#### Trabalho Final EDA

Upload via Moodle Entrevista conforme divulgado no Moodle

- EM DUPLAS (NÃO SERÃO ACEITOS TRABALHOS COM GRUPOS COM MAIS DE 2 MEMBROS)
- TODOS OS ARQUIVOS DEVERÃO SE ENTREGUES (H, C, PRJ, EXE, ETC.) BEM COMO UM PEQUENO RELATÓRIO IDENTIFICANDO OS AUTORES, A FUNCIONALIDADE DO SISTEMA, COMANDOS DE COMPILAÇÃO DE BIBLIOTECAS UTILIZADAS E COMENTÁRIOS SOBRE OS EXPERIMENTOS.
- OS TRABALHOS ENTREGUES FORA DO PRAZO TERÃO NOTA ZERO.
  OS CÓDIGOS DEVEM SER COMENTADOS.

**Tarefa:** construção de um programa de aplicação que lê um arquivo no modo texto, sem formatação e grava as palavras em uma ABB-AVL. A ABB-AVL servirá como estrutura de índice para a pesquisa de palavras.

#### 1) Descrição da árvore

Utilize uma ABB-AVL.

Cada nó da árvore armazena um descritor para uma lista de palavras. A lista conterá uma palavra ou todas as ocorrências da mesma palavra em locais distintos do texto (Figura 1).

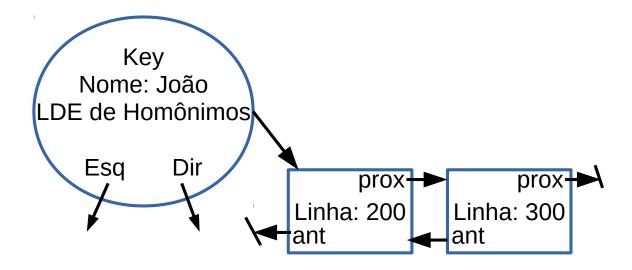


Figura 1: Nó da ABB-AVL e lista associada.

Arquivo de teste a ser utilizado		
100	João	
300	João	

### 1.1) Arquivo de dados

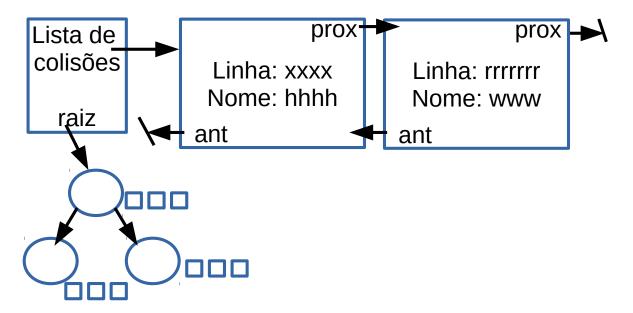
O arquivo para testes (arq.txt) está disponível no Moodle, logo abaixo do link de upload da tarefa.

#### 1.2) Ordenação da ABB-AVL

O valor da chave (uma cadeia de caracteres representando uma palavra) que ordena a árvore Pode ser calculada da seguinte forma:

Considere a string "abcd" cada caractere ocupa uma posição str[0]=='a', str[1]=='b', etc... Calcule o valor da chave somando os produtos entre cada código ascii do caractere e seu respectivo índice na string mais um: chave= int(str[0])\*1 + int(str[1])\*2+ int(str[2])\*3 + int(str[3])\*4

Palavras iguais fornecem o mesmo valor para a chave, porém, podem ocorrer colisões. Uma forma de tratar as colisões é criar uma segunda lista para conter esses casos. O descritor deve referenciar essa lista.



- 2) Funcionalidade do Sistema
  - 2.1) carga da ABB/Listas a partir do arquivo-txt;
  - 2.2) destruição da ABB/Listas;
  - 2.2) A busca/exibição da(s) ocorrência(s) de determinada palavra: o usuário fornece a palavra e o sistema informa a(s) linha(s) e coluna(s) da(s) ocorrência(s) do primeiro caractere da palavra no texto. A busca é realizada na ABB;
  - 2.3) \*OPCIONAL\* Remoção da(s) ocorrência(s) de determinada palavra: O usuário informa a tupla (palavra, linha, coluna) e o TDA utiliza o devido callback na ABB para detectar a posição da palavra procurada. Caso seja encontrada a ABB, a mesma é removida desta e da tabela. Caso haja uma única palavra, também remove-se o respectivo nó da ABB e da lista.
  - 2.4) \***OPCIONAL**\*: Atualização em disco: a remoção, busca, alteração ou inserção ocorrem na memória principal, portanto, não estão sendo efetivadas no arquivo. Para tanto o usuário precisará acionar a opção de atualização do arquivo txt fonte/original:

Cada palavra lida no arquivo é procurada na ABB, se existir tal palavra na árvore, a mesma é copiada do arquivo original para um *arqTemp.txt*. Ao final, usando a função *rename*, renomeie o *arquivo.txt* como *arquivoOLD.txt*, e após isso, renomeie *arqTemp.txt* como *arquivo.txt*.

2.5) \***OPCIONAL**\*: Alteração e inserção de nova palavra: O usuário informa a tupla (*palavra, linha, coluna*) e esta é procurada na ABB. Caso a posição já está ocupada por outra palavra, o usuário recebe msg e opta entre substituir a palavra existente, pela nova palavra (alteração), inserir a nova palavra na posição da atual sem removê-la, ou nada fazer.

## 3) Recomendações finais:

- Mantenha o encapsulamento entre aplicação e estrutura de dados (ABB/LDDE);
- No material disponível no site (Moodle) consta uma apostila com funções para leitura/escrita em arquivos "modo texto". Em especial há funções bastante úteis: fgets/fputs, fseek, rename, etc. Além de strcat, strcmp, etc;
- 3. Foque seus esforços nos objetivos do sistema, deixe os detalhes estéticos para o final. Não se preocupe com interfaces "bonitas" ou códigos extremamente rebuscados, mas sim com um projeto bem estruturado e que funcione;

- Reaproveite código tirando proveito dos códigos e bibliotecas de TDAs disponíveis (Árvore e LDDE). Altere esses códigos se achar conveniente;
- 5. <u>Em nenhuma hipótese</u> faça uso de variáveis globais! Use e abuse de funções da biblioteca padrão do C, mas tenha certeza que conhece o seu comportamento. Quaisquer dúvidas adicionais procure o professor ou o monitor da disciplina.

# 5) Fontes Bibliográficas

Moodle da disciplina;

http://www.lcad.icmc.usp.br/~nonato/ED/AVL/node67.html