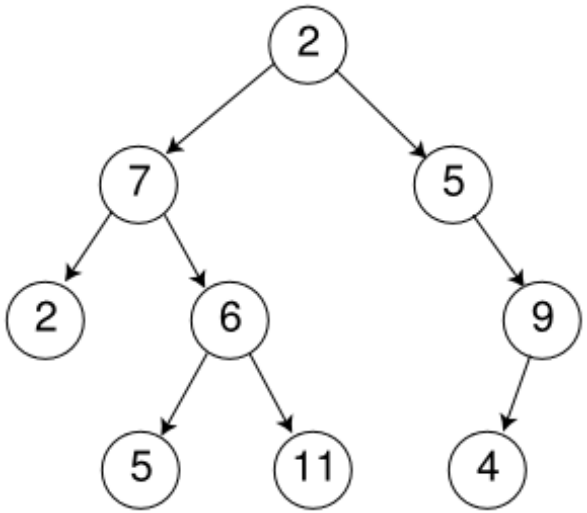
Warming-Up for Data Structures

Fall, 2006

Seoul National University
Internet Database Lab.

Introduction [1]

- In computer science, a data structure is a way of storing data in a computer memory so that it can be used efficiently
- Examples: lists, arrays, stacks, queues, trees, graphs, etc.



Binary tree, a simple
type of linked data structure

SNU Internet DataBase Lab.

Introduction [2]

- Importance of Data Structures
  - Different kinds of DSs are suited to different kinds of applications
    - e.g. B-trees are well-suited for implementation of databases
  - In the design of many types of programs, the choice of appropriate DSs is crucial
  - A carefully chosen DS ⇒ a more efficient algorithm

- 컴퓨터를 이용한 문제 해결 과정
  - 문제 정의와 분석
  - 자료구조의 설계
  - 알고리즘 고안
  - 프로그램 작성

SNU Internet DataBase Lab.

알고리즘 I

100만명을 대상으로 각자가 낸 납세액이 전체 납세액에서 차지하는 비율을 구하는 문제

1. 100만 명의 납세액을 입력 받는다. (1초)

2. 100만면중 첫번째 대상자의 납세액을 읽어 온다. (1/100만 초)

3. 100만 명의 납세액 총액을 구한다.
100만 * 1/100만 초+(100만 - 1)*1/100만 초 = 2 - 1/100만 초

4. 2의 값을 총합으로 나누어 납세 비중을 구한다. (1/100만 초)

5.아직 남은 대상자가 있으면 2~4의 과정을 반복한다.

총 소요시간 : 1+ (2 + 1/100만) * 100만 = 약 200만 2초= 555시간

SNU Internet DataBase Lab.

알고리즘 2

100만명을 대상으로 각자가 낸 납세액이 전체 납세액에서 차지하는 비율을 구하는 문제

1. 100만 명의 납세액을 입력 받는다. (1초)

2. 100만 명의 납세액 총액을 구한다.
100만 * 1/100만 초+(100만 - 1)*1/100만 초=2 - 1/100만 초

3. 100만명중 첫번째 대상자의 납세액을 읽어 온다. (1/100만 초)

4. 3의 값을 2에서 계산한 값으로 나누어 납세 비중을 구한다. (1/100만 초)

5.아직 남은 대상자가 있으면 3~4의 과정을 반복한다.

총 소요시간 : 1+ 2 – 1/100만 + (1/100만+1/100만)* 100만= 약 5초

SNU Internet DataBase Lab.

Performance Analysis

- The amount of computer memory and time needed to run a program
  - Space Complexity
    - The amount of memory it needs to run to completion
  - Time Complexity
    - The amount of computer time it needs to run to completion
- Good data structures must guarantee low complexities!

Data Structure Text Book

- Chapter 1: Java review
- Chapter 2-4: Complexity of algorithms
- Chapter 5-8: Linear List
- Chapter 9-11: Stack & Queue
- Chapter 12-16: Tree
- Chapter 17: Graph
- Chapter 18-22: Algorithm-Design Methods

SNU Internet DataBase Lab.

Linear Lists

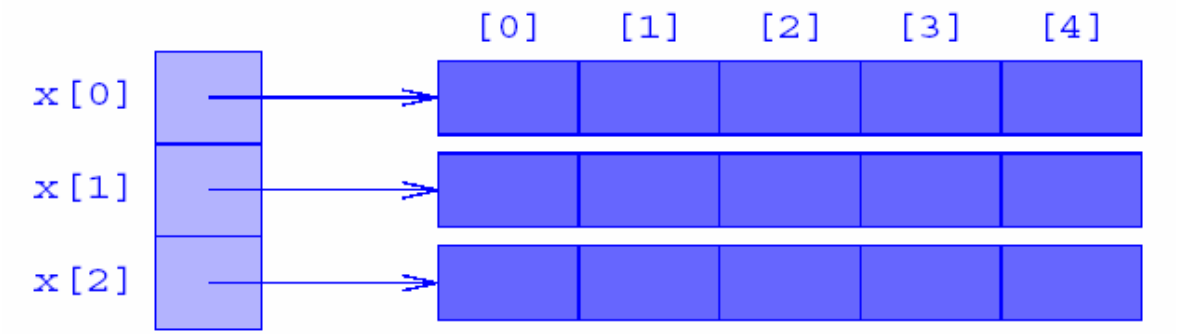
- Linear List (Ordered List)
  - An ordered collection of elements
  - Examples
    - 명단: (홍길동, 무대리, 임꺽정, …)
    - 성적: (98, 92, 90, 82, 77, …. 34)
    - A list of gold-medal winners in the Olympics ordered by year

Arrays

- Representations of a Multidimensional Array

0	1	2	3	4	5
6	7	8	9	10	11
12	13	14	15	16	17

(a) Row-major mapping



Memory structure for a two-dimensional array

Matrices

- Matrices
  - An *m* X *n* matrix is a table with *m* rows and *n* columns
  - *m*, *n*: dimensions of the matrix

	country			
asset	A	B	C	D
platinum	2	5	1	0
gold	6	2	3	8
silver	0	10	50	30

(a) asset

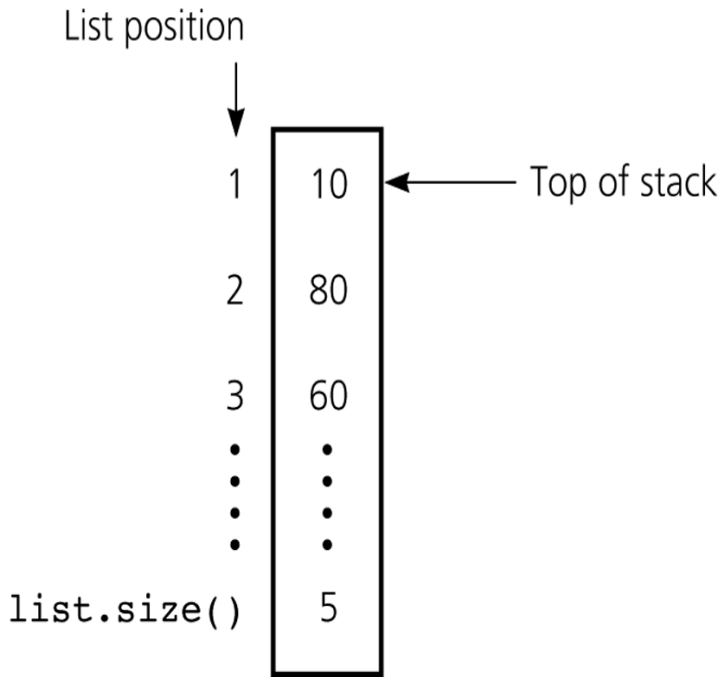
	scenario		
asset	1	2	3
platinum	20	15	50
gold	15	12	40
silver	1	1	2

(b) value

Asset and value matrices

Stacks

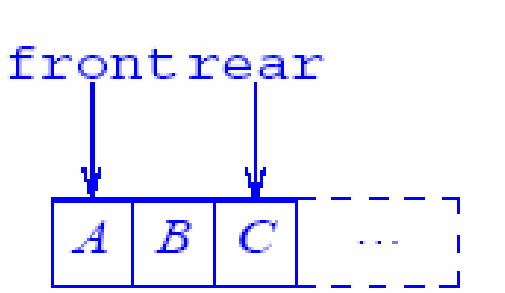
- A linear list in which insertions (called *pushes*) and removals (called *pops*) take place at the same end
- LIFO (Last In, First Out)

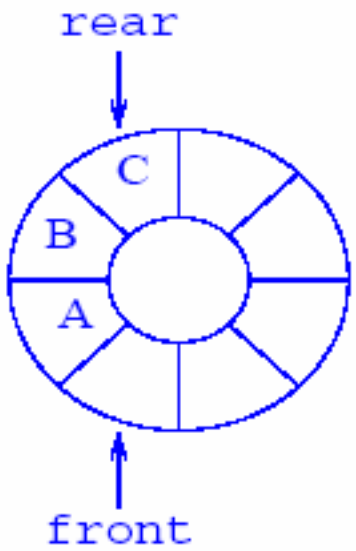


SNU Internet DataBase Lab.

Queues

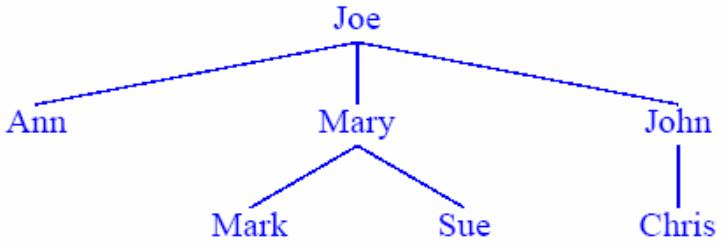
- A linear list in which insertions and deletions take place at different ends
- New elements are added at the *rear*; deleted, at the *front*
- FIFO (First In, First Out)



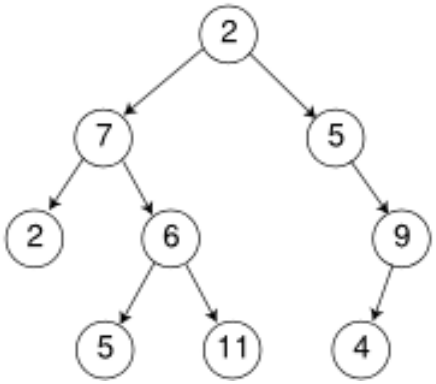


Various Trees (1)

- Tree
  - A special case of a graph which represents hierarchical data
  - root node, parent node, child node, etc.

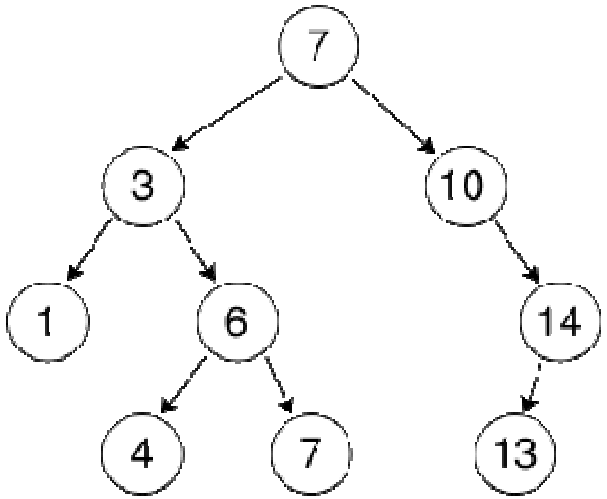


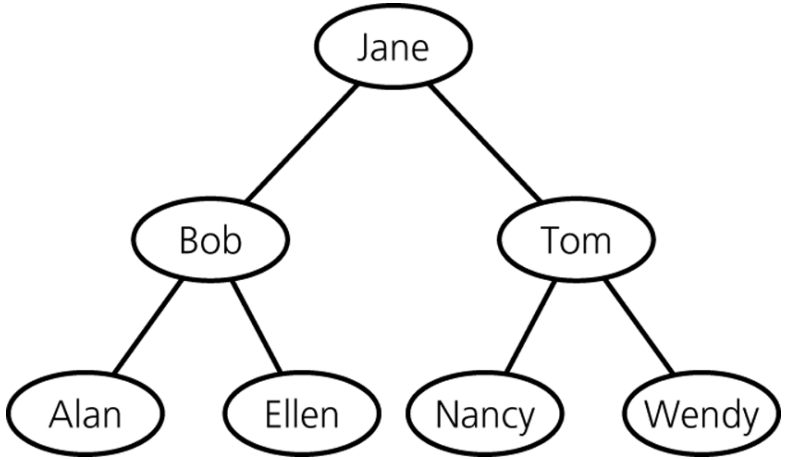
- Binary Tree
  - A tree in which each node has at most two children
  - e.g. a simple binary tree of size 9 and height 3 with a root node whose value is 2



Various Trees (2)

- Binary Search Tree: A binary tree in which
  - the left subtree of a node contains only values less than or equal to the node's value
  - the right subtree of a node contains only values greater than or equal to the node's value





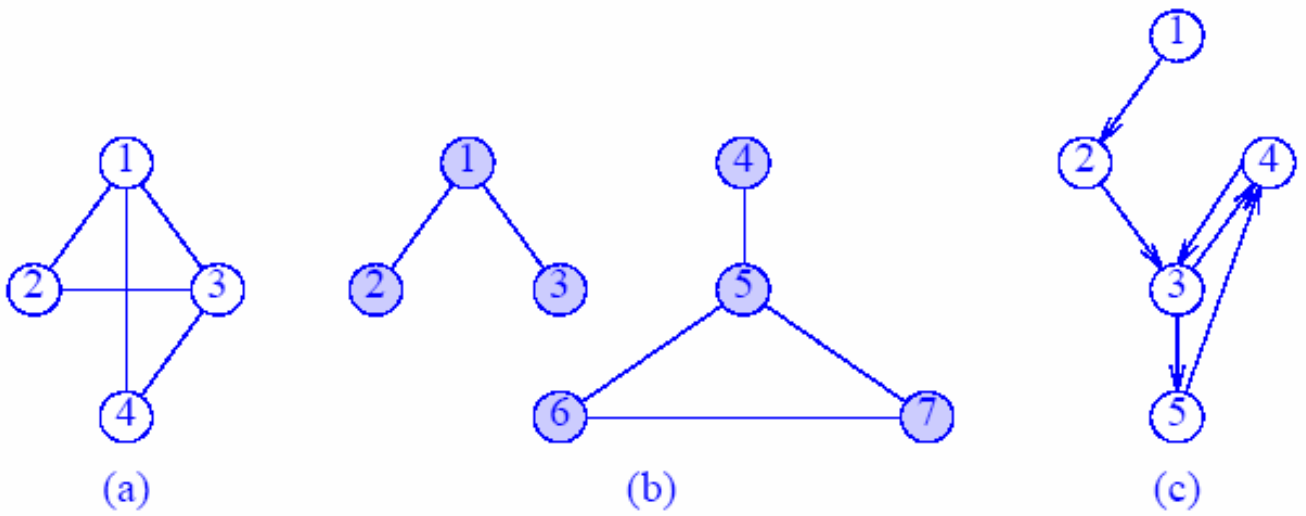
SNU Internet DataBase Lab.

Various Trees (3)

- AVL tree
- Balanced Tree
- Tournament Tree
- …….

Graphs [1/2]

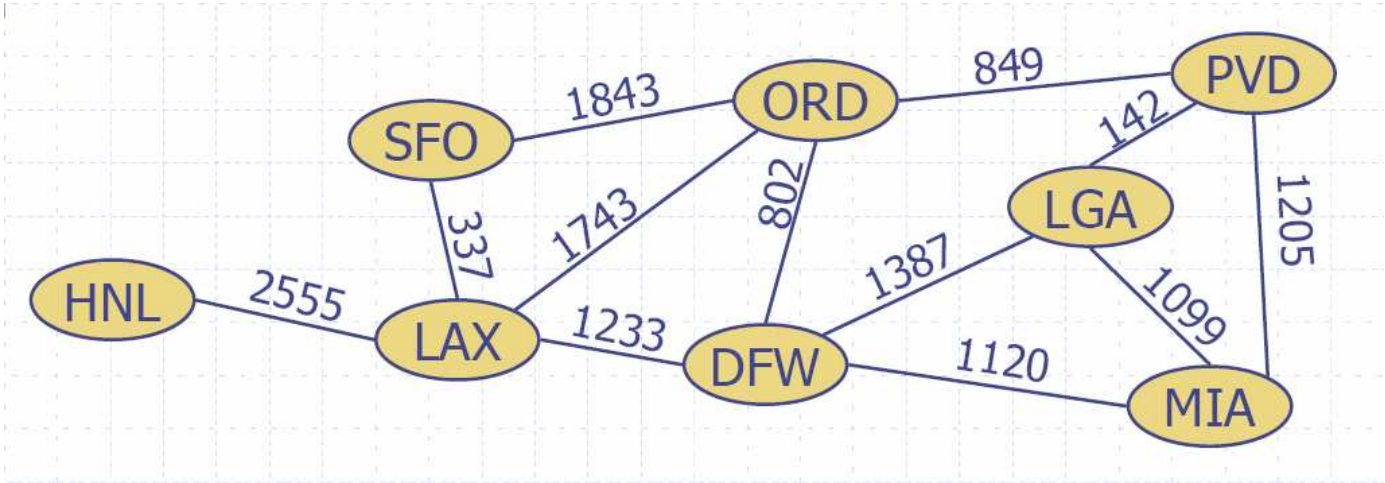
- An ordered pair of finite sets of V and E
  - V: vertices (nodes)
  - E: edges (arcs)



SNU Internet DataBase Lab.

Graphs [2/2]

- Examples
  - A vertex represents an airport and stores the three-letter airport code
  - An edge represents a flight route between two airports and stores the mileage of the route



Algorithm-Design Methods

- Algorithm
  - A procedure for accomplishing some task which, given an initial state, will terminate in a defined end-state
  - Design of a good algorithm is crucial to solving problems
- A Variety of Algorithms
  - The Greedy Method
  - Divide and Conquer
  - Dynamic Programming
  - Backtracking
  - Branch and Bound

Good Luck in Data Structures!