Search

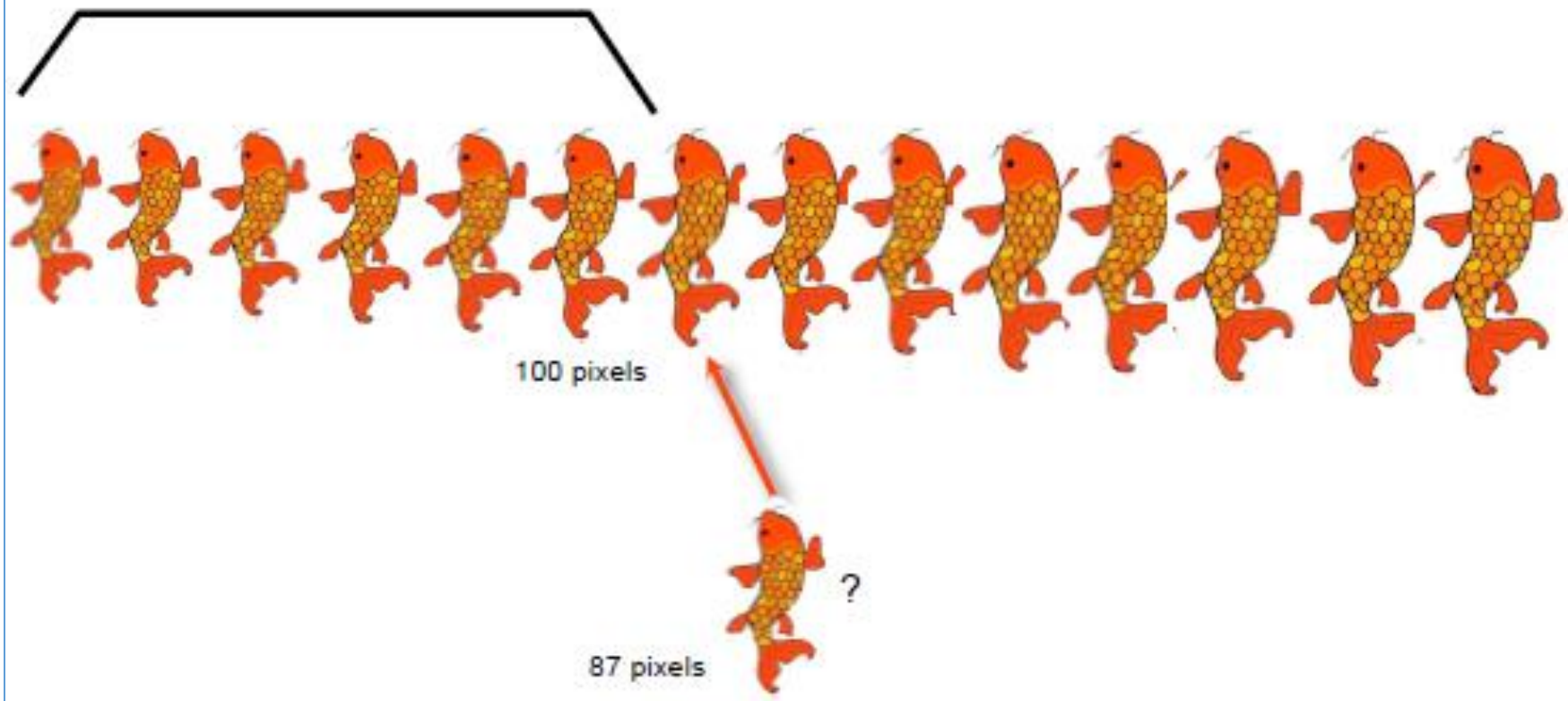
- Binary Search
- Binary Tree
- Binary Search Tree

Number guessing

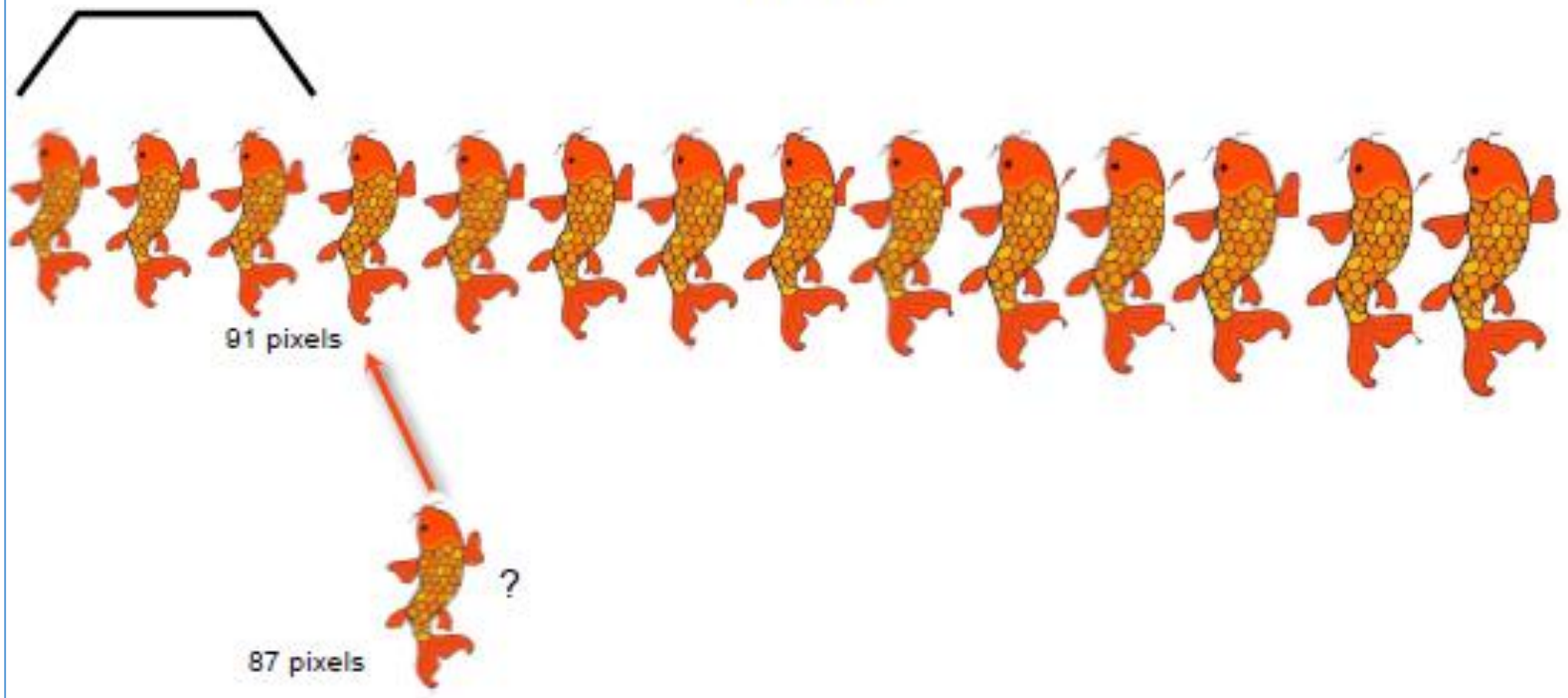
- I'm thinking of a number between 1 and 16
- You get to ask me, yes or no, is it greater than some number you choose
- How many questions do you need to ask?
- Which questions will you ask to get the answer quickest?

BINARY SEARCH

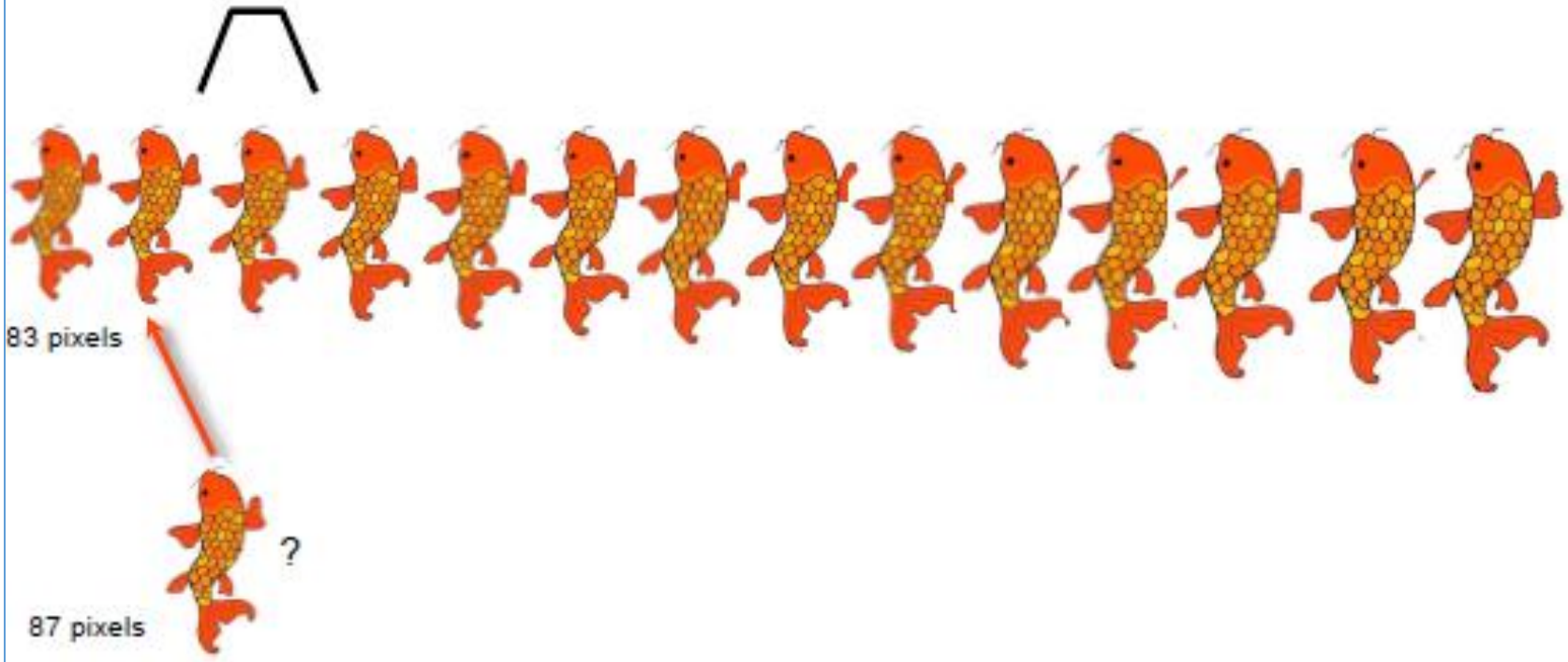
Binary Search in an Ordered List



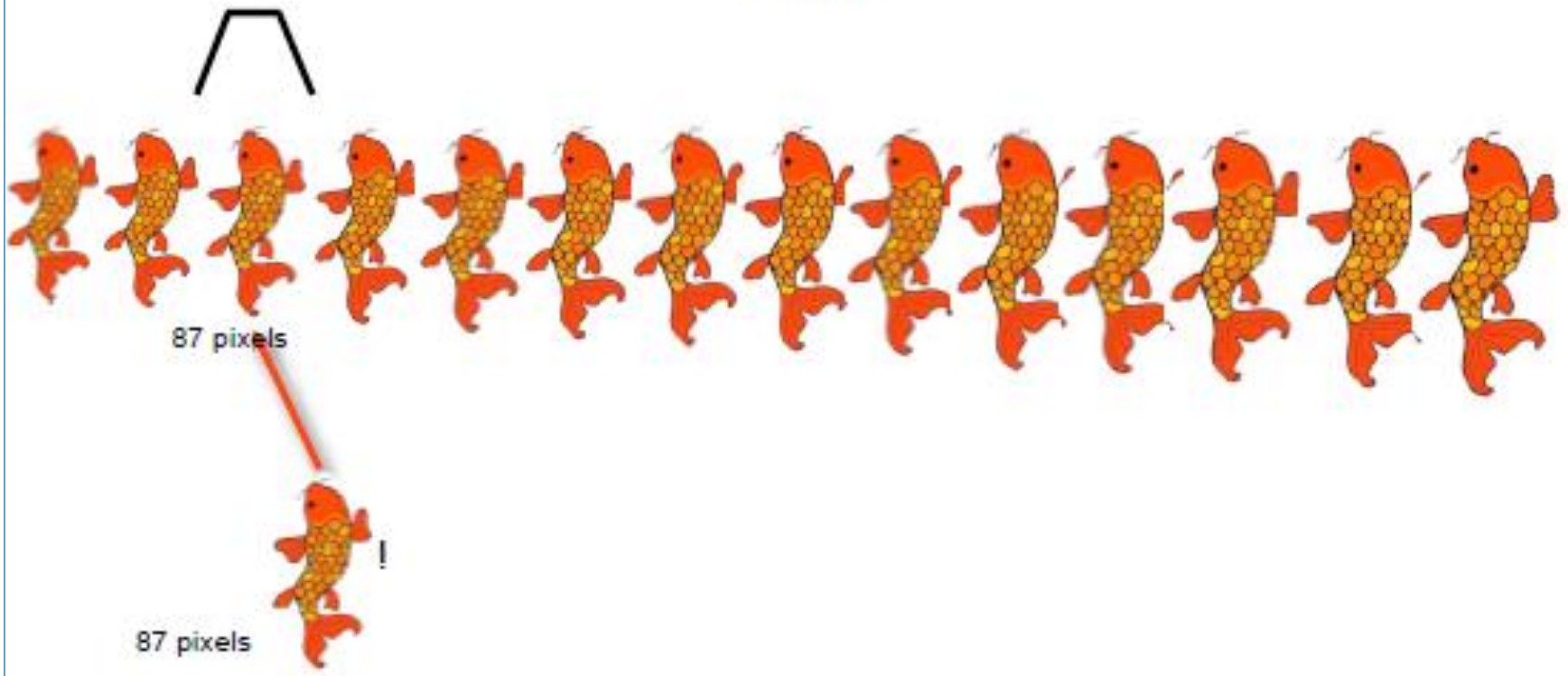
Binary Search in an Ordered List



Binary Search in an Ordered List



Binary Search in an Ordered List



Specification: the Search Problem

- Input: A list of n unique elements and a key to search for
  - **The elements are sorted in increasing order.**
- Result: The index of an element matching the key, or `None` if the key is not found.

Recursive Algorithm

BinarySearch(list, key):

1. Return None if the list is empty.
2. Compare the key to the middle element of the list
3. Return the index of the middle element if they match
4. If the key is less than the middle element then
   return BinarySearch(first half of list,key)
   Otherwise, return BinarySearch(second half of list,key).

Example 1: Search for 73

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	25	32	37	41	48	58	60	66	73	74	79	83	91	95

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	25	32	37	41	48	58	60	66	73	74	79	83	91	95

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	25	32	37	41	48	58	60	66	73	74	79	83	91	95

Found: return 9

Example 2: Search for 42

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	25	32	37	41	48	58	60	66	73	74	79	83	91	95

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	25	32	37	41	48	58	60	66	73	74	79	83	91	95

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	25	32	37	41	48	58	60	66	73	74	79	83	91	95

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	25	32	37	41	48	58	60	66	73	74	79	83	91	95

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	25	32	37	41	48	58	60	66	73	74	79	83	91	95

Not found: return None

Controlling the range of the search

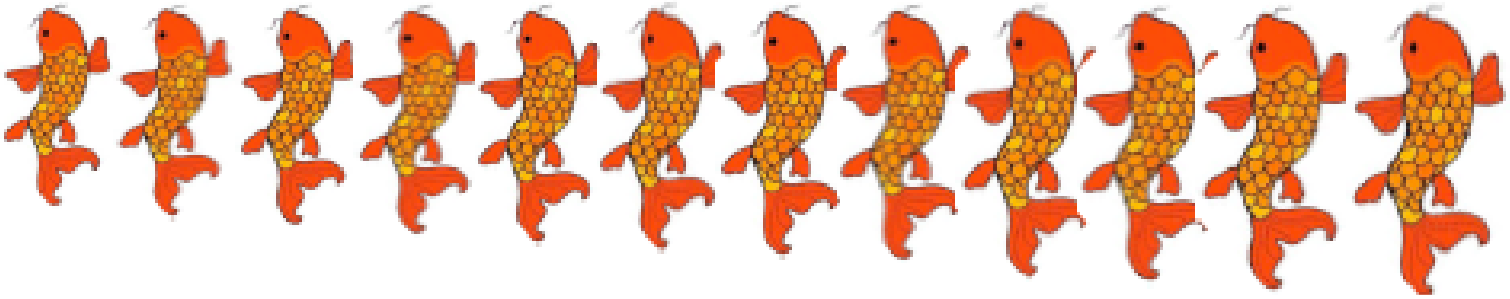
- Maintain three numbers: *lower*, *upper*, *mid*
- Initially *lower* is -1, *upper* is length of the list

lower = -1

upper = 12

0

11



Controlling the range of the search

- *mid* is the midpoint of the range:

  mid = (*lower* + *upper*) // 2 (**integer** division)

  Example: *lower* = -1, *upper* = 9

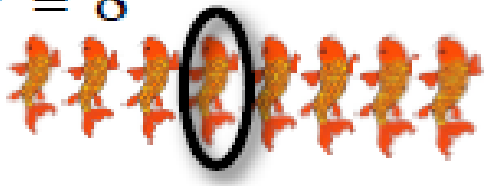
  (range has 9 elements)

  mid = 4

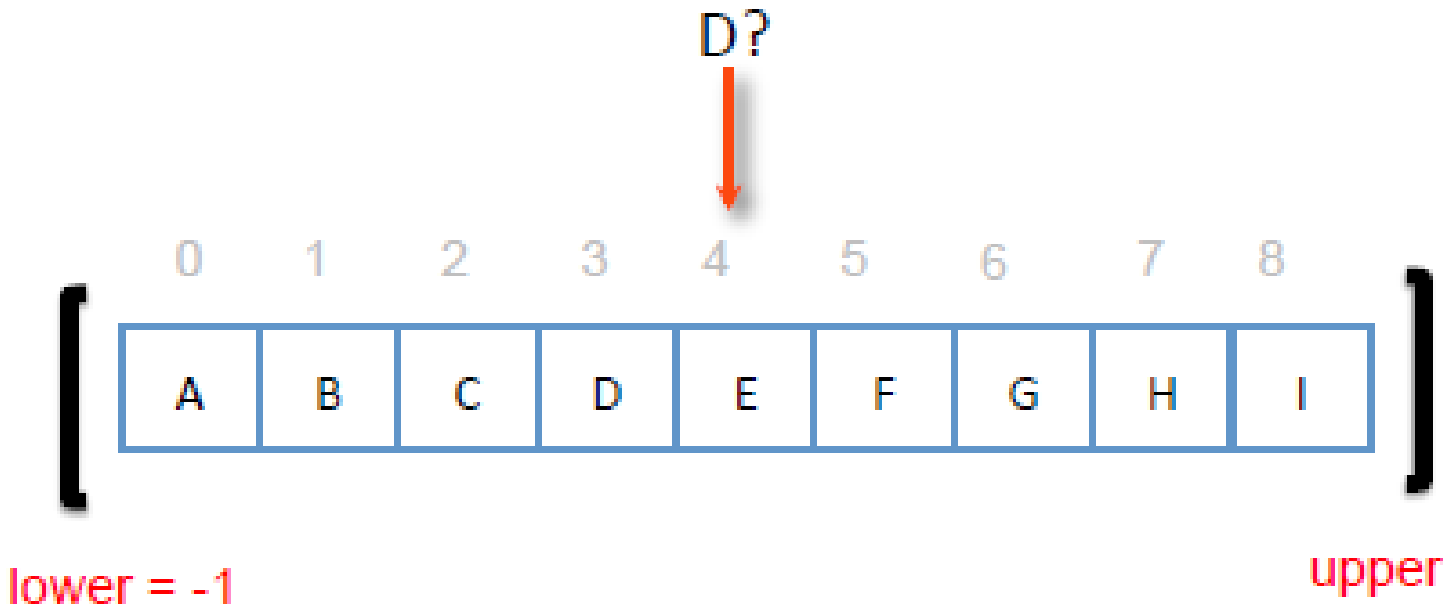
- What happens if the range has an even number of elements?

  Example: *lower* = -1, *upper* = 8

  mid = 3

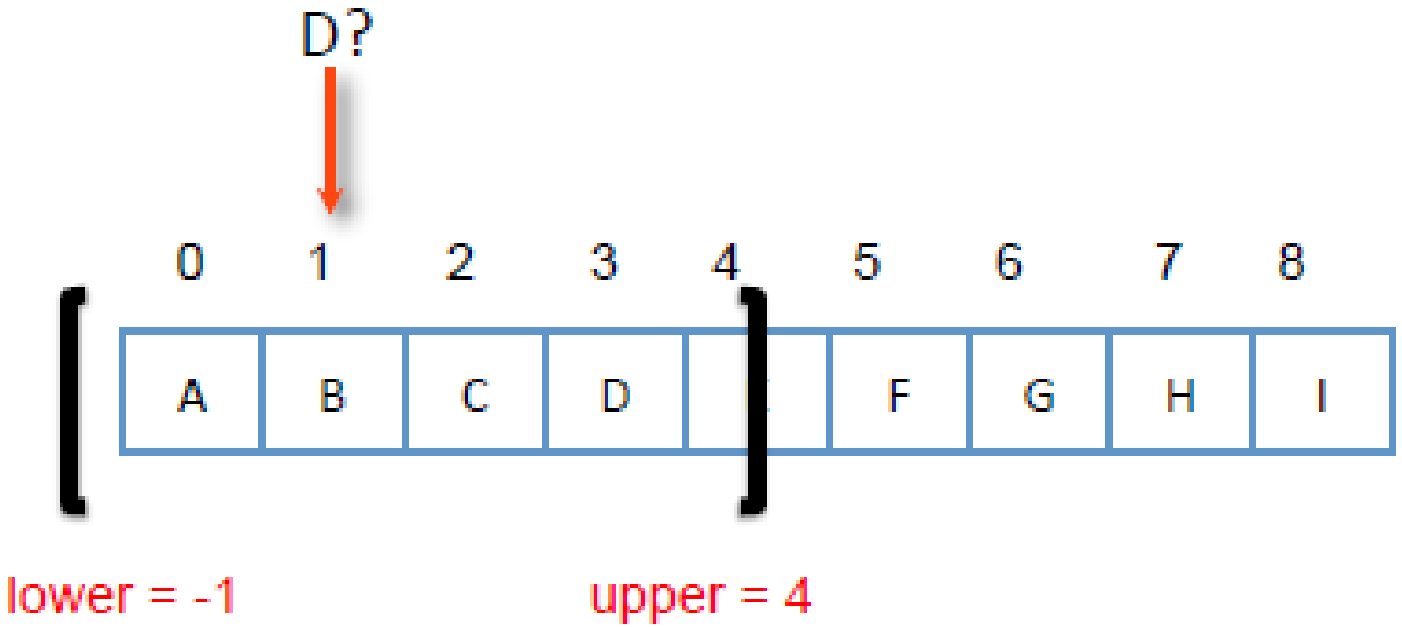


Example



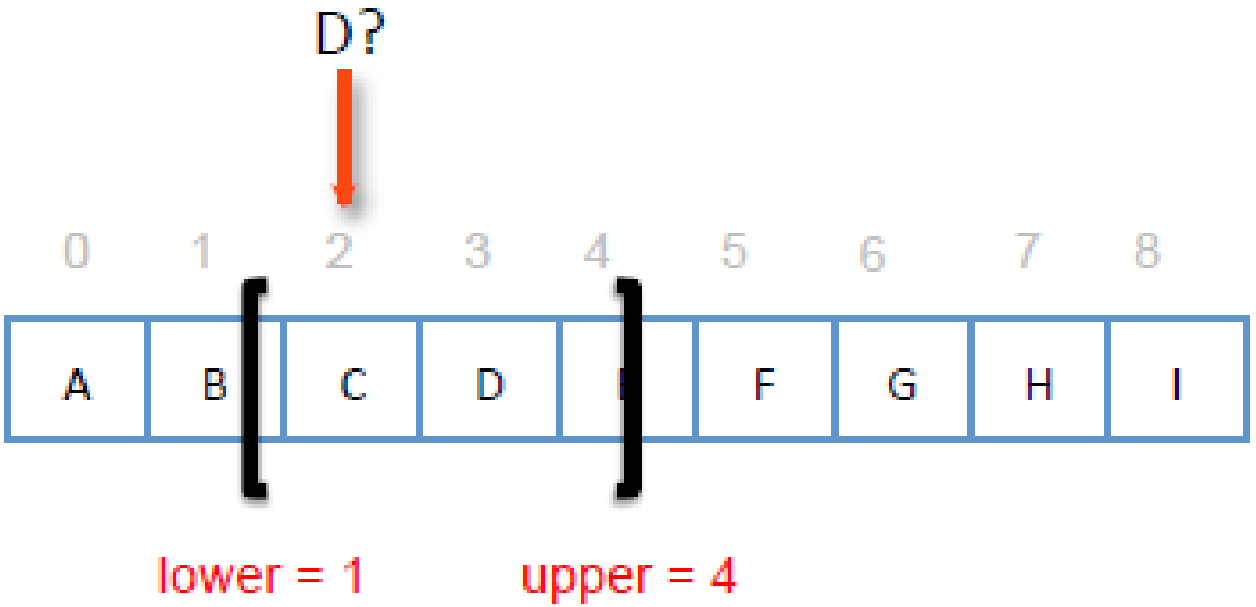
List sorted in ascending order.
Suppose we are searching for D.

Example



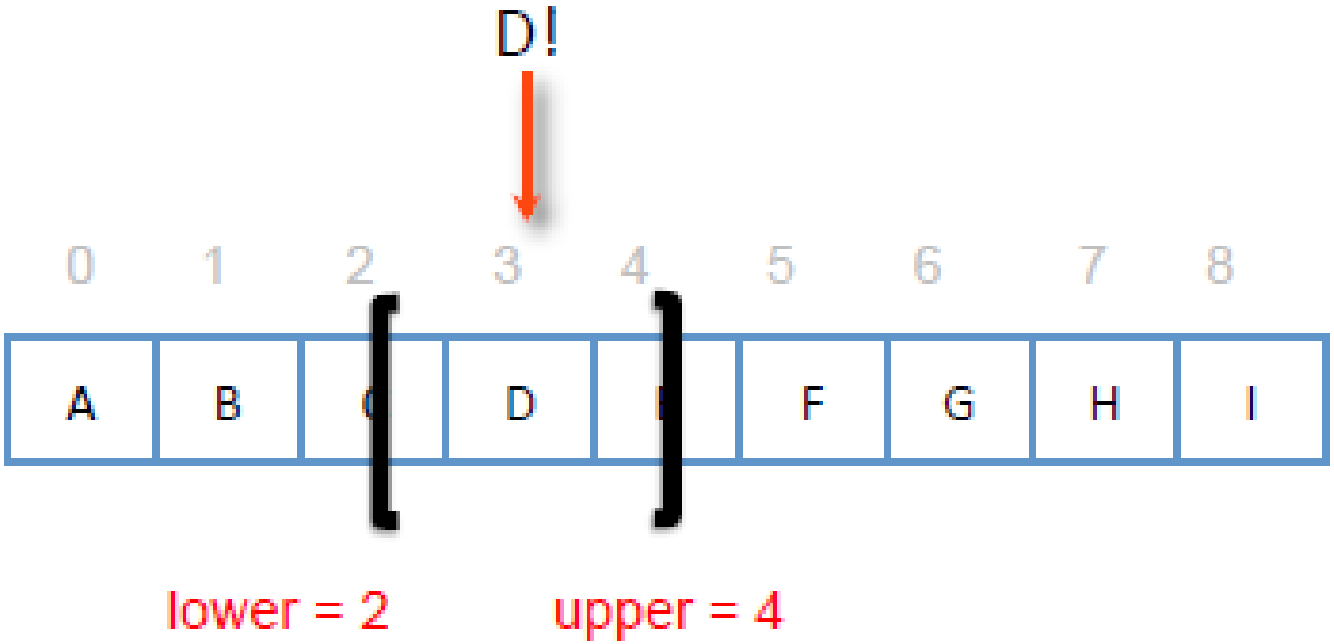
Each time we look at a smaller portion of the list within the window and ignore all the elements outside of the window

Example



Each time we look at a smaller portion of the array within the window and ignore all the elements outside of the window

Example



Each time we look at a smaller portion of the array within the window and ignore all the elements outside of the window

Recursive Binary Search in Python

first value for *upper*

first value for *lower*

```
# main function
def bsearch(items, key):
    return bs_helper(items, key, -1, len(items))

# recursive helper function
def bs_helper(items, key, lower, upper):
    if lower + 1 == upper: # Base case: empty
        return None
    mid = (lower + upper) // 2 # Recursive case
    if key == items[mid]:
        return mid
    if key < items[mid]: # Go left
        return bs_helper(items, key, lower, mid)
    else:                # Go right
        return bs_helper(items, key, mid, upper)
```

same value for *lower*

new value for *upper*

new value for *lower*

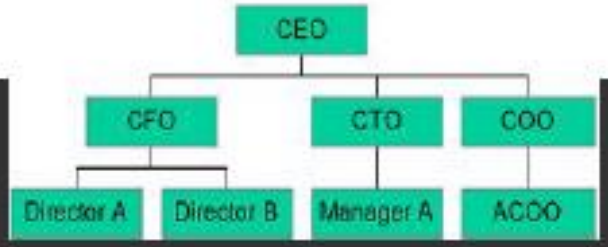
same value for *upper*

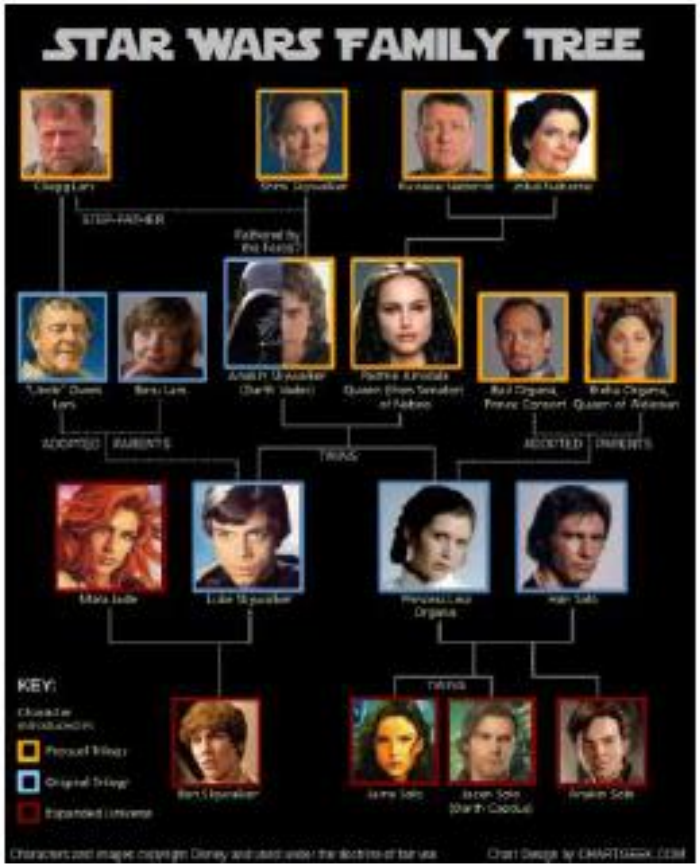
Search

- Binary Search
- Binary Tree
- Binary Search Tree

Hierarchical Data

NomE Inc.

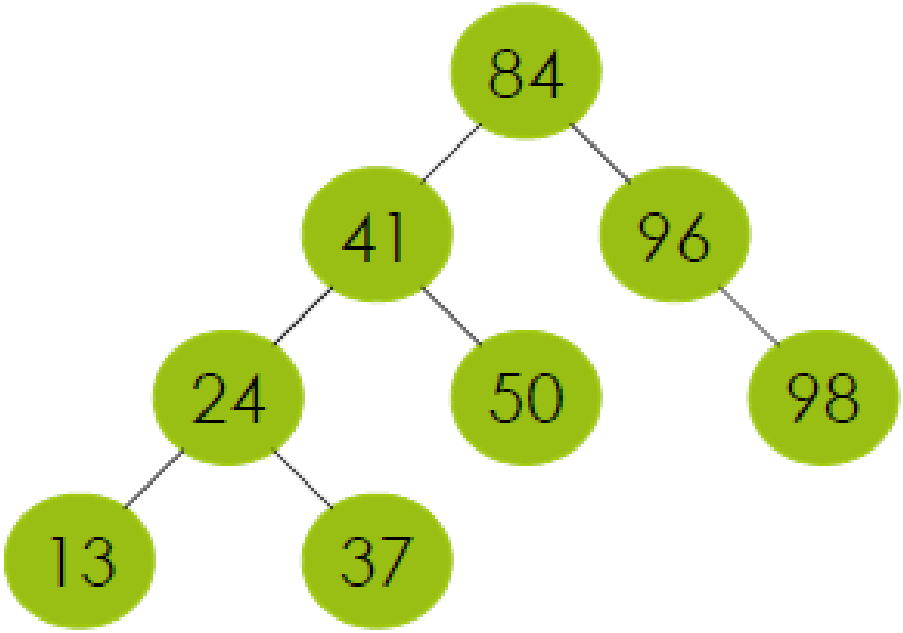




Trees

- A **tree** is a hierarchical data structure.
  - Every tree has a **node** called the **root**.
  - Each node can have 1 or more nodes as **children**.
  - A node that has no children is called a **leaf**.
- A common tree in computing is a **binary tree**.
  - A binary tree consists of nodes that have at most 2 children.
- Applications: data compression, file storage, game trees

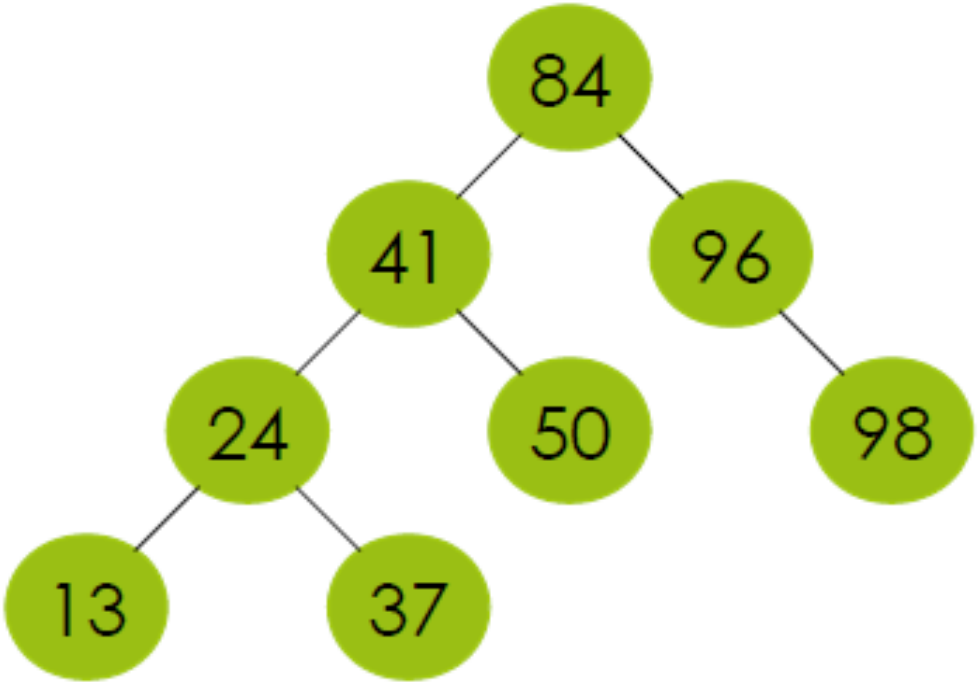
Binary Tree



In order to illustrate main ideas we label the tree nodes with the keys only.
In fact, every node would also store the rest of the data associated with that key. Assume that our tree contains integers keys.

Which one is the root?
Which ones are the leaves (external nodes)?
Which ones are internal nodes?
What is the height of this tree?

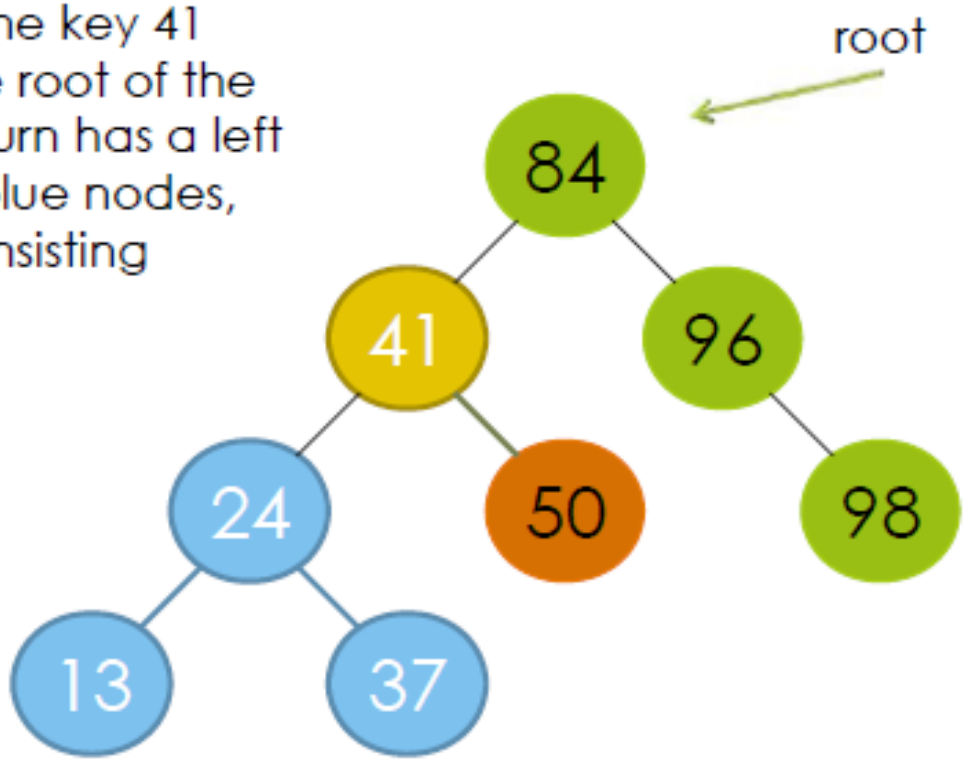
Binary Tree



The root contains the data value 84.
There are 4 leaves in this binary tree: nodes containing 13, 37, 50, 98.
There are 3 internal nodes in this binary tree: nodes containing 41, 96, 24
This binary tree has height 3 – considering root is at level 0,
the maximum level among all nodes is 3

Binary Tree

Note the **recursive** structure:
The yellow node with the key 41 can be viewed as the root of the left subtree , which in turn has a left subtree consisting of blue nodes, and a right subtree consisting of orange nodes.

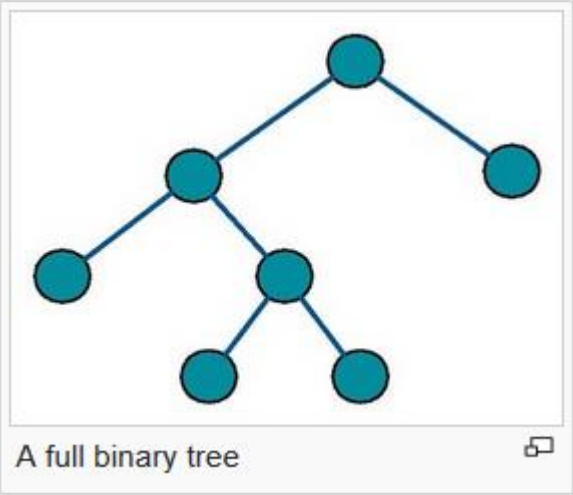


Terminology in Binary Tree

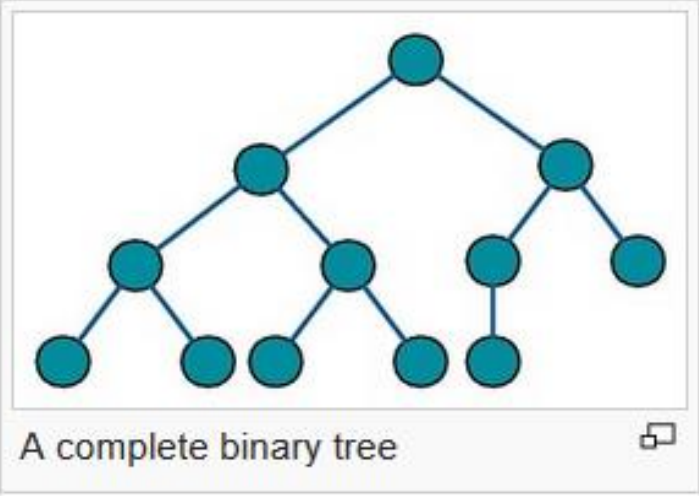
- A **rooted** binary tree has a root node and every node has at most two children.

- A **full** binary tree (sometimes referred to as a **proper** or **plane** binary tree) is a tree in which every node in the tree has either 0 or 2 children

- In a **complete** binary tree every level, *except possibly the last*, is completely filled, and all nodes in the last level are as far left as possible. It can have between 1 and 2^h nodes at the last level h.



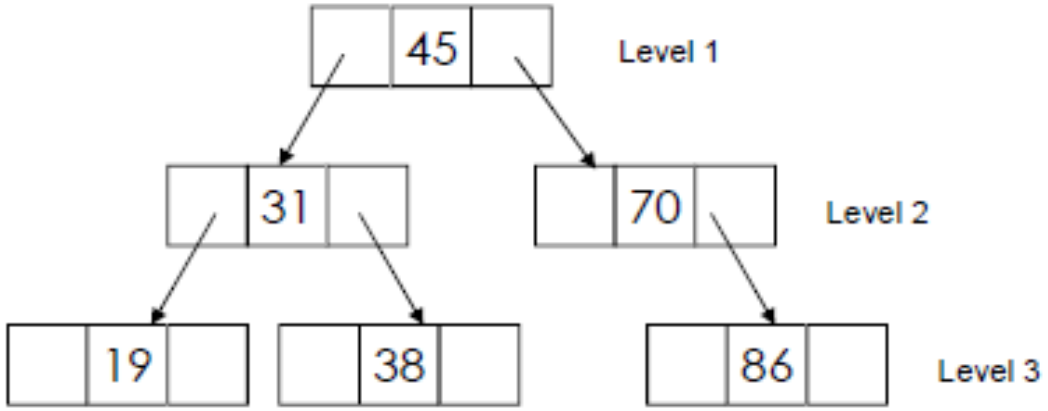
A full binary tree



A complete binary tree

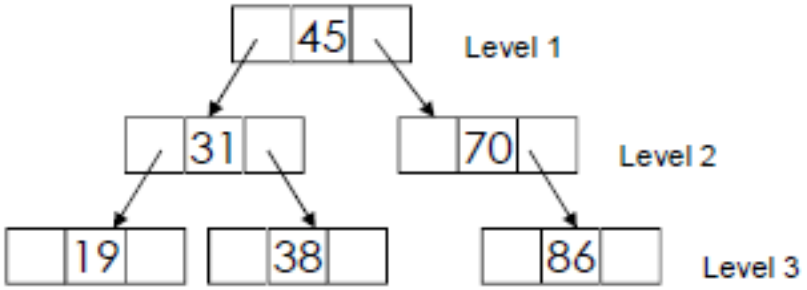
Binary Trees: Implementation

- One common implementation of binary trees uses nodes like a linked list does.
  - Instead of having a "next" pointer, each node has a "left" pointer and a "right" pointer.



Using Nested Lists

- Languages like Python do not let programmers manipulate pointers explicitly.
- We could use Python lists to implement binary trees. For example:




```
[45, left, right]

[45,[31,left,right],[70,left,right]]

[45, [31, [19,[],[]], [38,[],[]]],
      [70,[], [86, [], []]]
]
```

[ ] stands for an empty tree

Arrows point to subtrees

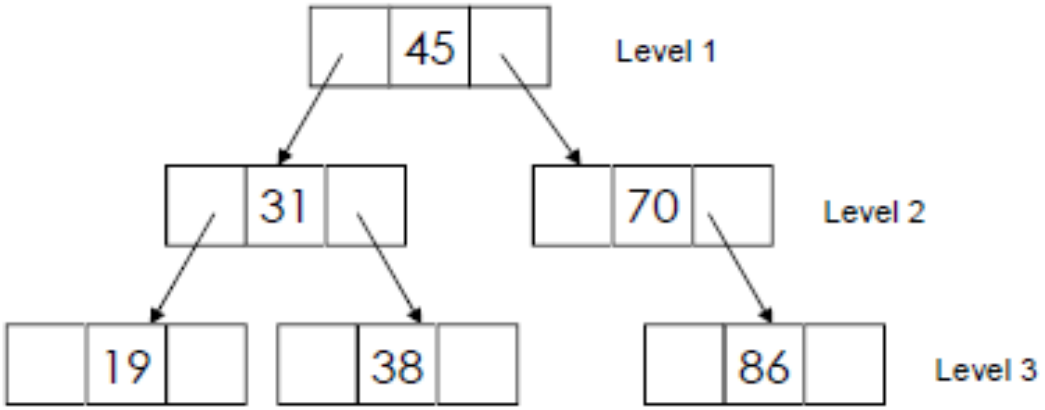
Building Binary Search Tree and Search in BST

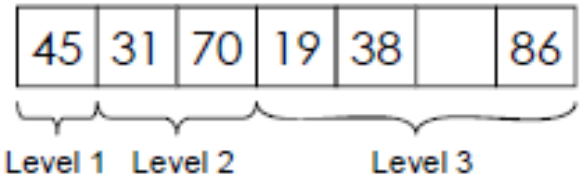
```
def bst_build(slist):

    if len(slist) == 0:
        return []
    if len(slist) == 1:
        return slist + [[]] + [[]]

    if len(slist) % 2 == 1:
            mid = len(slist) // 2
    else:
            mid = (len(slist) // 2) - 1

    print(" root: ", slist[mid]," left tree:  ", slist[0:mid], " right tree: ", slist[mid+1:])

    return [slist[mid]] + [bst_build(slist[0:mid])] + [bst_build(slist[mid+1:])]


def bst_in_list(lst):
    slist = sorted(lst)
    print("The sorted data: ", slist)
    return bst_build(slist)


def bst_search(bs_tree, key):
    if bs_tree == []:
        return print("Sorry, there is no such key in the BST")
    if bs_tree[0] == key:
        return print("Yes, The key is in the BST")
    if bs_tree[0] > key:
        bst_search(bs_tree[1], key)
    else:
        bst_search(bs_tree[2], key)
```

Insertions and Deletions are very very incovinient in this representation

Using One Dimensional Lists

- We could also use a flat (one-dimensional list).





Operations for Dynamic Data Set

- Insert
- Delete
- Search
- Find min/max
- ...

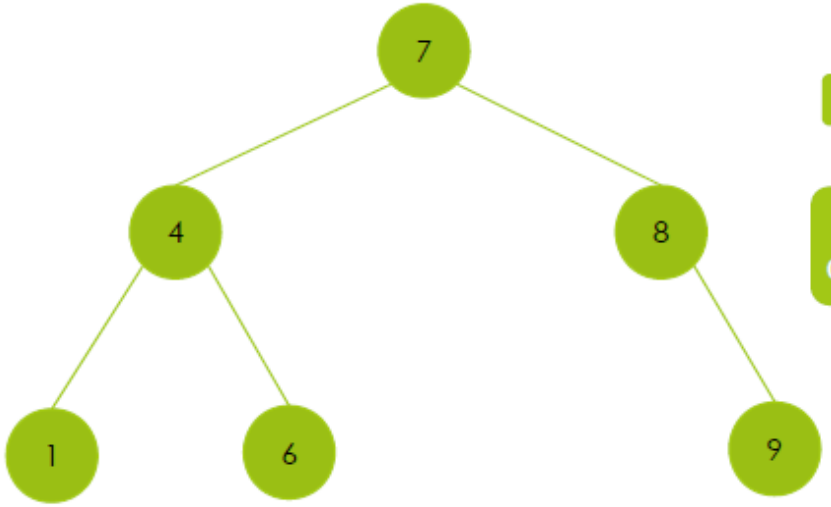
Choosing a specific data structure has consequences on which operations can be performed faster.

Search

- Binary Search
- Binary Tree
- Binary Search Tree

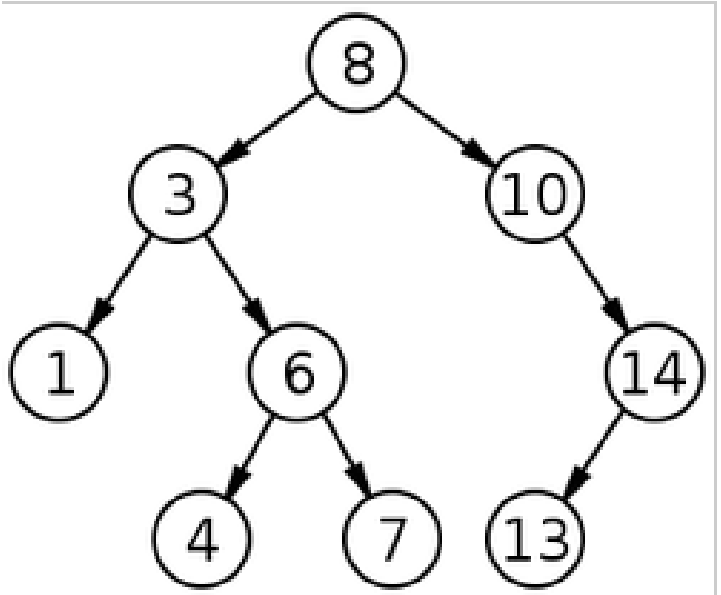
Binary Search Tree

BST ordering invariant: At any node with key *k*, all keys of elements in the left subtree are strictly less than *k* and all keys of elements in the rightsubtree are strictly greater than *k* (assume that there are no duplicates in the tree)



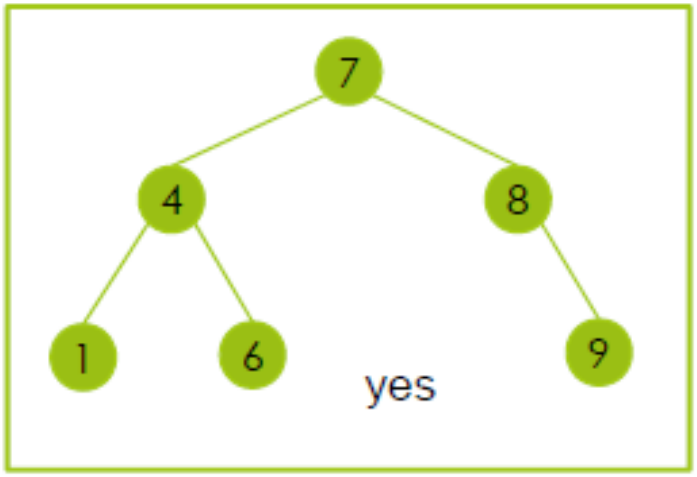
Binary tree

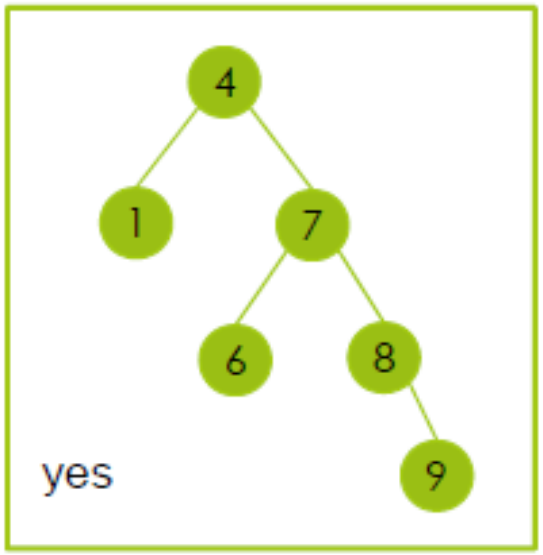
Satisfies the ordering invariant

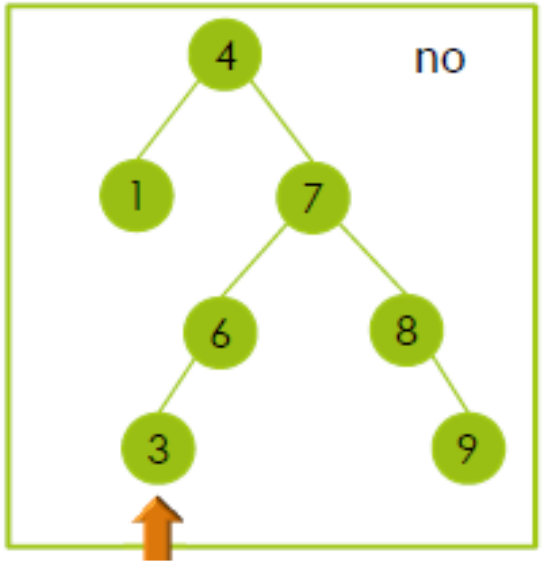


A binary search tree of size 9 and depth 3, with 8 at the root. The

Test: Is this a BST?





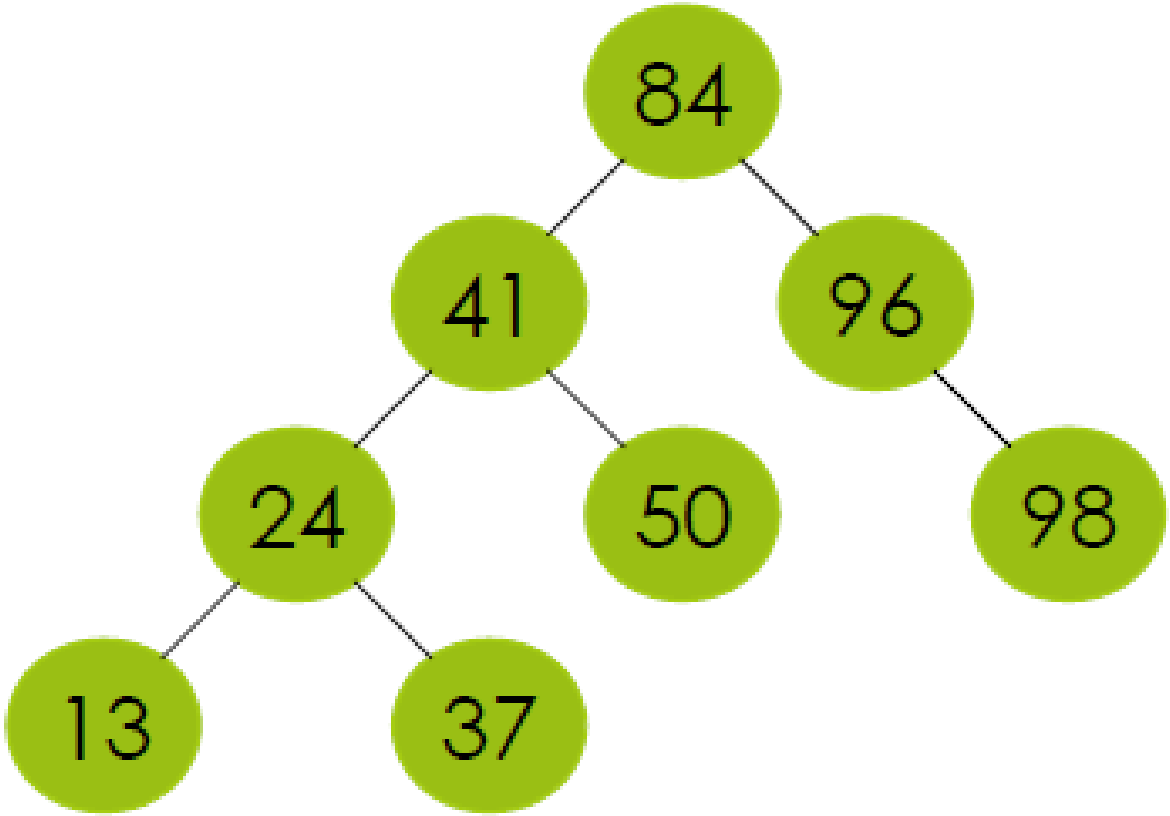


Inserting into a BST

- For each data value that you wish to insert into the binary search tree:
  - Start at the root and compare the new data value with the root.
  - If it is less, move down left. If it is greater, move down right.
  - Repeat on the child of the root until you end up in a position that has no node.
  - Insert a new node at this empty position.

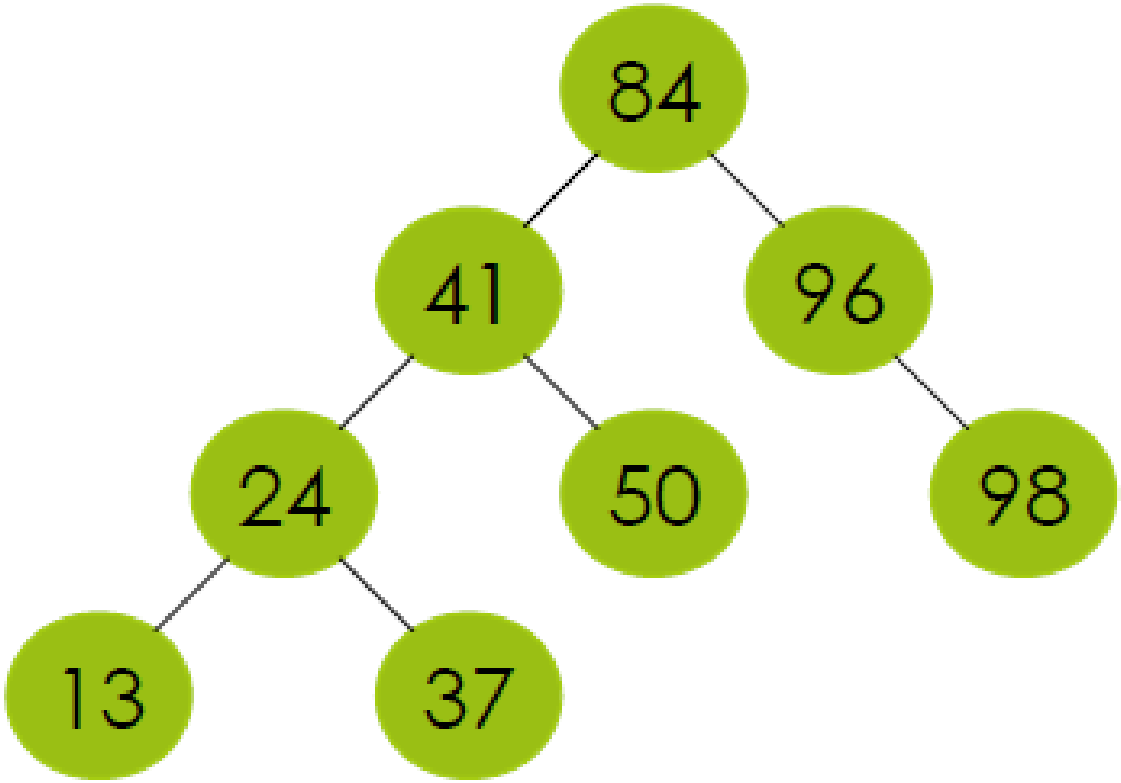
Example

- Insert: 84, 41, 96, 24, 37, 50, 13, 98



Using a BST

- How would you search for an element in a BST?

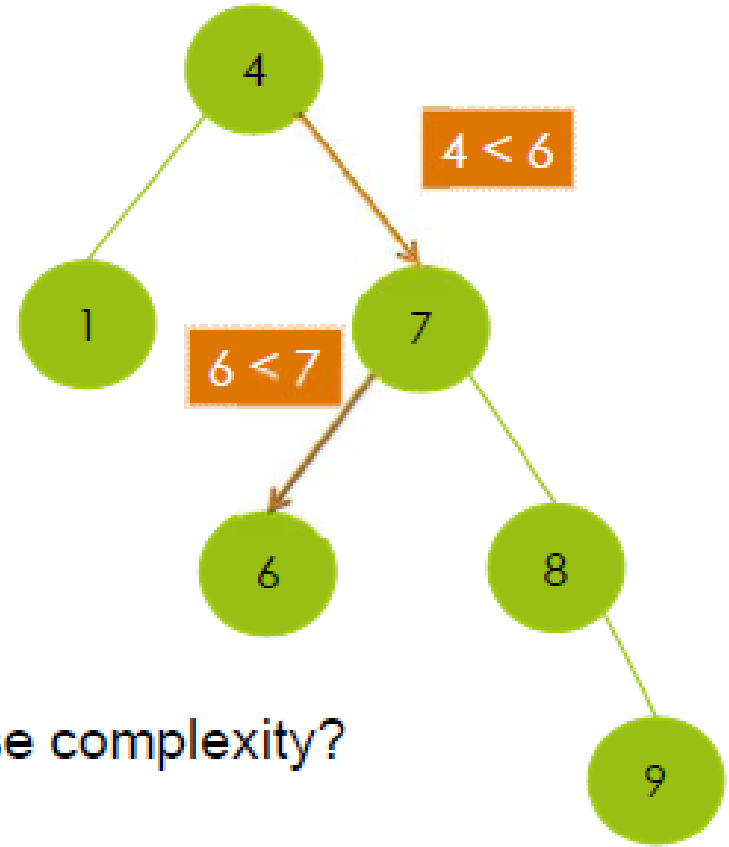


Searching a BST

- For the key that you wish to search
  - Start at the root and compare the key with the root. If equal, key found.
  - Otherwise
    - If it is less, move down left. If it is greater, move down right. Repeat search on the child of the root.
    - If there is no non-empty subtree to move to, then key not found.

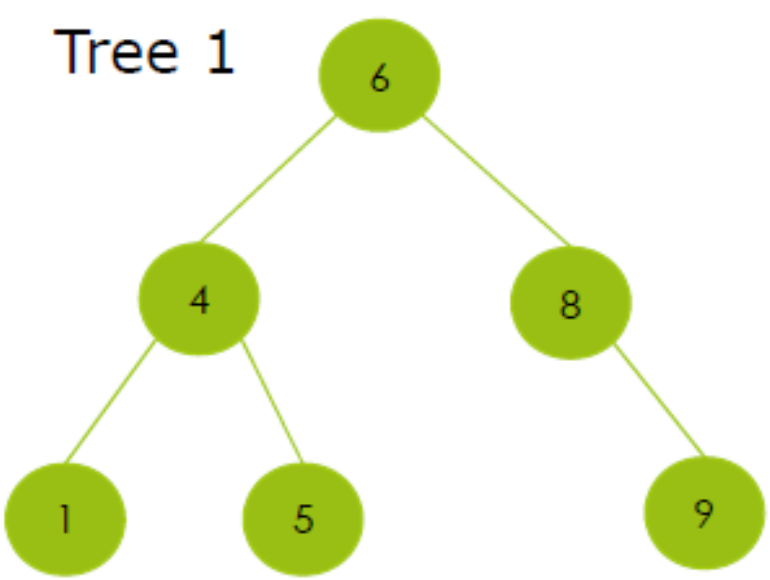
Searching the tree

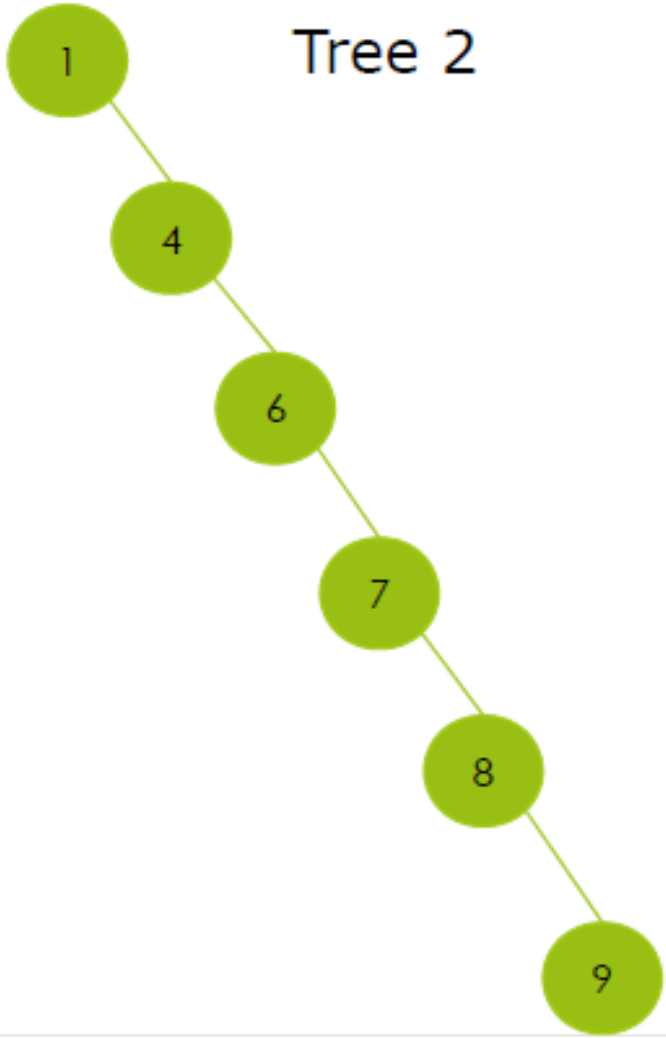
Example: searching for 6



Can we form a conjecture about worst case complexity?

Time complexity of search



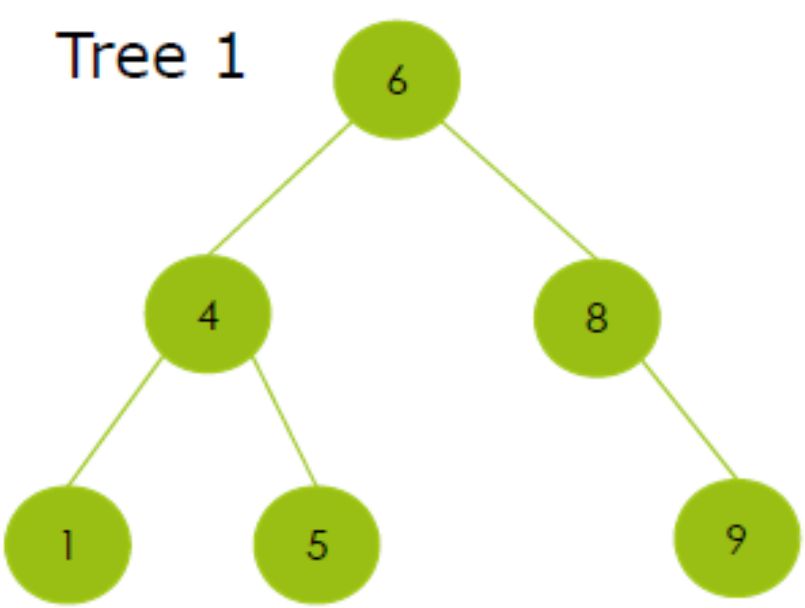


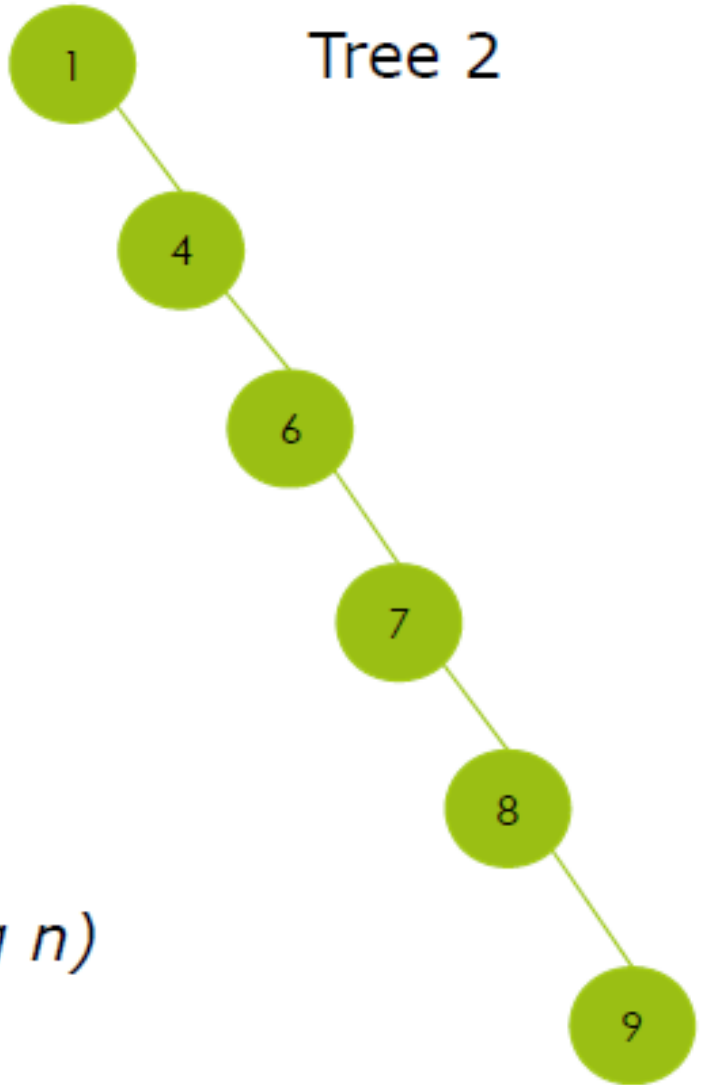
Number of nodes: *n*

Worst case: *O(height)*

Worst *height*: *n* ⇒ *O(n)*

Time complexity of search



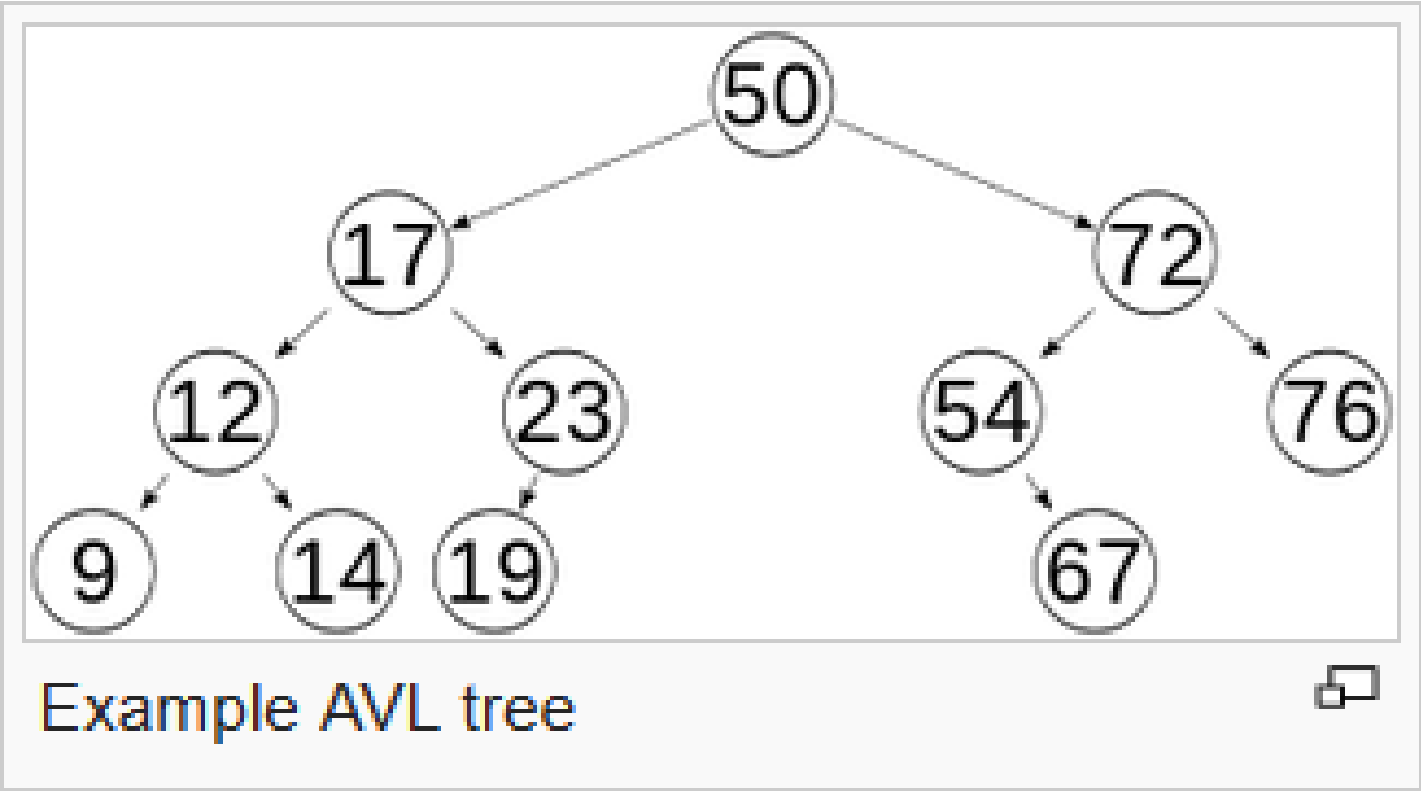


Number of nodes: *n*

What if we could always have balanced trees? ⇒ *O(log n)*

Balanced Binary Search Tree

- an **AVL tree** (Georgy Adelson-Velsky and Evgenii Landis) tree
- A self-balancing binary search tree
- 모든 Node의 left subtree와 right subtree의 height difference가 less than one!



Example AVL tree