Relatório 2º projecto ASA 2020/2021

Grupo: al086

Aluno(s): Allan Fernandes (97281)

Descrição do Problema e da Solução

Pelo enunciado, o problema a resolver consiste na modelação de uma árvore genealógica, tal que dado dois vertices P1 e P2 correspondentes a pessoas, o número de vértices n e arestas m, e igualmente uma sequência de m linhas correspondentes as arestas, seja possível obter o/os ancestral/is comum mais próximos entre P1 e P2, caso este exista. A solução desenvolvida foi implementada em c++, a árvore genealógica em questão foi tida como uma DAG que satisfaz um conjunto de propriedades adicionais, nomedamente: cada vértice só pode ter 0, 1 ou 2 progenitores, max(indegree) = 2.

A solução desenvolvida consiste em inicialmente verificar se a árvore fornecida (representada internamente como uma lista de adjacências) corresponde a uma árvore genealógica válida, tal foi feito, recorrendo primeiramente a um conjunto de DFS, com a diferença de que, caso se encontre um vértice marcado a **GRAY**, assume-se que um ciclo foi descoberto e altera-se o valor associado a uma flag, que posteriormente é verificado. Caso esta primeira verifcação não se assente, é feita uma segunda verificação, a qual procura-se encontrar o primeiro vértice que não respeita a restrição de max(indegree).

Quanto ao algoritmo desenvolvido a ideia passa por efetuar uma DFS ao vértice P1, tomando em conta lista de adjacências transposta, nesta fase todos os seus filhos (ancestrais na lista original) são marcados a **BLACK**, numa segunda fase é verificado se o vértice P2 encontra-se marcado a **BLACK**, caso esteja, temos já em mão a resposta (um vértice também pode ser visto como o seu ancestral mais próximo). Caso contrário efetua-se uma DFS ao vértice P2, caso se encontre um vértice previamente marcado com **BLACK**, altera-se a sua cor para **ORANGE**, e o vértice é adicionado a um set, a saída da segunda DFS marca-se os seus filhos não partilhados (não em comum com P1) a **RED**, numa terceira fase, já com os ancestrais comuns todos, é perguntado a cada um deles(vértices) em ciclo se possui qualquer filho no set, caso sim, o vértice é removido.

Análise Teórica

Nesta análise teórica, considera-se V como o número de vértices da árvore genealógica, E o número de arcos associados a mesma, N número de elementos no set, N << V.

A solução supracitada consiste nos seguintes passos:

- Leitura dos dados de entrada: simples leitura do input - usando o std::cin com o processo de sincronização a falso, com um ciclo a depender linearmente de E. Logo **Θ(E).**
- Verifcação da presença de ciclos: - usando uma DFS modificada. **O(E).**
- Verificação de indegree (caso o ponto anterior admita a não existência de ciclos): **O(V+ E)**
- Aplicação duma DFS ao vértice P1: **O(V + E)**
- Aplicação duma DFS ao vértice P2: **O(V + E)**

Relatório 2º projecto ASA 2020/2021

Grupo: al086

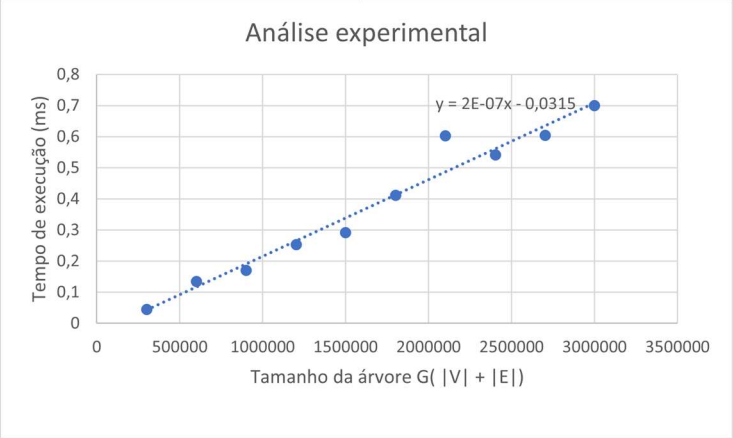
Aluno(s): Allan Fernandes (97281)

- Retirada de vertices cujos filhos estão presentes no set: um ciclo que depende do número de elementos do set (passível de variar ao longo tempo) **O(N)**, chamadas ao std::set::count, com complexidade de **O(logN),** retirada de elemento cujo filho foi detetado, normalmente **O(1)**, caso o balanço interno do set for destruído é preciso efetuar balanceamentos, estes associados a um fator de complexidade adicional de **log(N)** + std::distance(first,last). Logo, **O(NlogN**), assumindo que o balanço interno não for destruído.
- Apresentação dos dados. **O(1)**

- Complexidade global da solução: Θ(E) + O(E) + O(V+ E) + O(V + E) + O(V + E) + O(NlogN) + O(1) = O(V + E).

Avaliação Experimental dos Resultados

Com o intuito de validar os resultados teóricos acima apresentados e a a partir do gerador de instâncias fornecido foram geradas 10 “árvores genealógicas”, de tamanho incremental, com|V| de 10^5até 10^6, cujo tempo de execução associado foi cronometrado.



É possível observar o caráter linear do gráfico resultante, portanto, confirma-se assim a veracidade da análise teórica, isto é a relação linear entre o tempo de execução e o tamanho da árvore gerada.