

# Οργάνωση Βάσεων Βιοϊατρικών Δεδομένων Εξόρυξη Γνώσης Βιοϊατρικών Δεδομένων

Ακαδημαϊκό Έτος 2016-2017

## Τεχνολογίες Πληροφορικής στην Ιατρική και τη Βιολογία

### Άσκηση 2<sup>η</sup>

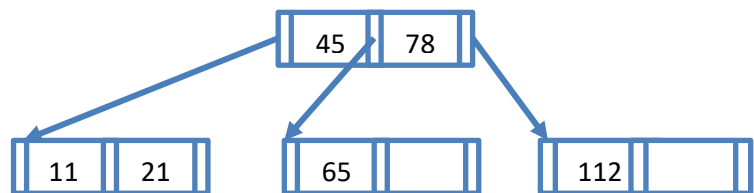
#### Δομές ευρετηρίων

##### Ερώτημα 1.

Σε κάθε υποερώτημα της παρούσας άσκησης σας δίνεται μια πράξη εισαγωγής ή διαγραφής στοιχείου σε ένα B-δένδρο. Να σχεδιάσετε το δένδρο που θα προκύψει μετά από κάθε πράξη και να δικαιολογήσετε τις επιλογές σας.

(Θεωρείστε ανεξάρτητες τις πράξεις μεταξύ τους, οπότε για κάθε μια πράξη, ξεκινάτε από το δένδρο που σας δίνεται στην εικόνα).

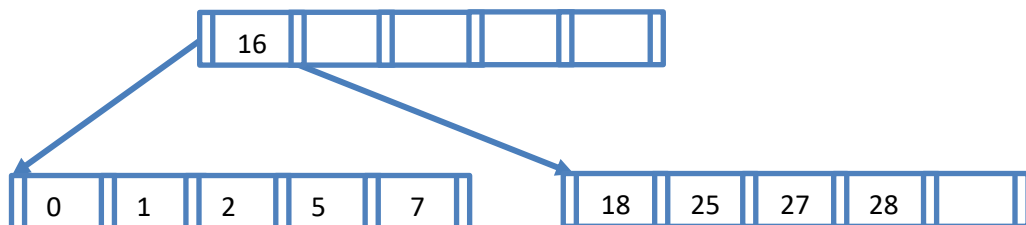
- Εισαγωγή του στοιχείου 75
- Εισαγωγή του στοιχείου 20
- Διαγραφή του στοιχείου 21
- Διαγραφή του στοιχείου 45
- Διαγραφή του στοιχείου 65
- Διαγραφή του στοιχείου 112



##### Ερώτημα 2.

Παρακάτω σας δίνεται ένα B-δένδρο τάξης 6.

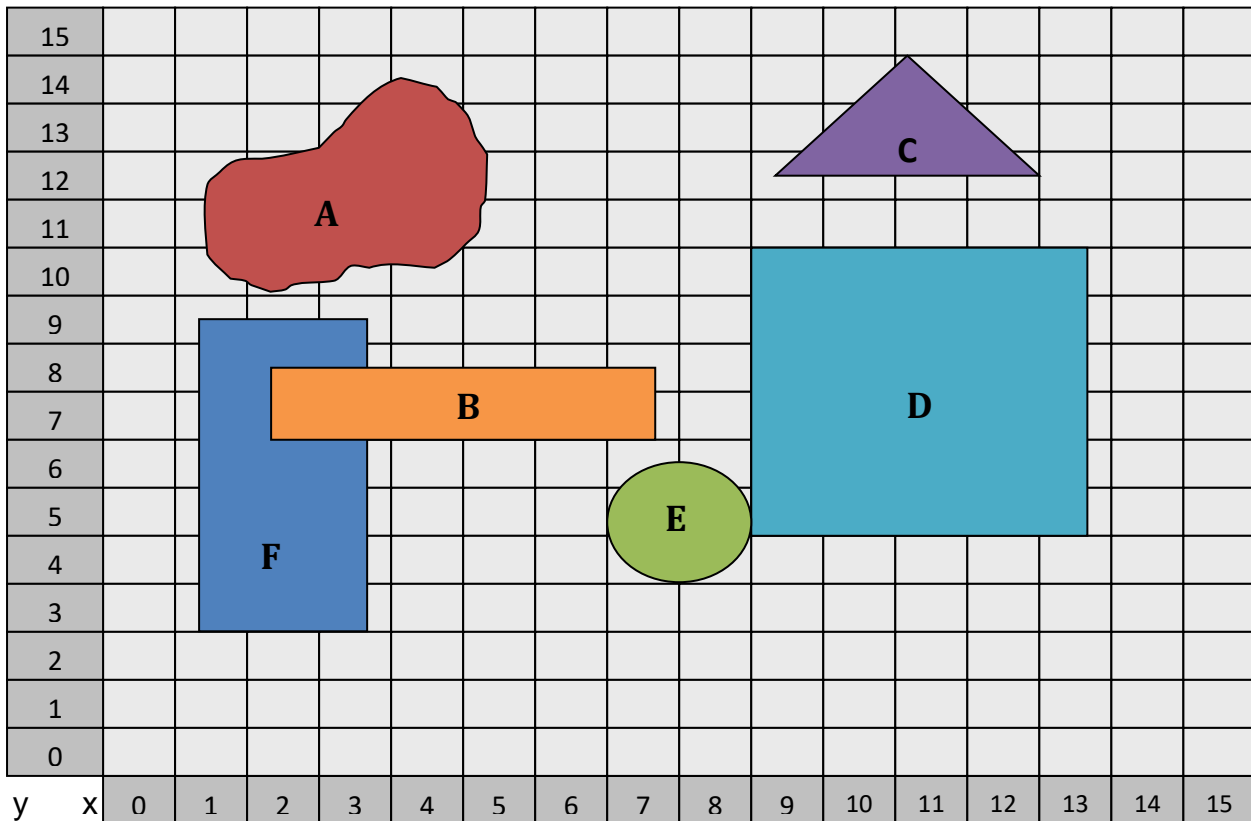
Δείξτε πώς εξελίσσεται το δένδρο κάνοντας τις παρακάτω πράξεις εισαγωγής ή διαγραφής με τη σειρά που δίνονται: (εισαγωγή του 4), (εισαγωγή του 30), (εισαγωγή του 24), (διαγραφή του 18), (διαγραφή του 24).



### Ερώτημα 3.

Παρακάτω δίνονται διάφορα αντικείμενα στο Καρτεσιανό σύστημα συντεταγμένων.

- Βρείτε τα MBRs (Minimum Bounding Rectangle) για κάθε αντικείμενο (και ορίστε τα με βάση τις συντεταγμένες).
- Υποθέτοντας ότι τα αντικείμενα τοποθετήθηκαν στο χώρο με αλφαβητική σειρά, φτιάξτε ένα R-δένδρο βαθμού 2, που μοντελοποιεί σωστά τις χωρικές σχέσεις που υπάρχουν στο χώρο. Περιγράψτε τα βήματα δημιουργίας του R-δένδρου αναλυτικά.



### Ερώτημα 4.

Θεωρήστε ένα σύστημα Βάσεων Δεδομένων που χρησιμοποιεί B-trees για την αποθήκευση των διατεταγμένων εγγραφών των σχέσεων στο δίσκο. Τα φύλλα του δέντρου περιέχουν τις ίδιες τις εγγραφές της σχέσης και όχι δείκτες προς αυτές. Ένα φύλλο καταλαμβάνει ένα block στο δίσκο και εκτός των εγγραφών της σχέσης περιέχει και ένα δείκτη στο επόμενο φύλλο. Οι εσωτερικοί κόμβοι του δέντρου καταλαμβάνουν επίσης ένα block και περιέχουν κλειδιά του δέντρου και δείκτες προς άλλους κόμβους. Επίσης θεωρήστε τα παρακάτω χαρακτηριστικά:

- Για κάθε block απαιτούνται 4096 bytes.
- Κάθε εγγραφή έχει μέγεθος 300 bytes.
- Για κάθε δείκτη σε άλλο κόμβο του δέντρου απαιτούνται 10 bytes.
- Για κάθε δείκτη σε εγγραφή της σχέσης απαιτούνται 12 bytes.
- Κάθε κλειδί έχει μέγεθος 8 bytes.
- Οι κόμβοι του δέντρου είναι κατά 85% γεμάτοι. Για παράδειγμα αν ένα φύλλο μπορεί να αποθηκεύσει μέχρι 100 εγγραφές, τότε αυτό περιέχει μέχρι 85. Ομοίως αν ένας εσωτερικός κόμβος μπορεί να αποθηκεύσει μέχρι 100 κλειδιά, ο κόμβος περιέχει μέχρι 85.
- Η σχέση R που οργανώνεται σαν B-tree περιέχει 1.000.000 εγγραφές.

**Απαντήστε στα παρακάτω ερωτήματα:**

- 1) Πόσες εγγραφές χωράνε σε ένα φύλλο;
- 2) Πόσα blocks καταλαμβάνουν τα φύλλα;
- 3) Πόσα κλειδιά χωράει ένας ενδιάμεσος κόμβος του δένδρου;
- 4) Ποιο είναι το ύψος του δένδρου;
- 5) Πόσα blocks χρειάζονται για την αποθήκευση των εγγραφών της σχέσης R (συμπεριλαμβανομένων των blocks του B-tree);

**Ημερομηνία Παράδοσης: 16/5/2017**

**Η άσκηση είναι ατομική.**

**Πρέπει να παραδοθεί στην διεύθυνση ηλ. ταχυδρομείου: [akrithara@iit.demokritos.gr](mailto:akrithara@iit.demokritos.gr)**