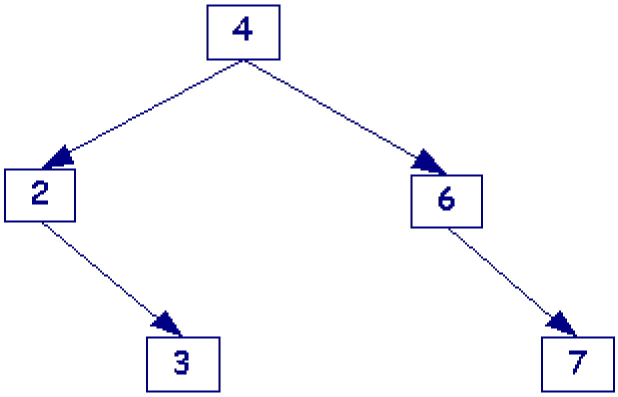
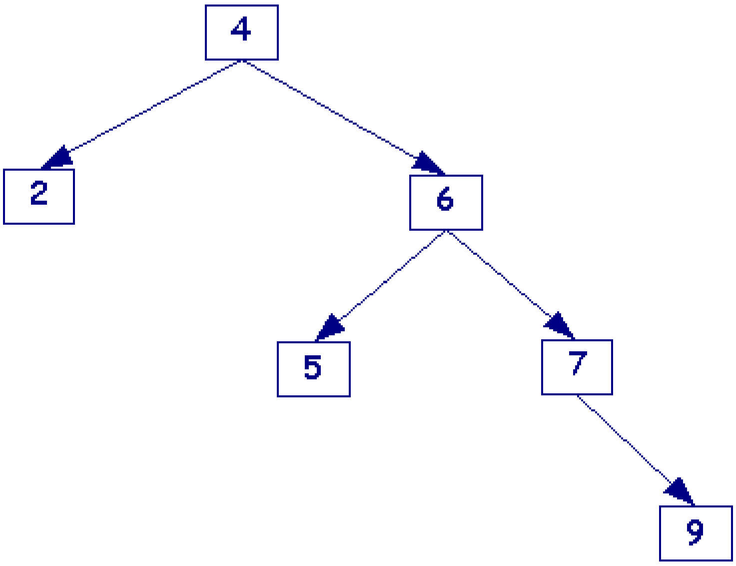
5.9 AVL-Δέντρα

Ένα **AVL-δέντρο** (AVL-tree) είναι ένα δυαδικό δέντρο αναζήτησης στο οποίο για κάθε κόμβο του ισχύει ότι |hl-hr|<=1, όπου hl & hr είναι τα ύψη του αριστερού και δεξιού υποδένδρου αντίστοιχα, δηλαδή για κάθε κόμβο το ύψος του αριστερού υποδέντρου διαφέρει από το ύψος του δεξιού υποδέντρου του κόμβου αυτού το πολύ κατά 1.

Για παράδειγμα, το δυαδικό δέντρο



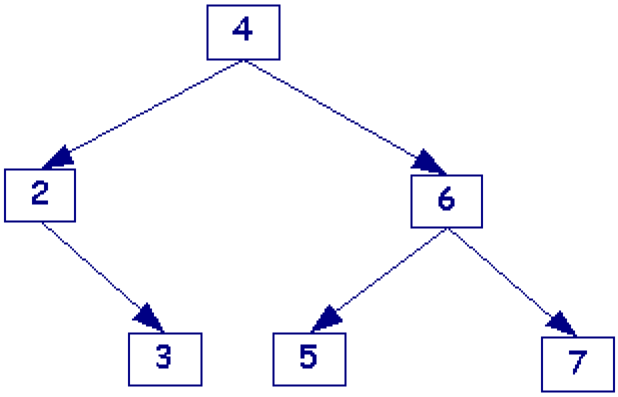
είναι AVL-δέντρο, ενώ το δέντρο



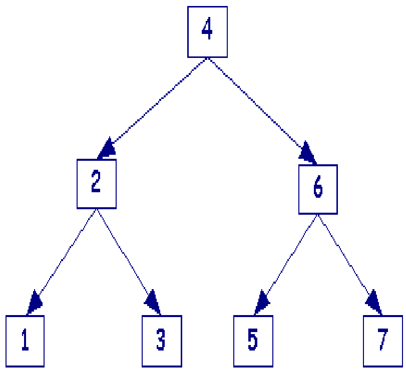
δεν είναι AVL-δέντρο, αφού το ύψος του αριστερού υποδέντρου της ρίζας είναι 1 ενώ το ύψος του δεξιού υποδέντρου είναι 3.

Εισαγωγή στοιχείου σε ένα AVL-δέντρο

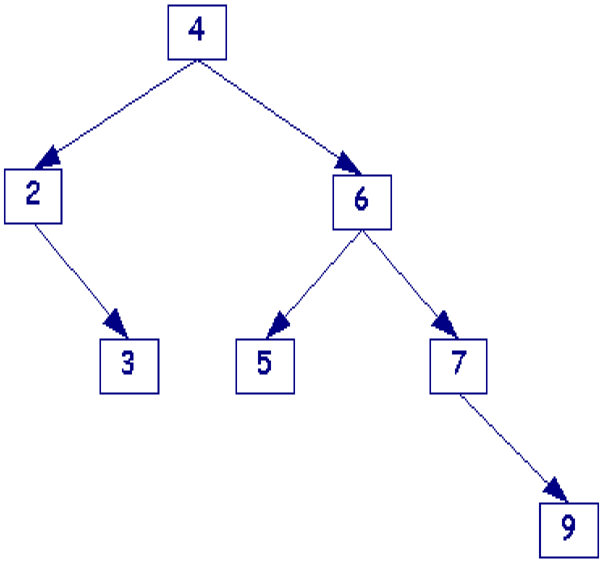
Εφόσον ένα AVL-δέντρο είναι ένα δυαδικό δέντρο αναζήτησης, θα ακολουθήσουμε τη διαδικασία της εισαγωγής στοιχείου σε δυαδικό δέντρο αναζήτησης, την οποία περιγράψαμε στο 5ο Κεφάλαιο. Το αποτέλεσμα της διαδικασίας αυτής μπορεί να είναι ένα δυαδικό δέντρο αναζήτησης που εξακολουθεί να είναι ένα AVL-δέντρο. Για παράδειγμα, αν εισάγουμε το 1 ή το 9 στο AVL-δέντρο



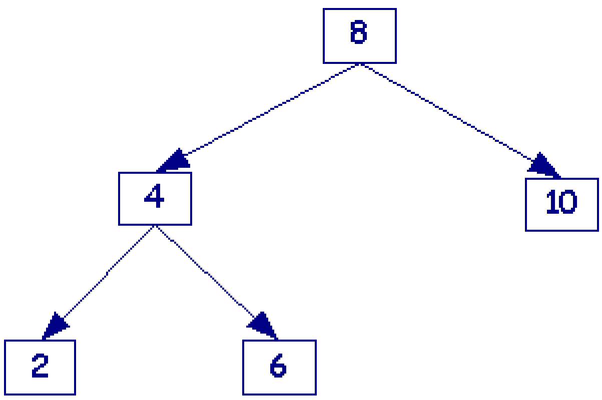
προκύπτουν αντίστοιχα τα δέντρα



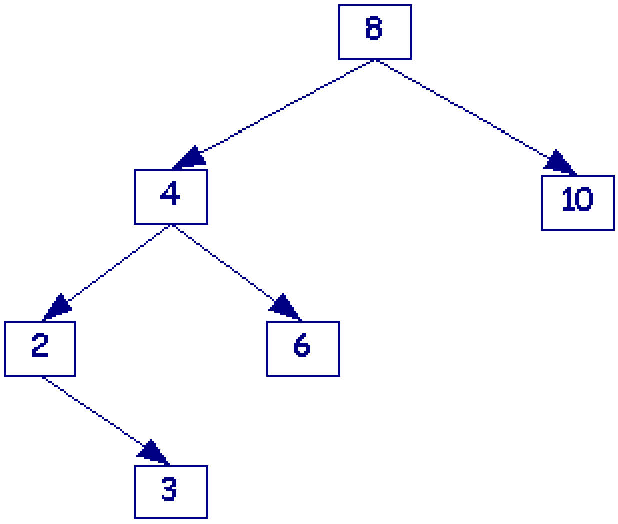
και



τα οποία είναι και τα δύο AVL-δέντρα. Ωστόσο, αν εισάγουμε τα στοιχεία 1, 3, 5 ή 7 στο AVL-δέντρο



παραβιάζουμε την ιδιότητα των AVL-δέντρων. Για παράδειγμα, η εισαγωγή του 3 έχει ως αποτέλεσμα το δέντρο



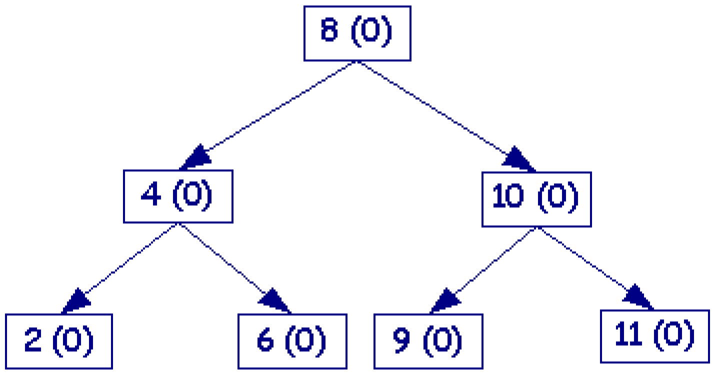
που δεν είναι AVL-δέντρο, αφού το ύψος του αριστερού υποδέντρου της ρίζας είναι 3 ενώ το ύψος του δεξιού υποδέντρου είναι 1.

Προκειμένου να αντιμετωπίσουμε το πρόβλημα της παραβίασης της ιδιότητας των AVL-δέντρων κατά την εισαγωγή ενός στοιχείου σε ένα AVL-δέντρο, είναι βολικό σε κάθε κόμβο του δέντρου να αντιστοιχίσουμε τον **παράγοντα ισοζύγισης** (*balance factor*), δηλαδή να χρησιμοποιήσουμε ένα επιπλέον πεδίο, έστω Balance, το οποίο θα έχει τιμή -1, 0 ή 1 που προκύπτει ως εξής:

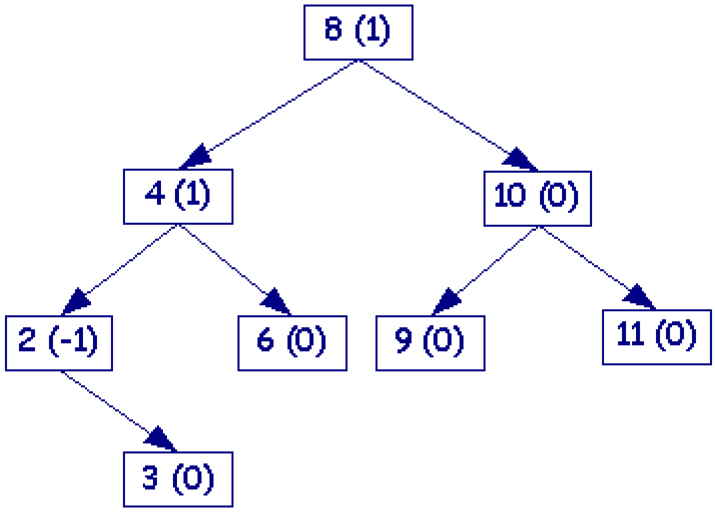
Ύψος του αριστερού υποδέντρου του κόμβου - Ύψος του δεξιού υποδέντρου

Όταν εισάγουμε ένα νέο στοιχείο σε ένα δυαδικό δέντρο αναζήτησης ξεκινάμε από τη ρίζα και διασχίζουμε μια ακολουθία κόμβων μέχρι να δημιουργήσουμε ένα νέο αριστερό ή δεξί παιδί του τελευταίου κόμβου της ακολουθίας και να το χρησιμοποιήσουμε για να αποθηκεύσουμε το νέο στοιχείο. Αυτή την ακολουθία των κόμβων την ονομάζουμε **μονοπάτι αναζήτησης** του νέου στοιχείου, αφού είναι εξάλλου το μονοπάτι που θα ακολουθούσαμε αν αναζητούσαμε το νέο στοιχείο.

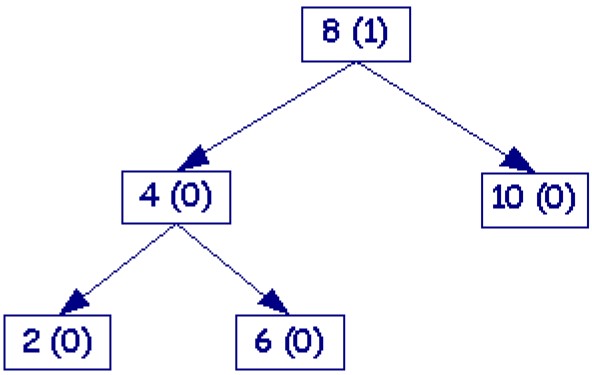
Την παραπάνω διαδικασία την εξειδικεύουμε για την περίπτωση ενός AVL-δέντρου. Συγκεκριμένα, αν για κάθε κόμβο του μονοπατιού αναζήτησης ενός νέου στοιχείου ισχύειBalance=0, τότε η εισαγωγή του στοιχείου θα έχει ως αποτέλεσμα την αλλαγή της τιμής του πεδίου Balance σε καθένα από τους κόμβους εκείνους, οι οποίοι θα πάρουν την τιμή -1 ή 1, έτσι ώστε το δέντρο να διατηρήσει την ιδιότητα των AVL-δέντρων. Για παράδειγμα, η εισαγωγή του στοιχείου με κλειδί 3 στο AVL-δέντρο (στο οποίο οι αριθμοί στις παρενθέσεις είναι οι τιμές του πεδίου Balance των κόμβων)



έχει ως αποτέλεσμα τη δημιουργία του εξής δέντρου



Στη συνέχεια ας θεωρήσουμε ότι τουλάχιστον ένας κόμβος στο μονοπάτι αναζήτησης έχει μη μηδενική τιμή στο πεδίο Balance. Έστω ότι ο κόμβος του μονοπατιού αναζήτησης με μη μηδενική τιμή που βρίσκεται πιο κοντά στο νέο κόμβο που θα χρησιμοποιούσαμε για την εισαγωγή σε ένα δέντρο δυαδικής αναζήτησης είναι ο κόμβος A. Ο κόμβος A ονομάζεται **κόμβος περιστροφής** (pivot node). Ας θεωρήσουμε, για παράδειγμα, το AVL-δέντρο



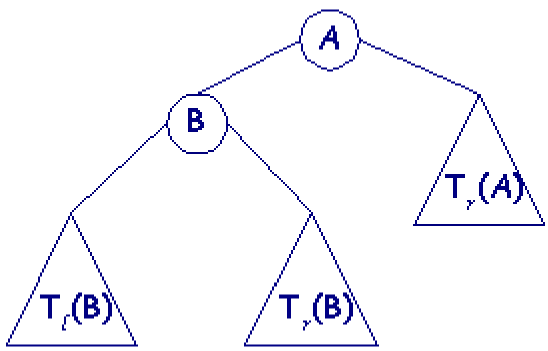
Για την εισαγωγή ενός στοιχείου με τιμή κλειδιού 5 "ακολουθούμε" το μονοπάτι αναζήτησης 8, 4, 6, ενώ για την εισαγωγή ενός στοιχείου με τιμή κλειδιού 9 το μονοπάτι αναζήτησης είναι

8, 10. Και στις 2 περιπτώσεις ο κόμβος A, στον οποίο βρίσκεται το στοιχείο με τιμή κλειδιού 8, είναι ο κόμβος περιστροφής.

Εφόσον το πεδίο Balance του κόμβου A έχει μη μηδενική τιμή τα υποδέντρα του έχουν διαφορετικό ύψος και στη συνέχεια θα αναφερόμαστε σε αυτά ως ψηλότερο και χαμηλότερο υποδέντρο. Είναι προφανές ότι αν εισάγουμε ένα στοιχείο στο χαμηλότερο από τα 2 υποδέντρα του A τότε το πεδίο Balance του A θα πάρει την τιμή 0 και το δέντρο που προκύπτει θα ικανοποιεί την ιδιότητα των AVL-δέντρων, χωρίς να χρειάζεται να κάνουμε καμία άλλη ενέργεια. Αυτό συμβαίνει αν εισάγουμε το 9 στο παραπάνω δέντρο.

Ας εξετάσουμε τώρα τι συμβαίνει αν η εισαγωγή πρέπει να γίνει στο ψηλότερο από τα 2 υποδέντρα του A. Αν εφαρμόζαμε τη διαδικασία της εισαγωγής, όπως αυτή εφαρμόζεται σε ένα δέντρο δυαδικής αναζήτησης, τότε η τιμή του πεδίου Balance θα άλλαζε σε 2 ή -2 και θα καταστρεφόταν η ιδιότητα των AVL-δέντρων. Στη συνέχεια θα δούμε πως αναδιατάσσουμε τα στοιχεία των κόμβων στο υποδέντρο με ρίζα A μαζί με το νέο στοιχείο, έτσι ώστε να δημιουργήσουμε ένα AVL-δέντρο.

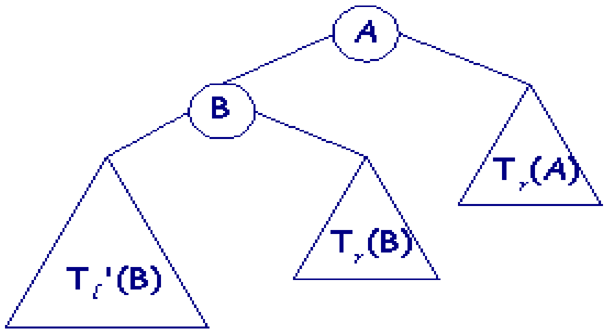
Ας υποθέσουμε ότι το ψηλότερο υποδέντρο του A είναι το αριστερό υποδέντρο, που σημαίνει ότι το πεδίο Balance του A έχει τιμή 1. Επίσης, ας υποθέσουμε ότι η ρίζα του συγκεκριμένου υποδέντρου είναι B: σύμφωνα με τον ορισμό του κόμβου περιστροφής A το πεδίο Balance του B θα έχει τιμή 0. Το υποδέντρο με ρίζα A μπορεί να αναπαρασταθεί σχηματικά, πριν από την εισαγωγή του νέου στοιχείου ως εξής



όπου Tr (A) είναι το δεξί υποδέντρο του A, ενώ Tl (B) και Tr (B) είναι το αριστερό και δεξί υποδέντρο αντίστοιχα του B. Εφόσον, για τον κόμβο A έχουμε Balance = 1 και για τον B Balance = 0, τα 3 υποδέντρα Tr (A), Tl (B) και Tr (B) έχουν όλα το ίδιο ύψος.

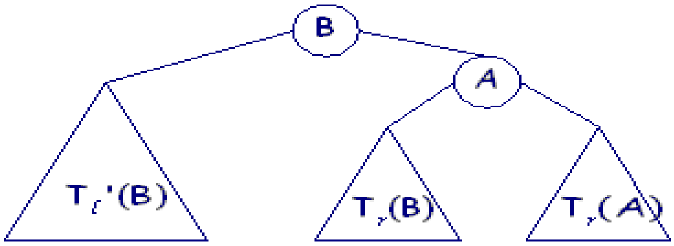
Οι περιπτώσεις που πρέπει να λάβουμε υπόψη είναι 2:

(1) στην 1η αναφερόμαστε ως **LL**: το νέο στοιχείο θα έπρεπε να εισαχθεί στο αριστερό υποδέντρο Tl (B) του B. Η απευθείας εισαγωγή όμως θα οδηγούσε στη δημιουργία ενός νέου αριστερού υποδέντρου Tl '(B) με ύψος μεγαλύτερο κατά 1 από το υποδέντρο Tl (B), και συνεπώς το υποδέντρο με ρίζα τον κόμβο A θα γινόταν



και για τον κόμβο A θα είχαμε Balance = 2, το οποίο δεν επιτρέπεται.

Σε αυτή την περίπτωση δημιουργούμε το δέντρο

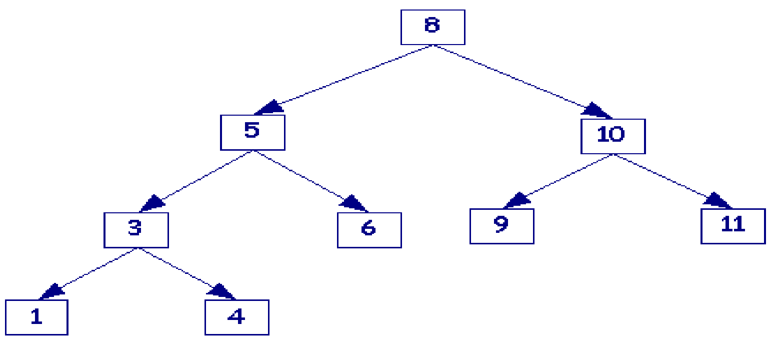


που ικανοποιεί την ιδιότητα AVL, αφού για τους κόμβους A και B έχουμε Balance = 0.

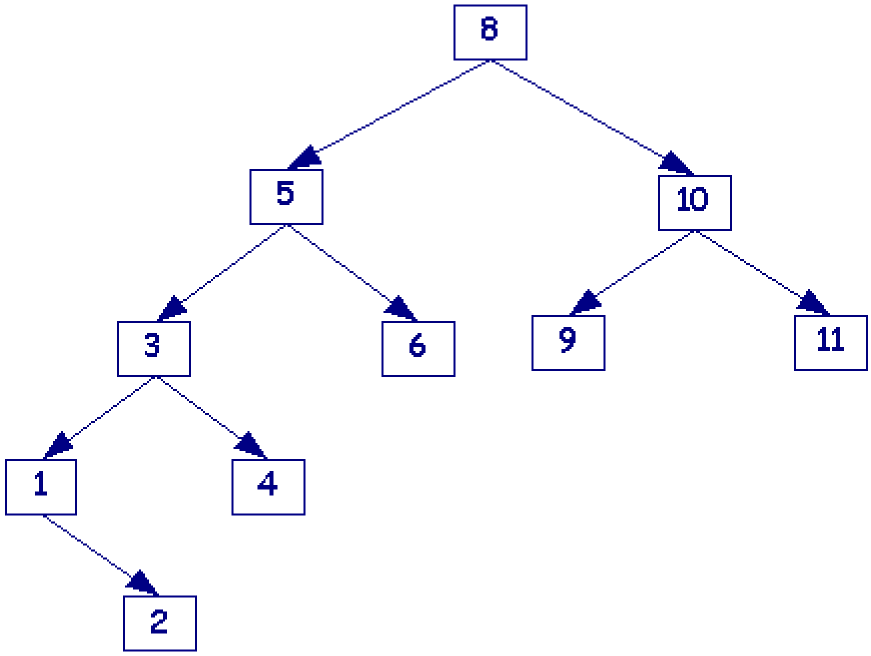
Η αναδιάταξη LL που συχνά αναφέρεται ως **LL-περιστροφή (LL-rotation)** μπορεί να περιγραφεί ως εξής:

- το αριστερό παιδί B του κόμβου περιστροφής A παίρνει τη θέση A, διατηρώντας το αριστερό υποδέντρο του
- ο κόμβος περιστροφής A γίνεται το δεξί παιδί του B, διατηρώντας το δεξί υποδέντρο του και "υιοθετώντας" ως αριστερό υποδέντρο το δεξί υποδέντρο του B.

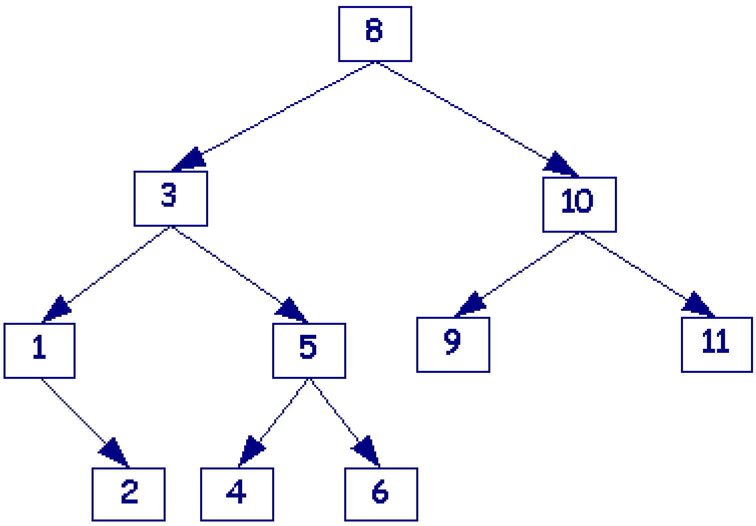
Ας θεωρήσουμε για παράδειγμα την εισαγωγή ενός κόμβου με τιμή κλειδιού 2 στο AVL-δέντρο



Το μονοπάτι αναζήτησης για το 2 είναι 8, 5, 3, 1, ο κόμβος περιστροφής A είναι ο κόμβος με τιμή κλειδιού 5 και το αριστερό παιδί B του A είναι ο κόμβος με τιμή κλειδιού 3. Η εισαγωγή, όπως πραγματοποιείται σε ένα δυαδικό δέντρο αναζήτησης, οδηγεί στο παρακάτω δέντρο, το οποίο δεν αποτελεί AVL-δέντρο

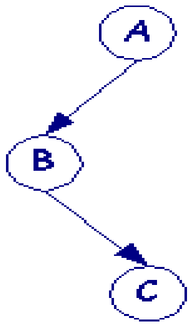


Εφόσον η εισαγωγή του 2 έγινε στο αριστερό υποδέντρο του B βρισκόμαστε στην περίπτωση LL, οπότε αναδιατάσσουμε το δέντρο με τον τρόπο που περιγράψαμε παραπάνω, δημιουργώντας το δέντρο

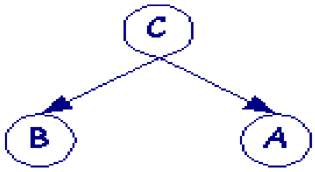


(2) η 2η περίπτωση αναφέρεται ως **LR**: στην περίπτωση αυτή το νέο στοιχείο πρέπει να εισαχθεί στο δεξί υποδέντρο Tr (B) του B. Στην συγκεκριμένη περίπτωση διακρίνουμε τρεις επιμέρους περιπτώσεις, οι οποίες αντιμετωπίζονται με παρόμοιο τρόπο.

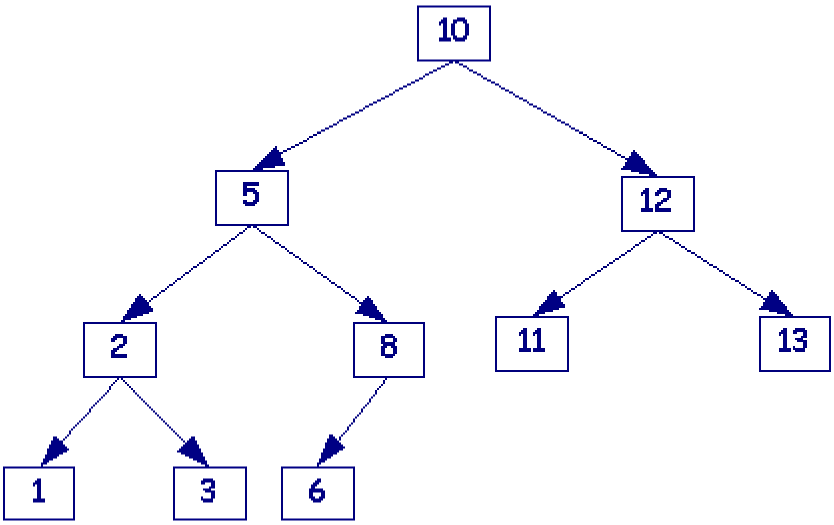
(2a) ο κόμβος B δεν έχει κανένα παιδί: σε αυτή την περίπτωση τα υποδέντρα Tl (B), Tr (B) και Tr (A) είναι όλα κενά. Σε αυτή την περίπτωση πρέπει να εισάγουμε ένα νέο κόμβο ως δεξί παιδί C του B και να αποθηκεύσουμε το νέο στοιχείο στο C, οπότε δημιουργείται το παρακάτω υποδέντρο με ρίζα τον κόμβο A



Στο δέντρο αυτό για τον κόμβο A ισχύει Balance = 2, που δεν επιτρέπεται, και για τον B Balance = 1. Για να αποκαταστήσουμε την ιδιότητα AVL αντικαθιστούμε το παραπάνω υποδέντρο με το εξής

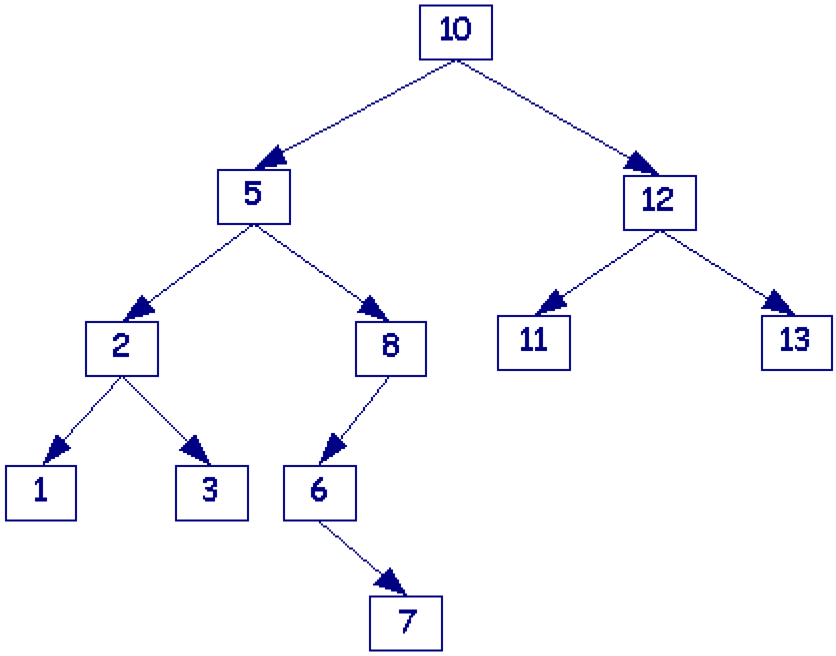


Την παραπάνω περίπτωση τη συναντάμε για παράδειγμα αν εισάγουμε ένα νέο κόμβο με κλειδί 7 στο AVL-δέντρο

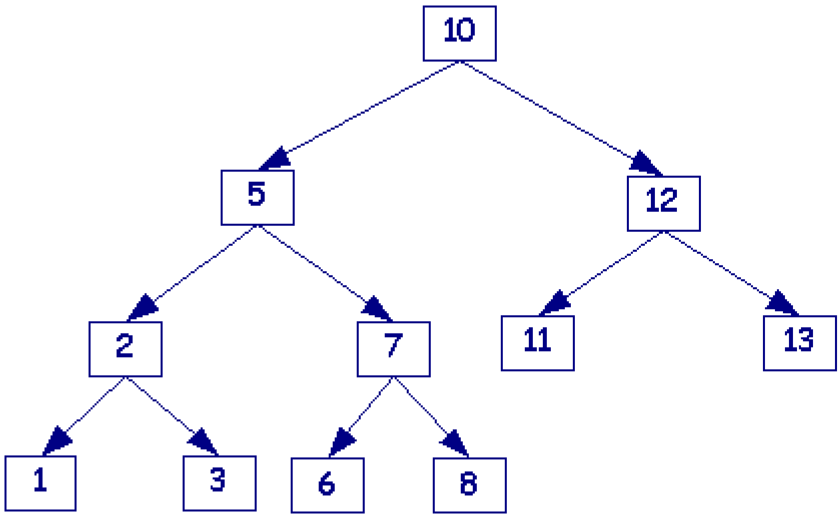


- το μονοπάτι αναζήτησης είναι 10, 5, 8, 6
- ο κόμβος περιστροφής A είναι ο κόμβος με τιμή κλειδιού 8
- ο κόμβος B, ή αλλιώς το αριστερό παιδί του A, είναι ο κόμβος με τιμή κλειδιού 6.

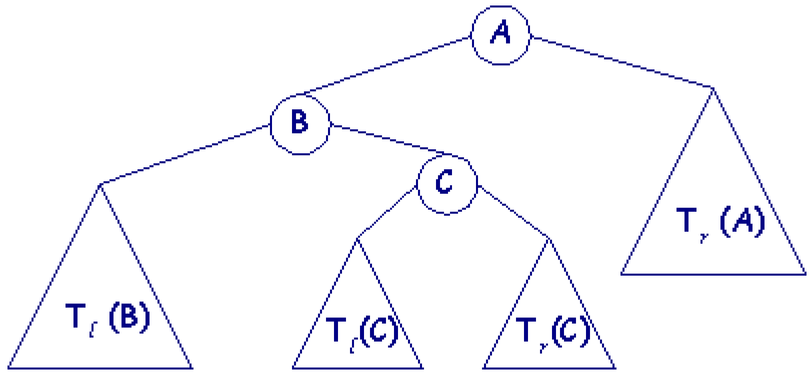
Αν εισάγουμε ένα νέο κόμβο με τιμή κλειδιού 7 ως δεξί παιδί του B προκύπτει το δέντρο



το οποίο δεν είναι AVL-δέντρο. Εφαρμόζοντας τη διαδικασία που περιγράψαμε παραπάνω προκύπτει το δέντρο

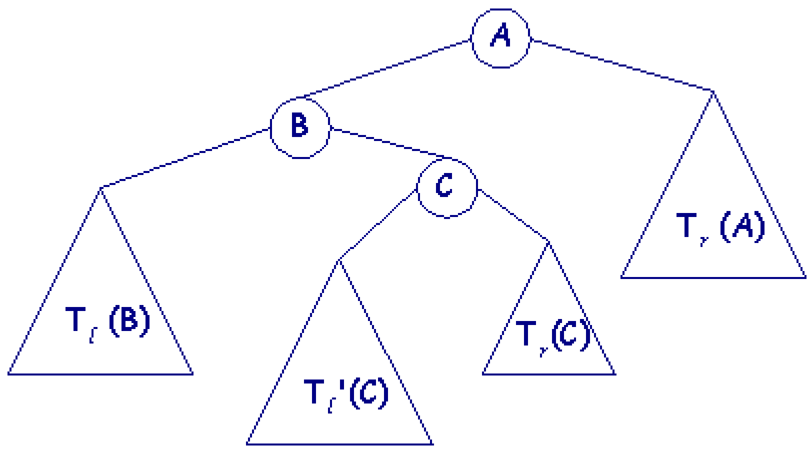


(2b) ας θεωρήσουμε τώρα την περίπτωση που το δεξί υποδέντρο του B (στο οποίο θα εισαχθεί το νέο στοιχείο) δεν είναι κενό, αλλά έχει κόμβο ρίζα - στην οποία αναφερόμαστε ως C - με αριστερό και δεξί υποδέντρο Tl (C) και Tr (C) αντίστοιχα (ενδεχομένως κενά). Η περίπτωση αυτή παρουσιάζεται σχηματικά ως εξής

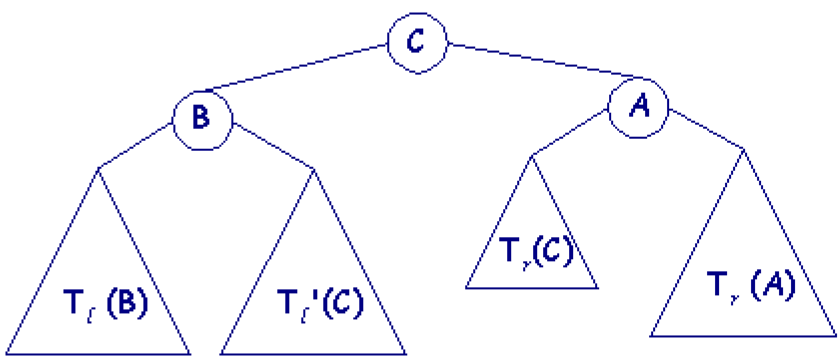


Εδώ, για τους κόμβους B και C έχουμε Balance = 0 εφόσον ο A είναι ο τελευταίος κόμβος στο μονοπάτι αναζήτησης με μη μηδενική τιμή στο πεδίο Balance. Έτσι, τα Tl (C) και Tr (C) έχουν το ίδιο ύψος, έστω h, και συνεπώς τα Tl (B) και Tr (A) έχουν ύψος h+1.

Έστω ότι το νέο στοιχείο πρόκειται να εισαχθεί στο Tl (C). Η διαδικασία της εισαγωγής σε ένα δυαδικό δέντρο αναζήτησης θα οδηγούσε σε ένα νέο αριστερό υποδέντρο Tl ΄(C) με ύψος κατά 1 μεγαλύτερο από το Tl (C). Αυτό θα είχε ως αποτέλεσμα την αλλαγή της τιμής του πεδίου Balance του κόμβου A σε 2, το οποίο δεν είναι αποδεκτό για ένα AVL-δέντρο. Η ιδιότητα AVL αποκαθίσταται αντικαθιστώντας το μη επιτρεπτό δέντρο

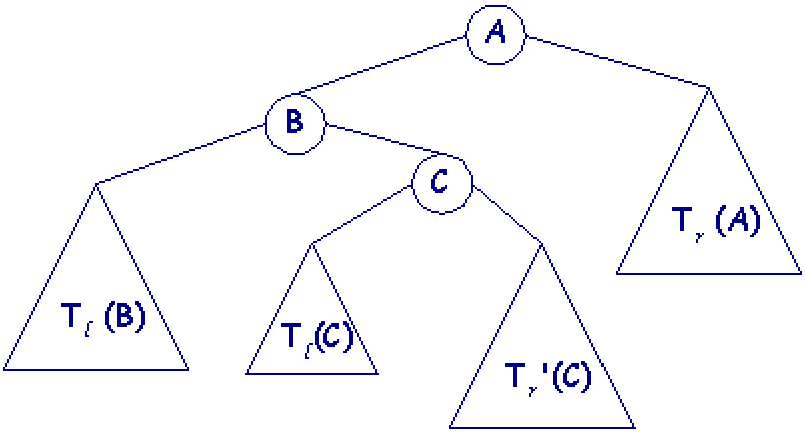


με το

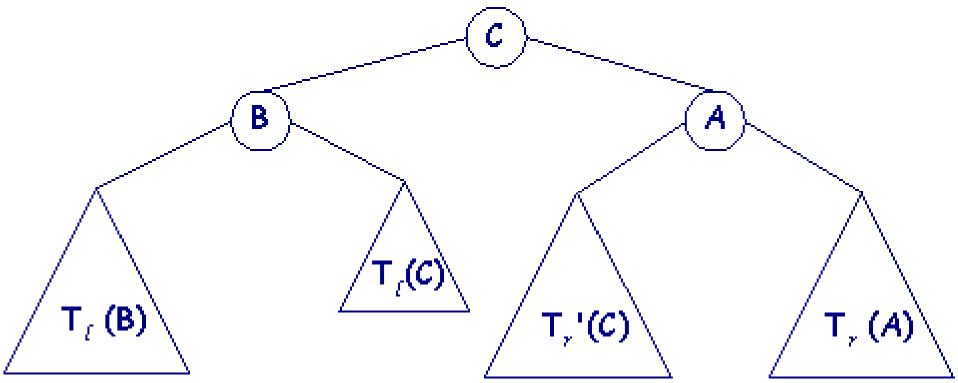


στο οποίο για τους κόμβους B και C έχουμε Balance = 0 και για τον A έχουμε Balance = -1.

(2c) ας υποθέσουμε ότι έχουμε να αντιμετωπίσουμε την περίπτωση (2b) με τη διαφορά ότι αυτή τη φορά το νέο στοιχείο πρέπει να εισαχθεί στο δεξί υποδέντρο Tr (C) του C. Η εισαγωγή, όπως πραγματοποιείται σε ένα δυαδικό δέντρο αναζήτησης, οδηγεί στο παρακάτω δέντρο που δεν είναι ένα AVL-δέντρο



όπου Tr'(C) είναι το δέντρο ύψους h+1 που προκύπτει από την εισαγωγή του νέου στοιχείου στο Tr (C). Σε αυτή την περίπτωση η ιδιότητα AVL αποκαθίσταται αντικαθιστώντας το παραπάνω δέντρο με το



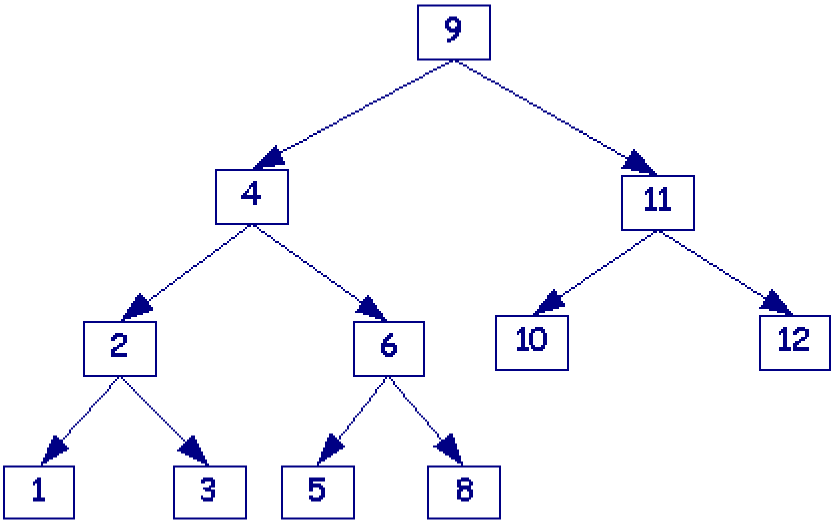
στο οποίο για τους κόμβους A και C ισχύει Balance = 0 και για τον B Balance = 1.

Οι 3 εκδόσεις της αναδιάταξης **LR-περιστροφή** (**LR-rotation**) μπορούν να περιγραφούν με τον ίδιο γενικό τρόπο. Όπως παραπάνω, θεωρούμε ότι A είναι ο κόμβος περιστροφής, B το αριστερό παιδί του και C το δεξί παιδί του B, οπότε:

- ο C παίρνει τη θέση του κόμβου περιστροφής A
- ο B γίνεται το αριστερό παιδί και ο A το δεξί παιδί του C
- ο B διατηρεί το αριστερό υποδέντρο του και έχει πλέον ως δεξί υποδέντρο το αριστερό υποδέντρο του C

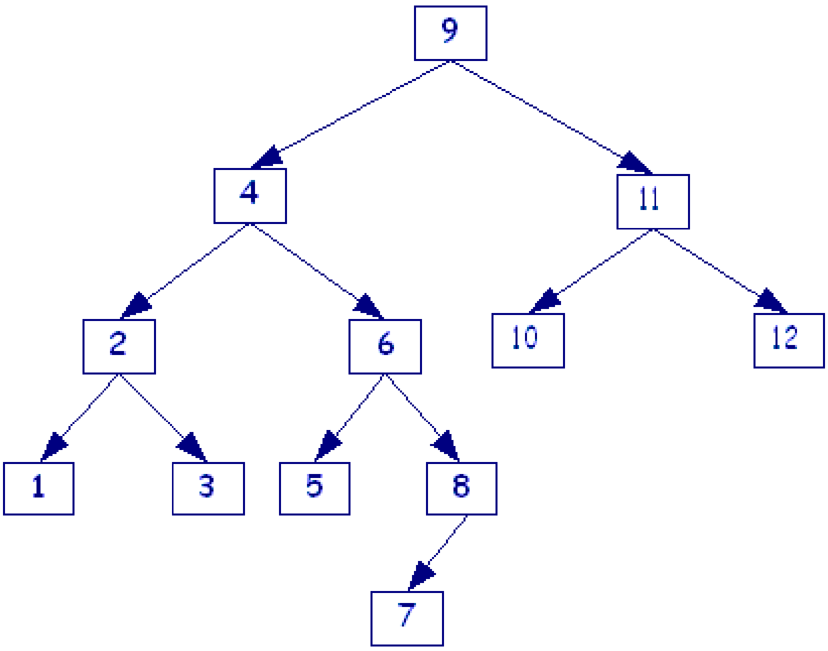
- ο A διατηρεί το δεξί υποδέντρο του και έχει πλέον ως αριστερό υποδέντρο το δεξί υποδέντρο του C.

Η περίπτωση (2c) παρουσιάζεται εξετάζοντας την εισαγωγή ενός κόμβου με κλειδί 7 στο AVL-δέντρο

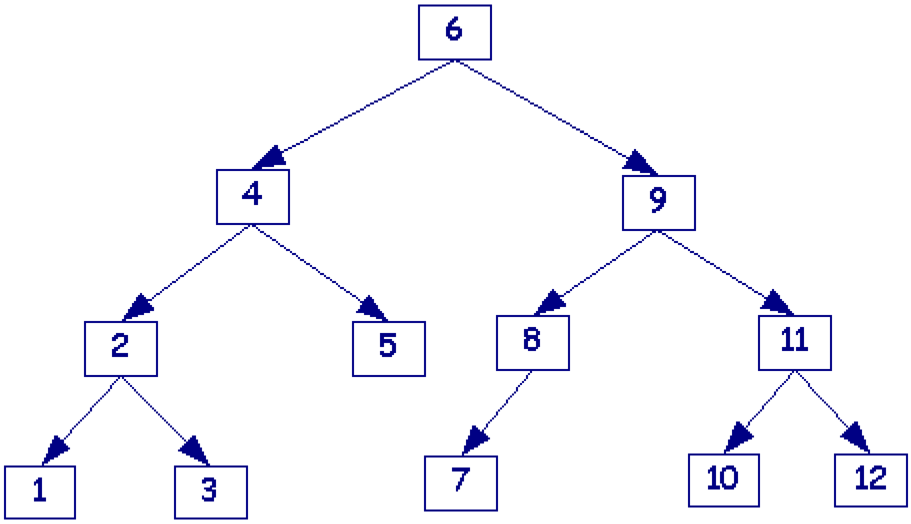


Το μονοπάτι αναζήτησης είναι 9, 4, 6, 8, ο κόμβος περιστροφής A είναι ο κόμβος με τιμή κλειδιού 9, ο B είναι ο κόμβος με τιμή κλειδιού 4 και ο κόμβος με κλειδί 7 πρέπει να εισαχθεί στο δεξί υποδέντρο του B. Επομένως, έχουμε να αντιμετωπίσουμε την περίπτωση LR(c).

Η εισαγωγή του στοιχείου οδηγεί στο παρακάτω δέντρο που δεν είναι AVL-δέντρο



και η αναδιάταξη που περιγράψαμε παραπάνω οδηγεί στο AVL-δέντρο



Μέχρι στιγμής αντιμετωπίσαμε μόνο την περίπτωση που ο κόμβος περιστροφής για μια συγκεκριμένη εισαγωγή έχει Balance = 1 και το νέο στοιχείο πρέπει να εισαχθεί στο αριστερό υποδέντρο του κόμβου αυτού. Οι περιπτώσεις που λάβαμε υπόψη ήταν 2:

- σύμφωνα με την 1η που την ονομάσαμε LL, το νέο στοιχείο πρέπει να εισαχθεί στο αριστερό υποδέντρο του αριστερού παιδιού και
- σύμφωνα με την 2η περίπτωση που την ονομάσαμε LR η εισαγωγή πρέπει να γίνει στο δεξί υποδέντρο του αριστερού παιδιού.

Επιστρέφουμε τώρα στην περίπτωση που για τον κόμβο περιστροφής ισχύει Balance = -1, οπότε το δεξί υποδέντρο του κόμβου περιστροφής είναι ψηλότερο και πρέπει να εισάγουμε το νέο στοιχείο σε αυτό το υποδέντρο. Πάλι υπάρχουν 2 περιπτώσεις, οι οποίες είναι συμμετρικές - με την έννοια που είναι και οι περιπτώσεις LL και LR που περιγράψαμε παραπάνω:

(1) RR-περιστροφή (RR-rotation): το νέο στοιχείο πρέπει να εισαχθεί στο δεξί υποδέντρο του δεξιού παιδιού B του κόμβου περιστροφής A. Σε αυτή την περίπτωση:

- ο B παίρνει τη θέση του A, διατηρώντας το δεξί υποδέντρο του
- ο A γίνεται το αριστερό παιδί του B, διατηρώντας το αριστερό υποδέντρο του, ενώ ως δεξί υποδέντρο έχει πλέον το αριστερό υποδέντρο του B.

(2) RL-περιστροφή (RL-rotation): το νέο στοιχείο πρέπει να εισαχθεί στο αριστερό υποδέντρο του B, δηλαδή στο αριστερό υποδέντρο του δεξιού παιδιού του κόμβου περιστροφής A. Αν C είναι το αριστερό παιδί του B τότε μετά την εισαγωγή, όπως αυτή πραγματοποιείται σε ένα δυαδικό δέντρο αναζήτησης,:

- ο C παίρνει τη θέση του A και έχει ως δεξί παιδί τον B και ως αριστερό παιδί τον A

- ο B διατηρεί το δεξί υποδέντρο του και έχει πλέον ως αριστερό υποδέντρο το δεξί υποδέντρο του C
- ο A διατηρεί το αριστερό υποδέντρο του και έχει πλέον ως δεξί υποδέντρο το αριστερό υποδέντρο του C.

Η μελέτη ενός μεγάλου όγκου εμπειρικών δεδομένων έχει δείξει ότι περισσότερες από τις μισές εισαγωγές σε ένα AVL-δέντρο δεν απαιτούν καμία αναδιάταξη, ενώ από εκείνες στις οποίες απαιτείται αναδιάταξη οι μισές είναι τύπου LL ή RR και οι άλλες μισές LR ή RL.