INSTITUTO FEDERAL
RIO GRANDE DO SUL

Estrutura de Dados II

Estrutura de dados Hierárquicas: **Árvores**

```
#include <iostream>

using namespace std;

int main()
{
    cout << "Hello world!" << endl;
    return 0;
}
```

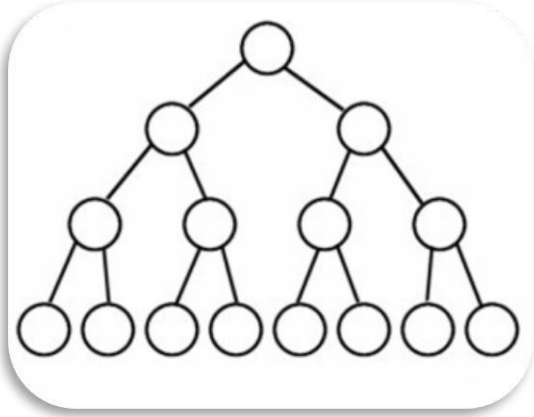
Árvores: Motivação

Em computação, há aplicações que necessitam de estruturas mais complexas que as estudadas até agora (lista, fila, pilha)

Um problema comum é como visualizar o conjunto de diretórios utilizando listas, pilhas e filas?

Estruturas de Árvores solucionam esse problema e além disso podem ser usados para modelar outros:

- Compressão de Dados
- Troca de Mensagens (ordenação)
- Comparações
- Pesquisa e ordenação de Dados.



Árvores: Motivação

Existem algoritmos eficientes para o tratamento de árvores.

Algumas árvores podem definir uma ordenação implícita, o que facilita a execução do algoritmo, com um tratamento condicional simples.

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores: Definição

Árvores são estruturas de dados que se caracterizam por uma organização hierárquica (**relação hierárquica**) entre seus elementos.

Essa organização permite a definição de algoritmos relativamente **simples, recursivos e de eficiência bastante razoável**.

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores: Definição

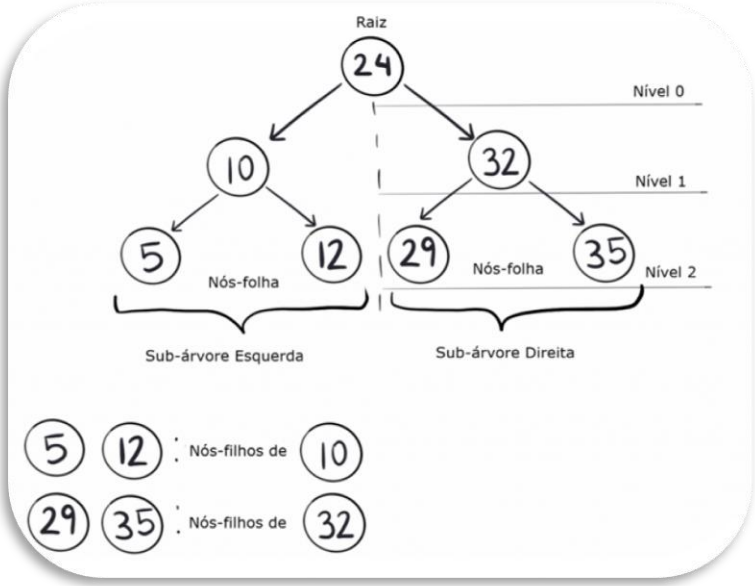
Uma árvore **N** é um conjunto finito de elementos denominados nós ou vértices tais que:

- **N = *0*** *é a árvore vazia*;
- Um único nó é uma árvore. **Este nó é raiz da árvore**.

Árvores: Subarvores

- Suponha que **N** é um nó e T_1, T_2, ..., T_k sejam árvores com raízes n_1, n_2, ... , n_k, respectivamente.

  Podemos construir uma nova árvore tornando **N** a raiz e T_1, T_2, ..., T_k **sejam subárvores da raiz**. Nós n_1, n_2, ..., n_k são chamados filhos do nó **N**.



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores: Caminho

- Um caminho de n_i a n_k, onde n_i é antecedente a n_k, é a sequencia de nós para se chegar de n_i a n_k.

- Se n_i é antecedente a n_k, n_k é descendente de n_i

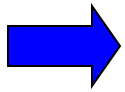
- O comprimento do caminho é o número de nós do caminho – 1.

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores: Representação

Uma Estrutura de Dados do tipo Árvore pode ser representada de forma:

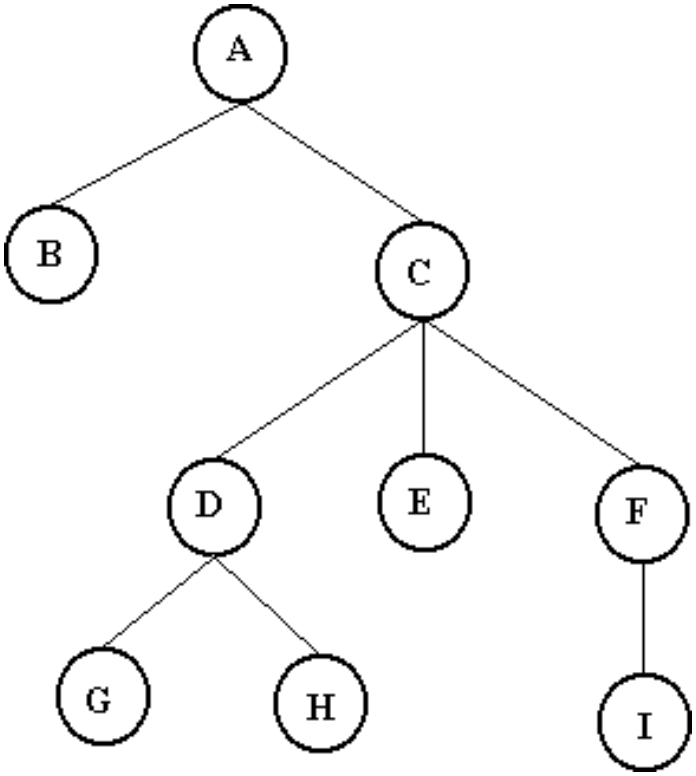
- Hierárquica
- Diagramas
- Expressão Parentetizada (parênteses aninhados)
- Expressão Não Parentetizada



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores: Representação

- **Hierárquica**

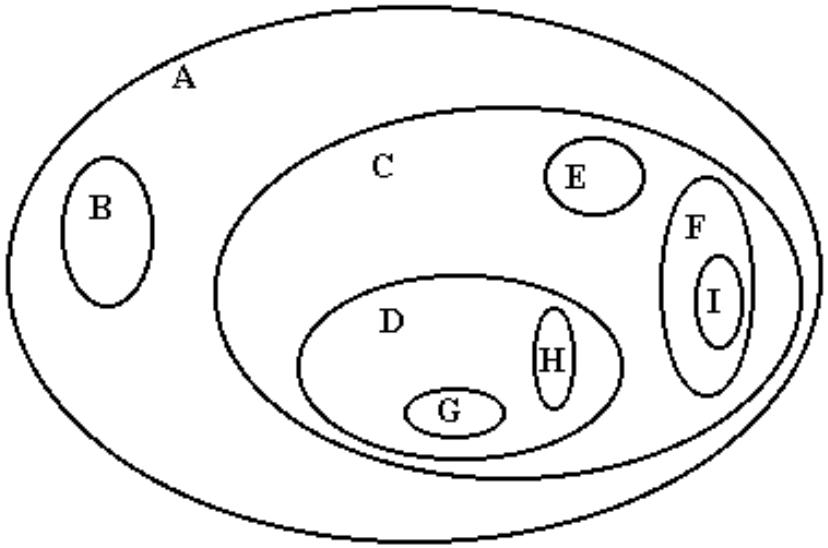


Árvores: Representação

- **Parênteses aninhados**

  (A (B) (C (D (G) (H)) (E) (F (I))))

- **Diagramas de inclusão**



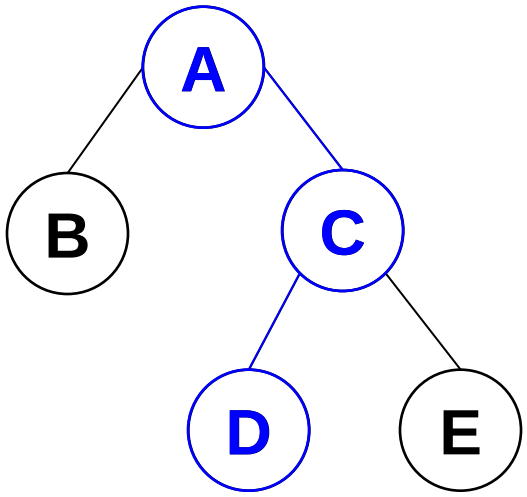
INSTITUTO FEDERAL
RIO GRANDE DO SUL

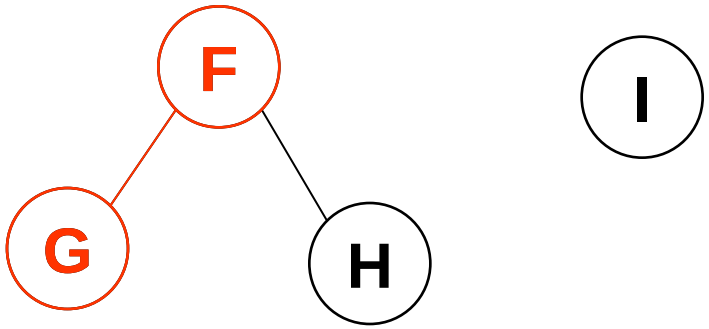
Árvore: Caminho

- **Caminho na árvore**
  - Seqüência de nós distintos v_1, v_2, ..., v_k, tal que existe sempre entre nós consecutivos (isto é, entre v_1 e v_2, entre v_2 e v_3, ... , $v_{(k-1)}$ e v_k) a relação "**é filho de" ou "é pai de**" .

- **Comprimento do caminho**
  - Um caminho que passa por v_k vértices é obtido pela seqüência de $_{k-1}$ pares. ***O valor $_{k-1}$ é o comprimento do caminho.***

Árvore: Caminho





Floresta T = T_A U T_F U T_I

- T_A = {A, B, C, D, E}
- T_F = {F, G, H}
- T_I = {I}

- **Caminhos:**
  - **Caminho1 = (A, C, D)**
  - **Caminho2 = (F, G)**

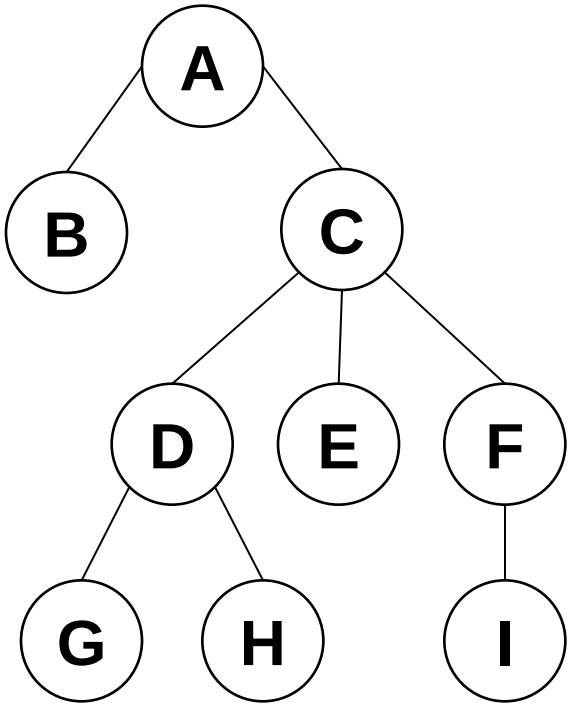
- **Comprimentos:**
  - **Comprimento (Caminho1) = 2**
  - **Comprimento (Caminho2) = 1**

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvore: Altura

- **Nível (ou profundidade) de um nó**
  - O nível ou profundidade de um nó é o número de nós do caminho da raiz até o nó (*raiz tem nível 1*)
- **Altura de um nó V**
  - Número de nós no maior caminho de v até um de seus descendentes (*folhas têm altura 1*)
- **Altura de uma árvore T ou h(T)**
  - Máximo nível de seus nós (a altura da sub-árvore de raiz v é representada por h(v))

Árvore: Altura



NÍVEIS	
A	1
B, C	2
D, E, F	3
G, H, I	4

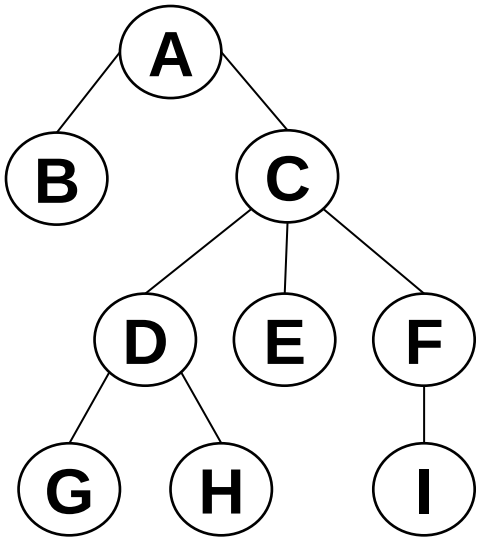
ALTURAS EM RELAÇÃO AO NÓ FOLHA COM MAIOR NÍVEL	
h(A)	4
h(B)	1
h(C)	3
h(D)	2
h(E)	1
h(F)	2
h(G)	1
h(H)	1
h(I)	1

h(T) = 4

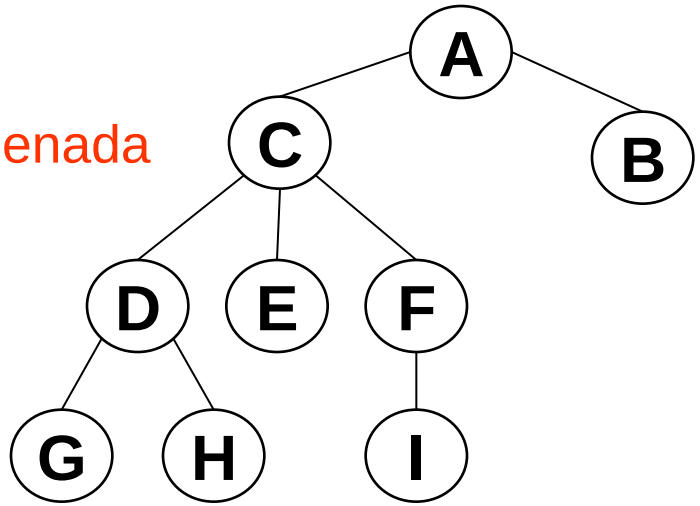
Árvore: Ordenada x Desornedada

Ordenada: Os filhos de cada nó estão ordenados (assume-se ordenação da esquerda para a direita)

ordenada



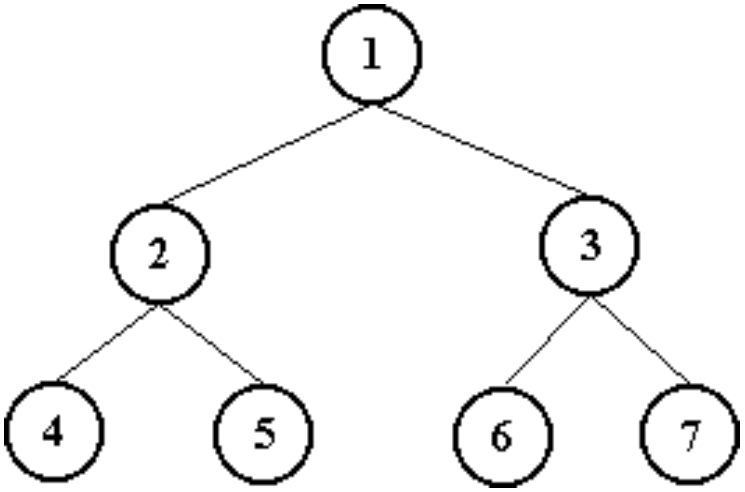
desordenada



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvore Cheia

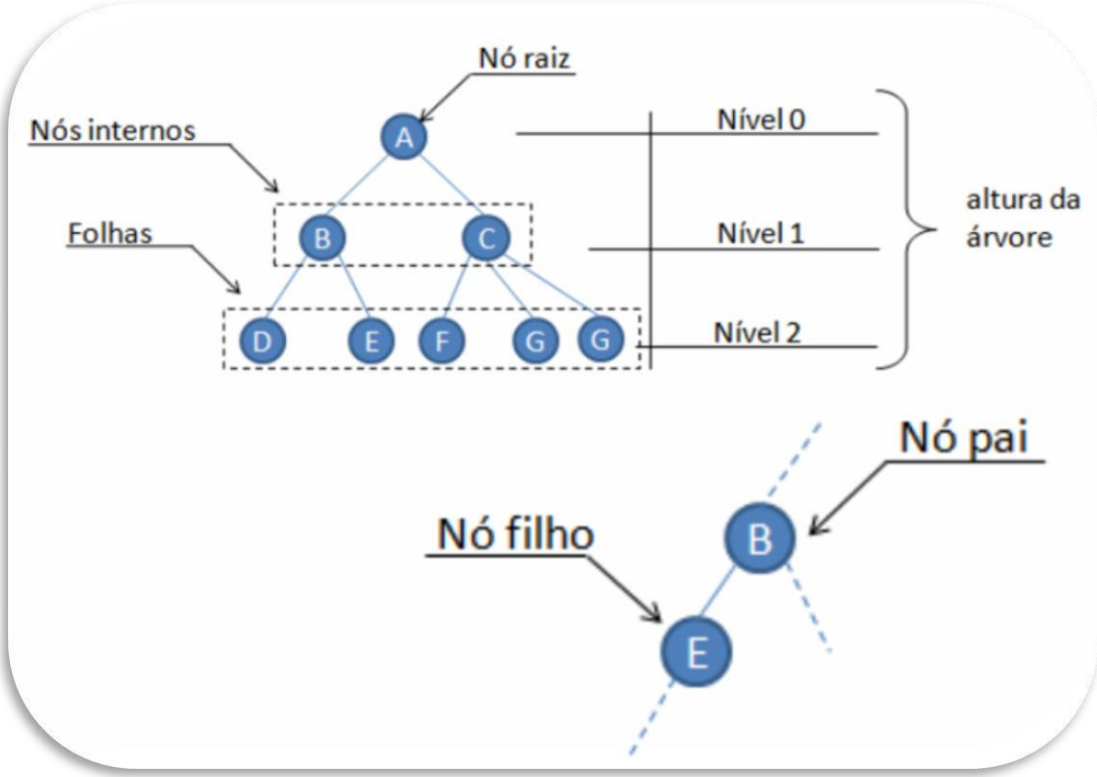
Uma árvore é cheia se possui o número máximo de nós, isto é, todos os nós tem número máximo de filhos exceto as folhas, e todas as folhas estão na mesma altura.



Árvore cheia de grau 2

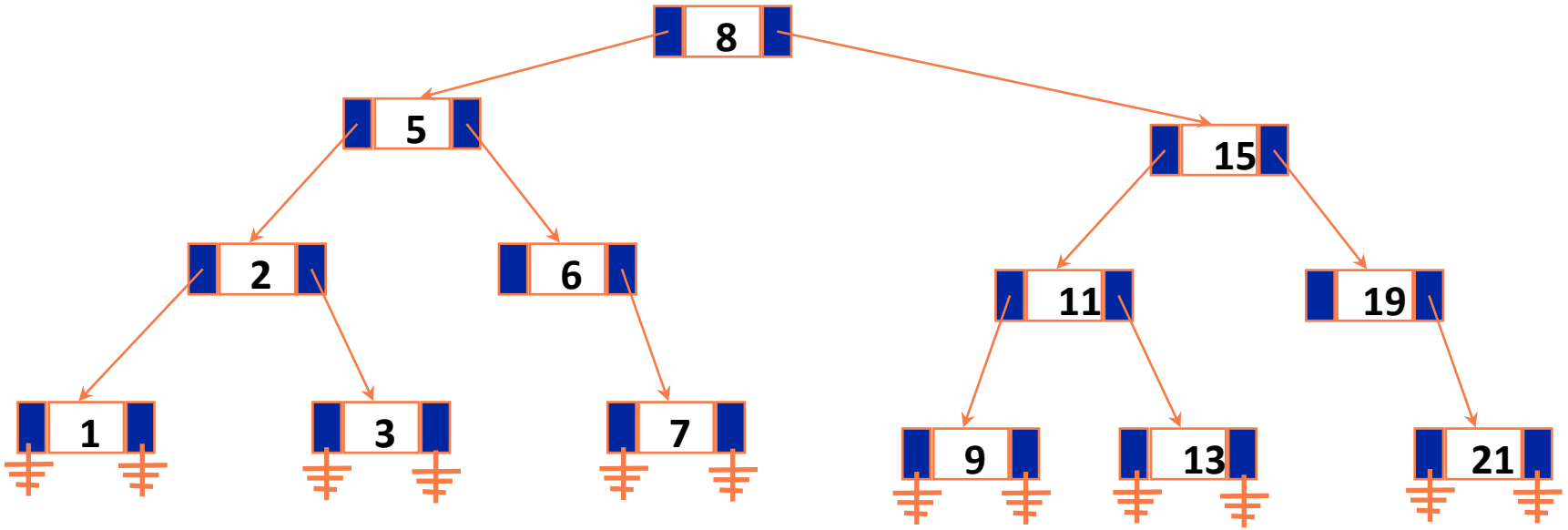
Classificação de Árvores

- Árvores Binárias de Busca
- Árvores Red Black
- Árvores AVL
- Árvores B



Árvores Binárias de Busca

Uma árvore que se encontra organizada de tal forma que, para cada nodo t_i, todas os valores da subárvore à esquerda de t_i são menores que (ou iguais a) t_i e à direita são maiores que t_i.



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores Binárias de Busca

Em uma árvore binária de busca é possível encontrar-se qualquer chave existente descendo-se pela árvore:

- Sempre à esquerda toda vez que a chave procurada for menor do que a chave do nodo visitado;
- Sempre à direita toda vez que for maior.

- A escolha da direção de busca só depende da chave que se procura e da chave que o nodo atual possui.

Observação: A busca de um elemento em uma árvore balanceada com n elementos toma tempo médio < log(n), sendo a busca então O(log n).

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores Binárias de Busca

Problema Deterioração:

- Quando inserimos utilizando a inserção simples, dependendo da distribuição de dados, pode haver deterioração;

- Árvores deterioradas perdem a característica de eficiência de busca.

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores AVL

Manter uma árvore binária de busca balanceada sob a presença de constantes inserções e deleções é ineficiente.

Para contornar esse problema foi criada a árvore AVL (**Adelson-Velskii e Landis**).

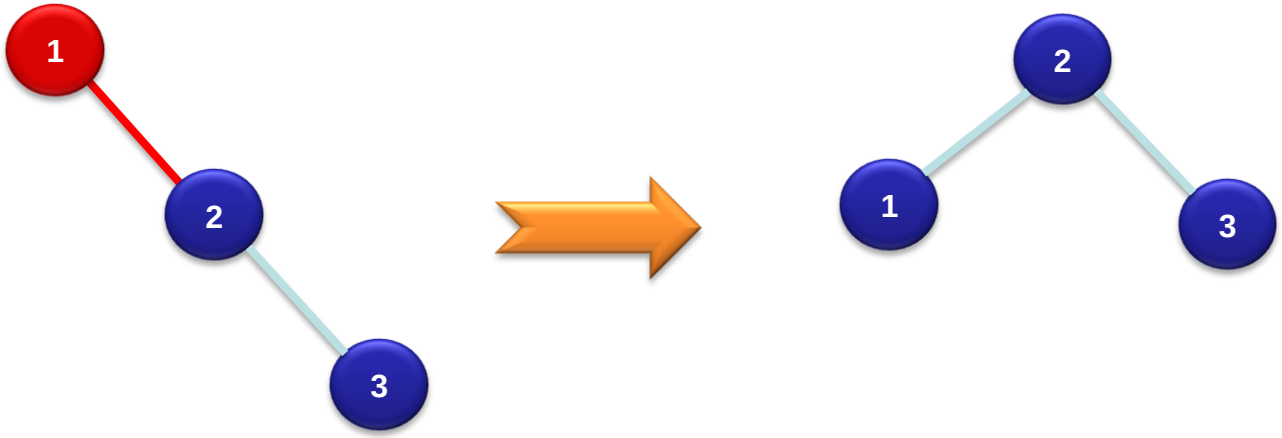
A árvore AVL é uma árvore binária com uma condição de balanço, porém não completamente balanceada.

Árvores AVL permitem inserção/deleção e rebalanceamento consideravelmente rápidos.

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores AVL

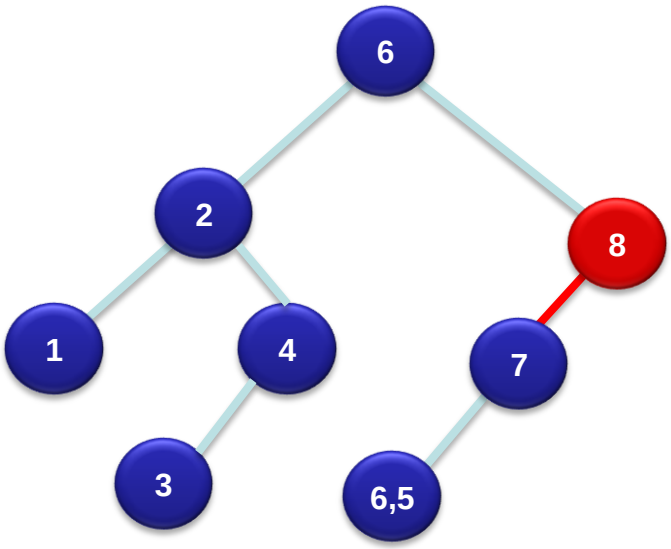
A árvore AVL é uma árvore binária com uma **condição de balanço**, porém não completamente balanceada.



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores AVL

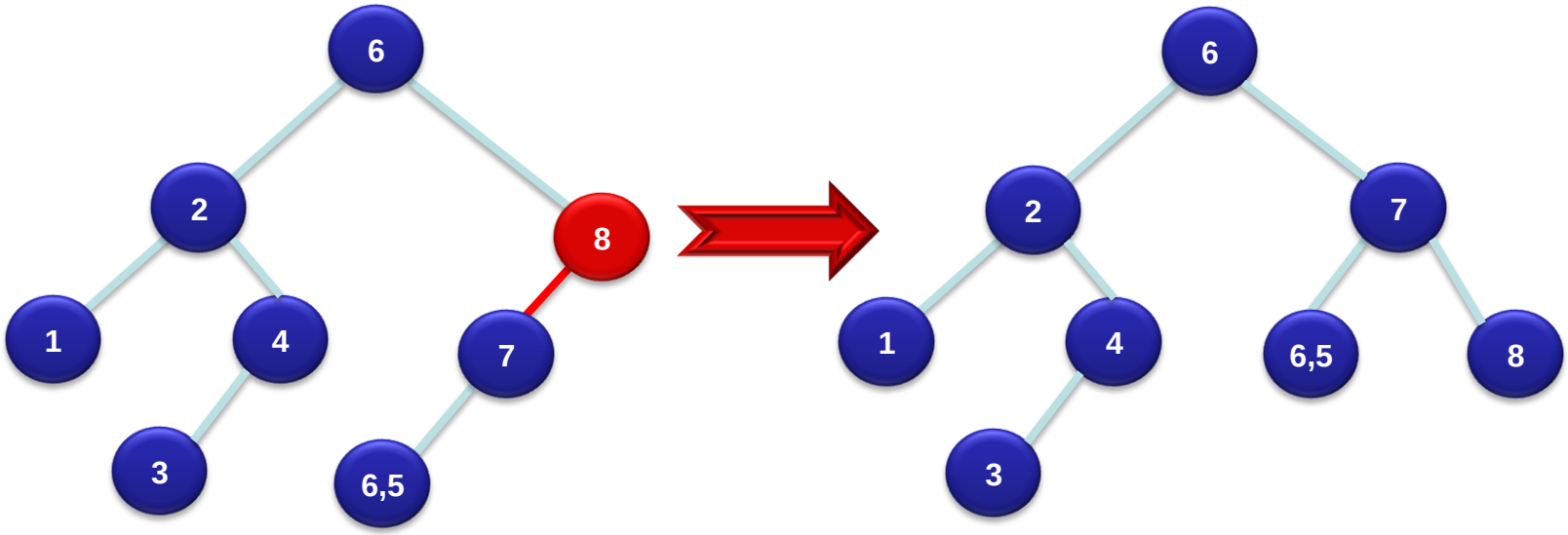
Dado um nodo qualquer, uma árvore está AVL-balanceada se as alturas das **duas subárvores deste nodo diferem de, no máximo, 1** **(Altura Esque. – Alt. Direit. < 2).**



Árvores AVL

Para rebalancear uma árvore após uma inserção, são utilizadas rotações de subárvores:

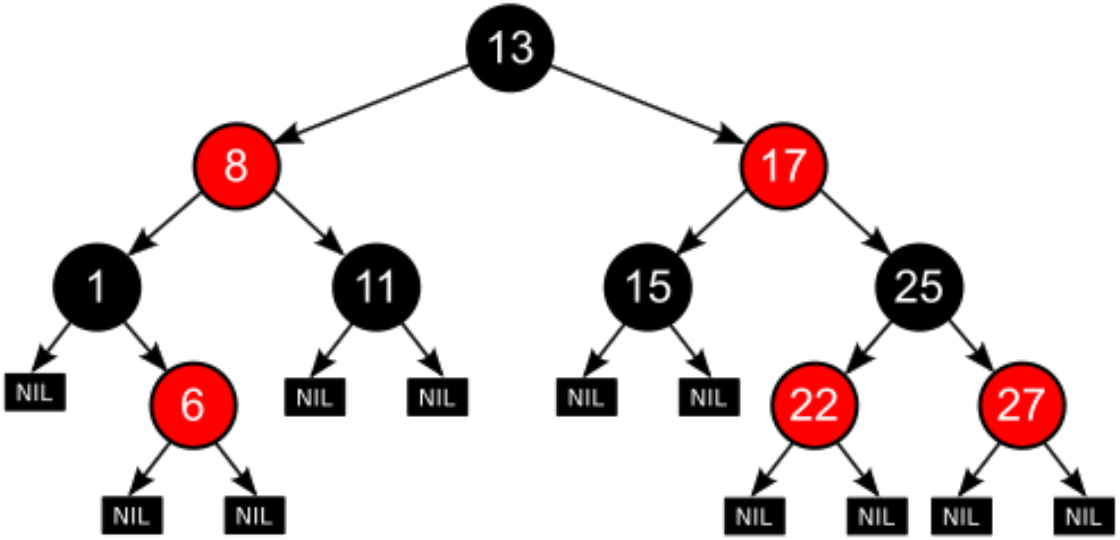
- rotações simples (normalmente).
- rotações duplas (em alguns casos).



Árvores Red Black

Árvores de pesquisa binária com um bit de informação adicional: **a cor**.

- Garante que nenhum caminho é mais que duas vezes mais longo que qualquer outro
- Árvore semibalanceada



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores Red Black

Caracteristicas:

- Todo nodo ou é rubro ou é negro
- A raiz é negra
- Toda folha é negra
  - **Folhas** são somente os nodos vazios (ponteiros nulos).
- Se um nodo for rubro, então ambos os filhos são negros.
- Para todo nodo, todos os caminhos até uma folha contém o mesmo número de nodos negros.

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores B

São árvores de pesquisa balanceadas **especialmente projetadas** para a pesquisa de informação em memória secundárias:

- minimizam o número de operações de movimentação de dados (escrita / leitura) numa pesquisa ou alteração;

Aumentam o número de ramificações na árvore diminuindo, assim, a altura .

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores B

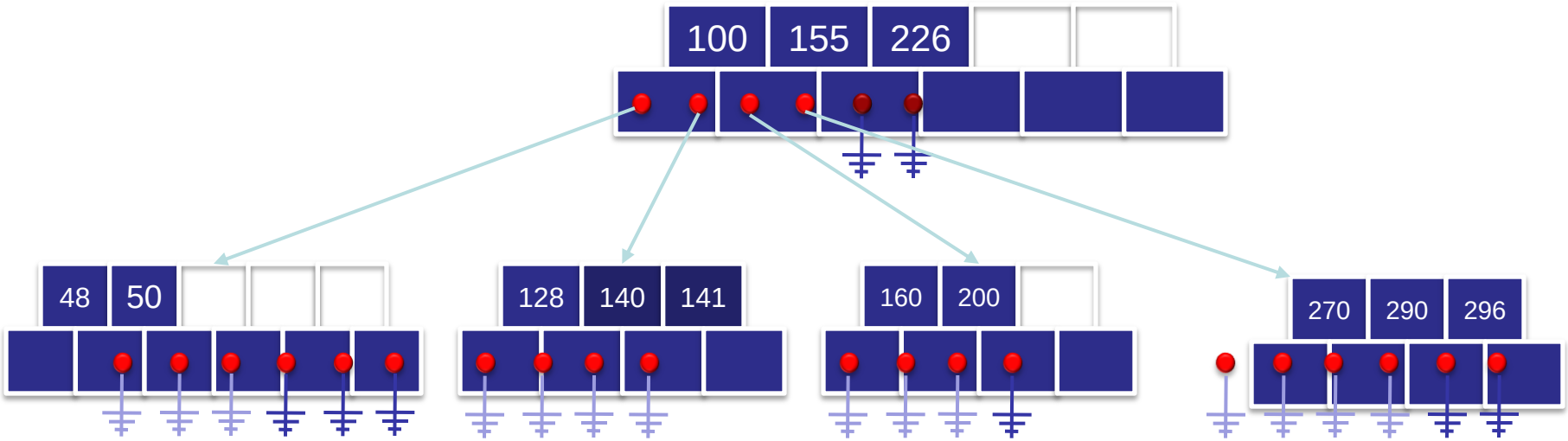
Uma árvore B possui as seguintes propriedades:

NChaves, Folha, P1, <K1, PR1>, P2,<K2,PR2>,...

- Descrição dos campos:
  - NChaves – número de chaves armazenadas no nodo;
  - Folha – flag indicando se o nodo é folha ou não;
  - P_i – ponteiro para o nodo filho i;
  - K_i – chaves;
  - PR_i – ponteiro para o elemento de dados na memória ou para um registro em disco onde se encontra a chave K_i;

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores B



Implementação

Assim como as listas lineares, estruturas de dados do tipo árvore podem ser implementadas:

- Através de arranjos (vetores)
- Através de ponteiros

```
typedef struct Nodo {
                    struct Nodo *esq;
                    char info;
                    struct Nodo *dir;
                   };
```

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Caminhamento

Diversas formas de percorrer ou caminhar em uma árvore listando seus nós, as principais:

- Pré-ordem (Pré-fixa)
- Central (Infixa)
- Pós-ordem (Pós-fixa)

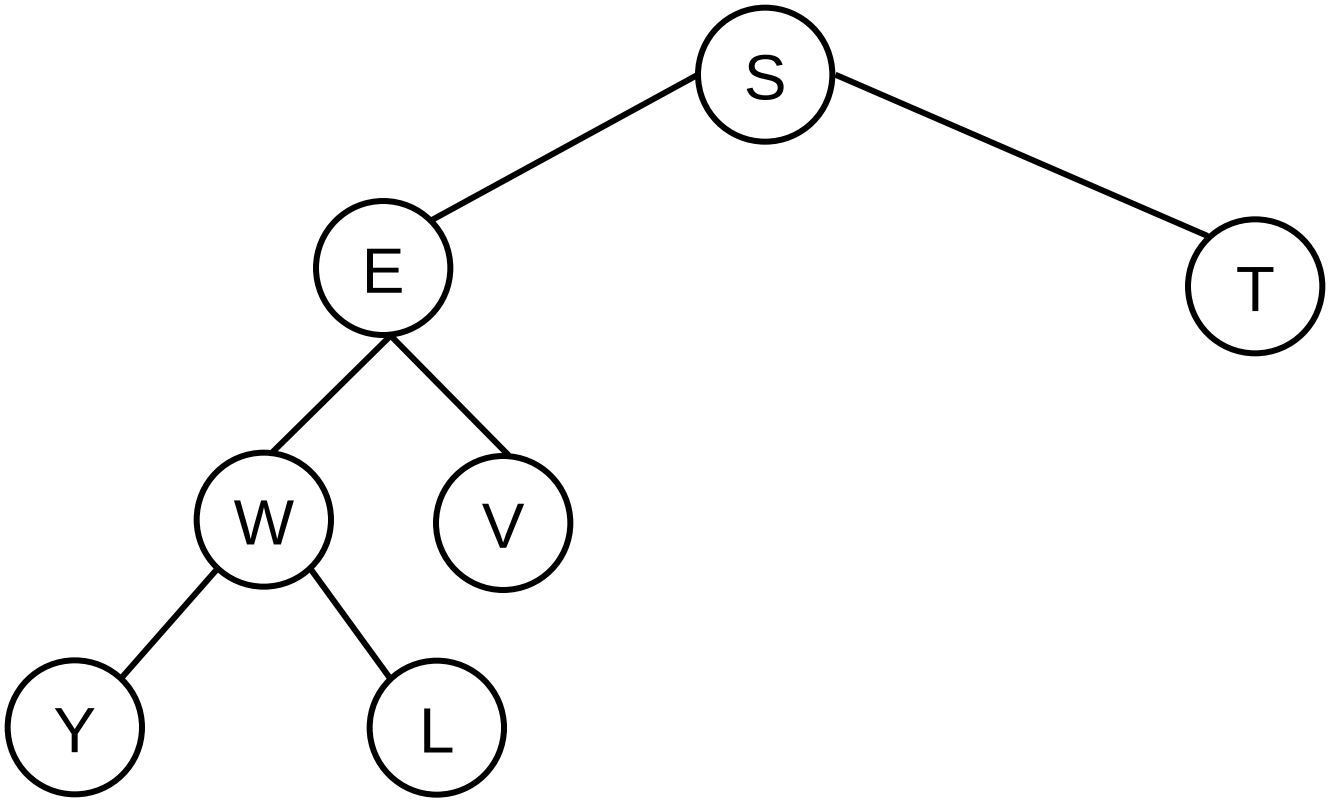
Pré-ordem

Pré-ordem: lista o nó raiz, seguido de suas subárvores (da esquerda para a direita), cada uma em pré-ordem.

```
void Caminhamento_Pre_Ordem(TArvore *a)
{
 if (!Vazia(a))
    {
     printf("%c ", a->info);
     Caminhamento_Pre_Ordem(a->esq);
     Caminhamento_Pre_Ordem(a->dir); sub_dir
    }
}
```

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Pré Ordem: S E W Y L V T



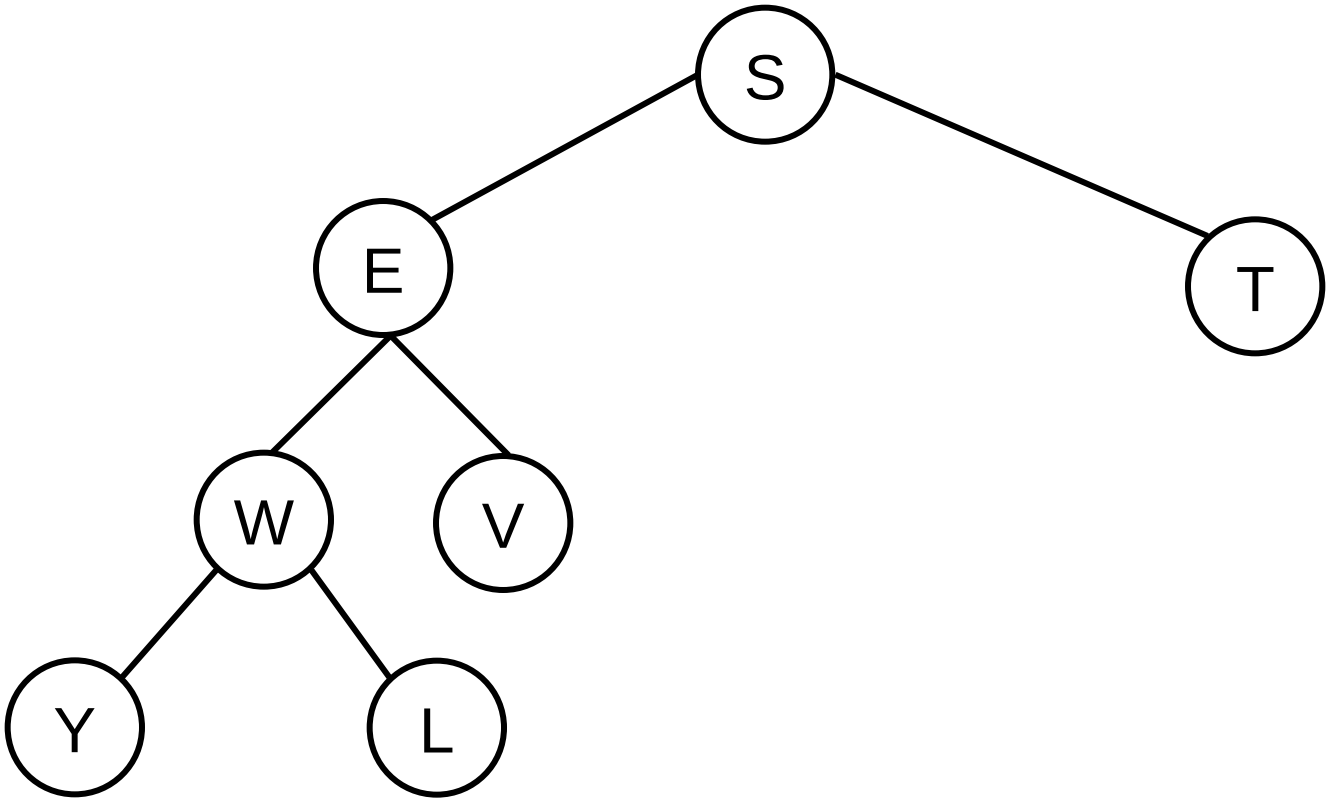
Infixa (Central)

Lista os nós da 1ª. subárvore à esquerda usando o caminhamento central, lista o nó raiz n, lista as demais subárvores (a partir da 2ª.) em caminhamento central (da esquerda para a direita)

```
void Caminhamento_In_Fixado(TArvore *a)
{
 if (!Vazia(a))
     {
       Caminhamento_In_Fixado(a->esq);
       printf("%c ", a->info);
       Caminhamento_In_Fixado(a->dir);
     }
}
```

INSTITUTO FEDERAL
RIO GRANDE DO SUL

CENTRAL: Y W L E V S T



INSTITUTO FEDERAL
RIO GRANDE DO SUL

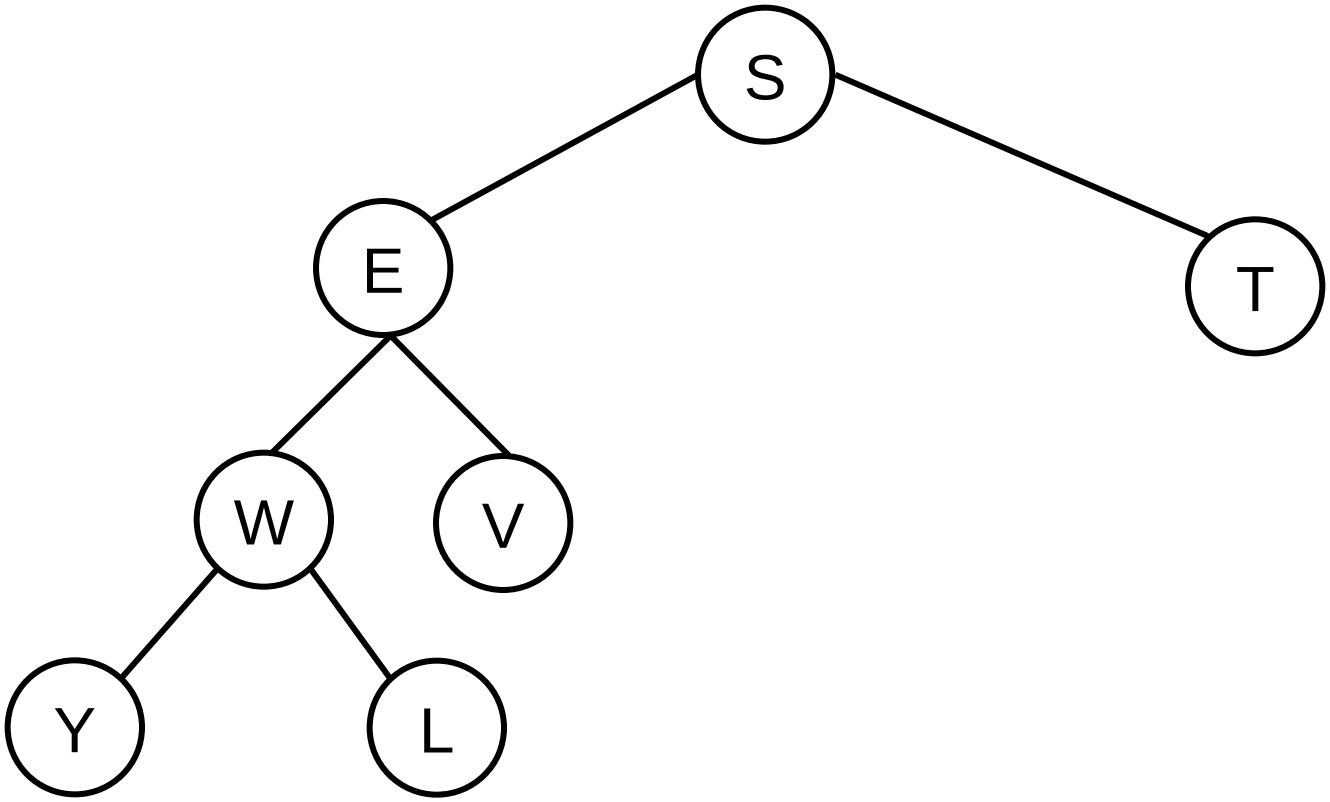
Pós-Ordem

- Lista os nós das subárvores (da esquerda para a direita) cada uma em pós-ordem, lista o nó raiz.

```
void Caminhamento_Pos_Fixado(TArvore *a)
{
 if (!Vazia(a))
    {
     Caminhamento_Pos_Fixado(a->esq);
     Caminhamento_Pos_Fixado(a->dir);
     printf("%c ", a->info);
    }
}
```

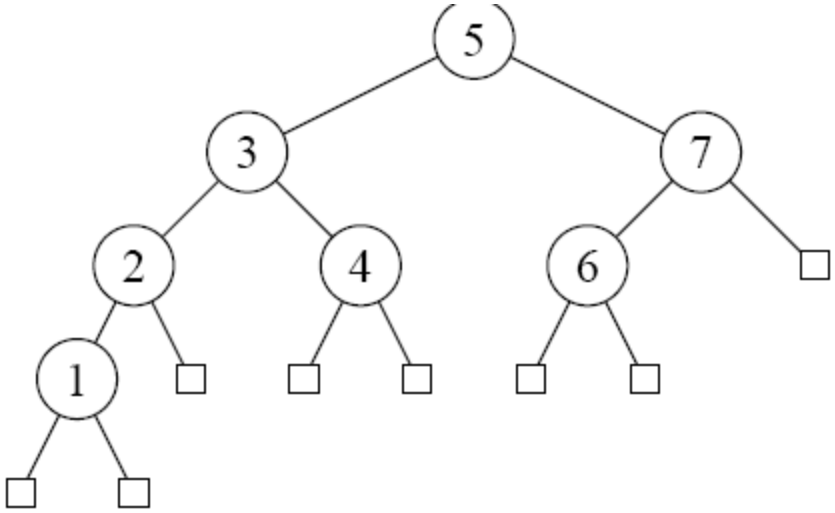
INSTITUTO FEDERAL
RIO GRANDE DO SUL

PÓS ORDEM: Y L W V E T S



Altura de Árvore

- O nível do nó raiz é 0.
  - Se um nó está no nível i então a raiz de suas subárvores estão no nível i + 1.
  - A altura de um nó é o comprimento do caminho mais longo deste nó até um nó folha.
  - A altura de uma árvore é a altura do nó raiz.



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Pesquisa

- **Para encontrar um registro com uma chave x:**
  - Compare-a com a chave que está na raiz.
  - Se x é menor, vá para a subárvore esquerda.
  - Se x é maior, vá para a subárvore direita.
  - Repita o processo recursivamente, até que a chave procurada seja encontrada ou um nó folha é atingido.
  - Se a pesquisa tiver sucesso então o conteúdo do registro retorna no próprio registro x.

INSTITUTO FEDERAL
RIO GRANDE DO SUL

Inserção

Onde inserir?

- Atingir um apontador nulo em um processo de pesquisa significa uma pesquisa sem sucesso.
- O apontador nulo atingido é o ponto de inserção.

Como inserir?

- Cria célula contendo registro
- Procura lugar na árvore
- Se registro não tiver na árvore, insere-o

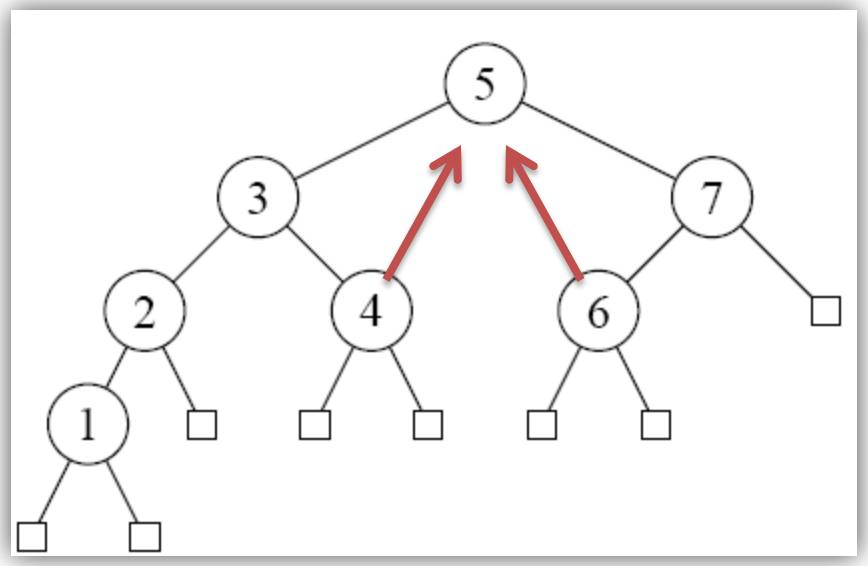
Remoção

Alguns comentários:

- **A retirada de um registro não é tão simples quanto a inserção.**
- Se o nó que contém o registro a ser retirado possui no máximo um descendente ⇒ **a operação é simples**.
- No caso do nó conter **dois descendentes** o registro a ser retirado deve ser primeiro substituído pelo registro mais à direita na subárvore esquerda ou pelo registro mais à esquerda na subárvore direita.

Remoção

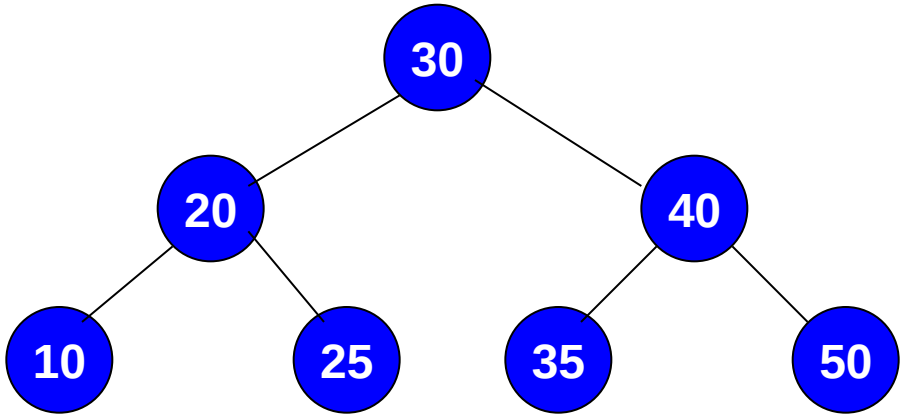
Ou seja, para retirar o registro com chave 5 da árvore basta trocá-lo pelo registro com chave 4 ou pelo registro com chave 6, e então retirar o nó que recebeu o registro com chave 5.



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores Balanceadas

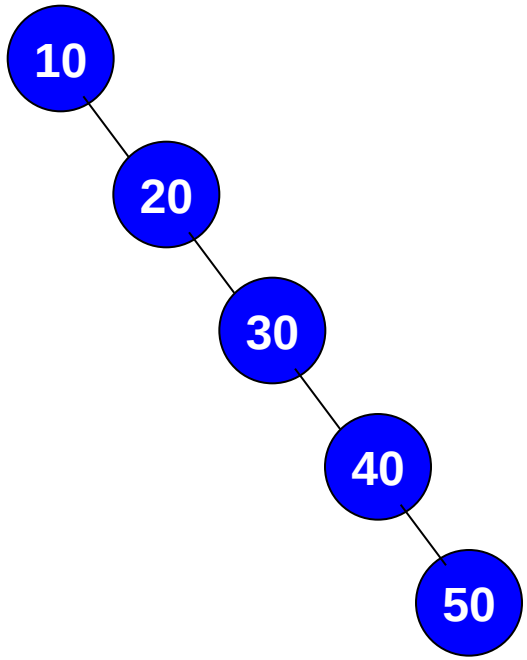
Inserindo os nós 30, 20, 40, 10, 25, 35 e 50 nesta ordem, teremos:



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores Balanceadas

Inserindo os nós 10, 20, 30, 40 e 50 nesta ordem, teremos:



INSTITUTO FEDERAL
RIO GRANDE DO SUL

Árvores Balanceadas

- A vantagem de uma árvore balanceada com relação a uma degenerada está em sua eficiência.
  - Por exemplo: numa árvore binária degenerada de 10.000 nós são necessárias, em média, 5.000 comparações (semelhança com arrays ordenados e listas encadeadas).
- Numa árvore balanceada com o mesmo número de nós essa média reduz-se a **14 comparações**.

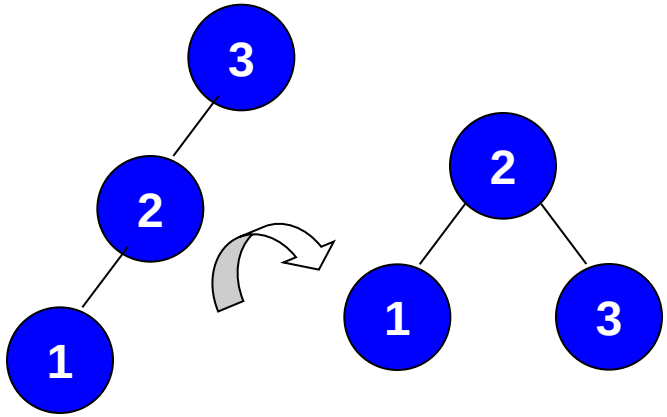
INSTITUTO FEDERAL
RIO GRANDE DO SUL

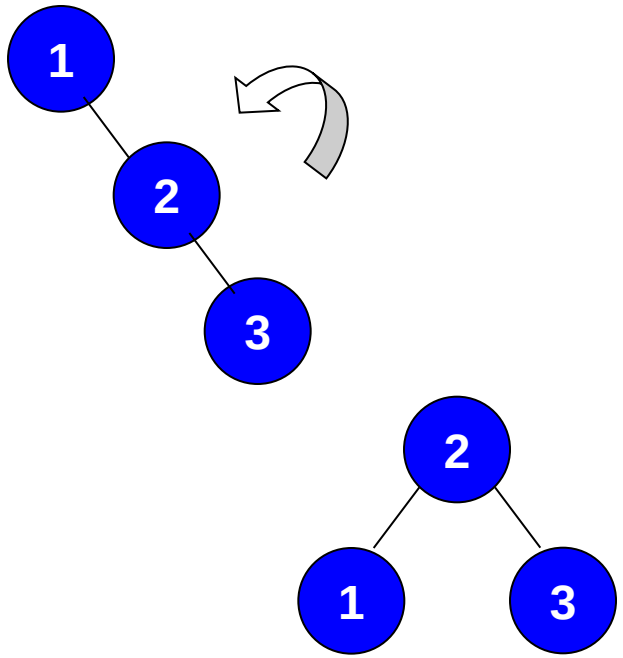
Árvores Balanceadas

- Inicialmente inserimos um novo nó na árvore normalmente.
- A inserção deste pode degenerar a árvore.
- A restauração do balanceamento é feita através de rotações na árvore no nó “pivô”.
- Nó “pivô” é aquele que após a inserção possui Fator de Balanceamento fora do intervalo.

Árvores de Balanceadas

- Rotação simples para a direita / Esquerda





INSTITUTO FEDERAL
RIO GRANDE DO SUL

Atividade

- Construa uma função para Medir a altura de uma Arvore.