

# 11/10 Database Notes

## I

attribute values  $\rightarrow$  tuple

tuple  $\rightarrow$  page

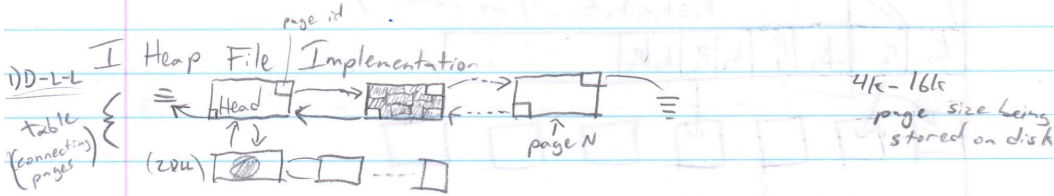
page  $\rightarrow$  File (table)

- $\rightarrow$  1. Heap File (No particular order)
- $\rightarrow$  2. Sorted File

File - connection of pages that hold a table

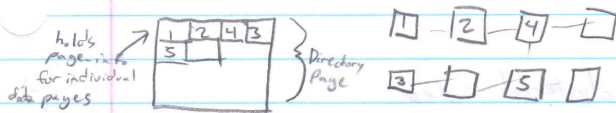
unsorted data inside

ordered data inside



	Search	Insertion	Deletion
Double linked list	$O(N)$	1 D-L-L $O(N)$ 2 D-L-L $O(2)$	$O(N)$

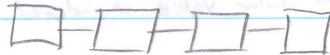
### 2) Directory-based



Main point is unordered implementations have to be read start to end to find, insert & delete pages

	Search	Insertion	Deletion
Directory Based	$O(N)$	$O(N)$	$O(N)$

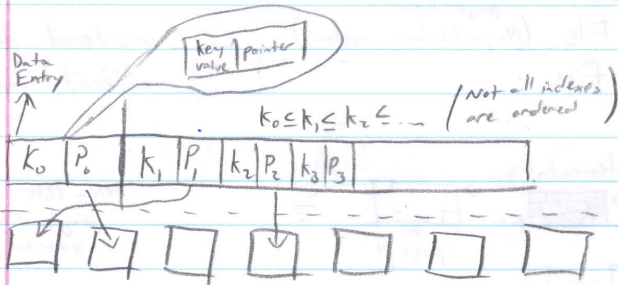
## II Sorted File Implementation - tuples are organized by the order of the values of attributes (can only be ordered one way)



	Search	Insertion	Deletion
Sorted	$O(\log N)$	$O(\log N)$	$O(\log N)$

## II Indexing

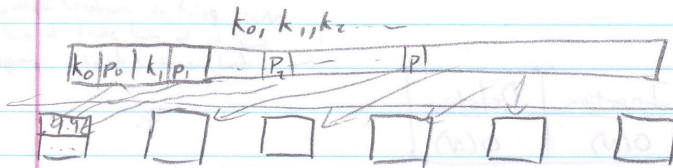
1) Index: A data structure for efficient searching.



2) Primary index vs. Secondary index

primary index is index on the primary key of the table (e.g. index on empid)

3) Clustered index vs. unclustered index



index shows where to find tuple. Cluster index gives an order, the order corresponds w/order of tuples in the data.

4) Sparse vs dense

↳ every particular value is indexed

