

Projeto 01

Disciplina: *Organização de Arquivos – Turma C*

Semestre: 2º/2017

Prof.: Flávio Vidal

Título: *Indexação de Arquivos e utilização de técnicas com Lista Invertida*

Entrega: *15/10/2017 – 23h50min (Relatório e Arquivos-Fontes via Aprender.unb.br)*

1. Objetivos

Implementar, utilizando obrigatoriamente a linguagem C/C++ ou Python, procedimentos de indexação de arquivos (geração de índices primários e secundários) com registros com campos de tamanho fixo, utilizando técnicas de Lista Invertida (nos arquivos de índices secundários) de forma a aplicar os conceitos básicos de técnicas de indexação. Este projeto é parte integrante da ementa da disciplina de Organização de Arquivos.

O projeto desta disciplina é uma atividade planejada de forma a complementar e reforçar o conteúdo programático da disciplina Organização de Arquivos. Espera-se que nas atividades de projeto os alunos desenvolvam sua capacidade de observação, análise e compreensão das metodologias de organização e estruturas de arquivos.

Desta forma cabe ao aluno, partindo da premissa que possui os requisitos para o curso, juntamente com o conteúdo adquirido nas aulas teóricas, desenvolver todas as etapas da implementação solicitada.

2. Introdução Teórica

Todos os aspectos teóricos podem ser encontrados no Capítulo 6 e 7 do livro texto, complementando com as transparências apresentadas em sala de aula sobre o assunto.

3. Metodologia

Deverá ser implementado utilizando linguagem de programação de alto nível (linguagem C para Linux/Unix de preferência) um programa que atenda os seguintes requisitos:

Requisito 1: A implementação deve ser capaz de realizar a leitura de arquivos do benchmark utilizado (arquivos texto disponibilizados no ambiente *aprender.unb.br*) e imprimir todas as operações que forem realizadas, tanto no arquivo de índices, quanto no arquivo de dados, no dispositivo de saída de vídeo (monitor) conforme descrito nos subitens abaixo:

Arquivo de Dados e de Índices: Deve-se utilizar forma tabular, em que a primeira linha contenha os rótulos indicativos de nome do campo e as demais linhas representando o conteúdo separadas por tabulações (tanto para o arquivo de dados, quanto de índice).

¹ MATRIC NOME	OP	CURSO	TURMA
² 024312 Leonardo Nunes de Souza	34	G	AB

Requisito 2: Criar arquivo de índice e de dados

A partir do arquivo de benchmarks disponibilizados (*lista1.txt* e *lista2.txt* disponibilizados no Aprender.unb.br), crie um arquivo de índices primário simples (*indicelista1.ind* e *indicelista2.ind*). Utilize para a criação deste índice os campos de MATRIC e NOME concatenados num total de 30(trinta caracteres) como chave primária para o primeiro campo. Desta forma serão minimizadas as possibilidades de chaves primárias repetidas. O segundo campo deverá conter a informação necessária para identificar qual o registro do arquivo de dados se refere no registro de índices primário. Deve-se também criar dois arquivos de índices secundários, tendo como campos a coluna OP e TURMA, correlacionados às suas respectivas chaves primárias. Estes arquivos de índices secundários deverão ser criados utilizando como princípio a técnica de Listas Invertidas, segundo apresentado em sala de aula.

Para ordenar o arquivo de índice utilize um algoritmo de ordenação HEAPSORT seguindo o modelo da literatura existente. Crie uma rotina para a visualização do arquivo de índice (primário e secundário) ordenado no monitor de vídeo a partir de um comando em menu de entrada e seleção de opção.

Requisito 3: Regravar arquivo de índice após o uso.

Defina um conjunto de rotinas básicas necessárias para após a utilização do arquivo de índice (primário e secundário), este seja regravado em disco após as operações sejam realizadas, tanto no arquivo de índice, quanto no arquivo de dados.

Requisito 5: Inclusão de registros

Deverão ser desenvolvidas rotinas que permitam a inclusão de novos registros no arquivo de dados e também no arquivo de índices (primário e secundário). Lembrando que o arquivo de índices deverá ser mantido ordenado. No processo de inclusão deverá ser mostrado no terminal de vídeo o arquivo de índice antes e depois da inclusão do novo registro.

Requisito 6: Excluir registros

Deverão ser desenvolvidas rotinas que permitam a exclusão de novos registros no arquivo de dados e também no arquivo de índices (primário e secundário). Lembrando que o arquivo de índices deverá ser mantido ordenado e metodologias de re-utilização do espaço disponibilizado. Mostrar no terminal de vídeo o arquivo de índice antes e depois da exclusão, como no item anterior.

¹ Cabeçalho da tabela a ser apresentada na saída de vídeo apresentando os rótulos relativos aos registros do arquivo.

² Conteúdo do registro seguindo o alinhamento proposto pelo cabeçalho, alinhado ao rótulo específico.

Requisito 7: Atualizar registros:

Definir rotinas nas quais implementem a atualização de registros nos arquivos de dados, e respectivamente no arquivos de índices (primário e secundário), seguindo as condições descritas abaixo:

Requisito 7.1: Com mudança de chave primária.

Requisito 7.2: Sem mudança de chave primária.

Deve-se ressaltar que estas rotinas, implementadas em cada requisito, deverão contemplar todos os aspectos relativos às características dos arquivos de índices, de forma que todas as operações realizadas, tanto no arquivo de índices, quanto as realizadas nos arquivos de dados (arquivos de *benchmarks*), devem ser apresentadas via terminal de vídeo, e ainda preservando o arquivo gerado em disco.

4. Grupos

Neste projeto será permitido a formação de grupos com no máximo 2(DOIS) alunos, sendo os alunos pertencentes à mesma turma. Não será permitida duplas em que os membros são de turmas diferentes. A partir do grupo formado, deverá ser indicado um líder, no qual este líder será o responsável pelo envio dos arquivos fontes para o sistema Aprender.unb.br. Somente serão aceitos os arquivos fontes enviado pelo líder do grupo. O relatório escrito deverá ser entregue em PDF na área específica da atividade no ambiente aprender.unb.br. Reitero que não será aceito nenhum arquivo (e/ou relatório) via e-mail do professor, independente de indisponibilidade do ambiente e/ou outro problema decorrente.

5. Relatório

O relatório deve demonstrar que a respectiva atividade de laboratório foi realizada com sucesso e que os princípios subjacentes foram compreendidos.

O relatório da atividade de laboratório é o documento gerado a partir do trabalho realizado seguindo as orientações exigidas na metodologia de laboratório. Este deve espelhar todo o trabalho desenvolvido nos processos de obtenção dos dados e sua análise. Apresentamos a seguir uma recomendação de organização para o relatório da atividade de laboratório. Deverá conter as seguintes partes:

i. Identificação: Possuir a indicação clara do título do experimento abordado, a data da sua realização, a identificação da disciplina/turma, os nomes dos componentes do grupo, número de matrícula e email.

ii. Objetivos: Apresentar de forma clara, porém sucinta, os objetivos do laboratório.

iii. Introdução: Deve conter a teoria necessária à realização da atividade de laboratório.

iv. Materiais e Métodos: É dedicada à apresentação dos materiais e equipamentos, descrição do arranjo experimental e uma exposição minuciosa do procedimento de laboratório realmente adotado.

v. Resultados: Nesta parte são apresentados os resultados das implementações efetuadas, na forma de tabelas e gráficos, sem que se esqueça de identificar em cada caso os parâmetros utilizados.

vi. Discussão e Conclusões: A discussão visa comparar os resultados obtidos e os previstos pela teoria. Deve se justificar eventuais discrepâncias observadas. As conclusões resumem a atividade de laboratório e destacam os principais resultados e aplicações dos conceitos vistos.

vii. Bibliografia: Citar as fontes consultadas, respeitando as regras de apresentação de bibliografia (autor, título, editora, edição, ano, página de início e fim).

O relatório do laboratório deverá ser confeccionado em editor eletrônico de textos, utilizando o padrão de formatação descrito no arquivo de exemplo, disponibilizado no website da disciplina (oficial e/ou mirror). Está disponibilizado um único padrão de formatação para editores científicos LATEX (arquivo extensão *.zip contendo arquivo de exemplo do uso do pacote), cabendo ao grupo a escolha de qual editor Latex será utilizado. Este modelo pode ser acessado no aprender da disciplina. Somente serão aceitos para avaliação relatórios em PDF feitos seguindo esta formatação.

Todo o código fonte (inclusive as diretrizes de compilação utilizadas) deverá ser entregue via *upload* no ambiente Aprender.unb.br, em arquivo *.zip completando a atividade designada ao laboratório correspondente (vide ambiente aprender.unb.br para maiores detalhes). O código fonte deverá ser “re-compilável” para que seja realizada a correção no ambiente computacional Linux. Não serão aceitos trabalhos entregues fora do prazo estipulado, sendo atribuída nota zero ao grupo. Não serão aceitos qualquer tipo de material (relatório e códigos-fonte) via email do professor. O único método de envio deverá ser feito pelo Aprender.unb.br.

Vale ressaltar que será atribuída nota zero, definida como atividade “incompleta”, ao grupo que não entregar o relatório e/ou código fonte implementado e devidamente identificado. Entende-se como atividade completa versão em PDF do relatório e arquivos-fonte corretamente enviados ao endereço eletrônico *Aprender.unb.br*.

a. Critérios Empregados na Correção do Relatório de Laboratório

A avaliação dos relatórios terá em consideração os seguintes itens:

No.	Item	Descrição	Peso (%)
1	Apresentação	Qualidade dos gráficos, impressão, tabelas, vocabulário, legendas, etc.	10%
2	Aspectos Teóricos	Apresentação e descrição da base teórica utilizada. Avaliação da bibliografia utilizada se necessário.	20%
3	Materiais e Métodos	Descrição de todos os procedimentos utilizados, contemplando dados técnicos, bem como a metodologia utilizada no decorrer do projeto.	10%
4	Resultados	Todos os resultados alcançados no projeto.	30%
5	Discussão e Conclusões	Discussão objetiva e devidamente explicada a respeito do projeto. Inclui-se também a pontuação por iniciativa.	30%

Dúvidas deverão ser encaminhadas ao fórum de discussão específico no ambiente *Aprender.unb.br*.