AVL Tree

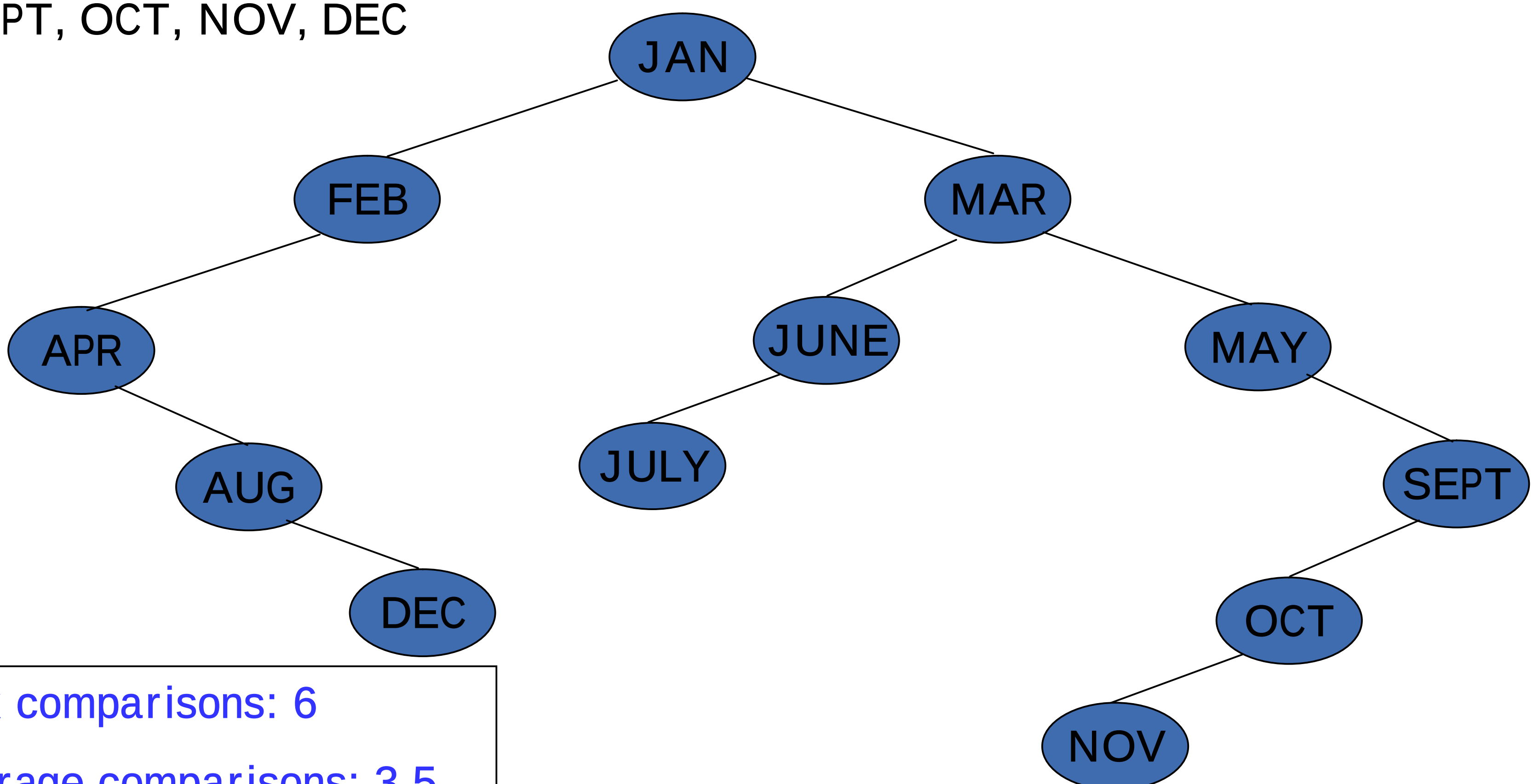
授課老師：詹寶珠教授

AVL Trees

- Dynamic tables may also be maintained as binary search trees.
- Depending on the order of the symbols putting into the table, the resulting binary search trees would be different. Thus the average comparisons for accessing a symbol is different.

Binary Search Tree for The Months of The Year

Input Sequence: JAN, FEB, MAR, APR, MAY, JUNE, JULY, AUG, SEPT, OCT, NOV, DEC



Max comparisons: 6

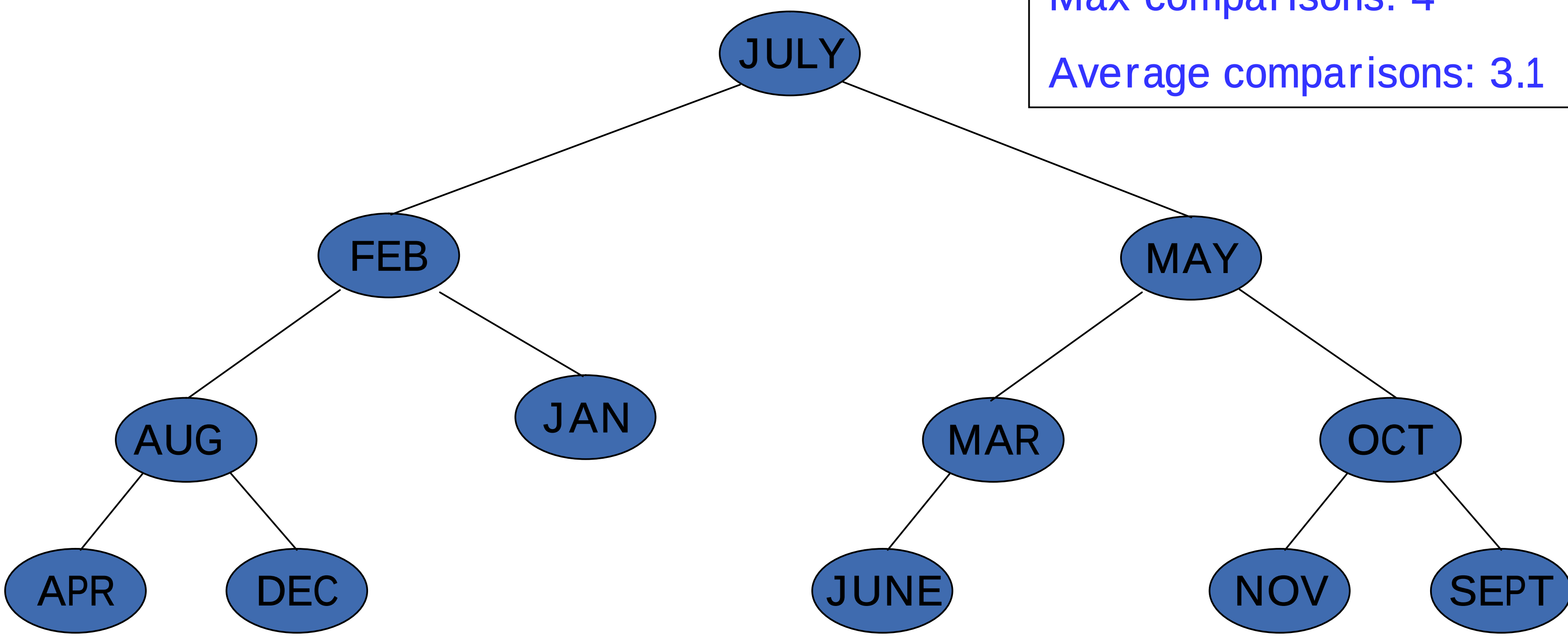
Average comparisons: 3.5

A Balanced Binary Search Tree For The Months of The Year

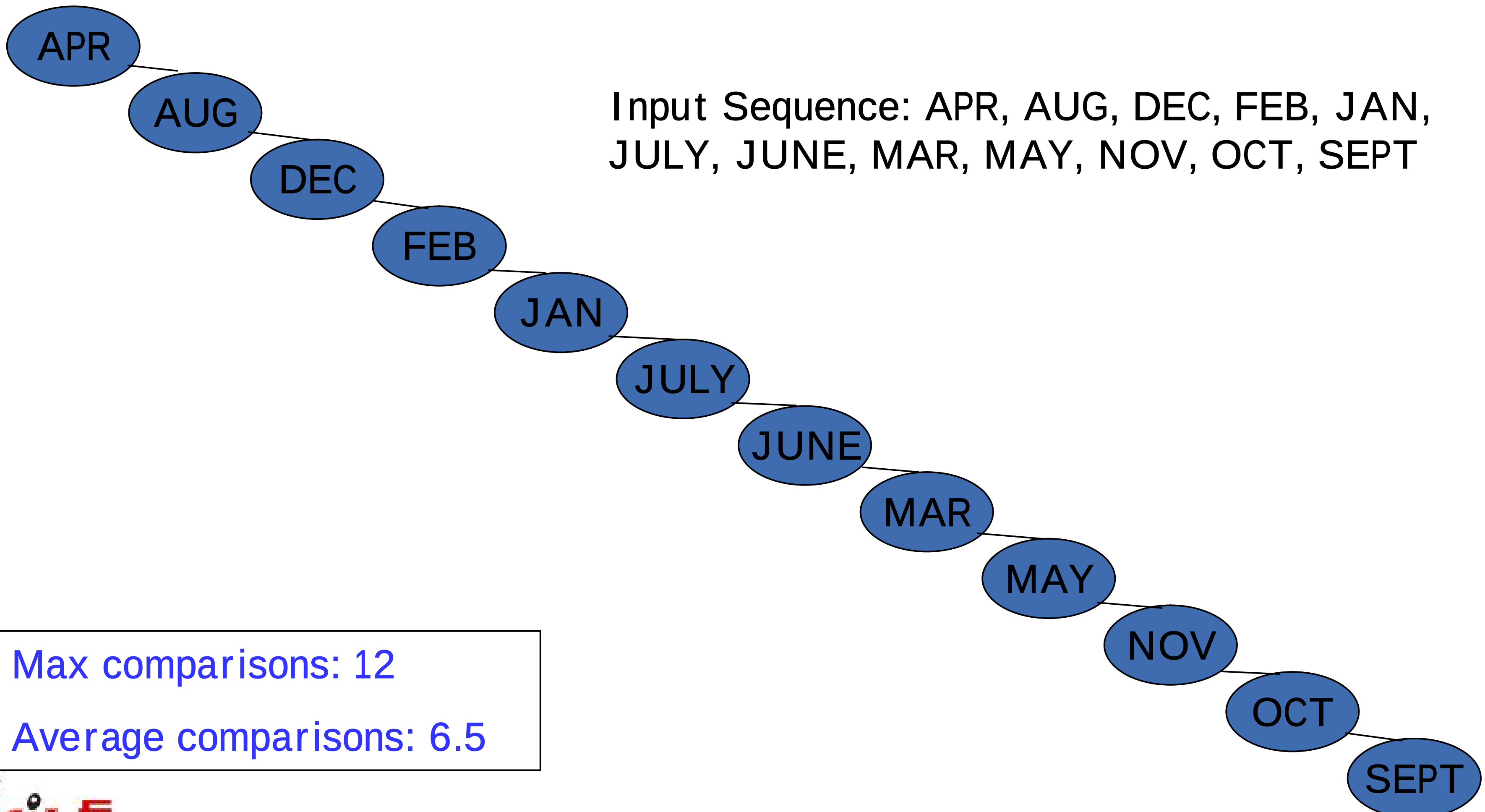
Input Sequence: JULY, FEB, MAY, AUG, DEC, MAR, OCT, APR, JAN, JUNE, SEPT, NOV

Max comparisons: 4

Average comparisons: 3.1



Degenerate Binary Search Tree



Input Sequence: APR, AUG, DEC, FEB, JAN, JULY, JUNE, MAR, MAY, NOV, OCT, SEPT

Max comparisons: 12

Average comparisons: 6.5

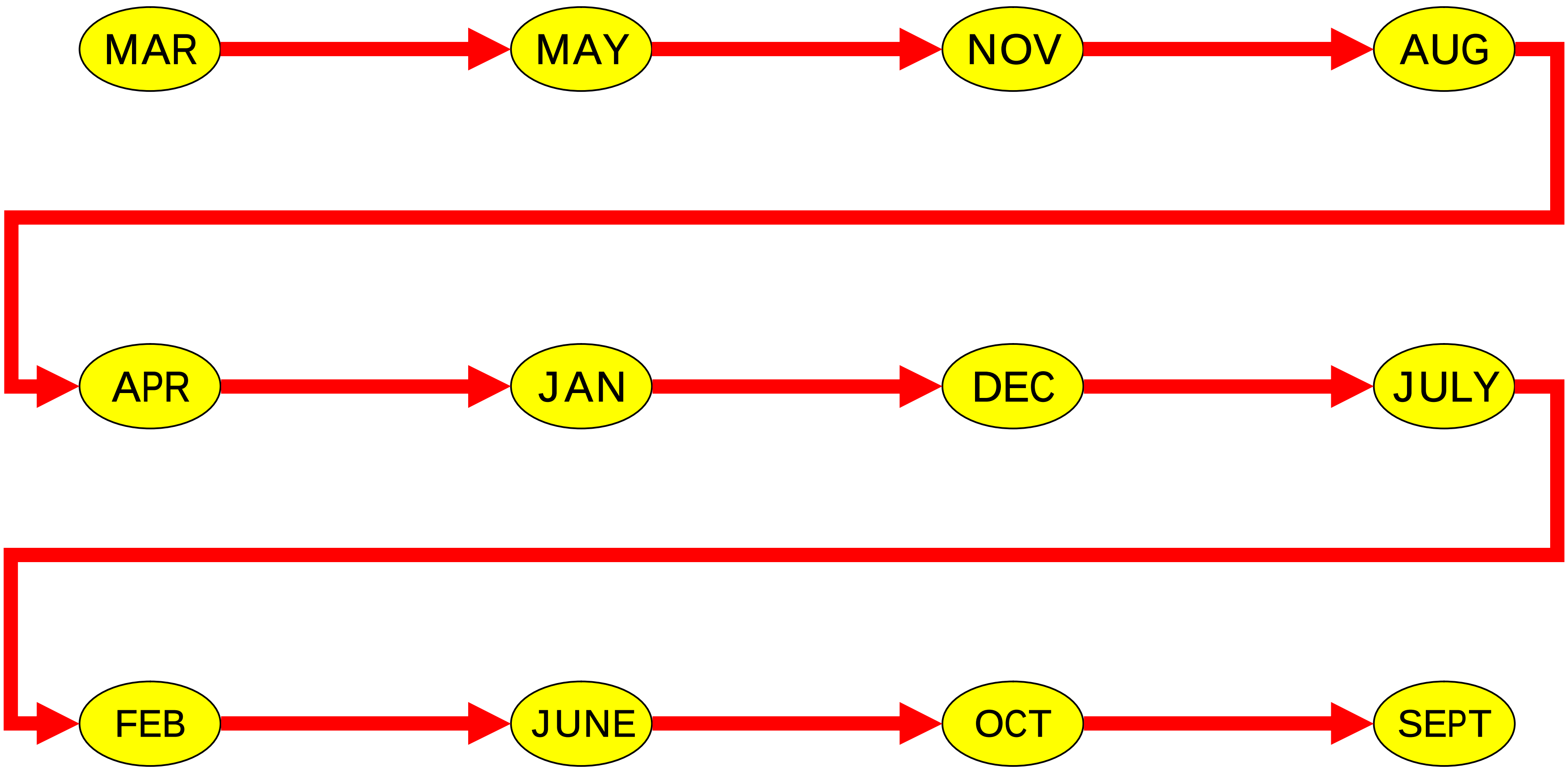
Minimize The Search Time of Binary Search Tree In Dynamic Situation

- From the above three examples, we know that the average and maximum search time will be minimized if the binary search tree is maintained as a complete binary search tree at all times.
- However, to achieve this in a dynamic situation, we have to pay a high price to restructure the tree to be a complete binary tree all the time.
- In 1962, Adelson-Velskii and Landis introduced a binary tree structure that is balanced with respect to the heights of subtrees. As a result of the balanced nature of this type of tree, dynamic retrievals can be performed in O(log n) time if the tree has n nodes. The resulting tree remains height-balanced. This is called an AVL tree.

AVL Tree

- **Definition: An empty tree is height-balanced. If *T* is a nonempty binary tree with T_L and T_R as its left and right subtrees respectively, then *T* is height-balanced iff**
  (1) T_L and T_R are height-balanced, and
  (2) $|h_L - h_R| \leq 1$ where h_L and h_R are the heights of T_L and T_R, respectively.
- **Definition: The *balance factor*, *BF(T)* , of a node *T* is a binary tree is defined to be $h_L – h_R$, where h_L and h_R, respectively, are the heights of left and right subtrees of *T*. For any node *T* in an AVL tree, *BF*(*T*) = -1, 0, or 1.**

Balanced Trees Obtained for The Months of The Year

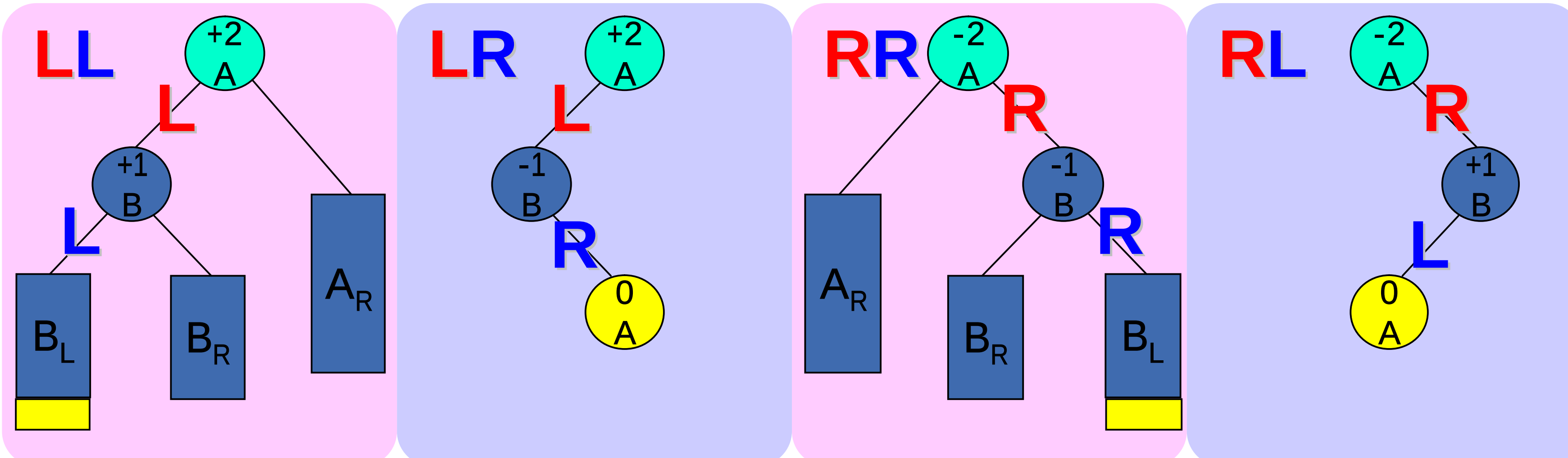


AVL Tree

- **The addition of a node to a balanced binary search tree could unbalance it.**
- **The rebalancing was carried out using four different kinds of rotations:**
  - **LL**: ***Y*** is inserted in the **left** subtree of the **left** subtree of ***A***
  - **LR**: ***Y*** is inserted in the **right** subtree of the **left** subtree of ***A***
  - **RR**: ***Y*** is inserted in the **right** subtree of the **right** subtree of ***A***
  - **RL**: ***Y*** is inserted in the **left** subtree of the **right** subtree of ***A***

These rotations are characterized by the **nearest ancestor**, ***A***, of the **inserted node**, ***Y***, whose **balance factor** becomes **±2**.

AVL Tree

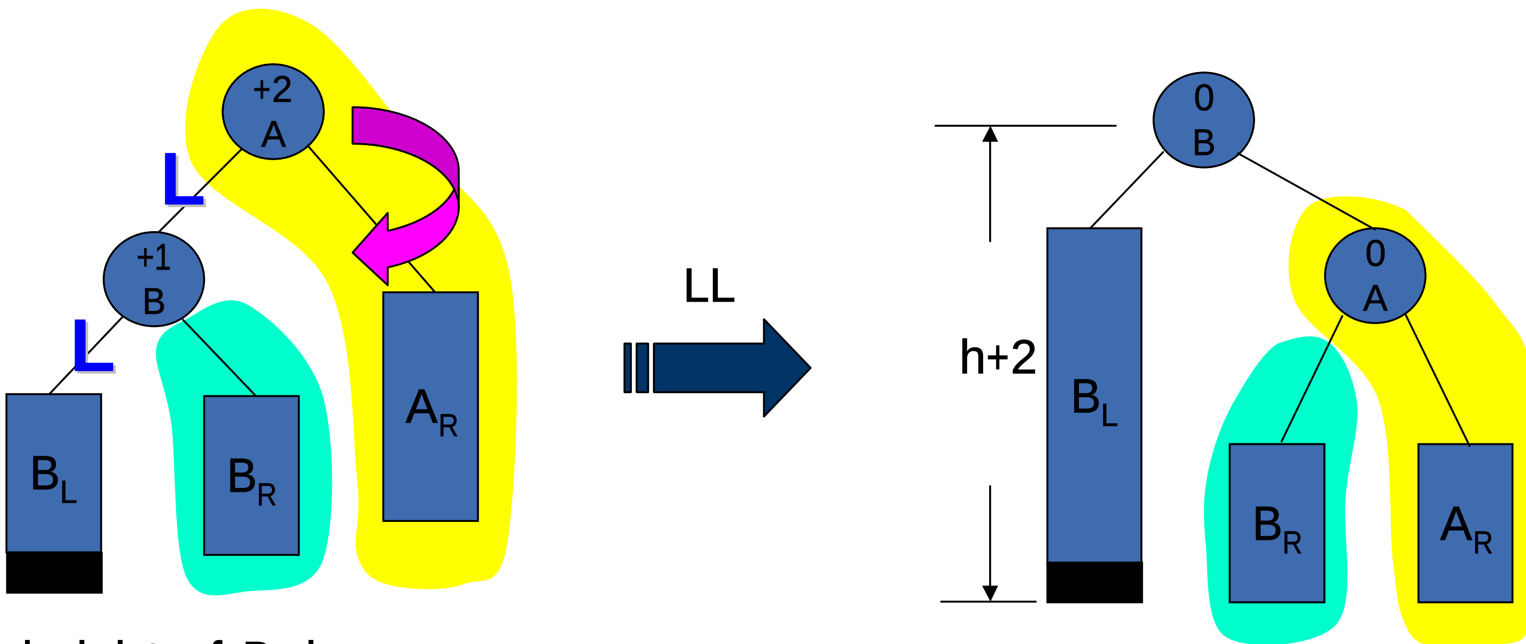


LL	LR	RR	RL
A clockwise	**A clockwise & Child anti-clockwise**	**A anti-clockwise**	**A anti-clockwise & Child clockwise**

★: **inserted node *Y***

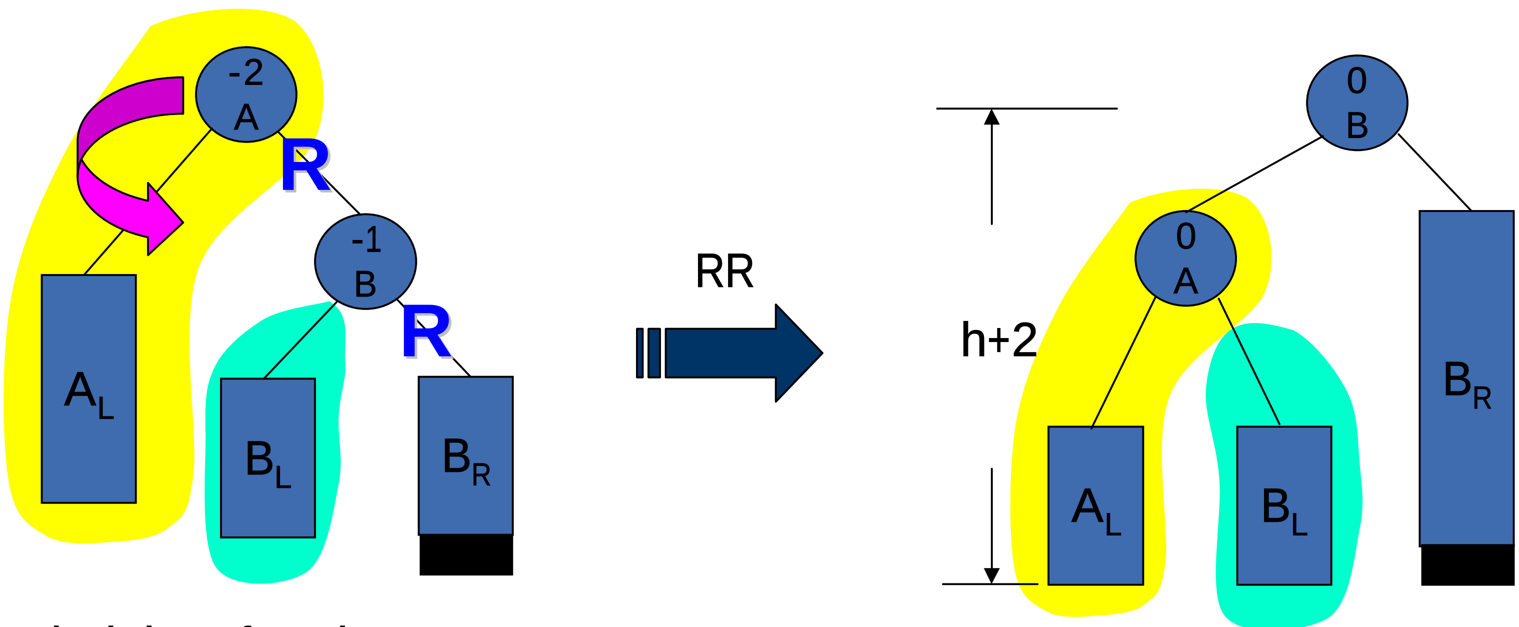
★: **nearest ancestor *A*, balance factor = ±2**

Rebalancing Rotation LL



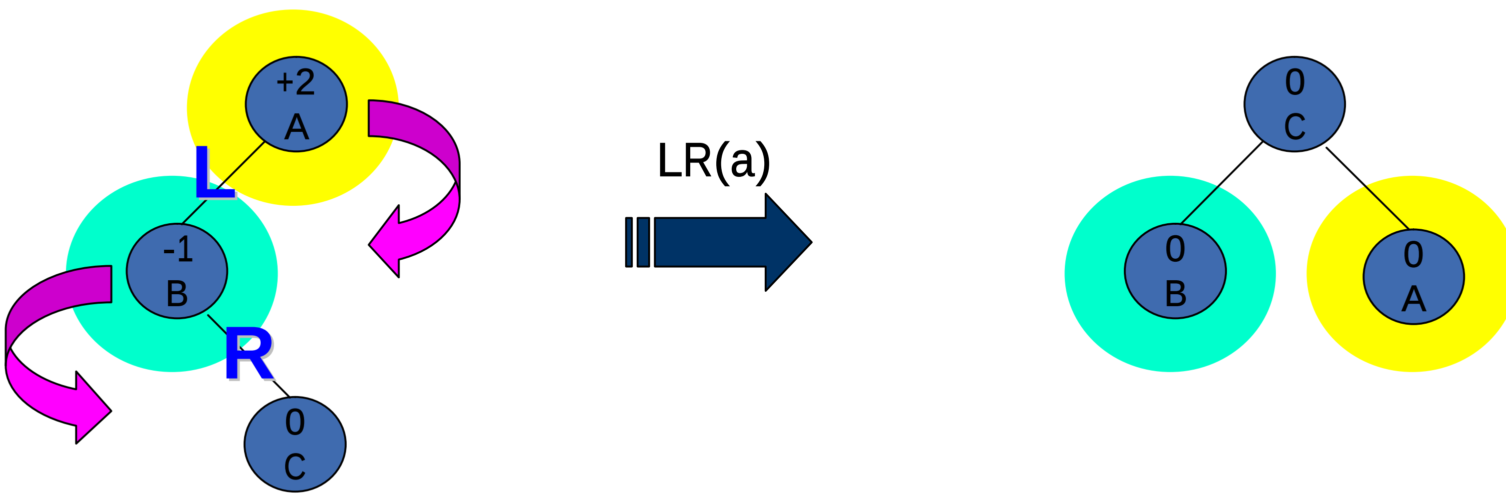
height of B_L increases to h+1

Rebalancing Rotation RR

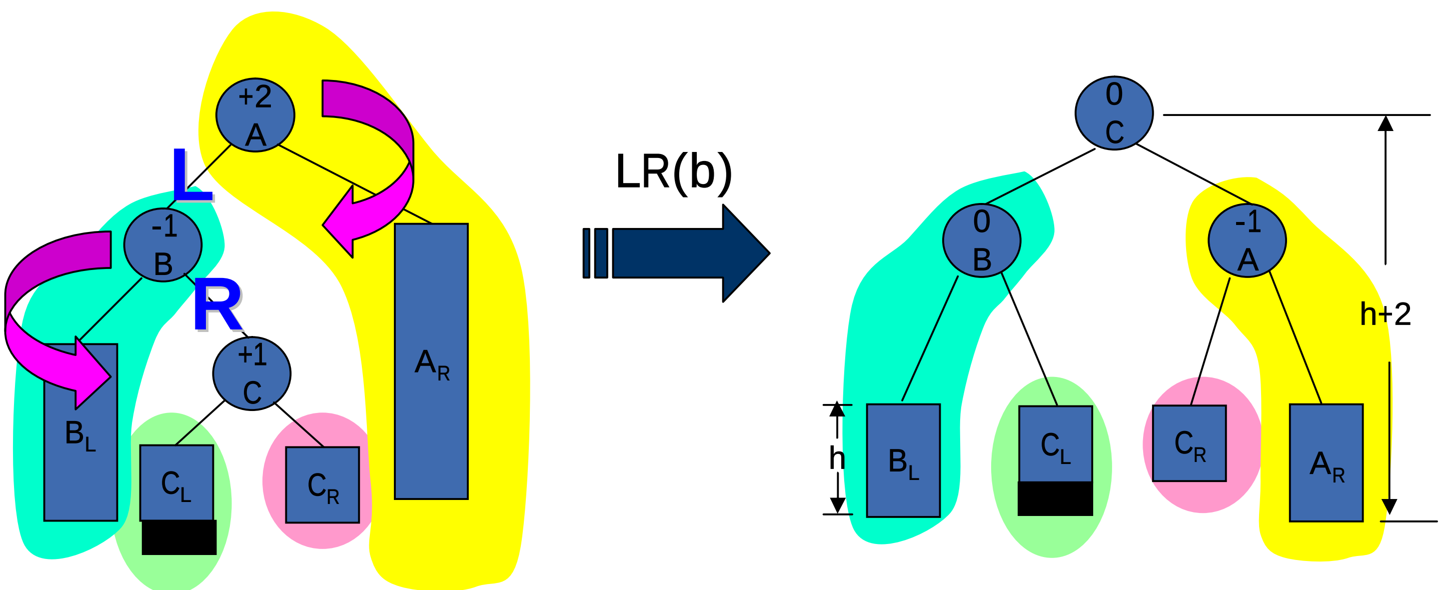


height of B_R increases to h+1

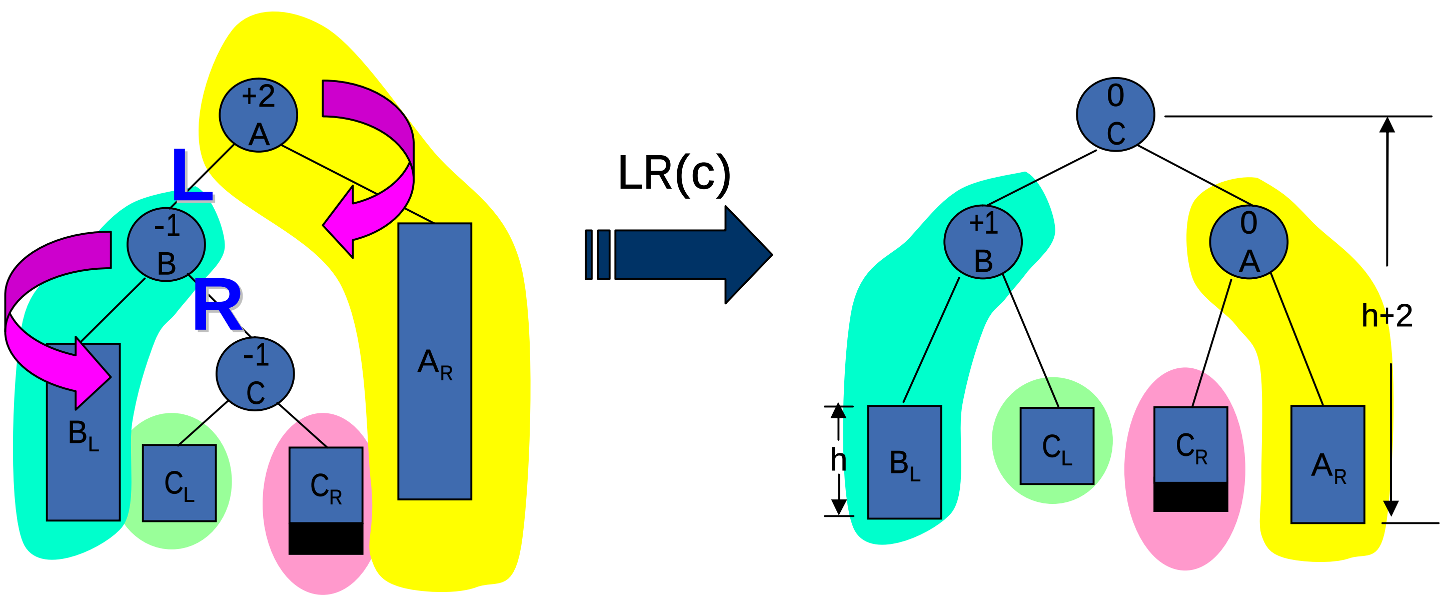
Rebalancing Rotation LR(a)



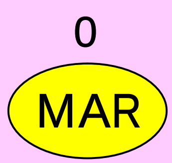
Rebalancing Rotation LR(b)



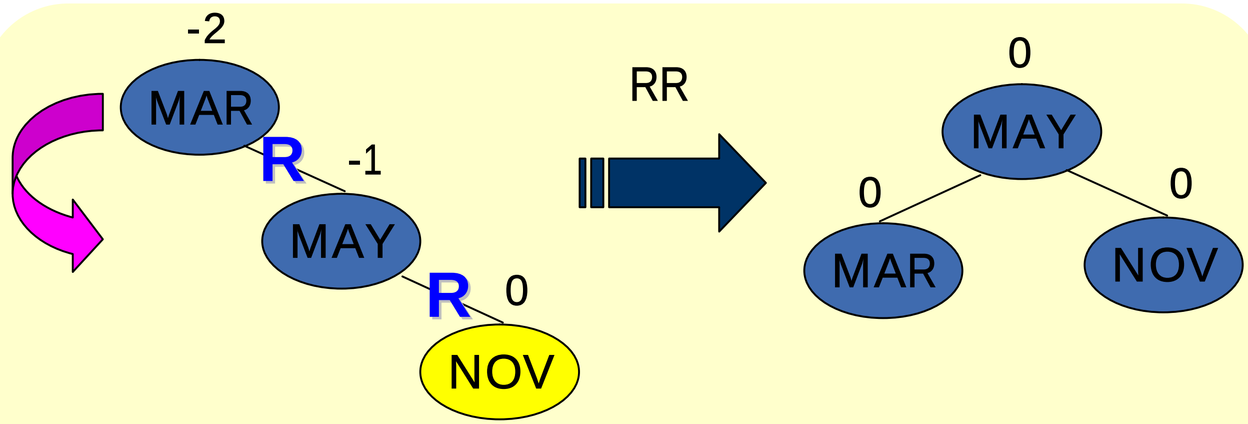
Rebalancing Rotation LR(c)



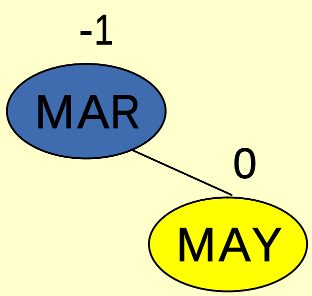
Balanced Trees Obtained for The Months of The Year



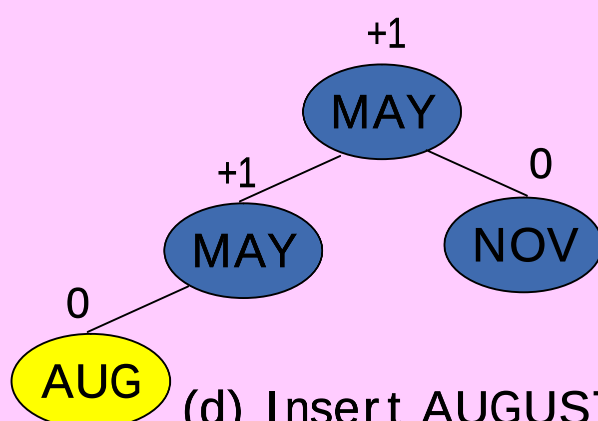
(a) Insert MARCH



(c) Insert NOVEMBER

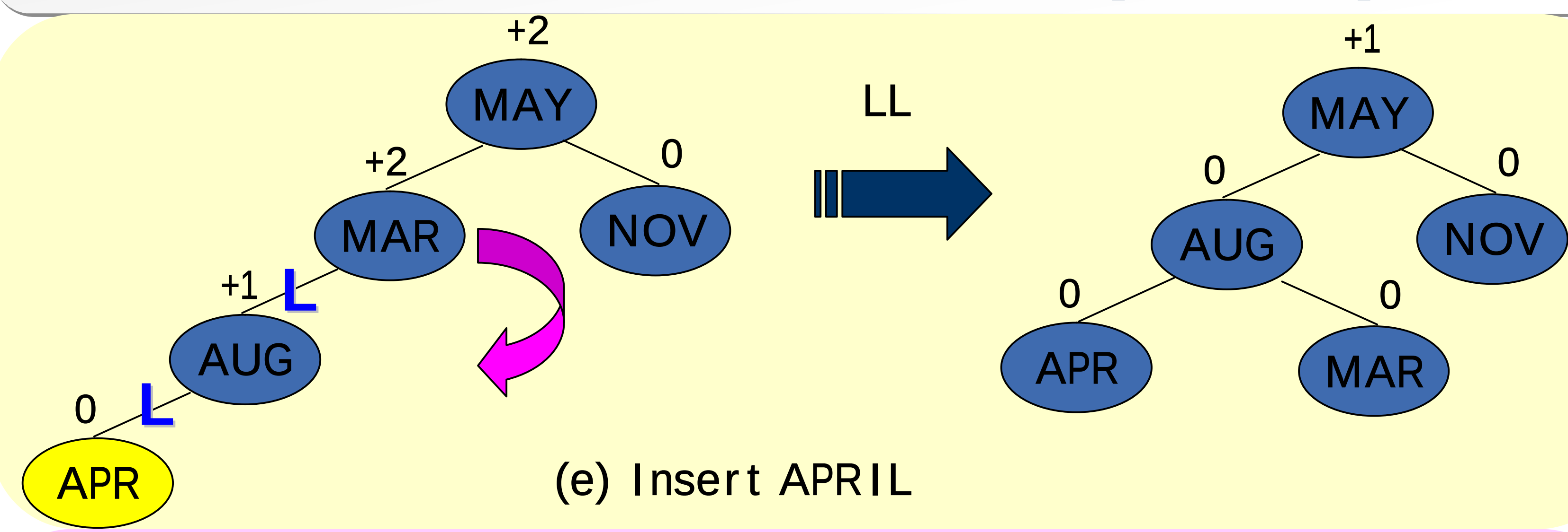


(b) Insert MAY

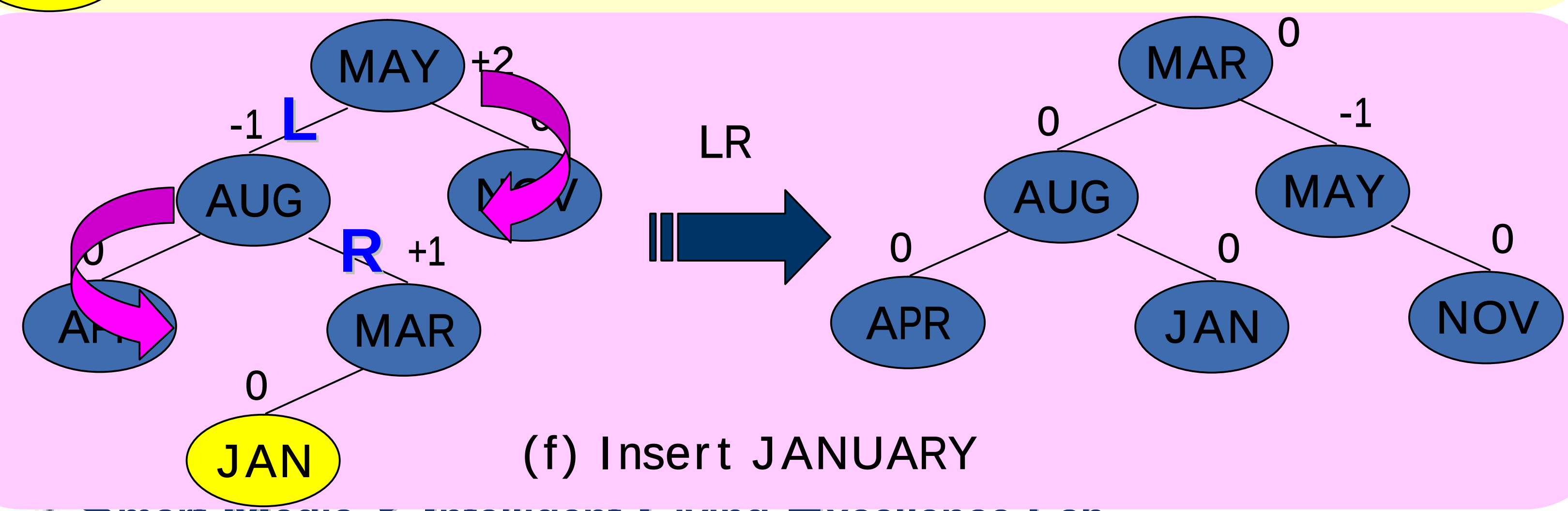


(d) Insert AUGUST

Balanced Trees Obtained for The Months of The Year (Cont.)

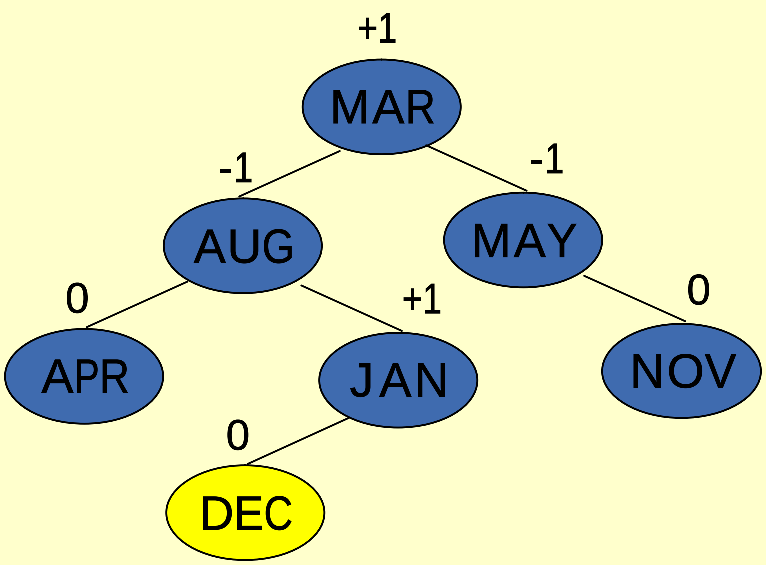


(e) Insert APRIL

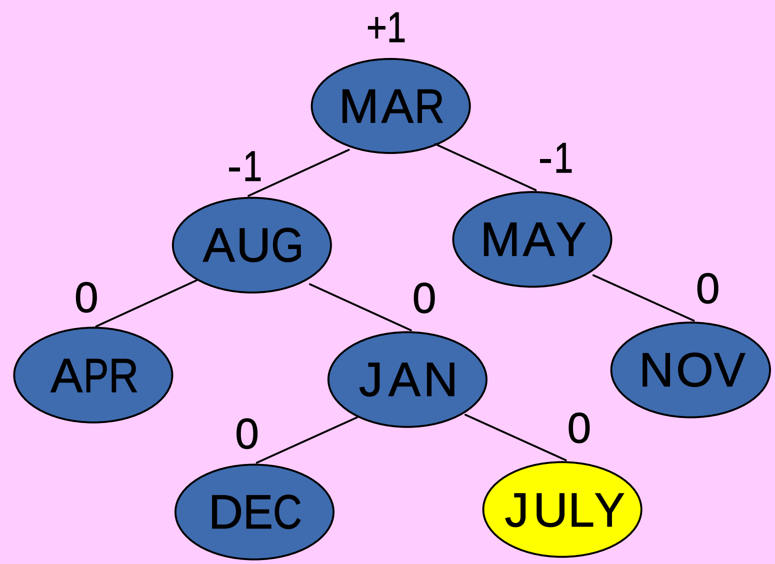


(f) Insert JANUARY

Balanced Trees Obtained for The Months of The Year (Cont.)

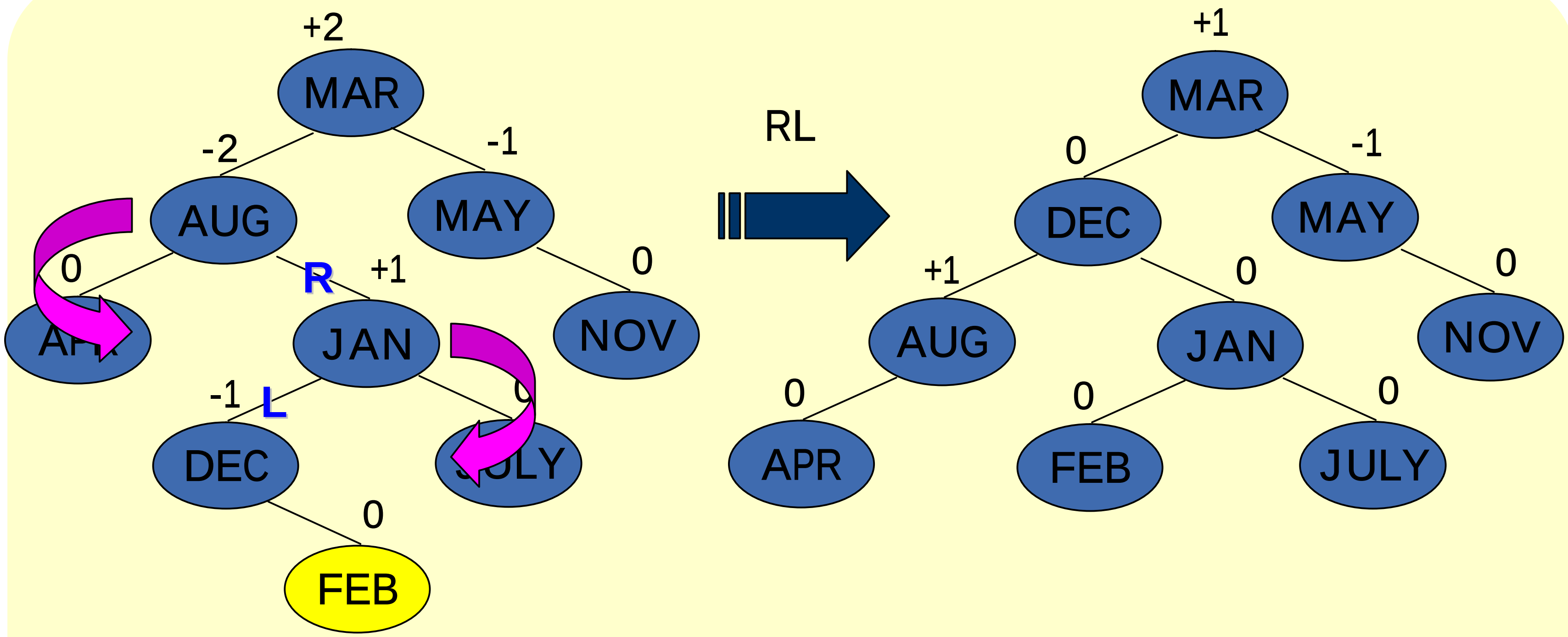


(g) Insert DECEMBER



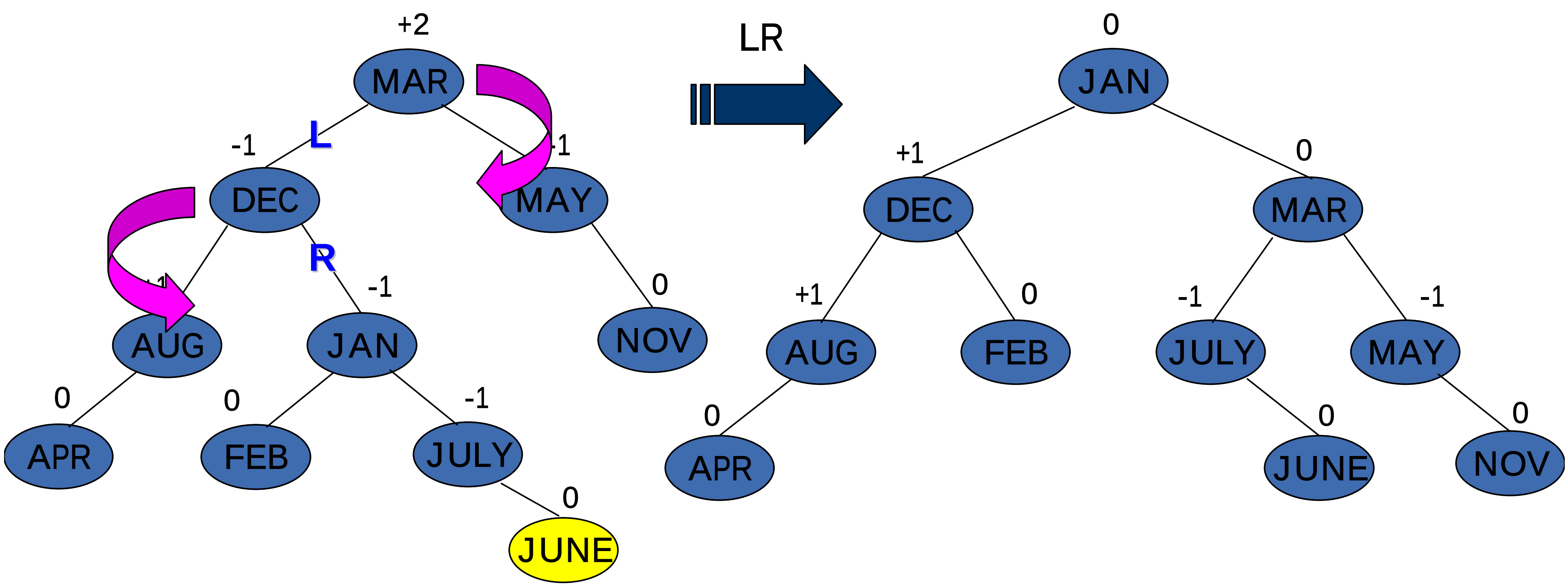
(h) Insert JULY

Balanced Trees Obtained for The Months of The Year (Cont.)



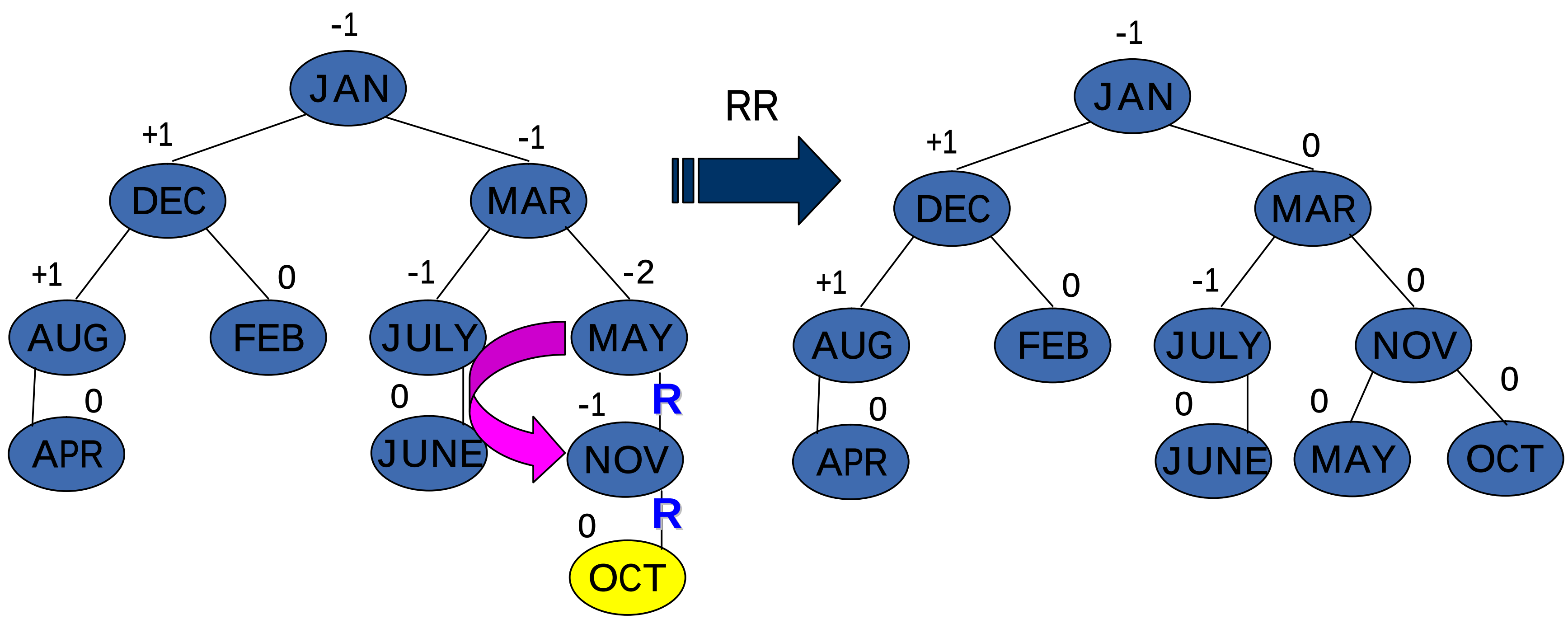
(i) Insert FEBRUARY

Balanced Trees Obtained for The Months of The Year (Cont.)



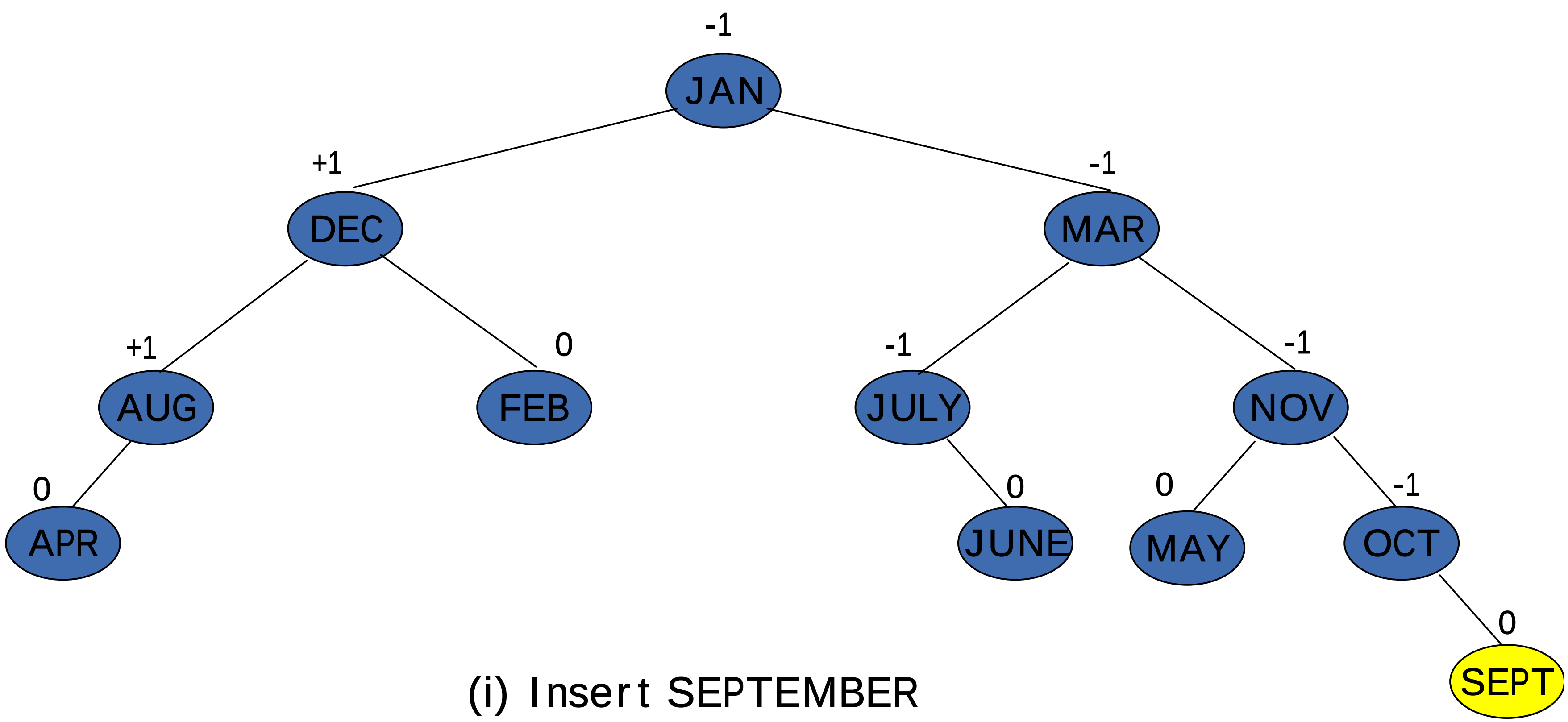
(j) Insert JUNE

Balanced Trees Obtained for The Months of The Year (Cont.)



(k) Insert OCTOBER

Balanced Trees Obtained for The Months of The Year (Cont.)



(i) Insert SEPTEMBER

AVL Trees (Cont.)

- **Once rebalancing has been carried out on the subtree in question, examining the remaining tree is unnecessary.**
- **To perform insertion, binary search tree with n nodes could have O(n) in worst case. But for AVL, the insertion time is O(log n).**

AVL Insertion Complexity

- **Let N_h be the minimum number of nodes in a height-balanced tree of height h. In the worst case, the height of one of the subtrees will be h-1 and that of the other h-2. Both subtrees must also be height balanced. $N_h = N_{h-1} + N_{h-2} + 1$, and $N_0 = 0, N_1 = 1,$ and $N_2 = 2$.**
- **The recursive definition for N_h and that for the Fibonacci numbers $F_n = F_{n-1} + F_{n-2}, F_0=0, F_1= 1$.**
- **It can be shown that $N_h = F_{h+2} - 1$. Therefore we can derive that $N_h \approx f^{h+2}/\sqrt{5} - 1$. So the worst-case insertion time for a height-balanced tree with n nodes is $O(\log n)$.**

Probability of Each Type of Rebalancing Rotation

- **Research has shown that a random insertion requires no rebalancing, a rebalancing rotation of type LL or RR, and a rebalancing rotation of type LR and RL, with probabilities 0.5349, 0.2327, and 0.2324, respectively.**

Comparison of Various Structures

Operation	Sequential List	Linked List	AVL Tree
Search for x	$O(\log n)$	$O(n)$	$O(\log n)$
Search for kth item	$O(1)$	$O(k)$	$O(\log n)$
Delete x	$O(n)$	$O(1)$[1]	$O(\log n)$
Delete kth item	$O(n - k)$	$O(k)$	$O(\log n)$
Insert x	$O(n)$	$O(1)$[2]	$O(\log n)$
Output in order	$O(n)$	$O(n)$	$O(n)$

1. Doubly linked list and position of x known.
2. Position for insertion known