자료구조(Data Structures) 7장. 트리(tree)

담당 교수 : 조 미경

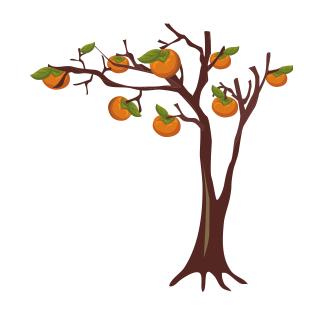
이번 장에서 학습할 내용



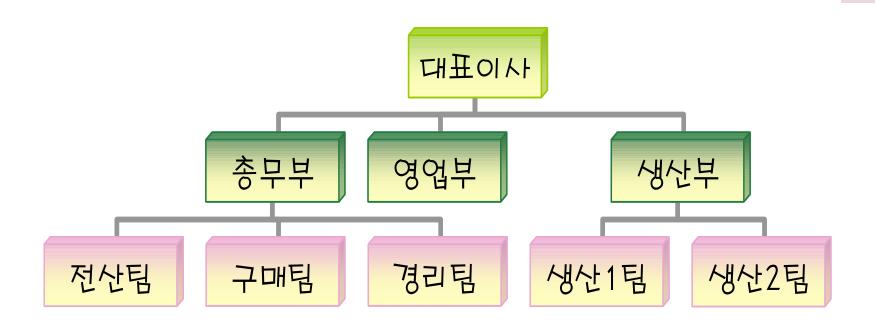
- * 트리란?
- * 이진 트리란?
- * 이진 트리 구현(배열, 연결리스트)
- * 이진 트리 순회(전위,중위,후위,레벨)
- * 스레드 이진 트리
- * 이진 탐색 트리
- * 이진 탐색 트리의 응용: 영어사전

트리(TREE)

- 트리: 계층적인 구조를 나타내는 자료구조
 - ▶ 리스트, 스택, 큐 등은 선형 구조
- 트리는 부모-자식 관계의 노드들로 이루어진다.
- 응용분야:
 - ▶ 계층적인 조직 표현
 - ▶ 컴퓨터 디스크의 디렉토리 구조
 - ▶ 인공지능에서의 결정트리 (decision tree)



회사의 조직



교과목: 자료구조 4/69

파일 디렉토리 구조

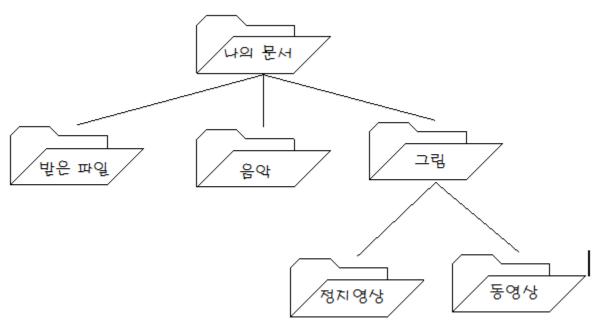


그림 7.2 트리의 예: 컴퓨터 디렉토리

교과목: 자료구조 5/69

결정 트리

• (예) 골프에 대한 결정 트리

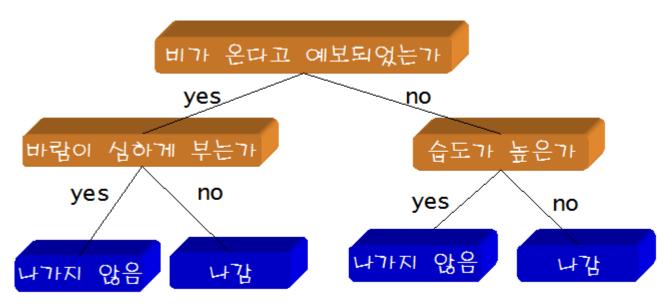
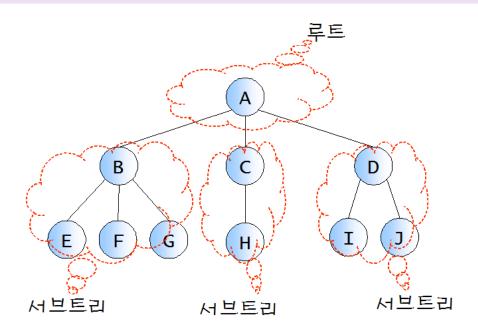


그림 7.3 트리의 예:결정 트리

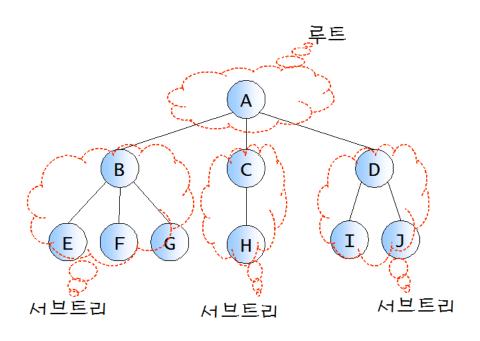
교과목: 자료구조 6/69

트리의 용어



- 노드(node): 트리의 구성요소
- 루트(root): 부모가 없는 노드(A)
- 서브트리(subtree): 하나의 노드와 그 노드들의 자손들로 이루어진 트리

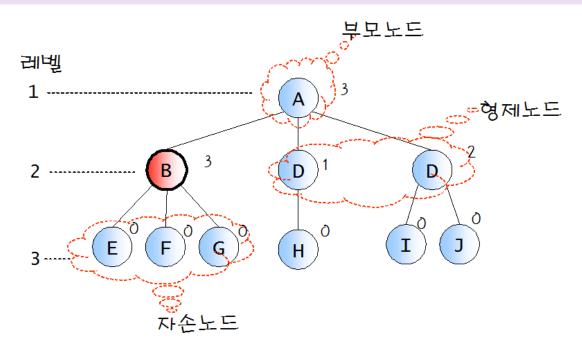
트리의 용어



- 단말노드(terminal node): 자식이 없는 노드(A,B,C,D)
- 비단말노드: 적어도 하나의 자식을 가지는 노드(E,F,G,H,I,J)

교과목: 자료구조 8/69

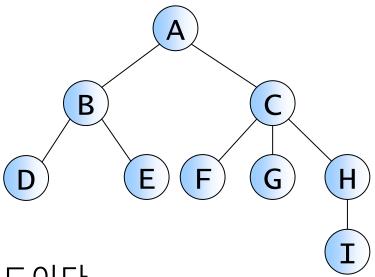
트리의 용어



- 자식, 부모, 형제, 조상, 자손 노드: 인간과 동일
- 레벨(level): 트리의 각층의 번호
- 높이(height): 트리의 최대 레벨(3)
- 차수(degree): 노드가 가지고 있는 자식 노드의 개수

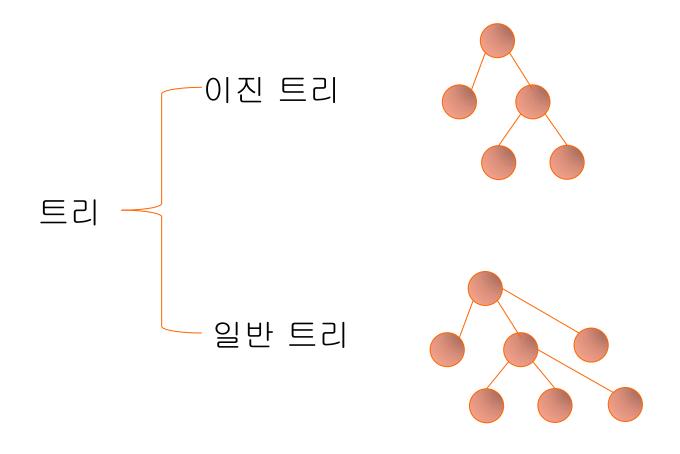
교과목: 자료구조 9/69

예제



- A는 루트 노드이다.
- B는 D와 E의 부모노드이다.
- C는 B의 형제 노드이다.
- D와 E는 B의 자식노드이다.
- B의 차수는 2이다.
- 위의 트리의 높이는 4이다.

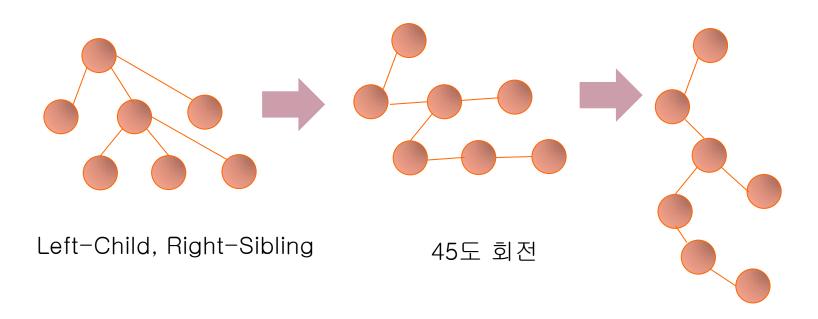
트리의 종류



교과목 : 자료구조 11/69

일반 트리를 이진 트리로 변환

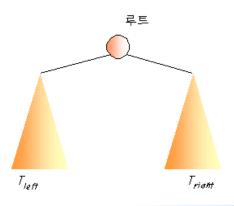
- 일반 트리를 이진 트리로 만드는 방법
 - ▶ 왼쪽노드는 자식 노드로, 오른쪽 노드들은 형제 노드로 배치(Left-Child, Right-Sibling)
- 모든 트리를 이진 트리 형태로 재구성 가능



교과목 : 자료구조 12/69

이진트리 (binary tree)

- 이진 트리(binary tree): 모든 노드가 2개의 서브 트리를 가지고 있는 트리
 - ▶ 서브트리는 공집합일수 있다.
- 이진트리의 노드에는 최대 2개까지의 자식 노드가 존재
- 모든 노드의 차수가 2 이하가 된다-> 구현하기가 편리함
- 이진 트리에는 서브 트리간의 순서가 존재



교과목 : 자료구조 13/69

이진 트리 검증

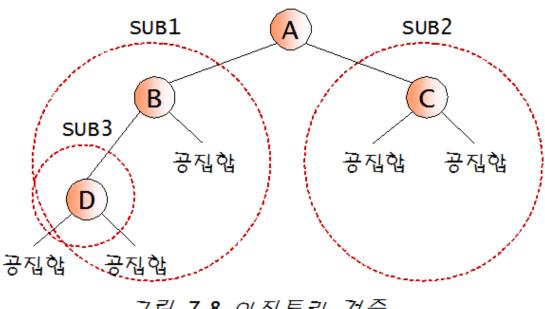


그림 7.8 이진트리 검증

- 이진 트리는 공집합이거나
- 루트와 왼쪽 서브 트리, 오른쪽 서브 트리로 구성된 노드들의 유한 집합으로 정의된다. 이진트리의 서브 트리들은 모두 이진 트리여야 한다.

이진트리의 성질

● 노드의 개수가 n개이면 간선의 개수는 n-1

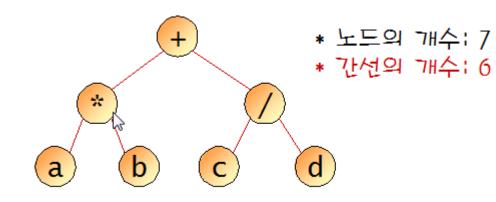
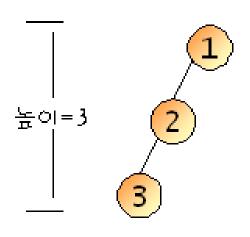


그림 7.10 노드의 개수와 간선의 개수와의 관계

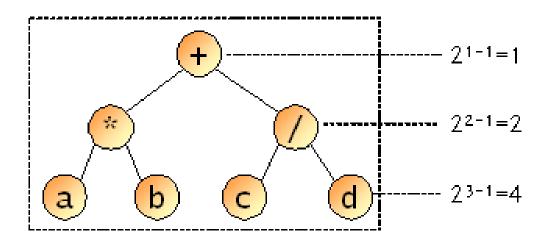
교과목: 자료구조 15/69

이진트리의 성질

• 높이가 h인 이진트리의 경우, 최소 h개의 노드를 가지며 최대 2^h-1 개의 노드를 가진다.



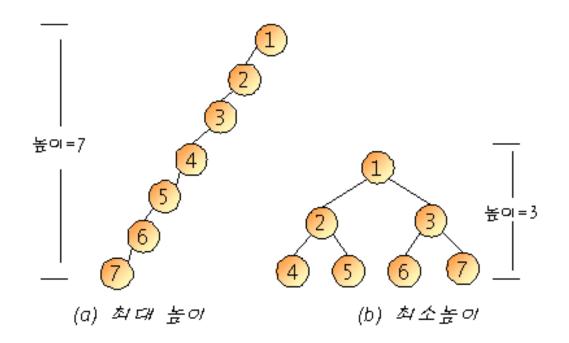
최소 노드 개수=3



교과목: 자료구조 16/69

이진트리의 성질

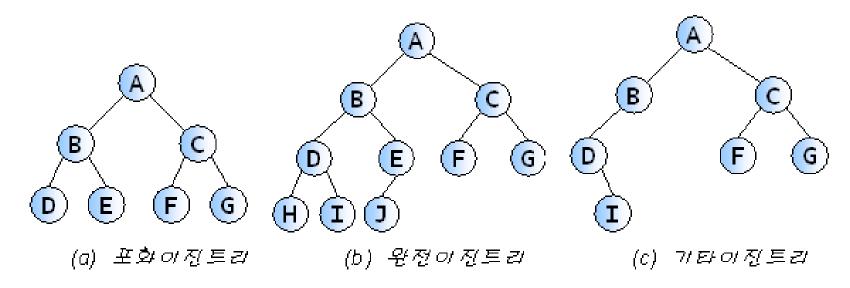
- n개의 노드를 가지는 이진트리의 높이
 - ▶ 최대 n
 - ▶ 최소 「log₂(n+1)]



교과목 : 자료구조 17/69

이진트리의 분류

- 포화 이진 트리(full binary tree)
- 완전 이진 트리(complete binary tree)
- 기타 이진 트리



교과목: 자료구조 18/69

포화 이진 트리

■ 용어 그대로 트리의 각 레벨에 노드가 꽉 차있는 이진트리를 의미한다.

전체 노드 개수 :
$$2^{1-1} + 2^{2-1} + 2^{3-1} + \dots + 2^{k-1} = \sum_{i=0}^{k-1} 2^i = 2^k - 1$$

■ 포화 이진 트리에는 다음과 같이 각 노드에 번호를 붙일 수 있다.

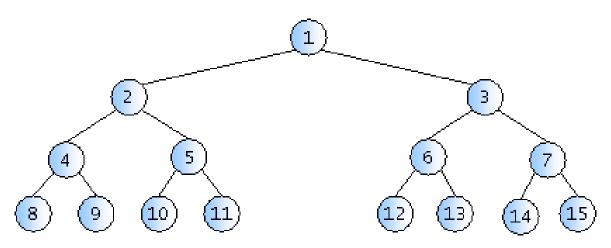
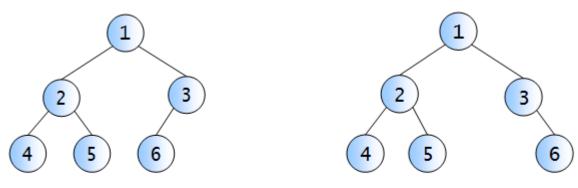


그림 7.15 포화이진트리에서의 노트의 번호

교과목: 자료구조 19/69

완전 이진 트리

- **완전 이진 트리(complete binary tree):** 레벨 1부터 k-1까지 는 노드가 모두 채워져 있고 마지막 레벨 k에서는 왼쪽부 터 오른쪽으로 노드가 순서대로 채워져 있는 이진트리
- 포화 이진 트리와 노드 번호가 일치



(a) 완전이진트리 (b) 완전이진트리가 아넴

그림 7.16 완전 이진트리의 예

교과목 : 자료구조 20/69

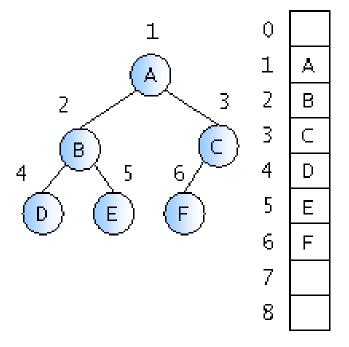
이진 트리의 표현

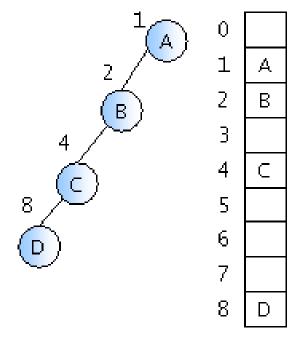
- 배열을 이용하는 방법
- 포인터를 이용하는 방법

교과목 : 자료구조 21/69

배열 표현법

 배열표현법: 모든 이진 트리를 포화 이진 트리라고 가정하고 각 노드에 번호를 붙여서 그 번호를 배열의 인덱스로 삼아 노 드의 데이터를 배열에 저장하는 방법



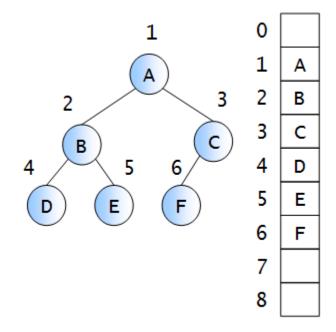


(a) <u>원전 이전 트리</u>

(b) 경사 <u>의진 투리</u>

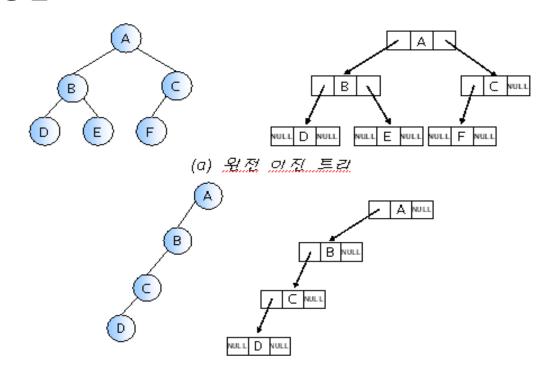
부모와 자식 인덱스 관계

- 노드 i의 부모 노드 인텍스 = i/2
- 노드 i의 왼쪽 자식 노드 인텍스 = 2*i
- 노드 i의 오른쪽 자식 노드 인텍스 = 2*i+1



링크 표현법

링크 표현법: 포인터를 이용하여 부모 노드가 자식 노드를 가리키게 하는 방법



(b) 경사 <u>이전</u> 트리

교과목 : 자료구조 24/69

링크의 구현

- 노드는 구조체로 표현
- 링크는 포인터로 표현

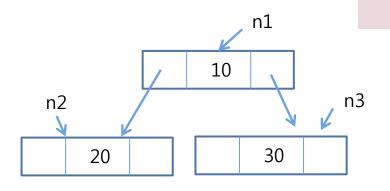
```
typedef struct TreeNode {
         int data;
         struct TreeNode *left, *right;
} TreeNode;
```

교과목 : 자료구조 25/69

링크 표현법 프로그램

26/69

```
#include <stdio.h>
#include <stdlib.h>
#include <memory.h>
typedef struct TreeNode {
         int data;
         struct TreeNode *left, *right;
} TreeNode;
void main()
 TreeNode *n1, *n2, *n3;
 n1= (TreeNode *)malloc(sizeof(TreeNode));
 n2= (TreeNode *)malloc(sizeof(TreeNode));
 n3= (TreeNode *)malloc(sizeof(TreeNode));
```



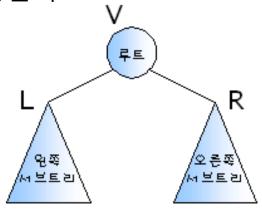
```
n1->data = 10;
n1->left = n2;
n1->right = n3;

n2->data = 20;
n2->left= n2->right = NULL;

n3->data = 30;
n3->left = n3->right = NULL;
}
```

이진 트리의 순회

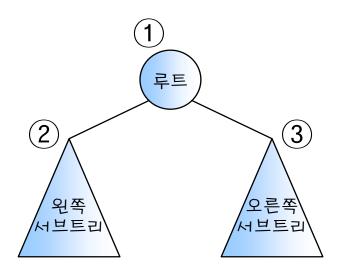
- 순회(traversal): 트리의 노드들을 체계적으로 방문하는 것
- 3가지의 기본적인 순회방법
 - ▶ 전위순회(preorder traversal) : VLR
 - ▶ 자손 노드보다 루트 노드를 먼저 방문한다.
 - ▶ 중위순회(inorder traversal) : LVR
 - ▶ 왼쪽 자손, 루트, 오른쪽 자손 순으로 방문한다.
 - ▶ 후위순회(postorder traversal): LRV
 - ▶루트 노드보다 자손을 먼저 방문한다.

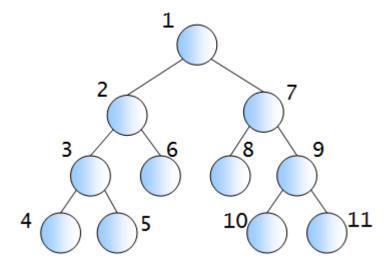


교과목 : 자료구조 27/69

전위 순회

- 1. 루트 노드를 방문한다
- 2. 왼쪽 서브트리를 방문한다
- 3. 오른쪽 서브트리를 방문한다

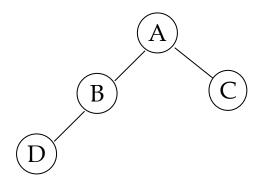




교과목 : 자료구조 28/69

전위순회 프로그램

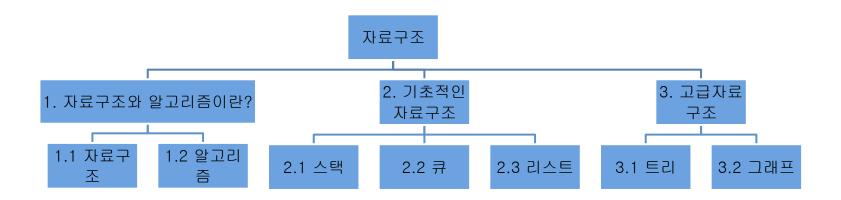
• 순환 호출을 이용한다.



교과목 : 자료구조 29/69

전위 순회 응용

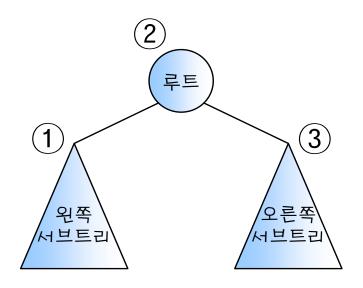
• (예) 구조화된 문서출력

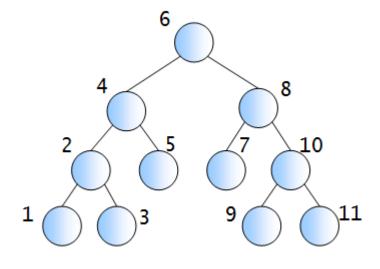


교과목 : 자료구조 30/69

중위 순회

- 1. 왼쪽 서브 트리를 방문한다
- 2. 루트 노드를 방문한다
- 3. 오른쪽 서브 트리를 방문한다

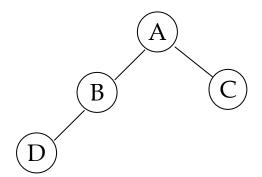




교과목 : 자료구조 31/69

중위순회 알고리즘

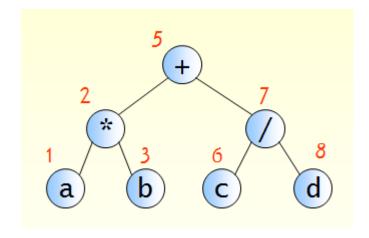
• 순환 호출을 이용한다.



교과목 : 자료구조 32/69

중위 순회 응용

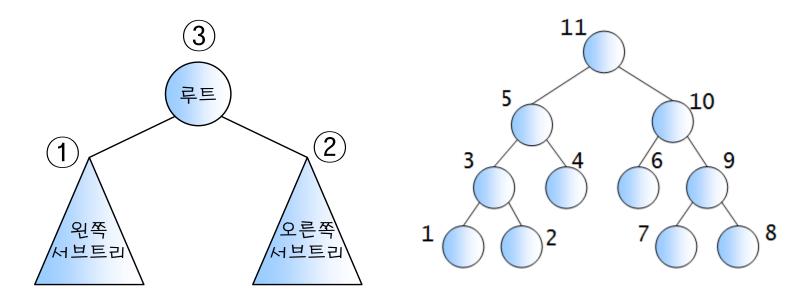
• (예) 수식 트리



교과목: 자료구조 33/69

후위 순회

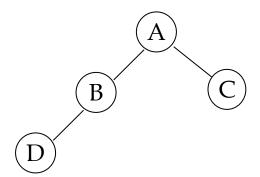
- 1. 왼쪽 서브트리를 방문한다
- 2. 오른쪽 서브트리를 방문한다
- 3. 루트 노드를 방문한다



교과목 : 자료구조 34/69

후위순회 알고리즘

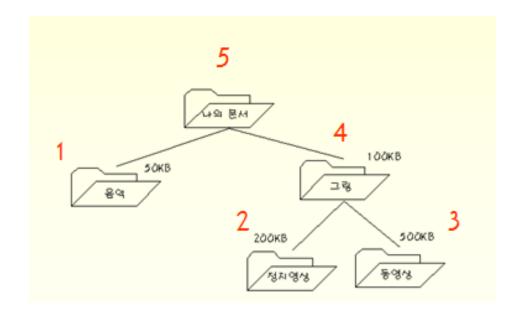
• 순환 호출을 이용한다.



교과목 : 자료구조 35/69

후위 순회 응용

• (예) 디렉토리 용량 계산



교과목 : 자료구조 36/69

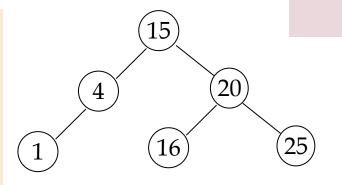
순회 프로그램

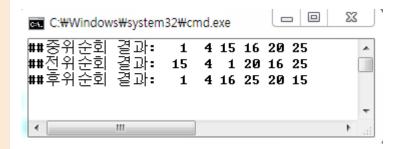
```
typedef struct TreeNode {
         int data;
         struct TreeNode *left, *right;
} TreeNode;
//
               15
                 16
                        25
TreeNode n1={1, NULL, NULL};
TreeNode n2={4, &n1, NULL};
TreeNode n3={16, NULL, NULL};
TreeNode n4={25, NULL, NULL};
TreeNode n5=\{20, &n3, &n4\};
TreeNode n6={15, &n2, &n5};
TreeNode *root= &n6;
```

```
// 중위 순회
void inorder( TreeNode *root ){
  if ( root ){
    inorder( root->left ); // 왼쪽서브트리 순회
    printf("%3d", root->data ); // 노드 방문
    inorder( root->right ); //오른쪽서브트리 순회
// 전위 순회
void preorder( TreeNode *root ){
  if (root){
    printf("%3d", root->data ); //노드 방문
    preorder( root->left ); //왼쪽서브트리 순회
    preorder( root->right ); //오른쪽서브트리 순회
```

순회 프로그램

```
// 후위 순회
void postorder( TreeNode *root ){
  if ( root ){
     postorder( root->left ); // 왼쪽서브트리 순회
     postorder( root->right ); // 오른쪽서브트리순회
     printf("%3d", root->data ); // 노드 방문
void main()
  printf("##중위순회 결과: "); inorder(root);
  printf("\n##전위순회 결과: "); preorder(root);
  printf("\n##후위순회 결과: "); postorder(root);
```

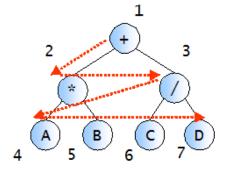


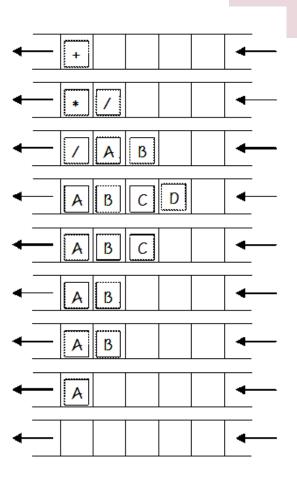


레벨 순회

■ 레벨 순회(level order)는 각 노드를 레벨 순으로 검 사하는 순회 방법

지금까지의 순회법이 스택을 사용했던 것에 비해 레벨 순회는 큐를 사용하는 순회법이다.





레벨 순회 알고리즘

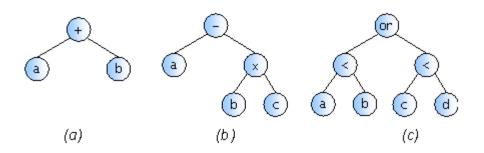
level_order(root)

```
    initialize queue;
    enqueue(queue, root);
    while is_empty(queue)≠TRUE do
    x← dequeue(queue);
    if(x≠NULL) then
    print DATA(x);
    enqueue(queue, LEFT(x));
    enqueue(queue, RIGHT(x));
```

교과목: 자료구조 40/69

수식 트리

- 수식트리: 산술식을 트리형태로 표현한 것
 - ▶ 비단말노드: 연산자(operator)
 - ▶ 단말노드: 피연산자(operand)
- 예)

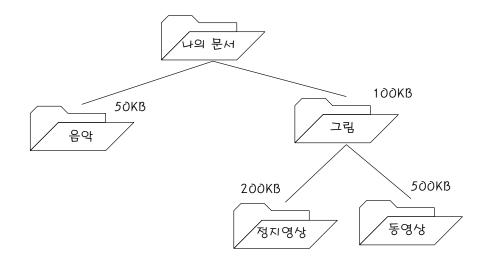


수식	a + b	a - (b × c)	(a < b) or (c < d)
전위순회	+ a b	- a × b c	or < a b < c d
중위순회	a + b	a - b × c	a < b or c < d
후위순회	a b +	abc×-	a b < c d < or

교과목 : 자료구조 41/69

디렉토리 용량 계산

• 디렉토리의 용량을 계산하는데 후위 트리 순회 사용



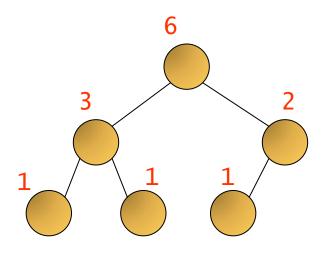
교과목 : 자료구조 42/69

디렉토리 용량 계산 프로그램

```
int calc_direc_size(TreeNode *root)
          int left_dir, right_dir;
          if (root){
                     left_size = calc_size( root->left );
                     right_size = calc_size(root->right);
                     return (root->data+left_size+right_size);
void main()
          TreeNode n4={500, NULL, NULL};
          TreeNode n5={200, NULL, NULL};
          TreeNode n3={100, &n4, &n5};
          TreeNode n2={50, NULL, NULL};
          TreeNode n1={0, &n2, &n3};
          printf("디렉토리의 크기=%d₩n",calc_direc_size(&n1));
```

이진 트리 연산: 노드 개수

- 탐색 트리안의 노드의 개 수를 계산
- 각각의 서브트리에 대하 여 순환 호출한 다음, 반환 되는 값에 1을 더하여 반 환

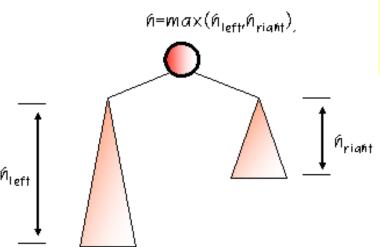


```
int get_node_count(TreeNode *node)
{
  int count=0;
  if( node != NULL )
    count = 1 + get_node_count(node->left)+
        get_node_count(node->right);
  return count;
}
```

교과목 : 자료구조 44/69

이진 트리 연산: 높이

 서브트리에 대하여 순환 호출하고 서브 트리들의 반환값 중에서 최대값을 구하여 반환

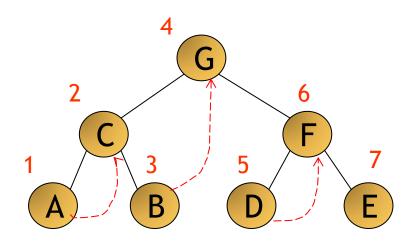


```
int get_height(TreeNode *node)
{
  int height=0;
  if( node != NULL )
    height = 1 + max(get_height(node->left),
        get_height(node->right));
  return height;
}
```

교과목: 자료구조 45/69

스레드 이진 트리

- 이진트리의 NULL 링크를 이용하여 순환호출 없이도 트리의 노드들을 순회
- NULL 링크에 중위 순회시에 후속 노드인 중위 후속자 (inorder successor)를 저장시켜 놓은 트리가 스레드 이 진 트리(threaded binary tree)



교과목 : 자료구조 46/69

스레드 이진 트리의 구현

 단말노드와 비단말노드의 구별을 위하여 is_thread 필드 필요

```
typedef struct TreeNode {
    int data;
    struct TreeNode *left, *right;
    int is_thread; //만약 오른쪽 링크가 스레드이면 TRUE
} TreeNode;
```

교과목 : 자료구조 47/69

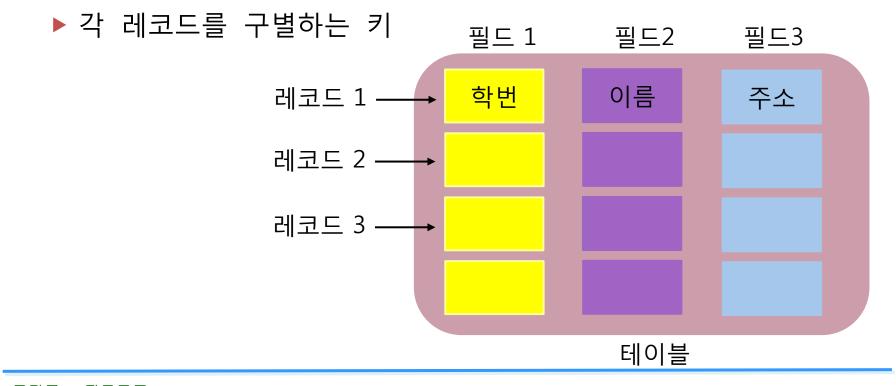
스레드 이진 트리의 구현

• 스레드 버전 중위 순회 함수 작성

48/69

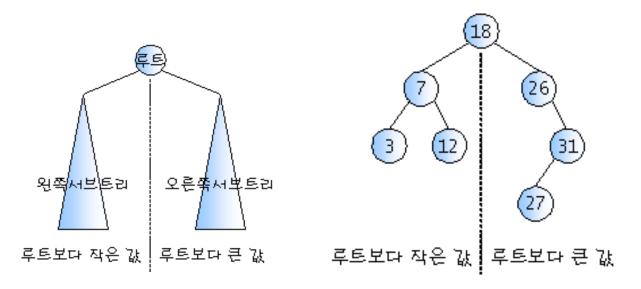
이진탐색트리

- 탐색작업을 효율적으로 하기 위한 자료구조
- 탐색은 레코드집합에서 특정한 레코드를 찾아내는 작업
- 주요키(primary key)



이진탐색트리

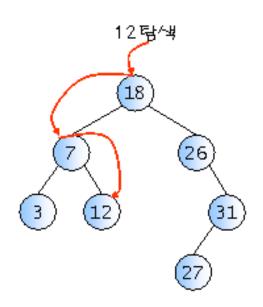
- 이진 탐색 트리의 정의
 - ▶ 모든 노드의 키는 유일하다.
 - ▶ 왼쪽 서브 트리의 키들은 루트의 키보다 작다.
 - ▶ 오른쪽 서브 트리의 키들은 루트의 키보다 크다.
 - ▶ 왼쪽과 오른쪽 서브 트리도 이진 탐색 트리이다.



교과목 : 자료구조 50/69

이진탐색트리에서의 탐색연산

- 비교한 결과가 같으면 탐색이 성공적으로 끝난다.
- 비교한 결과가, 주어진 키 값이 루트 노드의 키값보다 작으면 탐색은 이 루트 노드의 왼쪽 자식을 기준으로 다시 시작한다.
- 비교한 결과가, 주어진 키 값이 루트 노드의 키 값보다 크면 탐색은 이 루트 노드의 오른쪽 자식을 기준으로 다시 시작한다.



교과목 : 자료구조 51/69

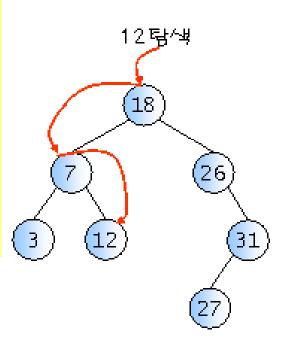
탐색을 구현하는 방법

- 순환적 방법
- 반복적 방법

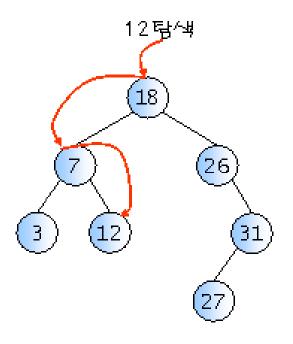
교과목: 자료구조 52/69

순환적인 방법

```
//순환적인 탐색 함수
TreeNode *search(TreeNode *node, int key)
{
    if ( node == NULL ) return NULL;
    if ( key == node->key ) return node; (1)
    else if ( key < node->key )
        return search(node->left, key); (2)
    else
        return sear ch(node->right, key); (3)
}
```

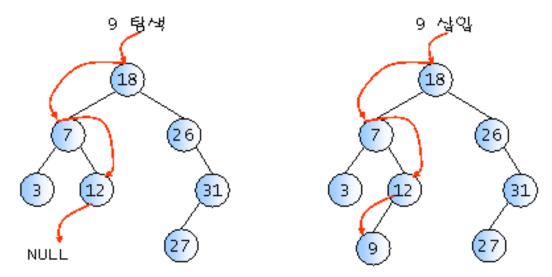


반복적인 방법



이진 탐색 트리에서의 삽입연산

- 이진 탐색 트리에 원소를 삽입하기 위해서는 먼저 탐색 을 수행하는 것이 필요
- 탐색에 실패한 위치가 바로 새로운 노드를 삽입하는 위 치



(a) 탐색을 먼저 수행

(b)탐색이 실패한 위치에 9를 섭임

교과목: 자료구조 55/69

이진 탐색 트리에서의 삽입연산

```
insert_node(⊤,z)
p←NULL;
t←root;
while t≠NULL do
 p←t;
 if z->key < p->key
    then t←p->left;
    else t←p->right;
if p=NULL
 then root←z;
               // 트리가 비어있음
 else if z->key < p->key
       then p->left<sup>←Z</sup>
       else p->right<sup>←</sup>Z
```

교과목 : 자료구조 56/69

이진 탐색 트리에서의 삽입연산

```
// key를 이진 탐색 트리 root에 삽입한다.
// key가 이미 root안에 있으면 삽입되지 않는다.
void insert_node(TreeNode **root, int key)
          TreeNode *p, *t; // p는 부모노드, t는 현재노드
          TreeNode *n: // n은 새로운 노드
          t = *root:
          p = NULL;
                                                       // key가 트리 안에 없으므로 삽입 가능
          // 탐색을 먼저 수행
                                                       n = (TreeNode *) malloc(sizeof(TreeNode));
          while (t != NULL){
                                                       if( n == NULL ) return;
            if( key == t->key ) return;
                                                       n->key = key; // 데이터 복사
                                                       n->left = n->right = NULL;
            p = t;
            if( key < t->key ) t = t->left;
                                                       // 부모 노드와 링크 연결
            else t = t->right;
                                                       if(p!= NULL)
                                                                 if( key < p->key )
                                                                           p->left = n;
                                                                 else p->right = n;
                                                       else *root = n:
```

- 3가지의 경우
 - 1. 삭제하려는 노드가 단말 노드일 경우
 - 2. 삭제하려는 노드가 하나의 왼쪽이나 오른쪽 서브 트리중 하나만 가지고 있는 경우
 - 3. 삭제하려는 노드가 두개의 서브 트리 모두 가지고 있는 경우
- CASE 1: 삭제하려는 노드가 단말 노드일 경우: 단말노드의 부모노드를 찾아서 연결을 끊으면 된다.

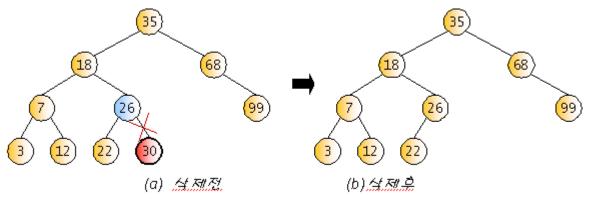


그림 7.42 이진탐색트라의 삭제연산: 삭제노트가 단말노트인 경우

교과목 : 자료구조 58/69

CASE 2:삭제하려는 노드가 하나의 서브트리만 갖고 있는 경우:삭제되는 노드가 왼쪽이나 오른쪽 서브 트리중 하나만 갖고 있을 때, 그노드는 삭제하고 서브 트리는 부모 노드에 붙여준다.

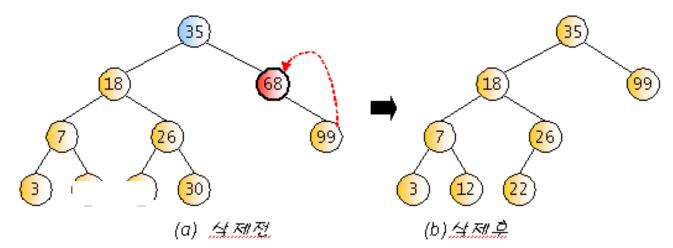
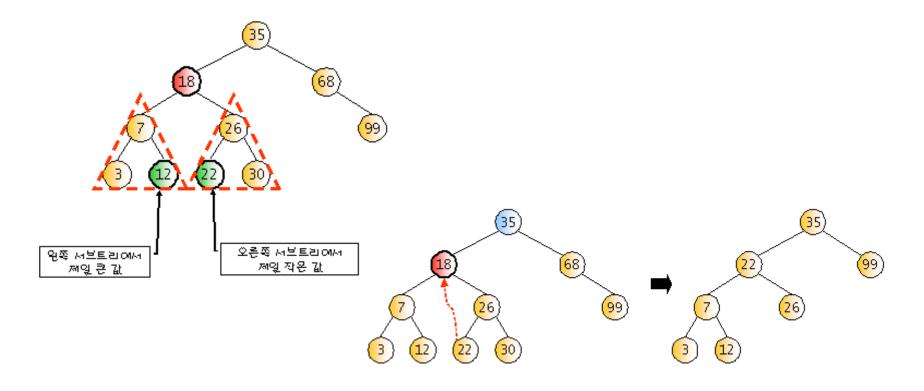


그림 7.43 이진탐색트리의 삭제면선: 삭제노드가 하나의 서보트리를 가지고 있는 경우

교과목: 자료구조 59/69

 CASE 3:삭제하려는 노드가 두개의 서브트리를 갖고 있는 경우: 삭제 노드와 가장 비숫한 값을 가진 노드를 삭제노드 위치로 가져온다.



교과목: 자료구조 60/69

```
// 삭제 함수
void delete_node(TreeNode **root, int key)
    TreeNode *p, *t, *child, *succ, *succ_p;
    // key를 갖는 노드 t를 탐색, p는 t의 부모노드
    p = NULL;
    t = *root;
    // key를 갖는 노드 t를 탐색한다.
    while( t != NULL && t->key != key ){
         p = t;
         t = ( key < t->key ) ? t->left : t->right;
    // 탐색이 종료된 시점에 t가 NULL이면 트리안에 key가 없음
    if(t == NULL) { // 탐색트리에 없는 키
          printf("key is not in the tree");
          return;
```

61/69

```
// 첫번째 경우: 단말노드인 경우

if( (t->left==NULL) && (t->right==NULL) ){

    if( p != NULL ){

        // 부모노드의 자식필드를 NULL로 만든다.

        if( p->left == t )

            p->left = NULL;

        else p->right = NULL;

    }

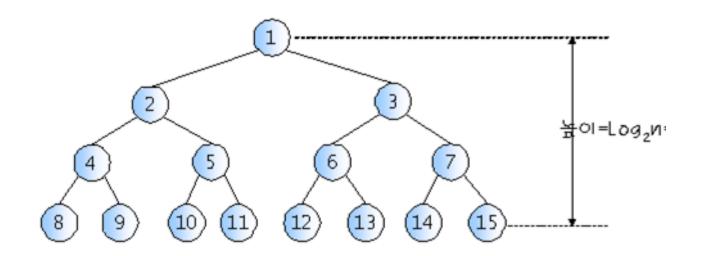
    else // 만약 부모노드가 NULL이면 삭제되는 노드가 루트

    *root = NULL;
}
```

```
// 세번째 경우: 두개의 자식을 가지는 경우
else{
          // 오른쪽 서브트리에서 후계자를 찾는다.
          succ_p = t;
          succ = t->right;
          // 후계자를 찾아서 계속 왼쪽으로 이동한다.
          while(succ->left != NULL){
                    succ_p = succ;
                    succ = succ->left;
          }
          // 후속자의 부모와 자식을 연결
          if( succ_p->left == succ )
             succ_p->left = succ->right;
          else
             succ_p->right = succ->right;
          t->key = succ->key; // 후속자가 가진 키값을 현재 노드에 복사
          t = SUCC; // 원래의 후속자 삭제
free(t);
```

이진 탐색 트리의 성능분석

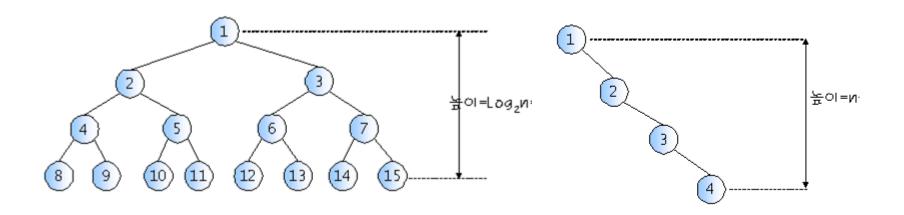
 이진 탐색 트리에서의 탐색, 삽입, 삭제 연산의 시간 복잡도는 트리의 높이를 h라고 했을때 h에 비례한다



교과목: 자료구조 64/69

이진 탐색 트리의 성능분석

- 최선의 경우
 - ▶ 이진 트리가 균형적으로 생성되어 있는 경우
 - \blacktriangleright h=log₂n
- 최악의 경우
 - ▶ 한쪽으로 치우친 경사이진트리의 경우
 - ► h=n
 - ▶ 순차탐색과 시간복잡도가 같다.



교과목 : 자료구조 65/69

이진 탐색 트리 응용: 영어사전구현

- 이진 탐색 트리를 이용하여 다음과 같은 메뉴의 기능을 제공하는 영어 사전을 구현해 보자.
- 이진 탐색 트리에 대한 연산
 - ▶입력
 - ▶ 삭제
 - ▶ 탐색
 - ▶ 사전 전체 출력(중위순회)
- 레코드 구조

이진 탐색 트리 응용: 영어사전구현

• 실행 결과

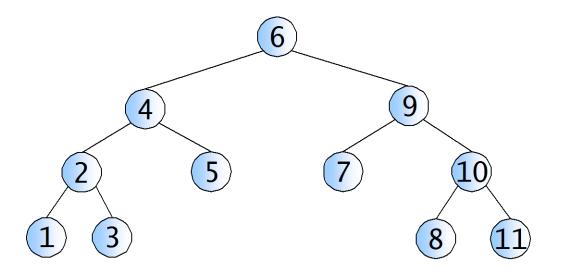
```
******
  사제
탐색
출력
종료
********
단어:suspect
의미:~일 것 같다고 생각하다.
  입력
삭제
탐색
출종
료
*****
단어:ascertain
희미:확인하다
******
a: 마시
s: 탐색
p: 출력
q: 종료
******
```

```
단어:exception
의미:예외
*****
i: 입학제생
d: 학자생
p: 출종
q:
********
((ascertain(exception))suspect)
*********
i: 입력
d: 삭제
s: 탐색
p: 출종
q: 종
********
단어:exception
*****
i: 입력
d: 삭제
s: 타색
p: 출력
q: 종료
********
((ascertain)suspect)
```

i: 입력 d: 삭제 s: 다색 p: 출종 q: ******** . 단어:**suspect** 의미:~일 것 같다고 생각한다. ***** i: 입력 d: 삭제 s: 탐색 p: 출력 q: 종료 ****** 단어:exception ****** d: 삭제 s: 탐색 p: 출목 q: ******

연습문제 1

- 다음의 이진 트리에 대해 다음 질문에 답하시오.
 - ▶ 전위, 후위, 중위, 레벨 순회한 결과는?



연습문제 2

- 다음의 전위 순회와 중위 순회 결과를 생성할 수 있는 이 진 트리를 그려보시오.
 - ▶ 전위 순회: A, B, D, E, C, F, G, H
 - ▶ 중위 순회: E, D, B, A, G, F, H, C

연습문제 3

다음 순서로 자료가 입력되었다고 가정하여 이진 탐색 트리를 생성하시오.

5,3,7,2,4,6,9,1,8,10