UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE INSTITUTO METRÓPOLE DIGITAL - IMD CURSO DE BACHARELADO EM TECNOLOGIA DA INFORMAÇÃO

IMPLEMENTAÇÃO DE ÁRVORE BINÁRIA DE BUSCA

Paulo Vitor Lima Borges
Dante Augusto Bezerra Pinto
Lucas Pires de Souza Marcolino

Natal-RN 2022

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE INSTITUTO METRÓPOLE DIGITAL - IMD CURSO DE BACHARELADO EM TECNOLOGIA DA INFORMAÇÃO

IMPLEMENTAÇÃO DE ÁRVORE BINÁRIA DE BUSCA

Paulo Vitor Lima Borges

Dante Augusto Bezerra Pinto

Lucas Pires de Souza Marcolino

Relatório técnico para a matéria de Estruturas de Dados Básicas II, descrevendo a implementação de uma Árvore Binária de Busca com operações a mais.

Sumário

1	INTRODUÇÃO	3
2	METODOLOGIA	4
2.1	Materiais utilizados	4
2.1.1	Ferramentas de programação	4
2.1.2	Árvore binária de busca	5
2.1.3	Métodos	6
2.1.3.1	Busca	6
2.1.3.2	Inserção	7
2.1.3.3	Remoção	8
2.1.3.4	Enésimo elemento	9
2.1.3.5	Posição	10
2.1.3.6	Mediana	11
2.1.3.7	Media	11
2.1.3.8	Écheia	13
2.1.3.9	É completa	14
2.1.3.10	Pré-ordem	14
2.1.3.11	Imprimir Árvore	15
3	ANÁLISES DE COMPLEXIDADES	17
3.0.1	Busca	17
3.0.2	Inserção	17
3.0.3	Remoção	17
3.0.4	Enésimo Elemento	17
3.0.5	Posição	17
3.0.6	Mediana	17
3.0.7	Média	18
3.0.8	É cheia	18
3.0.9	É completa	18
3.0.10	Pré-Ordem	18
3.0.11	Imprimir Árvore	18
4	CONCLUSÃO	10

1 Introdução

O presente trabalho tem como objetivo descrever a implementação de uma árvore binária de busca que guarde valores inteiros e tenha como métodos os convencionais, os de busca, inserção e remoção, e os métodos enunciados no anexo do trabalho.

Para tal implementação, o projeto foi desenvolvido com a linguagem de programação C++ criado a partir de uma classe para melhor visualização desse conceito abstrato. O próximo tópico irá mostrar mais detalhes de como isso foi realizado.

Por fim, todas as funções tiveram suas complexidades analisadas de acordo com sua implementação no presente trabalho.

2 Metodologia

2.1 Materiais utilizados

2.1.1 Ferramentas de programação

A implementação da Árvore Binária de Busca e as implementações de seus respectivos métodos foram feitos em C++ na versão 17.

Tanto no ambiente WSL/Windows quanto no GNU/Linux, os códigos foram compilados usando G++, através do CMake.

Para isso, foi usado o seguinte CMakeLists.txt:

Para compilação, foram usadas as seguintes linhas de comando:

cmake -S source -B build

cmake -build build

Já para execução, foi executada a seguinte linha:

./build/ABB <arquivo de descrição da árvore> <arquivo dos métodos>

2.1.2 Árvore binária de busca

Para implementar a árvore em si, a seguinte classe foi implementada:

```
#define ARVORE_H
    #include <stack>
10 using std::string;
12 class abb {
           abb* esq = nullptr;
           abb* dir = nullptr;
           int tamanho_esq = 0;
          int tamanho_dir = 0;
int valor;
           int altura = 1;
           abb();
           abb(int valor, abb* esq = nullptr, abb* dir = nullptr) : valor(valor), esq(esq) , dir(dir) {}
           int getValor();
          std::optional<abb*> busca(int x);
           bool inserir(int x);
            bool remover(int x);
            abb* remover(int x, abb* sub);
            std::optional<int> enesimoElemento(int n);
            std::optional<int> posicao(int x);
            std::optional<int> mediana();
            std::optional<double> media(int x);
            bool ehCheia();
            bool ehCompleta();
            string pre_ordem();
            void imprimeArvore(int s);
           void formato1(int qnt_tabs, int espaco, abb *no);
            void formato2(abb *no);
            void calcularAltura(abb* raiz);
            int qnt_nos();
```

Figura 1 – Árvore Binária de Busca

Analisando os atributos, vemos que esq e dir representam as sub-árvores à esquerda e à direita, tamanho_esq e tamanho_dir representam as quantidades de nós das mesmas, valor é o

que está armazenado no nó em questão, podemos dizer que o nó atual é a raíz da árvore, e altura, que é medida em relação a árvore toda.

2.1.3 Métodos

2.1.3.1 Busca

Esse método recebe como parâmetro o valor que queremos encontrar na árvore e retorna um ponteiro do nó em que o valor se encontra ou um valor vazio, caso não tenha um nó com esse valor. Assim, foi feito o seguinte código:

```
1 std::optional<abb*> abb::busca(int x){
2    if(x == valor)return this;
3    if(altura = 1){
4        return {};
5    }else{
6        if(valor < x){
7            if(tamanho_dir == 0) return {};
8            else return dir->busca(x);
9    }else{
10            if(tamanho_esq == 0) return {};
11            else return esq->busca(x);
12    }
13    }
14 }
```

Figura 2 – Método Busca

Dessa forma, note que esse método foi definido recursivamente e o caso base é quando encontramos o valor. Se não encontramos, checamos se o nó atual é um nó folha, o que implicaria que o valor não existe na árvore pois não está no nó atual e não há mais para onde descer. Caso não seja nó folha, é checado quem é menor, o valor do nó ou o que desejamos encontrar, e com base nisso fazemos a chamada recursiva na sub-árvore da esquerda, caso o valor do nó seja maior, ou na direita, caso seja menor, se a tal existir. Caso a sub-árvore em que devemos descer não exista, o valor não exite na árvore logo o retorno será vazio.

2.1.3.2 Inserção

```
bool abb::inserir(int x) {
       if(dir == nullptr) {
            tamanho dir = 1;
           inseriu = true;
        } else {
            inseriu = dir->inserir(x);
            tamanho_dir = dir->qnt_nos();
   if(valor > x) {
       if(esq == nullptr) {
           esq = new abb(x);
           tamanho_esq = 1;
            inseriu = esq->inserir(x);
            tamanho esq = esq->qnt nos();
    int alt_dir = 0;
   if(esq != nullptr) alt_esq = esq->altura;
    if(dir != nullptr) alt_dir = dir->altura;
    altura = 1 + std::max(alt esq, alt dir);
    return inseriu;
```

Figura 3 – Método Inserir

O método de inserção tem como objetivo inserir um elemento em sua devida posição (uma folha), mantendo as propriedades da árvore binária de busca. Para tanto, foi feito um procedimento similar ao de busca na árvore. Porém, ao chegar numa folha, é criado um novo nó à esquerda ou à direita, a depender do valor do nó a ser inserido.

O procedimento começa verificando se o nó da subárvore atual é maior ou menor que o elemento que se deseja inserir: caso seja maior, devemos inseri-lo na subávore à esquerda, caso seja menor, à direita. Caso ocorra que o valor a ser inserido já exista na árvore, o procedimento retorna o valor booleano "falso".

Para inserir o valor em uma subárvore, verificados se esta é nula à esquerda ou à direita, a depender do caso.

Além do trabalho de inserir, o método também trata de atualizar os atributos altura de cada nó e o tamanho de suas duas subárvore. Para isso, após inserir o valor, tais valores são calculados novamente.

2.1.3.3 Remoção

```
bool abb::remover(int x) {
   int nos_antes = qnt_nos();
   remover(x, this);
   return qnt_nos() < nos_antes;
}</pre>
```

Figura 4 – Método Principal de Remoção

O método de remoção tem como objetivo remover o nó que tem como chave aquele passado para o método. Para isso, foi feita uma função de remoção auxiliar, que recebe o valor a ser removido, a subárvore da qual se deseja remover e retorna tal subárvore já com o elemento removido. Portanto, o método remover apenas chama o auxiliar, passando a raiz da árvore para este.

```
abb* abb::remover(int x, abb* sub) {
    if(sub == nullptr) {
        return nullptr;
    }

    if(x < sub->valor) {
        sub->esq = remover(x, sub->esq);
    } else if(x > sub->valor) {
        sub->oir = remover(x, sub->dir);
    } else {
        if(sub->esq == nullptr) {
            return sub->dir;
        } else if(sub->oir == nullptr) {
            return sub->esq;
        } else {
        abb* novo_no = sub->esq;
        white(novo_no->dir != nullptr) {
            novo_no = novo_no->dir;
        }
        sub->valor = novo_no->valor;
        sub->esq = remover(sub->valor, sub->esq);
    }

int alt_esq = 0;
    int alt_dir = 0;
    if(sub->esq != nullptr) alt_esq = sub->esq->altura;
    if(sub->esq != nullptr) sub->tamanho_esq = sub->esq->qnt_nos();
    if(sub->esq != nullptr) sub->tamanho_dir = sub->dir->qnt_nos();
    if(sub->eir != nullptr) sub->tamanho_dir = sub->dir->qnt_nos();
    if(sub->dir != nullptr) sub->tamanho_dir = sub->dir->qnt_nos();
    return sub;
}
```

Figura 5 – Método Remover Auxiliar

O método auxiliar funciona da seguinte forma: Caso a subárvore da qual se deseja remover seja nula, o método retorna nulo.

Caso o valor que se deseja remover seja menor que a raíz da árvore, devemos remover da subárvore à esquerda, logo esta recebe o resultado da chamada recursiva de remover o valor de si mesma. Caso o valor seja maior, fazemos o mesmo procedimento para a subárvore à direita.

Ao chegarmos no nó que devemos remover, temos 3 casos:

- O nó é uma folha: nesse caso, apenas retornamos o valor null, pois ao remover este nó, a árvore resultante é vazia.
- O nó possui 1 subárvore: nesse caso, apenas retornamos a outra subárvore, uma vez que a árvore resultante após remover o nó em questão se trata do único nó que restará.
- O nó possui 2 subárvores: nesse caso, devemos procurar um nó para substituir a raiz que será removida. Na implementação do presente trabalho, tal nó é o mais à direita da subárvore à esquerda, que garante a manuntenção das propriedades da ABB.

2.1.3.4 Enésimo elemento

Esse método recebe um número inteiro e retorna o valor do nó que está naquela posição em ordem simétrica. Se o número for além da quantidade de nós ou não positivo, o retorno é vazio. Segue a imagem do código:

```
std::optional<int> abb::enesimoElemento(int n){
    if(n < 1 || n > (1 + tamanho_esq + tamanho_dir)) return
    {};
    if(n == tamanho_esq + 1)return valor;
    else{
        if(n <= tamanho_esq) return esq->enesimoElemento(n);
        else{
            return dir->enesimoElemento(n - tamanho_esq - 1);
        }
    }
}
```

Figura 6 – Método Enésimo Elemento

Como primeira condição, verificamos se o parametro é negativo ou se é maior que a quantidade geral de nós, pois em qualquer desses casos o retorno é vazio. A seguir, checamos se a posição é igual ao tamanho da sub-árvore a esqueda mais 1, pois essa é a posição do nó atual em ordem simétrica. Se for igual, encontramos o nosso retorno, se não devemos checar outros casos. Se a posição for menor ou igual que a quantidade de elementos da sub-árvore esquerda, então devemos chamar a recursão lá. Caso contrario, devemos chamar a recursão na sub-árvore direita mas a posição agora será diminuida uma quantidade igual a quantos nós a sub-árvore a esquerda tem mais um, que é o próprio nó que estamos. A subtração ocorre para casos como o seguinte: o quinto elemento da arvore binária de busca 1, 2, 3, 4, 5 é o mesmo que o segundo elemento da sub-árvore 4, 5.

2.1.3.5 Posição

Esse método recebe um valor e tem como retorno a posição em ordem simétrica do nó com aquele valor ou retorna vazio se a árvore não tiver um nó com aquele valor. Dessa forma, temos que o código é o seguinte:

Figura 7 – Método de posição

Note que esse código é muito semelhante ao de busca, pois o que está sendo feito é realmente buscar o nó com aquele valor. O diferencial seria o retorno, que é o tamanho da sub-árvore a esquerda mais 1, que é a posição do nó na ordem simétrica, como já foi dito. Com isso, ao chamarmos a recursão para a sub-árvore direita, nós primeiro checamos se ela tem valor, pois pode ser que não tenhamos encontrado o nó, e se tiver é necessário adicionar àquele valor a quantidade de nós da sub-árvore a esquerda mais 1 pois essa é a quantidade de elementos que antecedem o parametro na ordem simétrica.

2.1.3.6 Mediana

```
1 std::optional<int> abb::mediana(){
2    return enesimoElemento(ceil((double)(1+this->tamanho_esq+this->tamanho_dir)/(double)2));
3 }
```

Figura 8 – Método da mediana

O método da mediana em nossa Árvore Binária de Busca tinha como objetivo encontrar o elemento que está na metade do caminho do percorrimento em ordem simétrica. Pela própria descrição do problema, pode-se inferir o uso de um método implementado também pela árvore chamado enesimoElemento, que retorna o elemento presente na posição passada como parâmetro.

Decido parte do algoritmo, agora precisamos determinar em que parte estaria presente o elemento que representa na mediana. Para tal, basta sabermos quantos nós estão presentes na árvore e pegar a metade, arredondando pra cima, e passando esse número como parâmetro do método enesimoElemento.

2.1.3.7 Media

```
std::optional<double> abb::media(int x){
    auto aux = busca(x);
    if(aux.has_value()){
        abb* raiz{aux.value()};
        std::stack<abb*> s;
        int valor{0};
        int quantidade{1+raiz->tamanho_esq+raiz->tamanho_dir};
        s.push(raiz);
        while(not s.empty()){
   abb* atual{s.top()};
            s.pop();
            if(atual->dir != nullptr){
                s.push(atual->dir);
            valor += atual->valor;
            if(atual->esq != nullptr){
                s.push(atual->esq);
        return ((double)valor/(double)quantidade);
        return {};
```

Figura 9 – Método da média

O método da média em nossa Árvore Binária de Busca tinha como objetivo encontrar a média aritmética da subárvore a partir da chave - número inteiro - passada como parâmetro. Para isso, inicialmente precisamos fazer uma busca pela chave passada como parâmetro. Caso a chave

não seja encontrada, nada é retornado. Caso contrário, iremos inicialmente pegar a quantidade de nós que será contabilizado, somando a subárvore da esquerda, a da direita e o nó raiz. Após isso, percorreremos toda a árvore pegando seus valores e adicionando à uma variável que está guardando todos os valores dos nós que estamos passando ao longo do percurso. Ao final, basta retornar a divisão entre a soma resultante e a quantidade de nós contada previamente.

2.1.3.8 É cheia

O seguinte método é usado no é cheia e no é completa:

```
1 int abb::qnt_nos() {
2    return tamanho_esq + tamanho_dir + 1;
3 }
```

Figura 10 – Método da quantidade de nós

Agora, vamos para o método é cheia.

```
1 bool abb::ehCheia(){
2    return qnt_nos() == pow(2, altura) - 1;
3 }
```

Figura 11 – Método ehCheia

O método da ehCheia em nossa Árvore Binária de Busca tinha como objetivo dizer se uma árvore é cheia ou não. Para isso, temos que saber se a quantidade de nós na árvore é máxima para a altura do nó raiz. Para isso, inicialmente vamos saber quantos nós têm em cada nível de uma árvore. Por ser uma árvore binária, cada nível possui uma quantidade 2^n de nós, sendo n o nível da árvore.

Somando a quantidade de nós que há em cada nível temos 2^h-1 nós. Portanto, basta retornar se a quantidade de nós da árvore - somando a raiz e a quantidade de nós em cada subárvore - é igual à 2^h-1 .

2.1.3.9 É completa

```
bool abb::ehCompleta(){
   return qnt_nos() >= pow(2, altura - 1);
}
```

Figura 12 – Método ehCompleta

O método da ehCompleta em nossa Árvore Binária de Busca tinha como objetivo dizer se uma árvore é completa ou não. Para isso, temos que saber se a árvore é cheia até o penúltimo nível. Para isso, basta retornar se a quantidade de nós da árvore é maior ou igual a 2^{h-1} visto que se fosse igual, estaria com a quantidade máxima de nós até o penúltimo nível e com um nó apenas no último nível e, caso fosse maior, o limite seria de $2^h - 1$, sendo uma árvore cheia.

2.1.3.10 Pré-ordem

```
1 string abb::pre_ordem() {
2    string ordem = "";
3
4    std::stack<abb> pilha;
5    pilha.push(*this);
7    while(!pilha.empty()) {
9         abb topo = pilha.top(); pilha.pop();
10
11         ordem += std::to_string(topo.valor) + " ";
12
13         if(topo.dir != nullptr) pilha.push(*topo.dir);
14         if(topo.esq != nullptr) pilha.push(*topo.esq);
15    }
16
17    return ordem;
18 }
```

Figura 13 – Método de Pré-Ordem

Este método tem como objetivo criar uma string que contém os valores de todos os nós da subárvore em Pré-Ordem. Para isso, foi feito o percursso na árvore de forma iterativa, usando uma Pilha. Ao remover o topo da pilha, o valor deste é concatenado à uma string que será o retorno do método.

2.1.3.11 Imprimir Árvore

```
void abb::imprimeArvore(int s) {
   if(s == 1) {
      formatol(0, 36, this);
   } else {
      formato2(this);
   }
}
```

Figura 14 – Método de Imprimir Ávore

Este método tem como objetivo imprimir a árvore de forma hierárquica no formato informado à função. Tais formatos foram dados previamente pela descrição do projeto.

Caso o formato requerido seja o 1, o seguinte procedimento é executado:

```
void abb::formatol(int qnt_tabs, int espaco, abb *no) {
   std::cout << string(qnt_tabs, '\t');
   std::cout.width(espaco);
   std::cout.fill('-');
   std::cout << std::left << no->valor << std::endl;

if(no->esq != nullptr) {
   formatol(qnt_tabs + 1, espaco - 6, no->esq);
   }

if(no->dir != nullptr) {
   formatol(qnt_tabs + 1, espaco - 6, no->dir);
   }
}
```

Figura 15 – Formato de Impressão 1

Com as informações de quantas tabulações devem ser impressas e quantos espaços a linha de cada nó deve ocupar, o método primeira mente imprime tais informações, junto com o valor do nó atual, e logo após, caso haja subárvore à esquerda e à direita, o método é chamado com uma tabulação a mais e ocupando 6 caracteres de espaço a menos. Tal método executa uma pré-ordem.

Caso o formato requerido seja o 2, o seguinte procedimento é executado:

```
void abb::formato2(abb *no) {
   if(no != this) std::cout << " ";
   std::cout << "(";
   std::cout << no->valor;

   if(no->esq != nullptr) {
      formato2(no->esq);
   }
   if(no->dir != nullptr) {
      formato2(no->dir);
   }

std::cout << ")";
}</pre>
```

Figura 16 – Formato de Impressão 2

Neste formato, apenas com a informação do nó atual, antes de cada nó, é aberto um parêntese. Caso haja subárvore à esquerda ou à direita, o método é chamado recursivamente para tais. Após imprimir todas as subárvores, o parêntese é fechado.

3 Análises de complexidades

3.0.1 Busca

Nesse método, o algoritmo desce na árvore a partir da raiz. Dessa forma, caso o valor que desejamos buscar não esteja na árvore ou seja um nó folha, o número de passos realizados estará em função da altura. Em outras situações, teríamos valores menores que a mesma, tendo como menor número de passos quando queremos buscar o valor da raiz, o que nos leva a concluir que o algoritmo é O(h).

3.0.2 Inserção

Similar ao método de busca, o método de inserção depende diretamente da altura da árvore. Uma vez que ao chegar em um nível desta, decidimos se vamos para esquerda ou direita, até chegarmos ao último nível. Portanto, a complexidade de melhor se dá por $\Omega(h)$ e pior caso de O(h), onde h é a altura da árvore. Logo sua complexidade é $\Theta(h)$.

3.0.3 Remoção

No melhor caso, a raiz da árvore é o elemento que temos que remover e esta não tem nenhum filho. Portanto, no melhor caso, temos que a complexidade é $\Omega(1)$. Já no pior caso, devemos encontrar o nó a ser removido e este é uma folha, portanto devemos descer toda a árvore, assim como na inserção. Logo seu pior caso é O(h).

3.0.4 Enésimo Elemento

Semelhante ao método de busca, a procura pelo nó na posição enviada realiza passos que são no máximo em função da altura, quando o nó é uma folha e o algoritmo desce na árvore. Já o melhor caso é quando encontramos o enésimo elemento já na raiz. Assim, a complexidade é O(h).

3.0.5 Posição

Mais um algoritmo com a complexidade semelhante ao de busca e afins, pois descemos pela árvore em busca de um valor específico. Logo, a complexidade também será O(h).

3.0.6 Mediana

O método da mediana calcula, primeiramente, qual seria o elemento do meio caso a árvore binária de busca fosse representada de forma unidimensional com percurso de ordem simétrica.

Ao calcular o índice que está a mediana em $\Theta(1)$, é chamada a função de enésimo elemento que possui complexidade O(h). Logo, a complexidade da mediana seria a mesma complexidade de enésimo elemento, O(h).

3.0.7 Média

O método da média primeiramente efetua uma busca para o nó que possui como chave o valor passado no parâmetro. Caso não encontre o valor, a complexidade da média seria a mesma da busca, O(h). Caso encontre o valor, o algoritmo percorrerá todos os elementos que pertencem à subárvore que possui como chave o valor passado pelo parâmetro. Logo, no pior caso, temos que a complexidade é O(n) se o valor passado seja a raiz da árvore.

3.0.8 É cheia

O método é cheia tem complexidade $\Theta(1)$, visto que apenas, ao chamar uma função que apenas efetua uma soma em $\Theta(1)$ e logo após isso compara com o valor máximo de nós na árvore que também possui complexidade $\Theta(1)$.

3.0.9 É completa

O método é completa é análogo ao método é cheia, tendo complexidade $\Theta(1)$.

3.0.10 Pré-Ordem

O método de pré-ordem foi feito de maneira iterativa. Para imprimir todos os nós da árvore nesta ordem específica, ainda temos de passar por todos eles, logo a complexidade de melhor e pior caso são $\Omega(n)$ e O(n) respectivamente, logo sua complexidade é $\Theta(n)$, onde n é a quantidade de nós da árvore.

3.0.11 Imprimir Árvore

Assim como na pré-ordem, os métodos de imprimir a árvore deve passar por todos os nós da árvore. E suas impressões (de cada nó) são feitas em tempo constante. Assim, suas complexidades de melhor e pior caso são $\Omega(n)$ e O(n), logo, $\Theta(n)$.

4 Conclusão

De maneira geral, foi possível observar que a estrutura de dados Árvore Binária de Busca de fato melhora o desempenho de algumas operações em relação às estruturas de dados lineares. Visto que o tempo de fazer buscas na estrutura, uma vez que respeitadas suas propriedades, crescem proporcionalmente à sua altura, não ao número de elementos. Assim, dependendo de como o usuário insere valores nas árvores, a árvore pode ter uma altura de $\log_2 n$, mostrando o quão melhor o desempenho será.

Também foi possível perceber que algumas das funções extras requeridas no trabalho derivam diretamente das básicas, o que ajuda em sua implementação.