOrd(Arq *arq, int chave) {
  if ( arq->tamanho == 0 ||
                                                    // Arquivo vazio
     arg->itens[x] > chave ||
                                                    // Chave < 1° registro
     arg->itens[arg->tamanho-1] < chave )
                                                    // Chave > último registro
     return -1;
  int x = 0:
  while (arg->itens[x].chave < chave)
     X++;
  if (arq->itens[x].chave > chave)
     return -1; // Registro não foi encontrado
  return x; // Retorna posição do registro no arquivo
```

## Mantém os custos quando a chave é encontrada:

Melhor caso: C(n) = 1

Pior caso: C(n) = n

Caso médio: C(n) = (n+1)/2

## Mantém os custos quando a chave é encontrada:

Melhor caso: C(n) = 1

Pior caso: C(n) = n

Caso médio: C(n) = (n+1)/2

Ganho no caso médio (esperado) quando a pesquisa falha (chave não existe):

Melhor caso: C(n) = 1

Pior caso: C(n) = n

Caso médio: C(n) = (n+1) / 2

# Método de pesquisa

## Busca binária

## Pesquisa mais eficiente em vetores

Precisa que os registros estejam ordenados

## Pesquisa mais eficiente em vetores

Precisa que os registros estejam ordenados

#### **Funcionamento:**

Verifique o registro que está no meio da tabela (central)

## Pesquisa mais eficiente em vetores

Precisa que os registros estejam ordenados

#### **Funcionamento:**

Verifique o registro que está no meio da tabela (central) Se for a chave desejada, pára a busca e retorna a posição

## Pesquisa mais eficiente em vetores

Precisa que os registros estejam ordenados

#### **Funcionamento:**

Verifique o registro que está no meio da tabela (central)

Se for a chave desejada, pára a busca e retorna a posição

Se a chave for menor, deve-se procurar na 1ª metade da tabela (esquerda do registro atual)

## Pesquisa mais eficiente em vetores

Precisa que os registros estejam ordenados

#### **Funcionamento:**

Verifique o registro que está no meio da tabela (central)

Se for a chave desejada, pára a busca e retorna a posição

Se a chave for menor, deve-se procurar na 1ª metade da tabela (esquerda do registro atual)

Se a chave for maior, deve-se procurar na 2ª metade da tabela (direita do registro atual)

## Pesquisa mais eficiente em vetores

Precisa que os registros estejam ordenados

#### **Funcionamento:**

Verifique o registro que está no meio da tabela (central)

Se for a chave desejada, pára a busca e retorna a posição

Se a chave for menor, deve-se procurar na 1ª metade da tabela (esquerda do registro atual)

Se a chave for maior, deve-se procurar na 2ª metade da tabela (direita do registro atual)

Repita o processo sobre o novo espaço de busca até:

Chave ser encontrada (sucesso)

## Pesquisa mais eficiente em vetores

Precisa que os registros estejam ordenados

#### **Funcionamento:**

Verifique o registro que está no meio da tabela (central)

Se for a chave desejada, pára a busca e retorna a posição

Se a chave for menor, deve-se procurar na 1ª metade da tabela (esquerda do registro atual)

Se a chave for maior, deve-se procurar na 2ª metade da tabela (direita do registro atual)

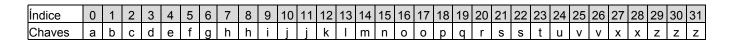
Repita o processo sobre o novo espaço de busca até:

Chave ser encontrada (sucesso)

Ficar apenas um registro cuja chave é diferente (falha)

Índice	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Chaves	а	b	С	d	е	f	g	h	h	i	j	j	k	ı	m	n	0	0	р	q	r	s	s	t	u	٧	٧	Х	х	Z	Z	z

Pesquisar 'r'

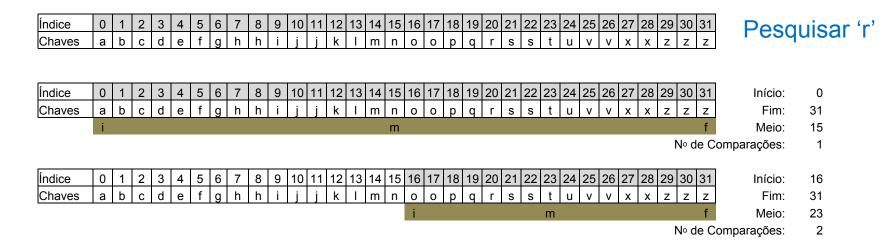


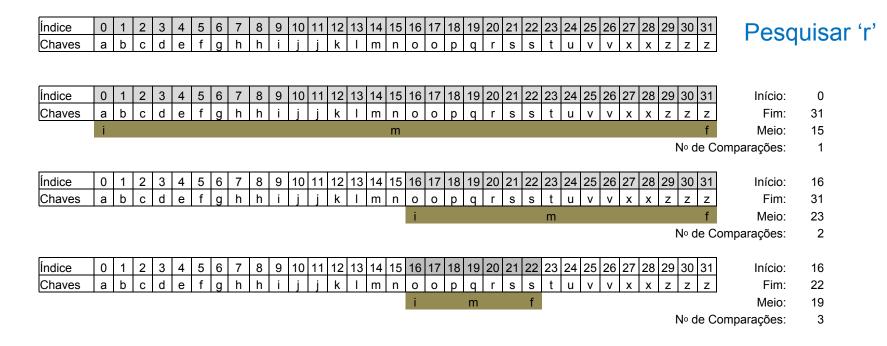
#### Pesquisar 'r'

0 31 15

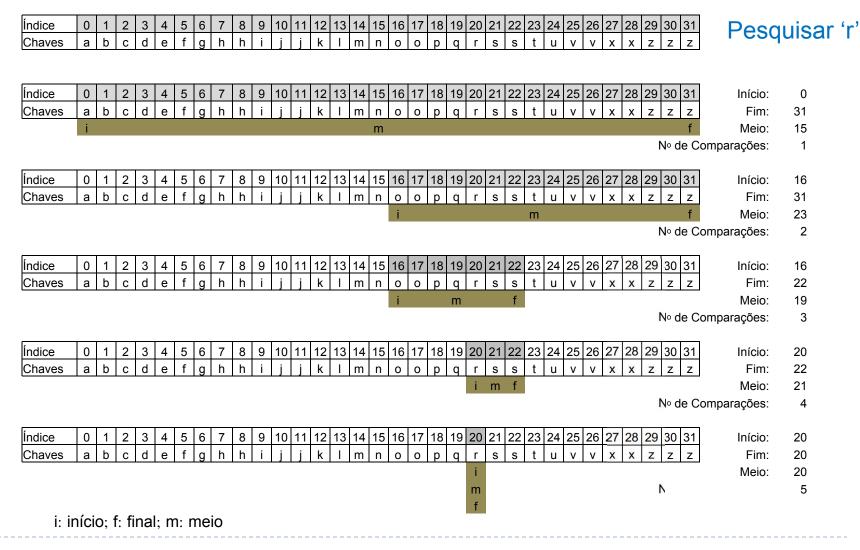
Índice	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Início:
Chaves	а	b	С	d	е	f	g	h	h	i	j	j	k	I	m	n	0	0	р	q	r	s	s	t	u	٧	٧	х	Х	z	z	z	Fim:
	i															m																f	Meio:

N∘ de Comparações: 1





Indice	Índice Chaves	0 a	1 b	2 c	3 d	4 e	5 f	6 g	7 h	8 h	9 i	10 j	11 j	12 k	13 I	14 m	15 n	16 0	17 0	18 p	19 q	20 r	21 s	22 s	23 t	24 u	25 v	26 v	27 x	28 x	29 z	30 z	31 z	P	Pesq	uisar
Indice			1				5				9	10	11		13							20			23									I		
No de Comparações: 1    Indice	Chaves		b	С	d	е	f	g	h	h	İ	L j	L j	k	1	m		0	0	р	q	r	S	S	t	u	V	٧	Х	Х	Z	Z	Z			
Indice		ı															m															N 1				-
Chaves a b c d e f g h h i j j k l m n o o p q r s s t u v v x x z z z Fim: 31    Imdice   O   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   16   17   18   19   20   21   22   23   24   25   26   27   28   29   30   31																																N∘ C	ie Co	ompara	içoes:	1
i m f Meio: 23 No de Comparações: 2    Indice	Índice	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	I	Início:	16
No de Comparações: 2    Indice	Chaves	а	b	С	d	е	f	g	h	h	i	j	j	k	ı	r	n	0	0	р	q	r	Ø	s	t	u	٧	٧	Х	Х	z	z	z		Fim:	31
Índice       0       1       2       3       4       5       6       7       8       9       10       11       12       13       14       15       16       17       18       19       20       21       22       23       24       25       26       27       28       29       30       31       Início:       16         Chaves       a       b       c       d       e       f       g       h       h       i       j       j       k       l       m       n       o       o       p       q       r       s       s       t       u       v       v       x       z       z       z       Fim:       22         Indice       0       1       2       3       4       5       6       7       8       9       10       11       12       13       14       15       16       17       18       19       20       21       22       23       24       25       26       27       28       29       30       31       Início:       20         Chaves       a       b       c       d       e       f <td></td> <td>i</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>m</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>f</td> <td></td> <td>Meio:</td> <td>23</td>																		i							m								f		Meio:	23
Chaves       a       b       c       d       e       f       g       h       h       i       j       j       k       l       m       n       o       o       p       q       r       s       s       t       u       v       v       x       z       z       z       z       Eim:       22         Meio:       19         No de Comparações:       3         Índice       0       1       2       3       4       5       6       7       8       9       10       11       12       13       14       15       16       17       18       19       20       21       22       23       24       25       26       27       28       29       30       31       Início:       20         Chaves       a       b       c       d       e       f       g       h       h       i       j       j       k       l       m       n       o       o       p       q       r       s       t       u       v       v       x       z       z       z       Fim:       22         Chaves       a <td></td> <td>N∘ d</td> <td>le Co</td> <td>ompara</td> <td>ações:</td> <td>2</td>																																N∘ d	le Co	ompara	ações:	2
Chaves       a       b       c       d       e       f       g       h       h       i       j       j       k       l       m       n       o       o       p       q       r       s       s       t       u       v       v       x       z       z       z       z       Eim:       22         Meio:       19         No de Comparações:       3         Índice       0       1       2       3       4       5       6       7       8       9       10       11       12       13       14       15       16       17       18       19       20       21       22       23       24       25       26       27       28       29       30       31       Início:       20         Chaves       a       b       c       d       e       f       g       h       h       i       j       j       k       l       m       n       o       o       p       q       r       s       t       u       v       v       x       z       z       z       Fim:       22         Chaves       a <td></td>																																				
i m f No de Comparações: 3    Início	Índice	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	I	Início:	16
No de Comparações: 3    Início   O   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   16   17   18   19   20   21   22   23   24   25   26   27   28   29   30   31   20   20   20   20   20   20   20   2	Chaves	а	b	С	d	е	f	g	h	h	i	j	j	k	ı	m	n	0	0	р	q	r	s	s	t	u	٧	٧	х	Х	z	z	z		Fim:	22
Índice       0       1       2       3       4       5       6       7       8       9       10       11       12       13       14       15       16       17       18       19       20       21       22       23       24       25       26       27       28       29       30       31       Início:       20         Chaves       a       b       c       d       e       f       g       h       h       i       m       n       o       o       p       q       r       s       s       t       u       v       v       x       z       z       z       Fim:       22         i       m       f       f       m       f       f       m       f       m       f       f       m       f	•																	i			m			f											Meio:	19
Índice       0       1       2       3       4       5       6       7       8       9       10       11       12       13       14       15       16       17       18       19       20       21       22       23       24       25       26       27       28       29       30       31       Início:       20         Chaves       a       b       c       d       e       f       g       h       h       i       m       n       o       o       p       q       r       s       s       t       u       v       v       x       z       z       z       Fim:       22         i       m       f       f       m       f       f       m       f       m       f       f       m       f																																N∘ d	le Co	ompara	ações:	3
Chaves       a b c d e f g h h i j j k l m n o o p q r s s t u v v x x z z z         i m f       Eim: 22																																			•	
Chaves a b c d e f g h h i j j k l m n o o p q r s s t u v v x x z z z Fim: 22 i m f	Índice	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		Início:	20
	Chaves	а	b	С	d	е	f	g	h	h	i	j	j	k	ı	m	n	0	0	р	q	r			t	u	٧	٧	х	х	z	z	z		Fim:	22
 N∘ de Comparações: 4																			•			i	m	f											Meio:	21
																																N∘ d	le Co	ompara	ações:	4



int busca\_bin(Arq \*arq, int chave) {

}

```
int busca_bin(Arq *arq, int chave) {
  if (arq->tamanho == 0) return -1; // Arquivo vazio
```

}

```
int busca_bin(Arq *arq, int chave) {
  if (arq->tamanho == 0) return -1; // Arquivo vazio
  int meio, ini = 0, fim = arq->tamanho-1;
  while (ini <= fim) {
```

```
int busca_bin(Arq *arq, int chave) {
  if (arq->tamanho == 0) return -1; // Arquivo vazio
  int meio, ini = 0, fim = arq->tamanho-1;
  while ( ini <= fim ) {
     meio = (ini + fim) / 2; // Divisão inteira
```

```
int busca_bin(Arq *arq, int chave) {
  if (arq->tamanho == 0) return -1; // Arquivo vazio
  int meio, ini = 0, fim = arq->tamanho-1;
  while ( ini <= fim ) {
     meio = (ini + fim) / 2; // Divisão inteira
     if ( arq->itens[meio].chave == chave )
       return meio; // Retorna posição do registro no arquivo
```

```
int busca_bin(Arq *arq, int chave) {
  if (arq->tamanho == 0) return -1; // Arquivo vazio
  int meio, ini = 0, fim = arq->tamanho-1;
  while ( ini <= fim ) {
     meio = (ini + fim) / 2; // Divisão inteira
     if ( arq->itens[meio].chave == chave )
       return meio; // Retorna posição do registro no arquivo
     else if ( arg->itens[meio].chave > chave )
       fim = meio-1:
```

```
int busca_bin(Arq *arq, int chave) {
  if (arq->tamanho == 0) return -1; // Arquivo vazio
  int meio, ini = 0, fim = arq->tamanho-1;
  while ( ini <= fim ) {
     meio = (ini + fim) / 2; // Divisão inteira
     if ( arq->itens[meio].chave == chave )
       return meio; // Retorna posição do registro no arquivo
     else if ( arg->itens[meio].chave > chave )
       fim = meio-1:
     else
       ini = meio+1:
```

```
int busca_bin(Arq *arq, int chave) {
  if (arg->tamanho == 0) return -1; // Arguivo vazio
  int meio, ini = 0, fim = arq->tamanho-1;
  while ( ini <= fim ) {
     meio = (ini + fim) / 2; // Divisão inteira
     if ( arq->itens[meio].chave == chave )
       return meio; // Retorna posição do registro no arquivo
     else if ( arg->itens[meio].chave > chave )
       fim = meio-1:
     else
       ini = meio+1:
  return -1; // Registro não foi encontrado
```

## Busca binária: análise

# Eficiência está relacionada com a quantidade de iterações

#### Número de comparações:

São realizadas no máximo 2 comparações: = e >

**Custo constante** 

#### Número de divisões:

Arquivo é dividido ao meio a cada iteração

No pior caso, o processo ocorre até não ser mais possível dividir o arquivo, ou seja, restar um único registro

Custo relacionado com o tamanho do vetor (fator crítico)

## Busca binária: análise

#### Quantas iterações precisamos para tratar um vetor?

Índice 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Chaves a b c d e f g h h i j j k l m n o o p q r s s t u v v v x x z z z			
Índice         0         1         2         3         4         5         6         7         8         9         10         11         12         13         14         15         16         17         18         19         20         21         22         23         24         25         26         27         28         29         30         31           Chaves         a         b         c         d         e         f         g         h         h         i         j         j         k         l         m         n         o         o         p         q         r         s         s         t         u         v         v         x         z         z         z         Fim:	0	Tamanho vetor	N = 32
i m f Meio:  Num. de Comparações:	15	Iteração	1
Indice 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31	16	Tamanho vetor	$\frac{N}{2} = 16$
Chaves       a       b       c       d       e       f       g       h       h       i       m       v       x       z       z       z       Fim:         i       m       m       f       Meio:         Num. de Comparações:	23	Iteração	2
Indice 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 Inicio:	16	Tamanho vetor	$\frac{N}{4} = 7$
Chaves       a       b       c       d       e       f       g       h       i       j       j       k       l       m       o       o       p       q       r       s       t       u       v       v       x       z <td>19</td> <td>Iteração</td> <td>3</td>	19	Iteração	3
Índice         0         1         2         3         4         5         6         7         8         9         10         11         12         13         14         15         16         17         18         19         20         21         22         23         24         25         26         27         28         29         30         31         Início:           Chaves         a         b         c         d         e         f         g         h         h         i         j         j         k         l         m         n         o         p         q         r         s         t         u         v         v         x         z         z         z         Fim:	20	Tamanho vetor	$\frac{N}{8} = 3$
i m f Meio:  Num. de Comparações:	21	Iteração	4
Índice         0         1         2         3         4         5         6         7         8         9         10         11         12         13         14         15         16         17         18         19         20         21         22         23         24         25         26         27         28         29         30         31         Início:           Chaves         a         b         c         d         e         f         g         h         h         i         j         j         k         l         m         n         o         p         q         r         s         t         u         v         v         x         z         z         z         Fim:	20	Tamanho vetor	$\frac{N}{16} = 1$
i Meio:  m Num. de Comparações:	20 20 5	Iteração	5

## Busca binária: análise

Se um vetor tem tamanho n, qual será seu tamanho após i divisões ao meio?

Se um vetor tem tamanho n, qual será seu tamanho após i divisõ

ses ao meio?	Tamanho vetor	Nº de divisões
	$n=\frac{n}{20}$	0

Se um vetor tem tamanho n, qual será seu tamanho

Tamanho vetor	Nº de divisões
$n = \frac{n}{2^0}$	0
$\frac{n}{2} = \frac{n}{2^1}$	1

Se um vetor tem tamanho n, qual será seu tamanho

elo?	Tamanno vetor	Nº de divisoes
• • •	$n = \frac{n}{2^0}$	0
2 (	$\frac{n}{2} = \frac{n}{2^1}$	1
2	$\frac{\frac{n}{2}}{2} = \frac{n}{4} = \frac{n}{2^2}$	2

Se um vetor tem tamanho n, qual será seu tamanho

neio?	Tamanho vetor	Nº de divisões
	$n = \frac{n}{2^0}$	0
<b>2</b> (	$\frac{n}{2} = \frac{n}{2^1}$	1
2 (	$\frac{\frac{n}{2}}{2} = \frac{n}{4} = \frac{n}{2^2}$	2
2 (	$\frac{\frac{n}{4}}{2} = \frac{n}{8} = \frac{n}{2^3}$	3

Se um vetor tem tamanho n, qual será seu tamanho

neio?	Tamanho vetor	Nº de divisões
• 2 (	$n = \frac{n}{2^0}$	0
<ul><li>2 €</li><li>2 €</li></ul>	$\frac{n}{2} = \frac{n}{2^1}$	1
	$\frac{\frac{n}{2}}{2} = \frac{n}{4} = \frac{n}{2^2}$	2
÷ 2 (	$\frac{\frac{n}{4}}{2} = \frac{n}{8} = \frac{n}{2^3}$	3
2 (	$\frac{\frac{n}{8}}{2} = \frac{n}{16} = \frac{n}{2^4}$	4

Se um vetor tem tamanho n, qual será seu tamanho

neio?	Tamanho vetor	Nº de divisões
	$n = \frac{n}{2^0}$	0
2	$\frac{n}{2} = \frac{n}{2^1}$	1
2 C	$\frac{\frac{n}{2}}{2} = \frac{n}{4} = \frac{n}{2^2}$	2
<b>+</b> 2 (	$\frac{\frac{n}{4}}{\frac{2}{3}} = \frac{n}{8} = \frac{n}{2^3}$	3
<b>2</b>	$\frac{\frac{n}{8}}{2} = \frac{n}{16} = \frac{n}{2^4}$	4
	•••	•••
2	$\frac{n}{2^i}$	i

Se um vetor tem tamanho n, qual será seu tamanho

<u>n</u>	
$2^{i}$	

reio?	Tamanho vetor	Nº de divisões
	$n = \frac{n}{2^0}$	0
<ul><li>2 €</li><li>2 €</li></ul>	$\frac{n}{2} = \frac{n}{2^1}$	1
	$\frac{\frac{n}{2}}{\frac{n}{2}} = \frac{n}{4} = \frac{n}{2^2}$	2
2 (	$\frac{\frac{n}{4}}{\frac{2}{2}} = \frac{n}{8} = \frac{n}{2^3}$	3
2 (	$\frac{\frac{n}{4}}{\frac{n}{2}} = \frac{n}{8} = \frac{n}{2^3}$ $\frac{\frac{n}{8}}{2} = \frac{n}{16} = \frac{n}{2^4}$	4
	•••	•••
<b>2</b>	$\frac{n}{2^i}$	i

Qual é o valor de *i* para que o vetor tenha tamanho 1 (não seja possível dividi-lo)?

$$\frac{n}{2^i}=1$$

Qual é o valor de *i* para que o vetor tenha tamanho 1 (não seja possível dividi-lo)?

$$\frac{n}{2^i}=1$$

Desenvolvendo a fórmula:  $n = 2^i$ 

Qual é o valor de *i* para que o vetor tenha tamanho 1 (não seja possível dividi-lo)?

$$\frac{n}{2^i}=1$$

Desenvolvendo a fórmula:  $n = 2^i$ 

Aplicando  $log_2$  nos 2 lados:  $log_2 n = log_2 2^i$ 

Qual é o valor de *i* para que o vetor tenha tamanho 1 (não seja possível dividi-lo)?

$$\frac{n}{2^i}=1$$

Desenvolvendo a fórmula:  $n = 2^i$ 

Aplicando  $log_2$  nos 2 lados:  $log_2 n = log_2 2^i$ 

$$log_2 n = i log_2 2$$

Qual é o valor de *i* para que o vetor tenha tamanho 1 (não seja possível dividi-lo)?

$$\frac{n}{2^i}=1$$

Desenvolvendo a fórmula:  $n = 2^i$ 

Aplicando  $log_2$  nos 2 lados:  $log_2 n = log_2 2^i$  $log_2 n = i log_2 2^{-1}$ 

Qual é o valor de *i* para que o vetor tenha tamanho 1 (não seja possível dividi-lo)?

$$\frac{n}{2^i}=1$$

Desenvolvendo a fórmula:  $n = 2^i$ 

Aplicando  $log_2$  nos 2 lados:  $log_2 n = log_2 2^i$   $log_2 n = i log_2 2^i$  $i = log_2 n$ 

### **Custo total:**

#### **Custo total:**

Comparações por iteração tem custo constante

#### **Custo total:**

Comparações por iteração tem custo constante Quantidade máxima de divisões (i) é log<sub>2</sub> n

#### **Custo total:**

Comparações por iteração tem custo constante Quantidade máxima de divisões (i) é log<sub>2</sub> n

$$C(n) = log_2 n$$

#### **Custo total:**

Comparações por iteração tem **custo constante** Quantidade máxima de divisões (*i*) é  $log_2 n$ 

$$C(n) = log_2 n$$

# Ressalva: não é recomendada para aplicações muito dinâmicas

Custo para manter a tabela ordenada é alto Envolve o deslocamento de elementos

## Exercícios

1. Construa um algoritmo para realizar a **análise empírica** da complexidade do caso médio da busca sequencial, considerando 100, 1.000 e 10.000 vetores de 15 posições (cada uma com um valor distinto de 1 a 15). Demonstre a complexidade usando o número 10 como chave. Ao final, o algoritmo deve apresentar o número de ocorrências do 10 em cada posição do vetor, a quantidade de comparações realizadas para encontrá-lo (detalhada como nos slides) e a média de comparações.

Implemente o algoritmo de busca binária de **forma recursiva** e compare sua complexidade com a versão iterativa, considerando 100 execuções sobre vetores de 1.000 inteiros distintos para encontrar o número 169.

## Bibliografia

Slides adaptados do material da Profa. Gina M. B. Oliveira e do Prof. Dr. Bruno Travençolo.

BACKES, A. Linguagem C Descomplicada: portal de vídeo-aulas para estudo de programação. Disponível em:

https://programacaodescomplicada.wordpress.com/indice/estrutura-de-dados/

CORMEN, T.H. et al. Algoritmos: Teoria e Prática, Campus, 2002

ZIVIANI, N. Projeto de algoritmos: com implementações em Pascal e C (2ª ed.), Thomson, 2004

MORAES, C.R. Estruturas de Dados e Algoritmos: uma abordagem didática (2ª ed.), Futura, 2003