

## Εργασία 3: Εξωτερική Ταξινόμηση

### Εξωτερική Ταξινόμηση

Οι βασικές λειτουργίες του αλγορίθμου εξωτερικής ταξινόμησης είναι η διαίρεση του αρχικού αρχείου σε μικρότερα που θα χωράνε στη μνήμη, η ταξινόμησή τους και η συγχώνευση των επιμέρους αρχείων σε ένα μεγαλύτερο ώστε να επιτευχθεί ταξινόμηση σε ολόκληρο το αρχείο.

Πιο συγκεκριμένα, για τη ταξινόμηση χρησιμοποιείται ο αλγόριθμος της quicksort. Επιλέχθηκε καθώς είναι απλός, αποδοτικός και δεν χρειάζεται επιπλέον μνήμη. Καλείται για όλα τα επιμέρους αρχεία που φτιάχτηκαν κατά τη διαμέριση.

Αφού ταξινομηθούν τα μικρά αρχεία, συγχωνεύουμε  $bufferSize - 1$  αρχεία τη φορά, παράγοντας ένα μεγαλύτερο ταξινομημένο αρχείο, έως ότου να έχουμε μόνο ένα αρχείου ίσου μεγέθους με το αρχικό αλλά ταξινομημένο.

### Δομή Αρχείου SR

#### Το πρώτο block

Το πρώτο block κάθε αρχείου ταξινόμησης περιέχει πληροφορίες (metadata) σχετικά με το αρχείο. Πιο συγκεκριμένα, περιέχει μια δομή `SR_Metadata` με έναν αναγνωριστικό αριθμό που επιβεβαιώνει πως πρόκειται για αρχείο ταξινόμησης και πληροφορίες που είναι απαραίτητες για τις λειτουργίες που υλοποιούνται, όπως το πλήθος των εγγραφών, το πλήθος των block και το πόσες εγγραφές χωράνε σε ένα block.

#### Block δεδομένων

Όλα τα επόμενα block του αρχείου περιέχουν τις εγγραφές που έχουν εισαχθεί.

Δεν κρατάμε παραπάνω metadata διότι εφόσον η εισαγωγή εγγραφής γίνεται πάντα στο τέλος μπορούμε εύκολα να συμπεράνουμε πως όλα τα block από το δεύτερο ως το προτελευταίο είναι γεμάτα και το τελευταίο έχει  $RecordCount \bmod BlockCapacity$  εγγραφές.

### Σχόλια σχετικά με τον κώδικα

#### `CALL_OR_ERROR` macro

Προστέθηκε ένα preprocessor macro, παρόμοιο με το `CALL_OR_DIE`, έτσι ώστε να αποφευχθεί η επανάληψη του ίδιου κώδικα πολλές φορές. Ουσιαστικά, για κάθε κλήση συνάρτησης από το `bf.h` η οποία επιστρέφει `BF_ErrorCode` αν η επιστρεφόμενη τιμή δεν είναι `BF_OK`, η συνάρτηση σταματάει και επιστρέφει `SR_ERROR`.

## Βοηθητικές Συναρτήσεις

Στα αρχεία πηγαίου κώδικα έχουμε προσθέσει το `sort_file_helping.c`, ένα αρχείο με βοηθητικές συναρτήσεις. Ταυτόχρονα έχουμε προσθέσει το αντίστοιχο αρχείο βιβλιοθήκης `sort_file_helping.h` στον φάκελο `include`, όπου περιέχονται και τα υπόλοιπα αρχεία βιβλιοθηκών που δόθηκαν.

Πιο συγκεκριμένα, υλοποιούνται οι εξής συναρτήσεις:

- `sort`
- `quicksort`
- `partition`
- `getRecord`
- `recordLessThanOrEqual`
- `swapElements`
- `getMetadata`
- `splitFile`
- `merge`
- `copyFile`

Περισσότερες πληροφορίες για τη λειτουργικότητα των παραπάνω συναρτήσεων υπάρχουν στο αντίστοιχο αρχείο βιβλιοθήκης `sort_file_helping.h`