

FILE ORGANIZATION

2nd Homework Assignment

Due on: Friday, December 3, 2004

- ❖ This homework is an extension of the first homework assignment. You will study with the same data file and the same operations on it, but this time you will organize the index using B-tree file rather than a simple index on main memory.

► B-Tree Index File Description

The B tree file contains a header record with the following fields

- *m* (1 byte), the order of the B tree
- *L* (1 byte), the number of levels of the B tree
- *Root* (4 byte integer) RRN of the root node in this file
- *Up-to-date flag* (1 byte boolean: true,false) We discussed the purpose of this flag before, so please refer to the corresponding lecture note for details.

Other records (nodes) in the B tree file are of fixed length and contain the following fields:

- *numKeys* (1 byte integer) number of keys stored at the node
- *keys[0], keys[1], keys[m-1]* *m* fields storing the keys
- *pt[0], pt[1], pt[m-1]* (4 byte integer each) *m* fields storing the pointers

► B-Tree Operations

You should handle the basic operations on the B-tree index file given in the following list:

1. Creating an empty B-Tree File: *L=0, root=-1, Up-to-date=true*.
2. Searching: returns the byte offset of the data file record containing the key or -1 if the search was not successful.
3. Adding: given a pair (*key, byteOffsetInDataFile*), inserts the pair into the B-Tree, using B-tree insertion algorithm studied in the class.
4. Deleting: given the key, deletes the record from the B-Tree and also from the data file.
(Do not forget to use AVAIL LIST both for B-tree and data file)

Note that all the B-tree index operations are done directly on the file.

► Implementation Details

Implement two B-trees, one for ISBN and one for Book Title. You should consider that ISBN is primary key and uniquely identifies any record on the data file, while book title is a secondary key and generally not unique. Basically your program should provide the same functionalities expected in the first program: adding/deleting book records, searching with respect to ISBN and Book Title.

Before you get into coding, make sure that you have already answered the following questions:

1. Is there any difference between B-tree nodes for Book Title and ISBN?
2. What is inverted list?
3. Assume that you open the index file for the first time and up-to-date flag is false, what does it mean? How should you react?

You should **provide a report** with your source code including a **detailed** explanation and a **flow chart** of the source code.

You will submit your source code, data file, index files, and the report to the following e-mail address: kurt@ce.itu.edu.tr as given below

