

ORGANIZAÇÃO DE ESTRUTURAS DE ARQUIVOS

PROVA 1 - 2022.1 - 21/06/2022

PROF. RENATO CAMPOS MAURO

Nome: _____

Questão 1 Nos arquivos texto no sistemas operacionais tipo UNIX, a quebra de linha é representada pelo caractere `\n`. Já nos sistemas originados a partir do DOS, a quebra de linha é representada pelos caracteres `\r` e `\n` (nessa ordem e um seguido do outro). Faça um programa que, dado um arquivo texto com quebra de linhas no formato UNIX converta-o para o formato DOS.

Questão 2 Alguns algoritmos de compactação de arquivos precisam saber a quantidade de ocorrência de cada byte no arquivo. Fazendo a leitura byte a byte, elabore um programa que conte a ocorrência de cada caractere. Crie um contador para cada byte, sabendo-se que este tipo varia entre 0 e 255.

Questão 3 Sistemas gerenciadores de banco de dados são extremamente eficientes e práticos para manuseio de arquivos grandes. Internamente são implementados métodos de manipulação de arquivos bastante sofisticados. Uma das operações mais comuns em banco de dados relacionais é a junção. Dependendo de como os dados estão organizados, algoritmos mais eficientes ou menos eficientes são executados. Quando não há nenhuma organização dos dados, o método realizado é o método *nested loop* (laço aninhado), onde todo mundo é comparado com todo mundo. Já quando as duas coleções estão ordenadas (ou quando não é dispendioso ordená-las), o método usado é o *merge join* (junção por intercalação).

Considere 2 arquivos (`candidatos2021.dat` e `candidatos2022.dat`) com registros no mesmo formato, cuja estrutura em linguagem C é apresentada no trecho abaixo. Considere que os dois arquivos estejam ordenados por CPF e

que em um mesmo arquivo não haja repetição de CPF. Faça um programa que EFICIENTEMENTE gere um novo arquivo contendo os dados dos candidatos que se inscreveram no concurso, tanto no ano de 2021 como no ano de 2022.

```
struct Registro {
    char cpf[15];
    char curso[20];
    char dataNascimento[11];
    char sexo;
    char email[40];
    char opcaoQuadro;
};
```

Questão 4 Usando a mesma estrutura da questão 3, escreva um programa que conte o número de candidatos inscritos para o concurso para cada curso da instituição. Caso na sua proposta de algoritmo seja necessário ordenar por um critério diferente da ordenação original (CPF), assumo que o arquivo cabe todo em memória.

Questão 5 Responda objetivamente.

- (a) Dados dois arquivos, como saber se os arquivos tem o mesmo tamanho?
- (b) Na questão 4 a hipótese era de que o arquivo cabia todo em memória. E se não coubesse?
- (c) Comente sobre a diferença de eficiência de juntar dois arquivos por intercalação e por loop aninhado (força bruta).
- (d) Que estrutura de dados poderia ser usada, no caso do problema da contagem de ocorrências, para após a contagem, estabelecer rapidamente qual caractere tem o maior número de ocorrência no arquivo?

BOA PROVA!!