

SSC0603 – Estrutura de Dados – Trabalho 3: E no final... A árvore decide. v1.1

Usando dados fictícios e reais, você deverá montar uma estrutura que possa auxiliar uma pessoa em tomadas de decisão. As situações poderão variar desde a classificação de tipos de flores a prognósticos de câncer em mulheres. Neste trabalho, você implementará uma árvore de decisão e resolverá alguns problemas reais, seguindo os critérios descritos abaixo:

- 1) A árvore de decisão deverá ser baseada em uma árvore binária, fazendo uso de ponteiros para conectar os nós.
- 2) Cada nó **não-folha** deve representar uma decisão a ser avaliada pela árvore. As decisões serão representadas pelo código abaixo:

TXXX “semântica_do_nó_entre_aspas” valor_teste nó_pai lado_da_conexão

- a) **T** representa o tipo do nó não-folha, podendo ser as opções:
 - i) **B** para verificação binária (se a entrada for **verdadeira**, a descida é feita pela esquerda, do contrário, pela direita).
 - ii) **N** para verificação numérica (se a entrada for **menor** que **valor_teste**, a descida é feita pela esquerda, do contrário, pela direita).
- b) **XXX** representa o índice do nó (e.g.: **N000** é o nó numérico, de índice 000), sendo **SEMPRE** três dígitos.
- c) “semântica_do_nó_entre_aspas” informa o sentido da avaliação, estando sempre entre aspas, com código 34 na tabela ASCII, sem caracteres da tabela ASCII estendida (e.g.: **N000** “salario anual” é o nó numérico, de índice 000, que irá avaliar o salário anual da pessoa).
- d) **valor_teste** representa o valor que será testado. Este campo existirá apenas para nós **numéricos**, sendo um número de ponto flutuante (e.g.: **N000** “salario anual” **36000.0** é o nó numérico, de índice 000, que irá avaliar se o salário anual da pessoa é menor que R\$ 36.000,00).
- e) **nó_pai** é uma entrada em formato **TXXX** que informará quem é o nó pai (e.g.: **N000** “salario anual” **36000.0** **B001** é o nó numérico, de índice 000, que irá avaliar

se o salário anual da pessoa é menor que R\$ 36.000,00, tendo como pai o nó B001). Caso o nó atual seja a raiz, este valor será representado por **R000**.

- f) **lado_da_conexão** é um caractere com possibilidades ‘E’ ou ‘D’ (esquerda ou direita) informará de qual lado do nó pai, o nó atual será inserido (e.g.: **N000** “salario anual” 36000.0 B001 E é o nó numérico, de índice 000, que irá avaliar se o salário anual da pessoa é menor que R\$ 36.000,00, sendo filho esquerdo do nó B001). Caso o nó seja o nó raiz, este valor será inexistente.

- 3) Cada nó **folha** deve representar uma classificação gerada pela árvore. As classificações serão representadas por um código, sendo:

CXXX “resultado_da_árvore_entre_aspas” nó_pai **lado_da_conexão** **nível_de_certeza**

- a) **C** (classe) é o código do nó folha.
- b) **XXX** representa o índice do nó (e.g.: **C000** é o nó de classificação, de índice 000), sendo **SEMPRE** três dígitos.
- c) “resultado_da_árvore_entre_aspas” fala qual foi a decisão da árvore, estando sempre entre aspas, com código 34 na tabela ASCII, sem caracteres da tabela ASCII estendida (e.g.: **C000** “nao liberar emprestimo” é o nó de classificação, de índice 000, que irá informar ao usuário para não liberar o empréstimo).
- d) nó_pai é uma entrada em formato **TXXX** que informará quem é o nó pai (e.g.: **C000** “nao liberar emprestimo” B000 é o nó de classificação, de índice 000, que irá informar ao usuário para não liberar o empréstimo, sendo filho do nó B000). Este campo é sempre presente, já que um nó de classificação não pode ser raiz.
- e) **lado_da_conexão** é um caractere com possibilidades ‘E’ ou ‘D’ (esquerda ou direita) informará de qual lado do nó pai, o nó atual será inserido (e.g.: **C000** “nao liberar emprestimo” B000 E é o nó de classificação, de índice 000, que irá informar ao usuário para não liberar o empréstimo, sendo filho esquerdo do nó B000).
- f) **nível_de_certeza** é um valor em ponto flutuante que representa a certeza da classificação. Este valor é gerado na etapa de treino, contabilizando a quantidade de classificações geradas pelo nós e considerando a quantidade de erros gerados pela classificação indevida (e.g.: **C000** “nao liberar emprestimo” B000 E **112.5** é o nó de classificação, de índice 000, que irá informar ao usuário para não liberar o empréstimo, sendo filho esquerdo do nó B000, com grau de certeza 112.5). Quanto maior este valor, maior é a certeza da decisão.

- 4) A montagem da árvore dar-se-á pela quantidade de nós da árvore e pelos códigos de cada um dos nós. Veja o exemplo na próxima página.
- 5) Após a montagem da árvore, o programa será testado com n conjuntos de entradas, exibindo cada uma das saídas, separados por uma quebra de linha. O padrão de entrada será:
 - a) Quantidade de entradas.
 - b) Sequência de caracteres que informarão se a entrada existe ou não.
 - i) 'E' para existe a informação, seguido de:
 - (1) 'V' ou 'F' para nós binários.
 - (2) Número de ponto flutuante para nós numéricos.
 - ii) 'N' para não existe.
 - iii) Assim que a execução encontrar o caractere 'N', o programa deverá imprimir a classificação de maior certeza, a partir do nó atual. Recomenda-se usar uma busca recursiva para encontrar o nó de maior certeza.
- 6) A saída do programa, para cada uma das entradas, será a descrição da classe, seguida por uma quebra de linha.
- 7) Qualquer caminho seguido pela árvore levará a alguma classificação.
- 8) Todas as regiões de memória alocadas dinamicamente **DEVEM** ser liberadas antes do encerramento da execução.
- 9) **COMENTAR O CÓDIGO!!!**

Exemplo:

Entrada:

7 // Quantidade de linhas que descrevem a árvore de decisão.

B000 "E estudante?" R000

C000 "Nao liberar o emprestimo." B000 E 258.08

B001 "Trabalha?" B000 D

C001 "Nao liberar o emprestimo." B001 D 235.4

N000 "Qual e o salario anual?" 24000.0 B001 E

C002 "Liberar ate R\$ 10.000,00." N000 E 115.85

C003 "Liberar ate R\$ 15.000,00." N000 D 42.31

4 // Quantidade de entradas de teste.

E V // A pessoa é estudante.

E F E F // A pessoa não é estudante e não trabalha.

E F E V E 32000.0 // A pessoa não é estudante, trabalha e ganha R\$ 32.000,00 por ano.

E F N // Só sabemos que a pessoa não é estudante.

Saída:

Nao liberar emprestimo. // Decisão para estudantes.

Nao liberar emprestimo. // Decisão para não estudantes sem trabalho.

Liberar ate R\$ 15.000,00. // Decisão para não estudantes com trabalho, que ganham salário anual maior ou igual a R\$ 24.000,00.

Nao liberar emprestimo. // Como apenas sabemos que a pessoa não é estudante, apresenta-se a classe de maior certeza: C001.

A estrutura do exemplo terá o seguinte formato:

