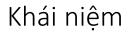
Chương 5 - Cây

Cây tổng quát, biểu diễn cây, duyệt cây, prefix tree, cây nhị phân, suffix tree

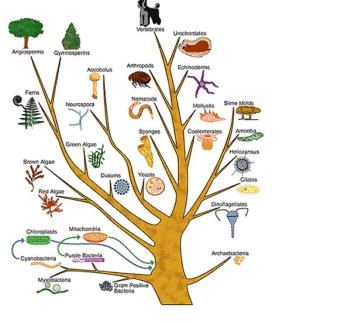
11/18/2021 hiepnd@soict.hust.edu.vn

Nội dung

- Khái niệm cơ bản
- Biểu diễn cây tổng quát
- Một số bài toán cơ bản trên cây tổng quát
- Cây nhị phân
- Ứng dụng của cây nhị phân
- Một số bài toán trên cây nhị phân
- Một số mở rộng của cây tổng quát

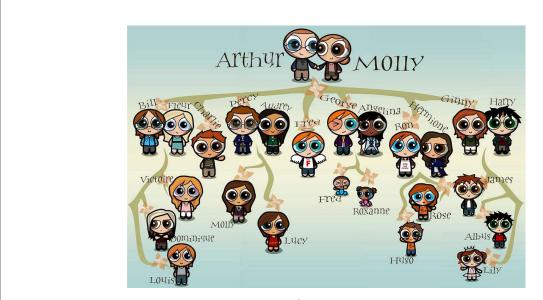


- Cây tree
- Là kiểu cấu trúc dữ liệu phi tuyến
- Biểu diễn mối quan hệ có thứ bậc giữa các đối tượng
- Quan hệ giữa các nút là quan hệ chacon (tổ tiên- con cháu) hoặc ngang hàng



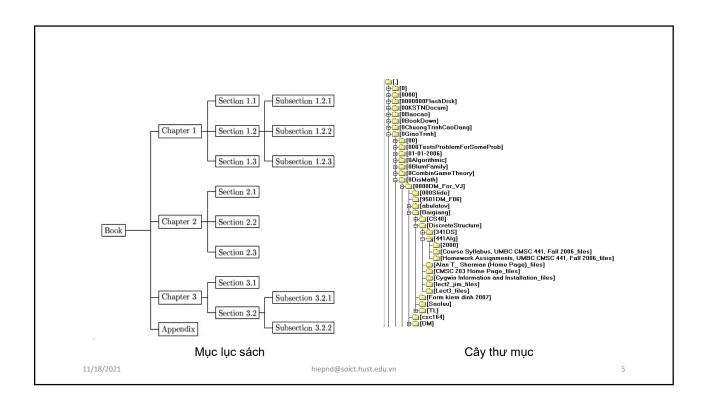
11/18/2021

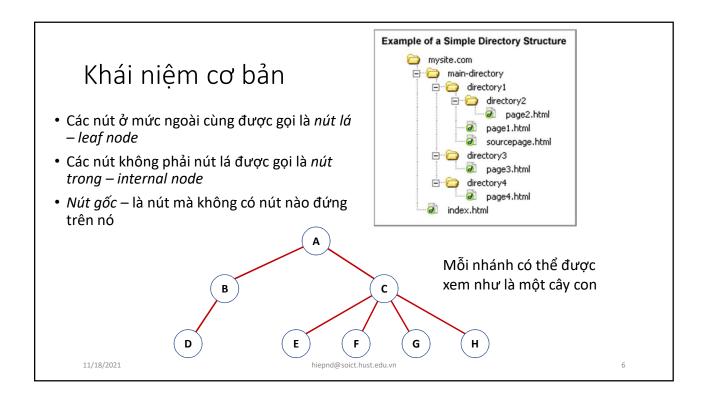
hiepnd@soict.hust.edu.vn



• Cây phả hệ - family tree

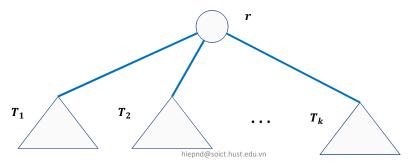
11/18/2021 hiepnd@soict.hust.edu.vn



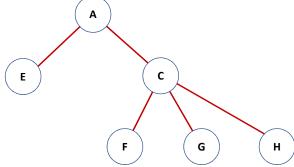


Khái niệm cơ bản

- Định nghĩa: Cây là một tập hợp gồm các nút.
 - Tập này có thể rỗng,
 - nếu không thì nó bao gồm một nút r được gọi là gốc và một tập rỗng hoặc khác rỗng các cây con T_1,T_2,\ldots,T_k mà có nút gốc được nối trực tiếp bằng một cạnh từ nút gốc r



Khái niệm cơ bản



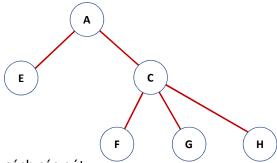
- Các nút cùng cha là anh em (sibling)
 - E và C, hoặc F, G và H
- Nút gốc: là nút không có nút nào đứng trên
 - A là gốc

11/18/2021

- Nút lá: nút không có nút con
 - E, F, G, H
- Nút trong: Nút có ít nhất 1 con
 - A,C

Cây tổng quát: 1 nút có thể không có con, hoặc có thể có vô số con

Khái niệm cơ bản

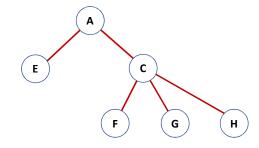


- Đường đi từ nút n_i tới nút n_j trên cây là một danh sách các nút $n_i, n_{i+1}, \ldots, n_j$ trong đó nút n_i là cha của nút n_{i+1} .
- Độ dài đường đi: là số cạnh trên đường đi.
 - Đường đi qua n đỉnh thì có độ dài n-1
 - Đường đi từ nút đến chính nó có độ dài là 0
- Đường đi từ A đến H là A, C, H có độ dài 2
- Đường đi từ C đến C có độ dài 0

11/18/2021

hiepnd@soict.hust.edu.vn

Khái niệm cơ bản



- Độ sâu/mức (depth/level) của nút: độ dài đường đi từ nút gốc đến chính nó
 - Độ sâu của nút gốc A là 0, của nút G, H là 2
- Độ cao (height) của nút: là độ dài đường đi lớn nhất từ nó đến một nút lá
 - Độ cao của nút lá E, F, G, H là 0, của nút A là 2
- Độ cao của cây: là độ cao của nút gốc
- Độ sâu của cây: là độ sâu lớn nhất của nút lá trên cây

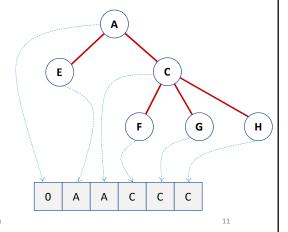
11/18/2021

hiepnd@soict.hust.edu.vn

Biểu diễn cây

• Biểu diễn cây: thông qua nhãn nút cha của nó (cấu trúc liên tiếp)

- Lưu dần các nút theo mức, mức 0 sẽ lưu đầu tiên
- Tìm nút cha của nút hiện tại nhanh
- Thêm/xóa nút phức tạp, thường phải dịch phần tử

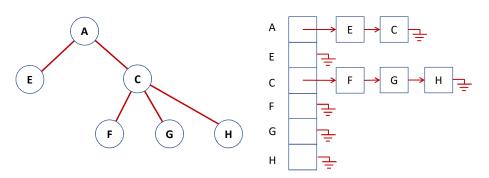


11/18/2021

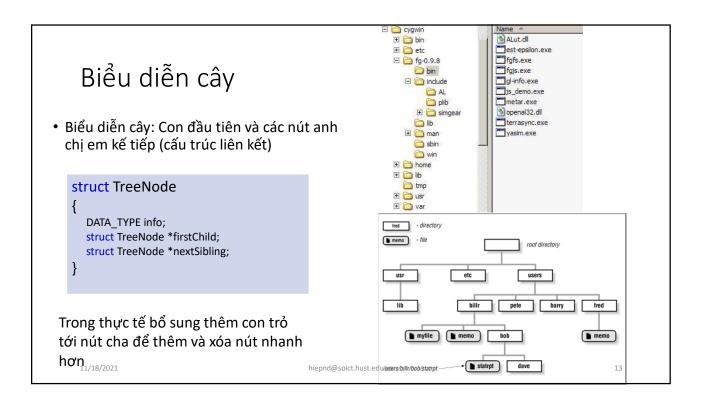
hiepnd@soict.hust.edu.vn

Biểu diễn cây

• Biểu diễn cây: danh sách các nút con (mảng của các danh sách liên kết)



- Danh sách các nút con của nút hiện tại biểu diễn bằng danh sách liên kết đơn
- Thông thường, con trái nhất sẽ đứng đầu danh sách
- Nút không có con thì danh sách tương ứng là rỗng soict.hust.edu.vn



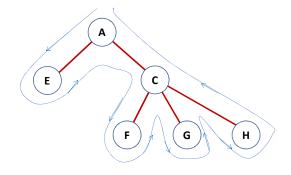
Duyệt cây

- Thuật toán duyệt cây:
 - Phân biệt bởi việc thăm nút con trước hay là xử lý dữ liệu tại nút hiện tại trước
 - Các nút con thông thường sẽ được thăm theo thứ tự con trái nhất trước sau đó đến các nút con kế tiếp
 - Cây tổng quát có các phương án duyệt cây thông dụng
 - Duyệt theo thứ tự trước gọi đệ quy hoặc dùng stack
 - Duyệt theo thứ tự sau gọi đệ quy hoặc dùng stack
 - Duyệt theo mức dùng hàng đợi

Duyệt cây

• **Duyệt cây theo thứ tự trước** (pre-order traversal): xử lý dữ liệu trên nút hiện tại trước, sau đó xử lý tiếp đến các nút con của nó

```
preOrder(root){
   if(root == NULL) return;
   processs(root);
   for each p = r<sub>1</sub>, r<sub>2</sub>, ..., r<sub>k</sub> {
       preOrder(p);
   }
}
```



In ra các đỉnh khi duyệt theo thứ tự trước: A, E, C, F, G, H

11/18/2021

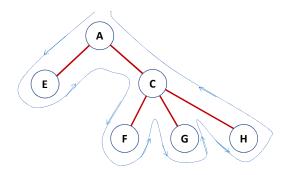
hiepnd@soict.hust.edu.vn

15

Khái niệm

• Duyệt cây theo thứ tự sau (post-order traversal): xử lý dữ liệu trên các nút con trước sau đó mới xử lý dữ liệu trên nút hiện tại

```
postOrder(root){
  if(root == NULL) return;
  for each p = r<sub>1</sub>, r<sub>2</sub>, ..., r<sub>k</sub> {
    postOrder(p);
  }
  processs(root);
}
```



11/18/2021

In ra sác đỉnh khi duyệt theo thứ tự sau: E, F, G, H, C, A

Thao tác với cây tổng quát

- find(r, label): tìm và trả về nút có nhãn label trên cây có gốc là r
- insert(r, label): tạo một nút có nhãn label, chèn vào cuối danh sách nút con của nút r
- height(r): trả về độ cao của nút r
- depth(r): trả về độ sâu của nút r
- parent(r): trả về nút cha của nút r
- count(r): trả về số nút trên cây có gốc là r
- countLeaves(r): trả về số nút lá trên cây có gốc là r

#include <stdio.h>

```
TreeNode* makeNode(char label)
         TreeNode* newNode = (TreeNode*)malloc(sizeof(TreeNode));
         newNode->label = label;
         newNode->firstChild = NULL;
         newNode->nextSibling = NULL;
         return newNode;
// chen nut da duoc tao newNode vao nut con moi nhat (phai nhat) cua nut root
TreeNode* insertNode(TreeNode* root, TreeNode* newNode)
         if (root == NULL) root = newNode;
         if (root->firstChild == NULL) root->firstChild = newNode;
                   TreeNode* rightmost = root->firstChild;
                   while (rightmost->nextSibling != NULL)rightmost = rightmost->nextSibling;
                   rightmost->nextSibling = newNode;
         return root;
      11/18/2021
                                                   hiepnd@soict.hust.edu.vn
                                                                                                                18
```

```
// chen nut da duoc tao newNode vao nut con moi nhat (phai nhat)
// cua nut root, ham khong can tra ve gia tri
void insertNode2(TreeNode** root, TreeNode* newNode)
          if (*root == NULL) *root = newNode;
          if ((*root)->firstChild == NULL) (*root)->firstChild = newNode;
                    TreeNode* rightmost = (*root)->firstChild;
                    while (rightmost->nextSibling != NULL)rightmost = rightmost->nextSibling;
                    rightmost->nextSibling = newNode;
}
                       // tao va chen nut moi thanh cay con phai nhat cua nut root
                       // Chu ý: root phai khac NULL
                       void insertNode_v2(TreeNode* root, char label)
                       {
                                  if (root == NULL) return;
                                  TreeNode* newNode = makeNode(label);
                                  if (root->firstChild == NULL) root->firstChild = newNode;
                                  else {
                                            TreeNode* rightmost = root->firstChild;
                                            while (rightmost->nextSibling != NULL) rightmost = rightmost->nextSibling;
                                            rightmost->nextSibling = newNode;
                                 }
                                                        hiepnd@soict.hust.edu.vn
       11/18/2021
```

```
// duyet cay theo muc
void levelTraversal(TreeNode* root)
          if (NULL == root) return;
                                                                      Ε
                                                                                             C
          queue<TreeNode*> Q;
          Q.push(root);
          while (Q.size() > 0)
                                                                                       F
                                                                                                    G
                    TreeNode* p = Q.front();
                    Q.pop();
                                                           int main()
                    printf("%c, ", p->label);
                    TreeNode* q = p->firstChild;
                                                                      TreeNode* root = makeNode('A');
                    while (NULL != q)
                                                                      insertNode_v2(root, 'E');
                                                                      insertNode_v2(root, 'C');
                              Q.push(q);
                                                                      insertNode_v2(root->firstChild->nextSibling, 'F');
                              q = q->nextSibling;
                                                                      insertNode_v2(root->firstChild->nextSibling, 'G');
                                                                      insertNode v2(root->firstChild->nextSibling, 'H');
                                                                      levelTraversal(root);
          printf("\n");
                                                                      return 0;
}
      11/18/2021
                                                      hiepnd@soict.hust.edu.vn
```

```
// duyet cay theo thu tu truoc
                                                            // duyet cay theo thu tu sau
   void preOrderTravesal(TreeNode* root)
                                                            void postOrderTravesal(TreeNode* root)
             if (NULL == root) return;
                                                                      if (NULL == root) return;
                                                                      TreeNode* p = root->firstChild;
             printf("%c, ", root->label);
             TreeNode* p = root->firstChild;
                                                                      while (NULL != p) {
             while (NULL != p) {
                                                                                postOrderTravesal(p);
                      preOrderTravesal(p);
                                                                                p = p->nextSibling;
                      p = p->nextSibling;
             }
                                                                      printf("%c, ", root->label);
                                                          int countLeaves(TreeNode* root) {
// dem so luong nut tren cay
int countNodes(TreeNode* root)
                                                                    if (root == NULL) return 0;
                                                                    int sum = 0;
                                                                    TreeNode* p = root->firstChild;
         if (NULL == root) return 0;
         TreeNode* p = root->firstChild;
                                                                    if (p == NULL) sum = 1;
         int sum = 0;
                                                                    while (p != NULL) {
         while (NULL != p) {
                                                                              sum += countLeaves(p);
                   sum += countNodes(p);
                                                                              p = p->nextSibling;
                   p = p->nextSibling;
                                                                    return sum;
         return sum + 1;
```

```
// tinh chieu cao cua nut (chiều cao la la 1)
                                                                         int height(TreeNode* root) {
                                                                                   if (root == NULL) return 0;
                                                                                   int maxh = 0;
                                                                                   TreeNode* q = root->firstChild;
// tim nut cha cua nut p tren cay co goc la root
                                                                                   while (q != NULL) {
TreeNode* parent(TreeNode* root, TreeNode* p) {
                                                                                             int h = height(q);
          if (root == NULL) return NULL;
                                                                                             if (h > maxh) maxh = h;
         TreeNode* q = root->firstChild;
                                                                                             q = q->nextSibling;
          while (q != NULL) {
                    if (p == q) return root;
                                                                                   return maxh + 1;
                    TreeNode* pp = parent(q, p);
                                                                         }
                    if (pp != NULL) return pp;
                    q = q->nextSibling;
          return NULL;
}
       11/18/2021
                                                       hiepnd@soict.hust.edu.vn
                                                                                                                        22
```

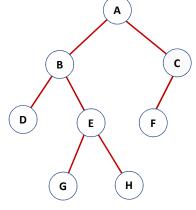
```
// tinh chieu cao cua nut
// coi nut la chieu cao la 0
int height(TreeNode* root) {
    if (root == NULL) return -1; // cay rong
        TreeNode* q = root->firstChild;
    if (NULL == q) return 0; // nut hien tai la la
    int maxh = 0;
    while (q!= NULL) {
        int h = height(q);
        if (h > maxh) maxh = h;
        q = q->nextSibling;
    }
    return maxh + 1;
}
```

```
// Tim do sau nut co nhan label tren cay goc la root
                                                                // Tim do sau nut p tren cay goc la root dung de quy
                                                                // d la do sau cua nut root
// dung de quy
int depth(TreeNode* root, char label, int d) {
          // d la do sau cua nut root
                                                                int depth(TreeNode* root, TreeNode* p, int d) {
                                                                           if (root == NULL) return -1;
          if (root == NULL) return -1;
          if (root->label == label) return d;
                                                                           if (root == p) return d;
          TreeNode* p = root->firstChild;
                                                                           TreeNode* q = root->firstChild;
          while (p != NULL) {
                                                                           while (q != NULL) {
                    if (p->label == label) return d + 1;
                                                                                     if (q == p) return d + 1;
                    int dv = depth(p, label, d + 1);
                                                                                     int dv = depth(q, p, d + 1);
                    if (dv > 0) return dv;
                                                                                     if (dv > 0) return dv;
                     p = p->nextSibling;
                                                                                     q = q->nextSibling;
                                                                          }
          return -1;
                                                                           return -1;
int calDepth(TreeNode* root, char v) {
                                                                int calDepth(TreeNode* root, TreeNode* v) {
          return depth(root, v, 0);
                                                                           return depth(root, v, 0);
       11/18/2021
                                                        hiepnd@soict.hust.edu.vn
```

```
// tinh do sau cua nut co nhan label, không de quy
            int depth(TreeNode* root, char label) {
                      if (root == NULL) return -1;
                      queue<pair<TreeNode*,int>> Q;
                      Q.push(make_pair(root,0));
                      int currentDepth;
                      TreeNode* p;
                      while (Q.size()>0) {
                                p = Q.front().first;
                                currentDepth = Q.front().second;
                               if (p->label == label) return currentDepth;
                                p = p->firstChild;
                                while (p != NULL) {
                                          Q.push(make_pair(p, currentDepth+1));
                                          p = p->nextSibling;
                      return -1;
                                               hiepnd@soict.hust.edu.vn
11/18/2021
```



- Cây nhị phân binary tree:
 - Là một tập rỗng, hoặc
 - Có một nút gốc với hai cây nhị phân con của gốc gọi là cây con trái (left subtree) và cây con phải (right subtree)

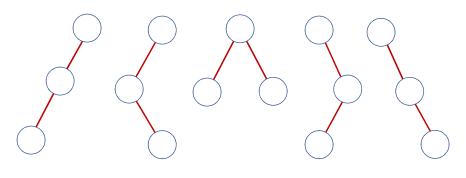


11/18/2021

hiepnd@soict.hust.edu.vn

Cây nhị phân

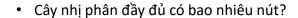
• Một số cây nhị phân gồm 3 nút khác nhau

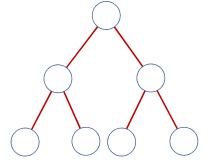


11/18/2021

hiepnd@soict.hust.edu.vn

- Cây nhị phân đầy đủ (full binary tree):
 - Các nút không phải là lá đều có 2 nút con
 - Các nút lá có độ sâu bằng nhau
- Cây nhị phân đầy đủ chiều cao h có bao nhiêu nút lá?

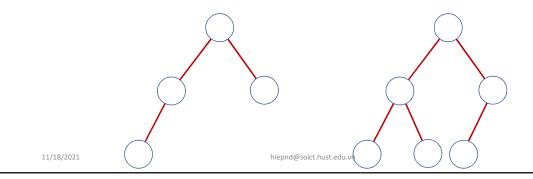




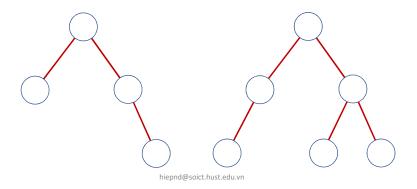
11/18/2021 hiepnd@soict.hust.edu.vn 29

Cây nhị phân

- Cây nhị phân hoàn chỉnh (complete binary tree) chiều cao h
 - Các nút từ mức 0 tới h-1 là cây nhị phân đầy đủ
 - Một số nút ở mức h-1 có thể có 0 hoặc 1 con
 - Nếu i, j là các nút ở mức h-1, thì i có nhiều con hơn j nếu i nằm ở bên trái j



- Cây nhị phân cân bằng (balanced binary tree): là cây nhị phân thỏa mãn
 - Với mọi nút thì chênh lệch chiều cao của cây con trái và cây con phải không quá 1
 - Mỗi cây nhị phân hoàn chỉnh là cây cân bằng

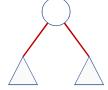


Duyệt cây – tree traversal

- Mỗi nút trên cây nhị phân gồm :
 - · Giá trị chứa tại nút
 - Cây con trái

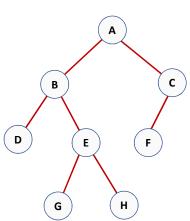
11/18/2021

• Cây con phải



- Duyệt cây theo thứ tự
 - Thứ tự trước preorder: Giá trị → con trái → con phải
 - Thứ tự giữa inorder : con trái → giá trị → con phải
 - Thứ tự sau postorder : con trái → con phải → giá trị

Duyệt cây – tree traversal



• Thứ tự trước: 1, 2, 3

• Thứ tự giữa : 2, 1, 3

• Thứ tự sau : 2, 3, 1



Thứ tự trước: A, B, D, E, G, H, C, D

Thứ tự giữa: D, B, G, E, H, A, F, C

Thứ tự sau: D, G, H, E, B, F, C, A

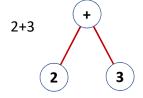
11/18/2021

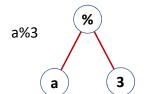
hiepnd@soict.hust.edu.vn

33

Cây biểu thức – expression tree

- Cây biểu thức expression tree: xây dựng từ các toán tử và toán hạng
 - Toán tử 2 ngôi: gốc là toán tử, 2 nút con lần lượt là toán hạng trái và phải





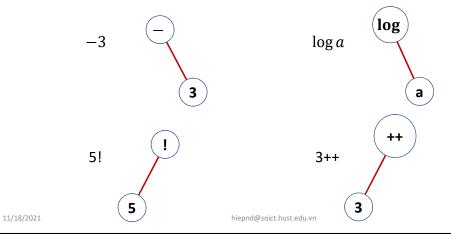
11/18/2021

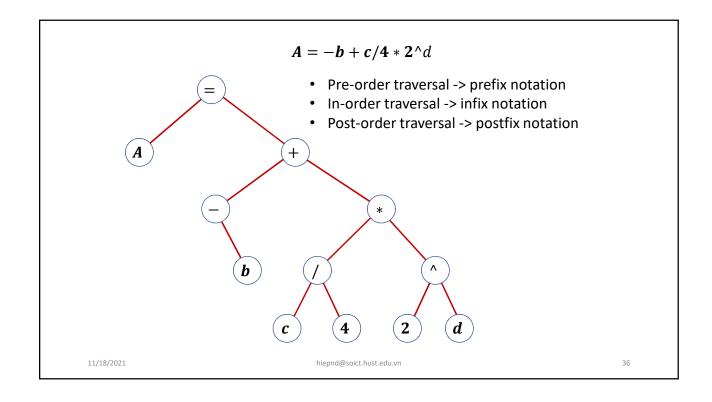
hiepnd@soict.hust.edu.vn

35

Cây biểu thức – expression tree

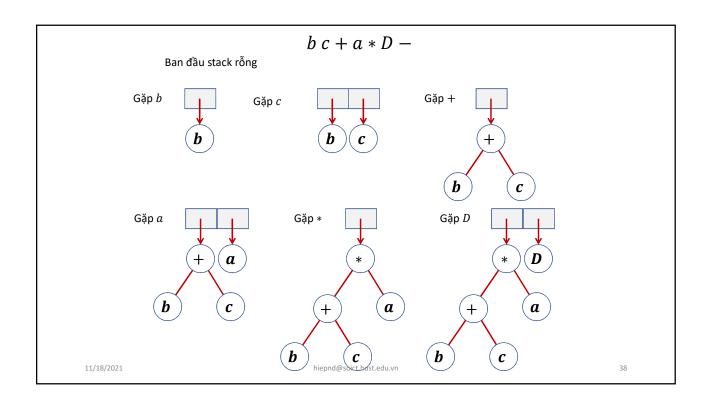
• Toán tử 1 ngôi: nút gốc biểu diễn toán tử, chỉ có 1 nút con biểu diễn toán hạng

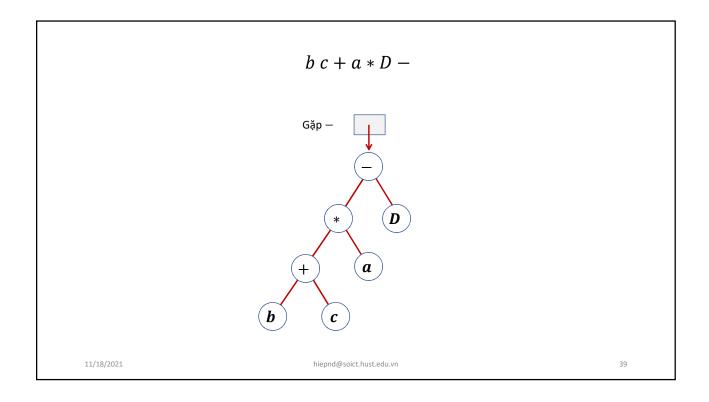


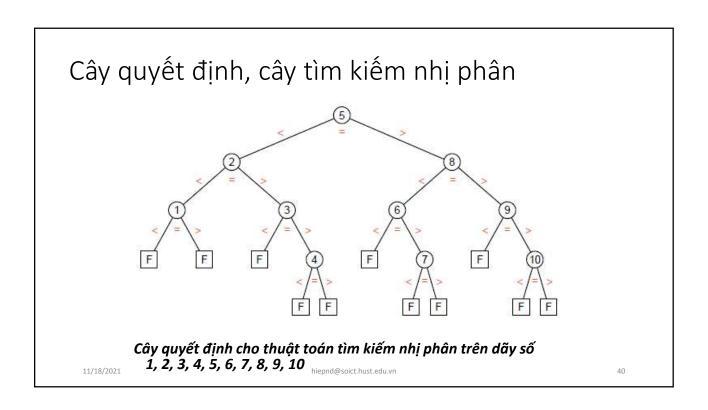


Cây biểu thức – expression tree

- Thuật toán xây dựng cây biểu thức từ biểu thức dạng hậu tố
- Đọc lần lượt các phần tử của biểu thức hậu tố
- Gặp toán hạng: tạo ra cây có 1 nút là toán hạng và đẩy vào stack
- Gặp toán tử:
 - Nếu là toán tử 1 ngôi: lấy ra 1 cây T trong stack, tạo ra 1 cây trong đó toán tử là nút gốc và cây T là nút con, sau đó đẩy cây này vào stack
 - Nếu là toán tử 2 ngôi: lấy ra 2 cây trong stack T_1 , T_2 (T_2 được lấy ra trước), tạo cây trong đó toán tử này là gốc và T_1 , T_2 lần lượt là con trái và con phải của nó, sau đó đẩy cây này vào stack
- Sau khi duyệt xong biểu thức hậu tố, cây biểu thức là cây có gốc nằm ở đỉnh của stack





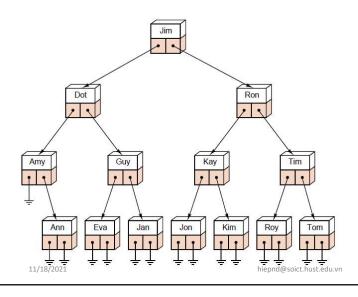


Biểu diễn cây nhị phân

- Biểu diễn cây nhị phân
- Dùng cấu trúc liên tiếp (mảng): số lượng nút trên cây chiều cao h là giới hạn ($\sum 2^i$)
 - Truy nhập nhanh O(1)
 - Lãng phí bộ nhớ lớn nếu cây chưa phải là dạng đầy đủ hoặc hoàn chỉnh
- Dùng cấu trúc liên kết: thường dùng hơn
 - Truy cập 1 nút chậm hơn
 - Tiết kiệm bộ nhớ hơn (trong trường hợp tổng quát)

11/18/2021 hiepnd@soict.hust.edu.vn 4

Biểu diễn cây nhị phân



```
struct TreeNode
{
    DATA_TYPE info;
    struct TreeNode *leftChild;
    struct TreeNode *rightChild;
}
```

• Biểu diễn bằng cấu trúc liên kết

Biểu diễn cây nhị phân

• Duyệt cây theo thứ tự trước/giữa/sau

```
void preorderTraversal(TreeNode* root)
{
    if(root==NULL) return;
    printf("%d ",root->data);
    preorderTraversal(root->leftChild);
    preorderTraversal(root->rightChild);
}
```

```
void postorderTraversal(TreeNode* root)
{
  if(root==NULL) return;
  postorderTraversal(root->leftChild);
  postorderTraversal(root->rightChild);
  printf("%d ",root->data);
}¹¹/¹8/²021
```

```
struct TreeNode
{
  int data;
  struct TreeNode *leftChild;
  struct TreeNode *rightChild;
}
```

```
void inorderTraversal(TreeNode* root)
{
   if(root==NULL) return;
   inorderTraversal(root->leftChild);
   printf("%d ",root->data);
   inorderTraversal(root->rightChild);
}
```

hiepnd@soict.hust.edu.vn

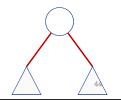
43

Cài đặt thuật toán trên cây nhị phân

- Cây nhị phân có 5 trường hợp khi xử lý đệ quy
 - Cây rỗng
 - · Nút gốc là nút lá
 - Nút gốc chỉ có con trái
 - Nút gốc chỉ có con phải
 - Nút gốc có đủ 2 con
- Tùy các bài toán mà có thể gộp các trường hợp lại với nhau







- Bài 1. Tìm nút có giá trị lớn nhất/ nhỏ nhất trên cây nhị phân
- Bài 2. Thêm/xóa nút trên cây nhị phân
- Bài 3. Kiểm tra 2 cây nhị phân đồng dạng
- Bài 4. Kiểm tra 2 cây nhị phân đối xứng
- Bài 5. Thêm nút mới vào cây theo mức lần lượt từ trái sang phải (tương tự trên cây tổng quát)
- Bài 6. In ra đường đi dài nhất trên cây từ gốc tới lá
- Bài 7. In ra nút lá nông nhất/sâu nhất trên cây
- Bài 8. Cho cây nhị phân mà giá trị tại nút là 0 hoặc 1 (gốc luôn là 1), in ra đường đi chứa toàn số 1 dài nhất từ gốc
- Bài 9. Định giá cây biểu thức

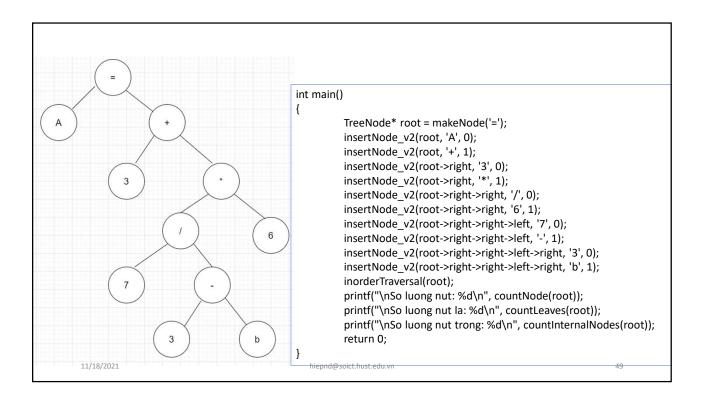
11/18/2021 hiepnd@soict.hust.edu.vn 45

Cây nhị phân

- Bài 10. Cho cây nhị phân trong đó mỗi nút trong có 2 con và nút lá, cho thứ tự duyệt cây trước là B C D E A F G và thứ tự giữa C B A E F D G, hãy đưa ra thứ tự duyệt sau
- Bài 11. Tìm nút tổ tiên gần nhất của 2 nút trên cây
- Bài 12. kiểm tra xem các nút lá trên cây có phải cùng mức
- Bài 13. Cho 2 cây T1 và T2, kiểm tra xem T2 có phải là cây con của T1
- Bài 14. Cho cây nhị phân với nút là số nguyên, in ra các đường đi từ gốc có tổng giá trị ≤k
- Bài 15. Loại các nút bị trùng lặp trên cây nhị phân

```
#include <stdio.h>
                                                         // tao va chen nut moi thanh cay con phai nhat cua nut root
#include <stdlib.h>
                                                         // Chu y: root phai khac NULL
typedef struct TREENode
                                                         // leftorRight = 0==> la con trai, 1 la con phai
                                                         void insertNode_v2(TreeNode* root, char label, int leftorRight)
         char label;
         struct TREENode* left;
                                                                   if (root == NULL) return;
         struct TREENode* right;
                                                                   TreeNode* newNode = makeNode(label);
} TreeNode;
                                                                  if (leftorRight == 0) root->left = newNode;
                                                                   else root->right = newNode;
TreeNode* makeNode(char label)
                                                         }
         TreeNode* newNode =
            (TreeNode*)malloc(sizeof(TreeNode));
                                                         void preorderTraversal(TreeNode* root)
         newNode->label = label;
         newNode->left = NULL;
                                                                  if (root == NULL) return;
         newNode->right = NULL;
                                                                   printf("%c ", root->label);
         return newNode;
                                                                   preorderTraversal(root->left);
}
                                                                   preorderTraversal(root->right);
                                                         }
     11/18/2021
                                                   hiepnd@soict.hust.edu.vn
```

```
void inorderTraversal(TreeNode* root)
                                                    int countNode(TreeNode* root)
{
          if (root == NULL) return;
                                                               if (NULL == root) return 0;
          inorderTraversal(root->left);
                                                               return 1 + countNode(root->left) + countNode(root->right);
          printf("%c ", root->label);
          inorderTraversal(root->right);
}
                                                    int countLeaves(TreeNode* root)
void postorderTraversal(TreeNode* root)
                                                               if (NULL == root) return 0;
          if (root == NULL) return;
                                                               if (NULL == root->left && NULL == root->right) return 1;
          postorderTraversal(root->left);
                                                               return countLeaves(root->left) + countLeaves(root->right);
                                                    }
          postorderTraversal(root->right);
          printf("%c ", root->label);
                                                    int countInternalNodes(TreeNode* root)
                                                               if (NULL == root) return 0;
                                                               if (NULL == root->left && NULL == root->right) return 0;
                                                               return 1 + countInternalNodes(root->left) +
                                                     countInternalNodes(root->right);
     11/18/2021
                                                     hiepnd@soict.hust.edu.vn
```



Mở rộng của cây tổng quát