

Data Management and File Organization

Lab Work

Linear Indexing

Linear indexes are sorted lists of key and record location values. These structures are very similar to the word index at the end of books. The main problem with a linear index is that if the index is too big to fit into the memory, we have to do the search in the index file. A search in a sorted list is generally performed using Binary search which requires $\log(n)$ access to find the record.

On the other hand we can search in sorted sequential file without using any index file using Binary search.

Answer the following questions giving reasons for your answers:

Assuming that the index file is too big to be loaded into the memory completely:

1. Is it still better to have a linear index rather than a sorted sequential file?
2. A sorted sequential file uses an overflow area to make insertions fast while we do not have any overflow in index files. Do you think insertion is faster in sorted sequential files than files with linear index? Why?
3. Is it a good idea to have overflow area in index files?
4. Is it better to mark deleted records in index files instead of deleting them physically?