

## 5.8 B-Δέντρα

Σ' ένα **B-δένδρο** κάθε κόμβος αποτελείται από μια ταξινομημένη ακολουθία κλειδιών και ένα σύνολο δεικτών. Το πλήθος των δεικτών είναι πάντα μεγαλύτερο από το πλήθος των κλειδιών. Έτσι B-δένδρο τάξης  $d$  ορίζεται το δένδρο με τα ακόλουθα χαρακτηριστικά:

- Κάθε κόμβος περιέχει το μέγιστο  $2d$  κλειδιά και  $2d + 1$  δείκτες.
- Κανένας κόμβος εκτός της ρίζας δεν μπορεί να έχει λιγότερο από  $d$  κλειδιά.
- Όλοι οι εξωτερικοί κόμβοι βρίσκονται στο ίδιο επίπεδο.

Ως παράγοντας διακλάδωσης (fan-out factor) ορίζεται το πλήθος των δεικτών που είναι αποθηκευμένα σε κάθε κόμβο δένδρου.

Αρα με βάση τα παραπάνω 3 χαρακτηριστικά σ' ένα B-δένδρο ο παράγοντας διακλάδωσης θα είναι μεταξύ  $d + 1$  και  $2d + 1$ .

Το B-δένδρο αυξάνεται καθώς γίνεται εισαγωγή νέων κλειδιών. Όταν ένα κλειδί εισάγεται σε κόμβο που είναι γεμάτος, τότε ο κόμβος διασπάται σε 3 φάσεις.

1. Το μεσαίο κλειδί ανεβαίνει ένα επίπεδο προς τα πάνω, στον κόμβο πατέρα
2. Όλα τα κλειδιά αριστερά του μεσαίου κλειδιού μένουν στον κόμβο,
3. Τα δεξιά κλειδιά του μεσαίου κλειδιού τοποθετούνται σε νέο κόμβο.

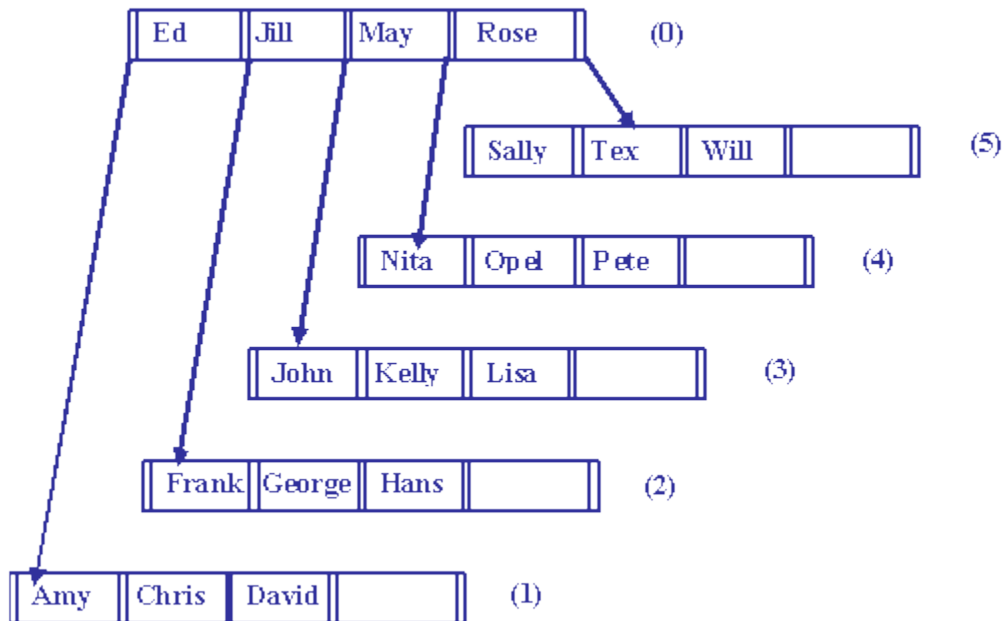
Κατά τη διαγραφή συμβαίνουν αντίστροφες διαδικασίες αυτής της εισαγωγής κλειδιού με το επιπλέον βήμα ότι ένας κόμβος δεν μπορεί να έχει λιγότερο από  $d$  κλειδιά. Αυτό επιτυγχάνεται με την περιστροφή κλειδιών μεταξύ κόμβων και τέλος με συνδυασμό κόμβων.

Μετά από εισαγωγές και διαγραφές κόμβων το B-δένδρο θα παραμείνει ισοζυγισμένο μια και παραμένει θαμνώδες (bushy) ή με μικρό ύψος λόγω του μικρού ορίου που επιτρέπεται για τον παράγοντα διακλάδωσης.

Οι εισαγωγές κλειδιών πραγματοποιούνται πάντα στους τερματικούς κόμβους. Η βασική διαδικασία συνίσταται στο να εντοπισθεί ο τερματικός κόμβος και να εισαχθεί το νέο στοιχείο του κόμβου (κλειδί και δείκτης). Επειδή υπάρχει το ενδεχόμενο ένας κόμβος να είναι ήδη γεμάτος εκείνο που επιχειρείται είναι να ανακατανείμουμε (redistribute) στοιχεία του κόμβου μεταξύ του πατέρα και του αδελφού κόμβου. Αν αυτό δεν μπορεί να συμβεί ο κόμβος διασπάται με το μεσαίο στοιχείο να ανεβαίνει ένα επίπεδο και να προστίθεται στον κόμβο πατέρα.

### Παράδειγμα

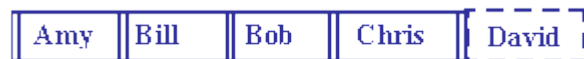
Έχουμε το παρακάτω B-δένδρο τάξης 2



Σχήμα 5.8.1. B-δένδρο 2ης τάξης.



Σχήμα 5.8.2. Ο κόμβος 1 μετά την εισαγωγή του στοιχείου "Bob".



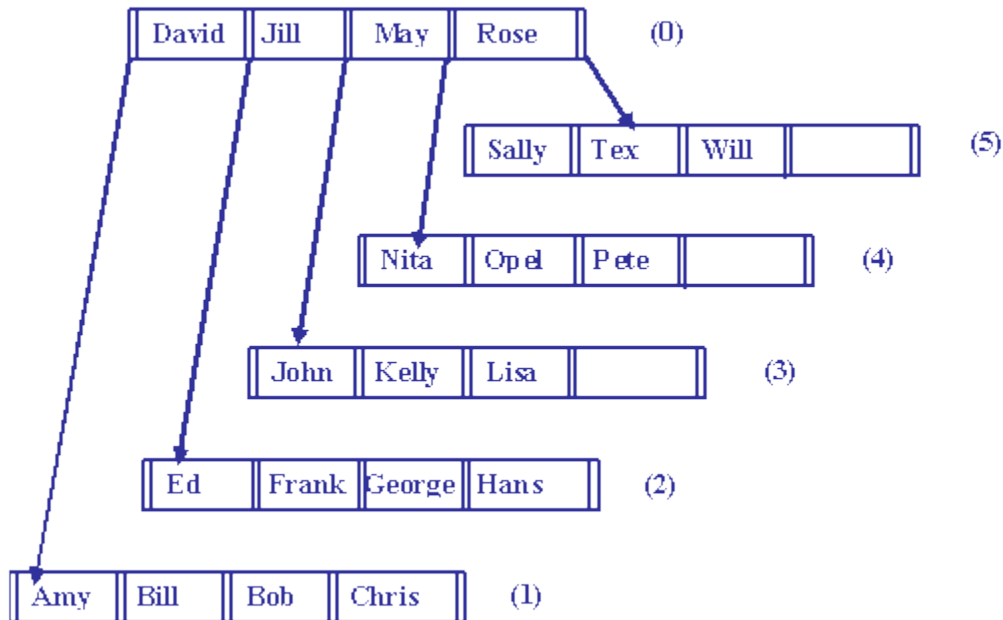
Σχήμα 5.8.3. Ο κόμβος (1) υπερχειλίζει μετά την εισαγωγή του "Bill".

Το 1ο σχήμα παρουσιάζει ένα B-δένδρο 2 επιπέδων. Θα εισάγουμε τα κλειδιά "Bob" και "Bill". Η εισαγωγή του "Bob" γίνεται χωρίς προβλήματα αφού ο κόμβος (1) δεν είναι πλήρης. Η εισαγωγή του κλειδιού "Bill" προκαλεί υπερχειλίση στον κόμβο (1). Η υπερχειλίση του κόμβου (1) διορθώνεται με την ανακατανομή. Ανακατανομή σημαίνει ότι τα πλεονάζοντα κλειδιά μοιράζονται σε γειτονικό αδελφό κόμβο που δεν είναι πλήρης. Στην περίπτωση μας ο κόμβος (2) δεν είναι πλήρης και επιλέγεται για ανακατανομή. Τα στοιχεία μεταξύ των 2 κόμβων μοιράζονται εξίσου. Στην περίπτωση μας  $(5 + 3) / 2 = 4$  στοιχεία σε κάθε κόμβο.

Όμως το κλειδί «David» δεν μπορεί να μετακινηθεί στον κόμβο (2) γιατί το κλειδί "Ed" του κόμβου (0) που χωρίζει τους κόμβους (1) και (2) θα ήταν λάθος τοποθετημένο. Η ανακατανομή γίνεται ως εξής: τίθενται σε λεξικογραφική διάταξη όλα τα κλειδιά των κόμβων (1) και (2) (αδελφών κόμβων) αλλά και το κλειδί του κόμβου (0) που τα χωρίζει. Έτσι έχουμε:

Any, Bill, Bob, Chris, **David**, Ed, Frank, George, Hans

Το μεσαίο κλειδί “David” αυτής της ακολουθίας αντικαθιστά το κλειδί “Ed” του κόμβου (0). Τα κλειδιά που βρίσκονται πριν το μεσαίο κλειδί πάνε στον κόμβο (1) ενώ τα δεξιά του στον κόμβο (2). Έτσι το δένδρο παρουσιάζεται στο Σχήμα 5.8.4



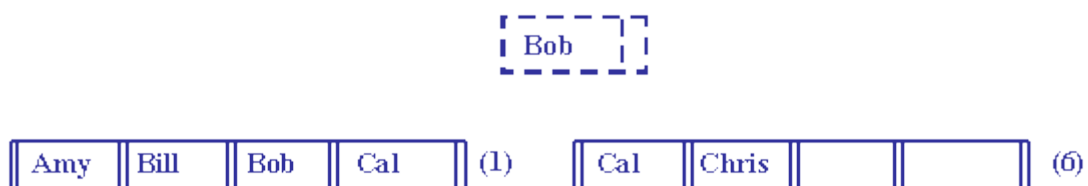
**Σχήμα 5.8.4.** Το B-δένδρο μετά την ανακατανομή των κλειδιών μεταξύ των κόμβων 0, 1, 2.

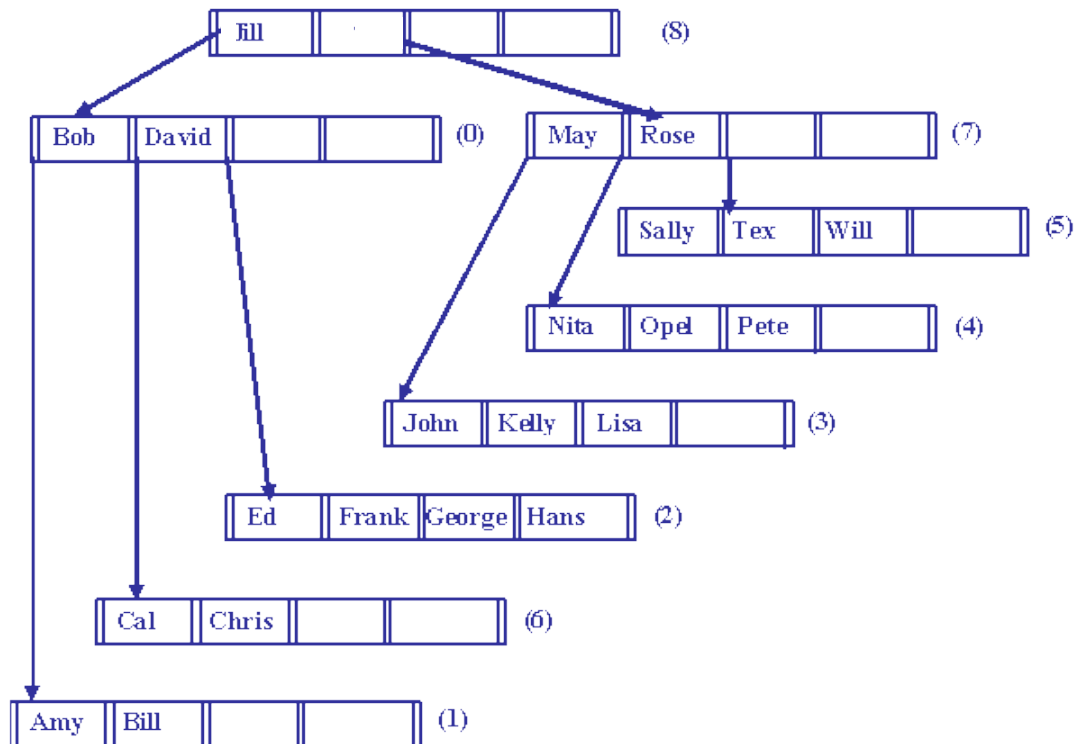
Το επόμενο κλειδί που προστίθεται είναι το “Cal” το οποίο ανήκει στον (1) κόμβο.



**Σχήμα 5.8.5.** Ο κόμβος (1) υπερχειλίζει με την εισαγωγή του “Cal”.

Και πάλι η εισαγωγή του κλειδιού δημιουργεί υπερχείλιση όπως φαίνεται και στο σχήμα 5.8.5. Σ’ αυτήν την περίπτωση δεν υπάρχει γειτονικός αδελφός κόμβος μη πλήρης έτσι δεν είναι δυνατόν να εφαρμόσουμε ανακατανομή. Αλλά είναι αναγκαίο να διασπάσουμε τον κόμβο (split). Έτσι το μεσαίο στοιχείο ανεβαίνει ένα επίπεδο, στον πατέρα κόμβο, τα αριστερά στοιχεία του παραμένουν στον αρχικό κόμβο και τα δεξιά του στοιχεία τοποθετούνται σε νέο κόμβο. Έτσι το στοιχείο “Bob” πρέπει να εισαχθεί στον κόμβο (0) που όμως με την σειρά του είναι πλήρης. Είναι ανάγκη να δημιουργηθεί νέος κόμβος. Είναι η μόνη περίπτωση που το δένδρο αυξάνει κατά ένα επίπεδο. Ο νέος κόμβος είναι ο (8) όπως φαίνεται στο σχήμα 5.8.8



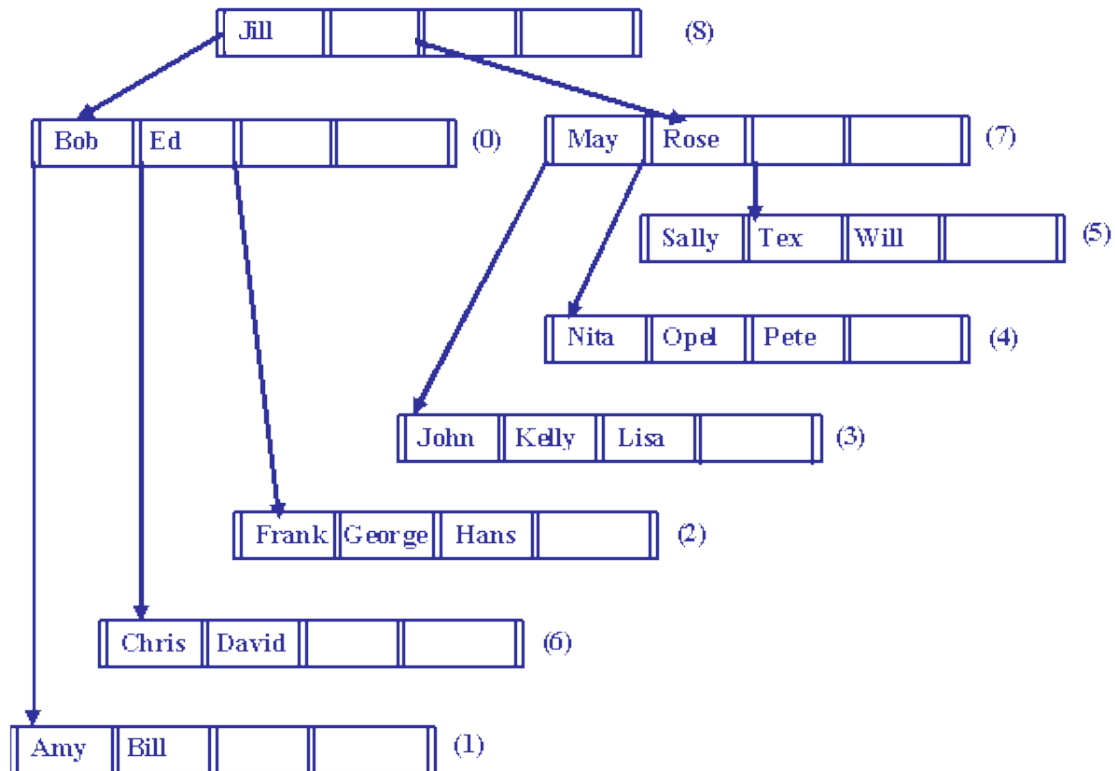
**Σχήμα 5.8.6.** Ο κόμβος (1) μετά τον διαχωρισμό.**Σχήμα 5.8.7.** Ο κόμβος (0) υπερχειλίζει μετά την εισαγωγή του στοιχείου "Bob".**Σχήμα 5.8.8.** Το B-δένδρο μετά τον διαχωρισμό των κόμβων 0 και 1.

### Διαγραφή

Η διαγραφή θα λέγαμε ότι είναι σχεδόν αντίστροφη της εισαγωγής. Η μόνη διαφορά είναι όταν το στοιχείο που διαγράφεται δεν ανήκει σε τερματικό κόμβο. Όταν συμβούν διαγραφές σε τερματικό κόμβο, όπως συνήθως συμβαίνει, το στοιχείο απομακρύνεται από τον κόμβο. Μετά απ' αυτό χρειάζεται να γίνει έλεγχος μήπως συμβεί υποχείλιση (underflow). Αυτό συμβαίνει όταν οποιοσδήποτε κόμβος έχει λιγότερα στοιχεία από  $d$  (εκτός της ρίζας).

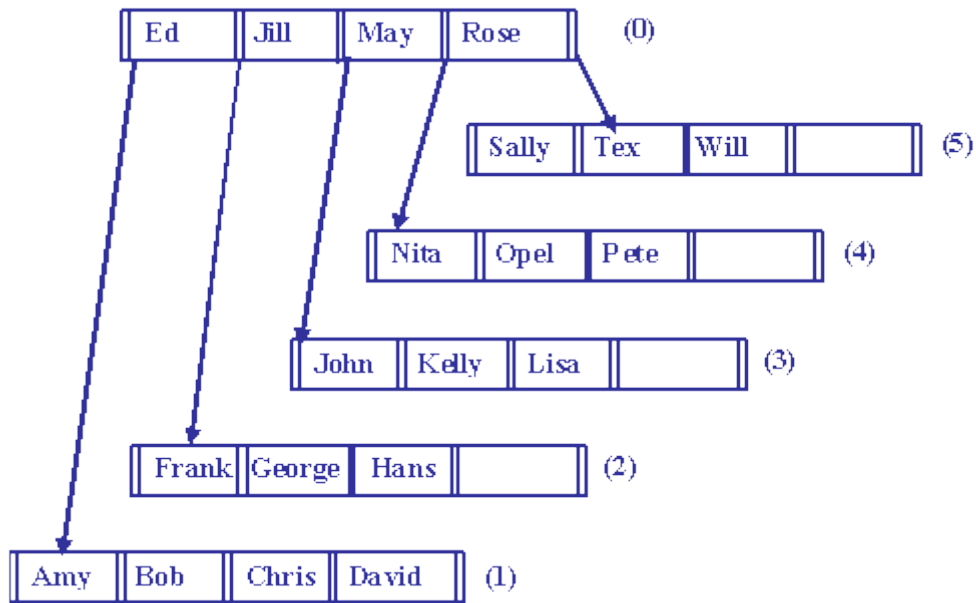
Το πρώτο διέξοδο σε περίπτωση υποχείλισης είναι η ανακατανομή. Αν κάποιος γειτονικός αδελφός κόμβος έχει περισσότερα από  $d$  στοιχεία, τα στοιχεία μοιράζονται μεταξύ αυτών των 2 κόμβων, όπως συνέβη και στην εισαγωγή. Η ίδια διαδικασία που χρησιμοποιήθηκε για την ανακατανομή κατά την εισαγωγή μπορεί να χρησιμοποιηθεί και κατά την διαγραφή. Όπως και προηγουμένως στοιχείο του κόμβου πατέρα αλλάζει. Αν δεν υπάρχει γειτονικός αδελφός κόμβος η ανακατανομή αποτυγχάνει και ο κόμβος πρέπει να συνδυασθεί με τον αδελφό κόμβο. Είναι το αντίστροφο της διάσπασης κόμβου.

Π.χ. αν από το B-δένδρο του σχήματος 5.8.8 αφαιρεθεί το στοιχείο "Cal" τότε ο κόμβος (6) υποχειλίζει. Ο (6) είναι γειτονικός του (2) που έχει περισσότερα από  $d = 2$  στοιχεία. Έτσι χρησιμοποιείται ανακατανομή για να ισοζυγιστούν οι κόμβοι (6) και (2). Έτσι ο "David" μετακινείται στον κόμβο (6) και το "Ed" εναλλάσσεται στον κόμβο (0).



**Σχήμα 5.8.9.** Το B-δένδρο μετά τη διαγραφή του "Cal" και την ανακατανομή.

Διαγραφή του "Bill" επίσης δημιουργεί υποχείλιση στον κόμβο (1). Τώρα δεν υπάρχει γειτονικός κόμβος με περισσότερα από  $d$  στοιχεία, οπότε οι κόμβοι πρέπει να συνδυαστούν. Αν συνδυαστούν οι κόμβοι (1) και (6), ο πατέρας στοιχείο που τους χωρίζει "Bob" επίσης περιλαμβάνεται στον συνδυασμένο κόμβο. Η αφαίρεση του "Bob" από τον κόμβο (0) δημιουργεί υποχείλιση. Και σ' αυτή την περίπτωση δεν υπάρχει γειτονικός αδελφός κόμβος με περισσότερα από  $d$  στοιχεία, οπότε είναι αναγκαίο να συνδυασθούν κόμβοι. Οι κόμβοι (0) και (7) συνδυάζονται και ο συνδυασμένος κόμβος περιλαμβάνει και το κλειδί "Jill" που μεταφέρθηκε από τον πατέρα κόμβο. Οπότε ο κόμβος ρίζα γίνεται κενός, οπότε διαγράφεται και ο κόμβος (0) γίνεται ρίζα.



**Σχήμα 5.8.10.** Το B-δένδρο μετά τη διαγραφή του "Bill" και τον συνδυασμό κόμβων.

Όταν συμβεί να διαγραφεί στοιχείο σε εσωτερικό κόμβο τότε αντικαθίστανται από ένα στοιχείο τερματικού κόμβου. Οπότε τροποποιείται το παραπάνω πρόβλημα σε διαγραφή στοιχείου τερματικού κόμβου. Όπως αναφέρθηκε παραπάνω το τερματικό στοιχείο μπορεί να είναι είτε το δεξιότερο στοιχείο του αριστερού υποδένδρου ή το αριστερότερο στοιχείο του δεξιού υποδένδρου. Από το δένδρο του σχήματος 5.8.10 αν αφαιρεθεί το στοιχείο "Jill" τότε μπορεί να μετακινηθεί είτε το "Hans" είτε το "John" για να το αντικαταστήσουν. Αυτό μπορεί να οδηγήσει σε υποχείλιση που αντιμετωπίζεται όπως προηγούμενα.