

TRABALHO PRÁTICO PARTE 1

Alunos:

Beatriz Cunha Rodrigues
Daniel Ribeiro Lavra

201776038
201735042

Relatório parcial

Nesta etapa do trabalho foram propostas as tarefas de manipular adequadamente arquivos em modo texto, representar as entidades dessa base de dados escolhendo uma estrutura de dados mais adequada e armazená-las em arquivos binários. Em geral, as tarefas foram divididas entre os membros do grupo, mas como a falta de hábito com a linguagem dificultava bastante em alguns aspectos, mesmo com as tarefas definidas para cada um, nenhuma foi desenvolvida de forma exclusiva, no entanto, logo abaixo será definido qual membro e sua responsabilidade no código. A atualização dos arquivos no Git era feito pela Beatriz devido a sua intimidade com a plataforma, e o Daniel sempre trabalhando em paralelo devido a ferramenta da IDE utilizada (Visual Studio Code).

Etapa 1: Processamento de dados

Toda estrutura principal do programa onde contém, o carregamento inicial do arquivo csv no buffer e verificação se foi lido corretamente, a impressão desse arquivo na tela, toda a estrutura que o usuário utilizará para interagir com o sistema escolhendo as funcionalidades, no arquivo “main.cpp”. Além disso, o arquivo “Data.h”, a classe que especifica a estrutura em que os registros serão salvos posteriormente, também foi feita pelo Daniel.

A parte especificada para a Beatriz, começou no tratamento da leitura no “Processamento.h”, onde ocorre a separação de fato dos campos de registros para serem salvos na estrutura de dados citada anteriormente e o tratamento adequado desses registros, além disso, esses registros já tratados foram salvos em arquivo binário(.bin).

Etapa 2: Importação de Registros e Módulo de Teste

Nessa etapa do trabalho, ficou dividido para que cada membro realizasse uma tarefa. Para escolha de um i-ésimo registro solicitado pelo usuário, o Daniel ficou responsável por essa função que acessasse o devido registro no arquivo binário, a função `acessaRegistro(i)`; A função de teste ficou pela responsabilidade da Beatriz, na qual seleciona registros aleatórios do arquivo bin conforme especificado pelo usuário, ou seja, o usuário define a quantidade N de registros que serão exibidos, verificando ainda se o

usuário prefere apenas uma exibição em tela ou que esses arquivos sejam salvos num arquivo texto, conforme especificado pela função `testeImportacao()`;

Decisões de implementação

Vale destacar primordialmente que, apesar das tarefas terem sido especificadas para cada integrante, grande parte do trabalho foi feita em conjunto, utilizando a ferramenta Live Share no Visual Studio Code da Microsoft. Essa ferramenta permitiu a colaboração em equipe em tempo real, que na maioria das vezes, entrávamos e desenvolvíamos de forma colaborativa.

A leitura do arquivo '.csv' foi realizada linha a linha visto que um dos registros (*review_text*) tinha um comportamento diferente dos outros. Os registros eram divididos por vírgula e, quando estivesse no último (*posted_date*), havia uma quebra de linha obrigatória. Na leitura do *review_text*, houve uma diferença para os outros registros onde o mesmo poderia possuir vírgula ou quebra de linha no meio do texto. Caso houvesse vírgula ou quebra de linha, o registro iniciava com " e terminava com ". Sendo assim, foi tratado adequadamente na leitura utilizando as funções *get()* e *peek()*. Com isso, a maneira mais prática encontrada foi a manipulação do arquivo linha a linha para uma melhor organização da separação dos dados.

Para o armazenamento, foi utilizado um vetor do tipo *Data.h* pois facilitaria o acesso dos dados, possui tamanho variável e seria de fácil remoção dos mesmos após a sua utilização. Essa estrutura foi utilizada também para salvar o arquivo binário, fazendo a conversão dos elementos para char. Além disso, o vetor teve seus dados deletados após o uso com a função *clear()*.

Para a leitura binária, é necessário um índice no qual serve para posicionar no arquivo onde irá iniciar a leitura. Os dados foram alocados em variáveis (char e int) e a função retorna um objeto do tipo *Data.h* que é utilizado em comum para as funções de acessar registro, teste importação (saída em console ou em arquivo .txt) e a função teste importação de N registros (também é possível a saída em console ou em arquivo.txt).

Como executar o programa

Especificado e detalhado no arquivo "README.md".

<https://github.com/beatrizcunharo/estrutura-de-dados-2>

Referências

- [1] BANSAL, Shivam. 3.5 Million Tiktok Mobile App Reviews (2021). Disponível em: <https://www.kaggle.com/shivamb/35-million-tiktok-mobile-app-reviews> (acessado em 25 de outubro de 2021).
- [2] Ferramenta utilizada para colaboração em equipe. Disponível em: <https://visualstudio.microsoft.com/pt-br/services/live-share/>