



Universidade do Minho
Licenciatura em Engenharia Informática

Relatório Fase 1 – Grupo 7
Ano Letivo 2022/2023

Vicente Castro, A91677
Gabriel Silva, A97363
Maurício Pereira, A95338

Conteúdo

1. Introdução	3
2. Fase 1	3
2.1 Enunciado e explicação	3
2.2 Planeamento e estratégias a usar	4
2.3 Principais desafios	4
2.4 Queries	5

1. Introdução

No âmbito da unidade curricular Laboratórios de Informática III, o presente relatório tem como objetivo apresentar os métodos de trabalho usados na Fase 1 do projeto e expor certos problemas que possam ter condicionado o desempenho do programa e, além disso, também os procedimentos e técnicas usadas para colmatar esses mesmos defeitos.

Deste modo, esta primeira fase tinha como meta criar um programa na linguagem C que lesse determinados ficheiros (formato CSV), fizesse o parsing dos dados guardando-os em “catálogos” e ainda a possibilidade da realização de algumas queries sobre os mesmos.

2. Fase 1

2.1 Enunciado e explicação

Nesta primeira fase podemos dividir a estrutura do programa em algumas partes.

Uma primeira onde trata do parsing dos dados, onde é realizada a leitura e tratamento dos três ficheiros de entrada, e modulação dos dados que guarda todos os dados provenientes dos ficheiros fornecidos.

Temos também a parte onde criamos as hashtable e a parte do batch onde tivemos alguns obstáculos, posteriormente referidos no ponto 2.3.

E ainda a implementação das queries e as suas funções auxiliares correspondentes.

2.2 Planeamento e estratégias a usar

Após uma análise sobre a arquitetura da aplicação fornecida no enunciado do trabalho prático, decidimos usar uma abordagem relativamente simples.

Criámos uma estrutura de dados (Driver, Ride, User) para cada ficheiro de entrada. À medida que líamos o ficheiro, linha-a-linha, íamos fazendo o parsing para uma dessas estruturas e adicionávamos a uma hashtable correspondente.

No fim, teríamos três hashtables (users,drivers,rides) uma para cada ficheiro para um melhor armazenamento dos dados e para, mais tarde, um melhor acesso na hora da realização das queries.

Todas as estruturas de dados usadas neste trabalho ou foram criadas por nós ou pertencem à biblioteca GLIB, nomeadamente Hash Tables.

2.3 Principais desafios

Sendo um projeto de programação em larga escala, foi necessário recorrer a princípios mais avançados de programação, como o uso de estruturas de dados eficientes para armazenar e consultar grandes quantidades de informação, garantindo sempre o encapsulamento dos dados, no nosso caso nomeadamente as Hash Tables.

O maior desafio, numa fase inicial, foi encontrar a melhor solução para a criação das hashtables, especificamente, qual o “value” que guardávamos na qual optamos por guardar um gpointer para uma struct.

Depois fomos tendo pequenos problemas e dificuldades como por exemplo na parte do batch, na criação de variáveis para alocar as path, não estávamos a perceber o porquê do erro “double free”, surgiram alguns problemas como uso excessivo de memória, na qual resolvemos com a criação de variáveis com memória estática, e otimização do código, acreditamos que hajam maneiras de tornar o código mais eficiente e menos pesado para a memória do computador assim como para o CPU mas foram as soluções que encontramos e estamos satisfeitos com elas.

2.4 Queries

Nesta primeira fase foi nos dada a possibilidade de escolhermos apenas três das nove queries para implementar, das quais optamos as que nós consideramos de mais fácil implementação, visto que queríamos focar a nossa atenção mais no encapsulamento e modulação dos dados.

Optamos então por implementar as queries 1, 4 e 5.