

---

## 1ª Lista de Exercícios

---

- ① Explique a diferença entre arquivo lógico e arquivo físico.
- ② Descreva o que acontece quando um arquivo já existente é aberto por um aplicativo (p.ex., um programa) para escrita. E se o mesmo ocorrer com um arquivo não existente?
- ③ No que consiste a operação de posicionamento (seeking) em um arquivo? Qual a sua utilidade?
- ④ Explique o que é um cilindro, e a razão para a organização de arquivos em cilindros.
- ⑤ Por que os discos são considerados o gargalo de um sistema computacional? Explique como este problema pode ser minimizado.
- ⑥ O que são “buffers” de E/S (ou I/O)?
- ⑦ Descreva as operações fundamentais que podem ser realizadas em um arquivo. Descreva as funções que executam estas operações na linguagem de programação que você usa, e como elas são utilizadas. Por que existem vários modos de abrir um arquivo?
- ⑧ Discuta as diferenças e semelhanças entre a memória principal (RAM) e a memória secundária (os arquivos).
- ⑨ Como são organizados fisicamente os discos? De que forma os discos armazenam os arquivos?
- ⑩ Como os arquivos são organizados em uma fita magnética?
- ⑪ No que consiste um sistema de armazenamento terciário?
- ⑫ Quais parâmetros são considerados para calcular o tempo de leitura de um arquivo em fita?
- ⑬ O que são “buffers” de E/S (ou I/O)? Quais os passos executados para que ler um byte do disco de forma que ele possa ser utilizado por um programa?
- ⑭ Quais as vantagens e desvantagens de fitas sobre discos com dispositivos de memória secundária?
- ⑮ Explique o que é um cilindro, e a razão para a organização de arquivos em cilindros.
- ⑯ Explique o que é um *cluster* e o que é um *extent*.

- 17 O que é a fragmentação de um arquivo no disco? Quais os tipos de fragmentação do arquivo, porque elas ocorrem e quais seus efeitos?
- 18 Discuta as vantagens e desvantagens de organizar arquivos em blocos de tamanho definido pelo usuário, ao invés de em setores de tamanho fixo.
- 19 Considere armazenar em fita 1.000.000 de registros com 100 bytes cada. Suponha fita com 6.250 bpi, com intervalo entre blocos (suponha 1 bloco = 1 registro / 1 bloco = 50 registro) de 0.3 polegadas. Quanto de fita é necessário?
- 20 As aplicações usualmente armazenam as informações em arquivos organizando as em campos e registros. Explique as diferentes maneiras pelas quais um campo pode ser armazenado em um arquivo para posterior recuperação.
- 21 Explique as diferentes estratégias que podem ser utilizadas para separar um registro de outro. Discuta as vantagens e desvantagens de cada uma delas.
- 22 Explique o que é fragmentação de campos e registros. Quando e por que ela ocorre?
- 23 Se a separação entre registros e campos é feita por delimitadores, quais as restrições para a escolha desses delimitadores. Descreva uma situação que exemplifique sua resposta.
- 24 Como um registro é identificado para acesso aleatório? Qual operação permite localizar um registro no arquivo em C?
- 25 Explique como é possível melhorar o desempenho de um acesso sequencial a todo o conteúdo de um arquivo. Tal solução também garante um melhor desempenho de uma sequência arbitrária de acessos aleatórios? Discuta.
- 26 Quais os comandos em C para: (i) reposicionar o indicador de posição em um arquivo, (ii) retornar a posição atual em um arquivo? Explique-os.
- 27 Considere um arquivo composto de registros de tamanho fixo de 64 bytes com 3 campos de tamanho fixo: RA (10 bytes), nome (40 bytes) e curso (14 bytes). Desconsidere os bytes usados como delimitadores. Qual será o comando para acessar o primeiro byte do campo nome do RRN=3? (considerando que o RRN do primeiro registro é 1).
- 28 Quais as vantagens e as desvantagens de utilizar arquivos organizados em registros de tamanho fixo?
- 29 O que é RRN? Como é possível fazer acessos aleatórios em arquivos a registros de tamanho variável?
- 30 Aplique a compressão de dados no seguinte trecho de texto para obtenção do código de Huffman "fdgfggsdgaasggsagsdfg".  
(a) construa a árvore de Huffman  
(b) calcule o número de bytes economizados.

- 31) Como se faz a alocação de registros de tamanho variável tendo em vista a existência de espaços vazios, existem 3 abordagens. Indique quais são e explique suas vantagens e desvantagens.
- 32) É vantajoso manter um arquivo separado para armazenar apenas as chaves e os byte offsets ou RRNs, dos registros no arquivo de dados? Como isto afeta a inserção de um novo registro?
- 33) Como um registro pode ser eliminado de um arquivo?
- 34) Explique a diferença entre fragmentação externa e interna?
- 35) O que são modelos abstratos de dados, para que são utilizados e quais as suas vantagens?
- 36) Por que é interessante utilizar cabeçalhos nos arquivos?
- 37) Considere um arquivo contendo os 3 registros de tamanho fixo apresentados abaixo e mostre como fica o arquivo nas seguintes situações.
- Quando o 2.<sup>o</sup> registro é excluído.
  - Quando o arquivo é compactado após a exclusão do 2o. registro.

```
Carla|Guimarães|Rua Riachuelo 123|Jardim América|033|720|.....
Nivaldo|Soares|Rua Moraes Barros 100|Centro|145|162|.....
Djavan Carlos| Moura|Av D. Pedro I 1234|Vila Independência|033|730|
```

- 38) Considere um arquivo contendo registros de tamanho variável apresentados abaixo e mostre como fica a lista Dispo e o arquivo após as exclusões/inclusões:
- Excluir o segundo registro
  - Incluir Gabriel|Goes|Av Dois 123|Vila Zac|033|124| permitindo a fragmentação interna
  - Incluir Gabriel|Goes|Av Dois 123|Vila Zac|033|124| sem fragmentação interna
  - Excluir o primeiro registro

```
57 Carla|Guimarães|Rua Riachuelo 123|Jardim América|033|720|51 Nivaldo|Soares|
Rua Moraes Barros 100|Centro|145|162|65 Djavan Carlos|Moura|Av D.Pedro I
1234|Vila Independência|033|730|50 Aline|Miranda|Rua Treze de Maio 10|Cen-
tro|010|320|
```

- 39) Em princípio, não é possível fazer busca binária em um arquivo de dados com registros de tamanho variável. Por que a indexação do arquivo torna a busca binária possível? Com um arquivo com registros de tamanho fixo é possível fazer busca binária. Isto significa que indexação não é necessária para arquivos de registros de tamanho fixo?
- 40) Qual o propósito em deixar um indicador de desatualizado (flag "out-of-date") no cabeçalho de um índice utilizado num ambiente de multiprogramação.? Dê um exemplo de situação em que manter esse indicador poderia ajudar.

- 41 Quando um registro é atualizado num arquivo, as chaves primárias e secundárias do índice podem ser alteradas ou não, dependendo do arquivo ter registros de tamanho fixo ou variável e dependendo do tipo de alteração que foi feita no registro de dados. Faça uma lista das situações diferentes que podem ocorrer devido a atualizações e explique como cada uma pode afetar os índices.
- 42 O que é uma lista invertida? Quando ela é útil? Esquematize o conteúdo de um índice secundário organizado como lista invertida para um arquivo de dados hipotético.
- 43 Suponha que você tenha um arquivo de dados de CDs, como descrito no exemplo em aula, muito grande, com um índice pela chave primária e índices pelas chaves secundárias organizados por compositor e gravadora. Suponha que uma lista invertida é usada para as chaves secundárias. Explique, passo a passo, como um programa poderia responder às seguintes solicitações:
- (a) Liste todas as gravações de Chico Buarque;
  - (b) Liste todas as gravações de Milton Nascimento e gravadora Emi-Odeon.
- 44 Por que, para a eliminação de um registro do arquivo de dados, é possível eliminar o registro apenas do índice primário, e não do secundário?
- 45 Explique o conceito de late binding e early binding. Quando é aconselhável o *early binding* e por que?
- 46 Considere um arquivo armazenando os registros: M A C F D E R G B. Suponha uma memória interna com capacidade para 3 registros. Faça: calcule o número de corridas e seu tamanho, depois o tamanho dos buffers de entrada, mostre o passo-a-passo da ordenação por k-vias e o arquivo final ordenado.
- 47 Suponha que um arquivo de dados com 6.000 registros, mantido em disco, deve ser ordenado em um computador cuja memória interna acomoda no máximo 600 registros por vez, usando o procedimento de intercalação. Considere que serão geradas 10 corridas de 600 registros cada, e que será realizada uma intercalação em 10 vias. Essa mesma área de memória interna é usada como buffer de entrada para leitura de dados do disco, e o sistema conta com um buffer de saída adicional que acomoda 200 registros.
- (a) durante a intercalação, quantos registros serão lidos de cada corrida cada vez que ela é acessada? Justifique.
  - (b) quantos seeks serão realizados para ler dados durante o processo de intercalação (excluindo a fase de geração de corridas)? Quantos serão realizados para escrever dados? Justifique.
- 48 Qual a complexidade do método de ordenação externa de intercalação por k-vias? Explique.
- 49 O desempenho de um processo de ordenação é, geralmente, medido em função do número de comparações necessárias. Porque? Explique por que o número de comparações não é uma medida adequada para avaliar o desempenho de métodos de ordenação de grandes arquivos que não cabem em memória.