

Nome: Laura Sarto
Nº de matrícula: 202120154
Nome: Dominique Antunes
Nº de matrícula: 202020377

Relatório Técnico

1. Descrição das Estruturas de Dados Utilizadas

Struct Atributos

A estrutura Atributos é definida para representar os registros lidos do arquivo CSV e armazenados no arquivo binário. Contém vários campos de caracteres (char) para armazenar diferentes tipos de informações: "id", "latitude", "longitude", "descricao", "zip", "title", "timeStamp", "twp", "addr" e "e". Cada campo é um vetor de caracteres com um tamanho específico adequado para os dados que irá armazenar. O construtor inicializa todos os campos com o caractere nulo ('\0'), garantindo que não haja dados residuais.

Streams de Arquivo

fstream é usado para operações de leitura e escrita em arquivos (ios::in | ios::out) de forma a manipular os arquivos CSV e binários (call911_1.csv, call911_1.bin, f1.bin, f2.bin, s1.bin, s2.bin).

2. Descrição de Alto Nível dos Programas

ManipulaArquivoBinario.cpp (Manipulação de Arquivo Binário)

Este programa fornece um menu interativo para manipular um arquivo binário (call911_1.bin). As opções incluem adicionar, ler, alterar e trocar registros, imprimir registros e ordenar o arquivo. O arquivo binário é manipulado usando operações de leitura e escrita em posições específicas, identificadas por deslocamentos calculados com base no tamanho dos registros (sizeof(Atributos)). A opção de ordenação do arquivo inicia um processo de ordenação externa, envolvendo a divisão e a fusão de arquivos auxiliares.

OrdenadorArquivoBinario.cpp (Ordenação Externa)

Este código implementa uma ordenação externa, uma técnica utilizada para ordenar dados que não cabem na memória principal. Inicia-se o processo de merge sort, dividindo o arquivo principal em dois arquivos (f1.bin e f2.bin), realiza a ordenação intercalada desses arquivos em blocos de tamanho crescente, e finalmente mescla os arquivos ordenados de volta ao arquivo original (call911_1.bin). A ordenação é feita com base em critérios definidos (como "descricao" e "id" dos registros). Este processo envolve várias leituras e escritas em arquivos temporários (s1.bin e s2.bin).

ConversorCSVParaBinario.cpp (Conversão de CSV para Binário)

Este código lê um arquivo CSV (call911_1.csv) e converte os registros em formato binário, armazenando-os em call911_1.bin. Utiliza a estrutura Atributos para mapear os campos do CSV para um formato binário. Cada linha do CSV é lida, dividida em campos individuais, convertida para a estrutura Atributos, e então escrita no arquivo binário. Mantém a contagem do número de registros e grava essa contagem no início do arquivo binário.

Atributos.h (Armazena a Struct)

Código utilizado para armazenar um trecho de código comum a todos os programas anteriormente citados, a estrutura de dados dos arquivos com os campos definidos. É chamado nos códigos anteriores através de #include.

Considerações Finais

Os programas em conjunto fornecem um sistema para manipular e ordenar dados de um arquivo CSV em um formato binário, o que pode ser útil para processamento e análise de grandes volumes de dados. O uso de arquivos temporários para a ordenação externa é uma técnica chave para lidar com restrições de memória.