

Trabalho 01

Estrutura de Dados
Prof. Roberto Cabral
10 de Outubro de 2017

1. Problema

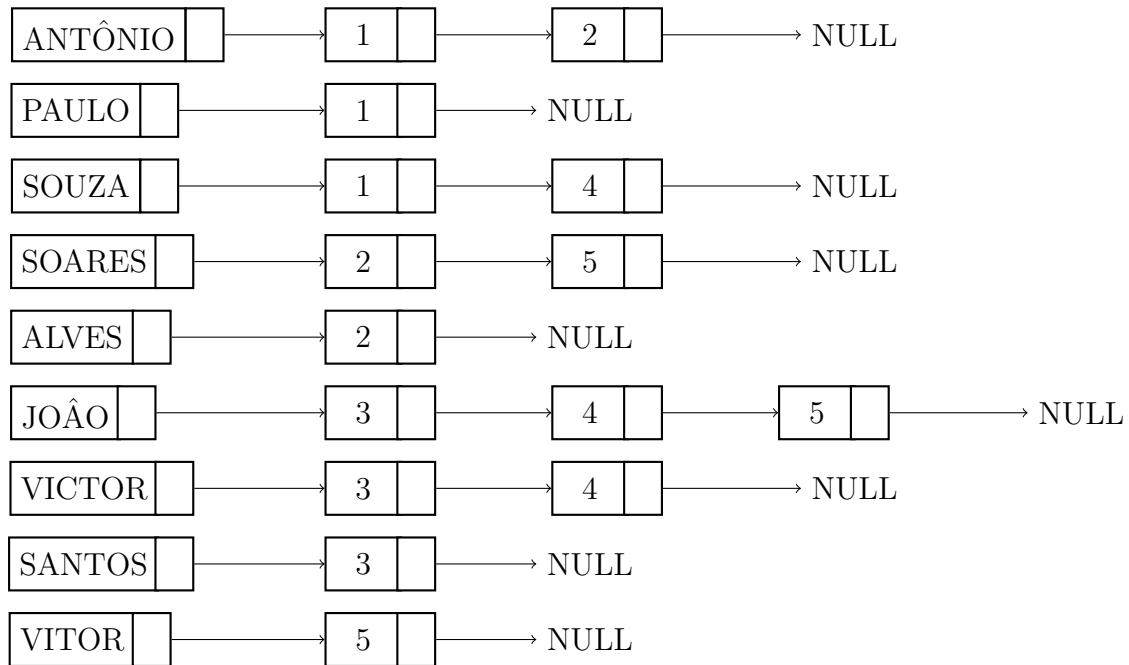
Nesta atividade, deve ser implementada uma pequena parte de um sistema de consulta de documentos antigos. Este sistema deve processar, dentre outras coisas, informações relacionadas com o registro de pessoas em livros antigos e permitir que sejam realizadas consultas que indicam em quais páginas de um dado livro um determinado nome (i.e., nomes, pré-nomes, sobrenomes) aparece.

Para otimizar as consultas, você deve criar um índice que relaciona todos nomes (i.e., nomes, pré-nomes, sobrenomes) encontrados em um livro com as páginas em que eles aparecem. Nesta atividade, este índice deve ser implementado utilizando uma lista invertida.

Em termos práticos, uma lista invertida pode ser vista como uma de coleção de listas encadeadas que, por sua vez, armazenam as ocorrências (e.g., número da página do livro) de cada termo indexado (e.g., nome).

| Página | Nome |
|--------|----------------------|
| 1 | ANTÔNIO PAULO SOUZA |
| 2 | ANTÔNIO SOARES ALVES |
| 3 | JOÃO VICTOR SANTOS |
| 4 | JOÃO VICTOR SOUZA |
| 5 | JOÃO VITOR SOARES |

Como exemplo considere o caso ilustrado na tabela com 5 nomes completos ocorrendo nas páginas 1, 2, 3, 4 e 5 de um livro qualquer. O índice, neste caso, é composto pelas 6 (seis) listas seguintes (uma para cada nome encontrado):



Observe que é simples realizar consultas de ocorrências de um conjunto de termos utilizando a lista invertida. Para tanto, deve-se selecionar as listas de todos os termos envolvidos na consulta e calcular a interseção destas listas. Para responder uma consulta com os termos “João Victor”, por exemplo, deve-se calcular a intersecção entre as listas dos termos “João” e “Victor”. Em nosso exemplo, esta suposta consulta revelaria que os termos “João Victor” aparecem nas páginas três e quatro do livro. Observe que lista resultante desta interseção tem dois elementos com referência para a terceira e quarta página.

2. **Entrada** A primeira linha da entrada contém um número n que indica o número de nomes completos encontrados no livro. Em cada uma das n linhas seguintes, há um número inteiro p e uma cadeia de caracteres T que representam, respectivamente, um número de página e um nome completo. Em seguida, são especificadas as consultas que devem ser realizadas. A primeira linha a seguir contém um número inteiro k que indica o número de consultas a serem realizadas. Em cada uma das k linhas seguintes, há uma cadeia de caracteres Q com os termos a serem consultados.

Exemplo:

```
7
13 Laura Costa Azevedo
11 Gabrielly Rodrigues Almeida
7 Júlia Lima Gomes
3 Nicolas Barbosa Costa
5 Mateus Souza Santos
17 Júlia Rocha Barros
31 Leonardo Costa Almeida
5
Júlia
Julia
Gabrielly Almeida
Costa
Leonardo Barros
```

Para fins de implementação, você pode assumir que todas as cadeias de caracteres consideradas nesta atividade tem no máximo 100 caracteres.

3. **Saída** Sua solução deve interpretar as consultas e exibir as saídas conforme elas são produzidas.

Exemplo:

```
7 e 17.
Não há ocorrência.
11.
3, 13 e 31.
Não há ocorrência.
```

Observe que se o termo consultado não for encontrado no livro, então a mensagem “Não há ocorrência” deve ser apresentada.

4. Organização do Código

Seu trabalho deve ser dividido da maneira seguinte:

- `lista.h` que deve conter as assinaturas de funções que implementam o TAD lista encadeada.
- `lista.c` que deve conter as implementações das funções descritas no arquivo `lista.h`.
- `indice.h` que deve conter as assinaturas de funções que implementam o TAD lista invertida em memória primária.
- `indice.c` que deve conter as implementações das funções descritas no arquivo `indice.h`.

- `main.c` que deve tratar a entrada, realizar chamadas de funções contidas em `lista.c` e `indice.c` e escrever os dados de saída.

5. Informações adicionais:

- Deverá ser submetido, juntamente com o código, um relatório técnico explicando tudo que foi feito no trabalho. Dentre outras coisas, o relatório deve ter:
 - Uma descrição do trabalho realizado, contendo a especificação completa das estruturas de dados utilizadas;
 - Uma seção descrevendo como o trabalho foi dividido entre as duplas;
 - Um pequeno manual que explique o funcionamento do sistema;
 - Uma seção de dificuldades encontrada.
- O trabalho deverá ser feito em duplas (**as notas serão individuais**);
- O trabalho deverá ser entregue até o dia **22 de Outubro**;
- A apresentação do trabalho será feita em horário definido pelo professor.

Obs.: Qualquer indício de plágio resultará em nota ZERO para todos os envolvidos.