

Databases Class Notes 11/10/09

Little Review

How to put data into physical media?

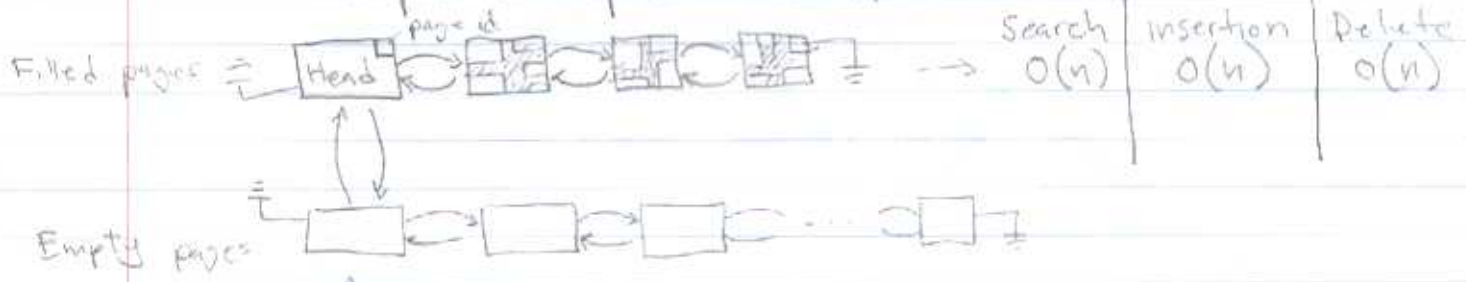
attribute values \rightarrow tuple (oriented storage)
 tuple \rightarrow page (how to put tuples in pages)
 pages \rightarrow file (how to connect pages to a file)

2 ways to do this

1. Heap file - no particular order
2. Sorted file - some kind of organization

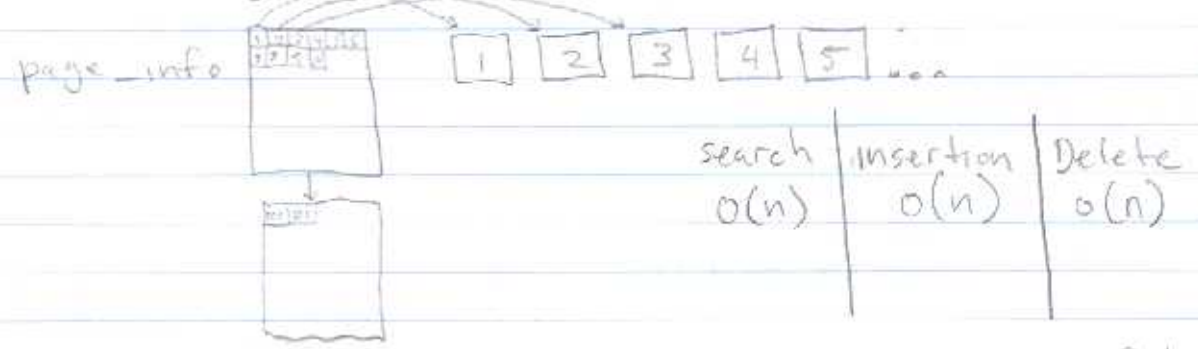
1. Heap file implementation

Time Restraints



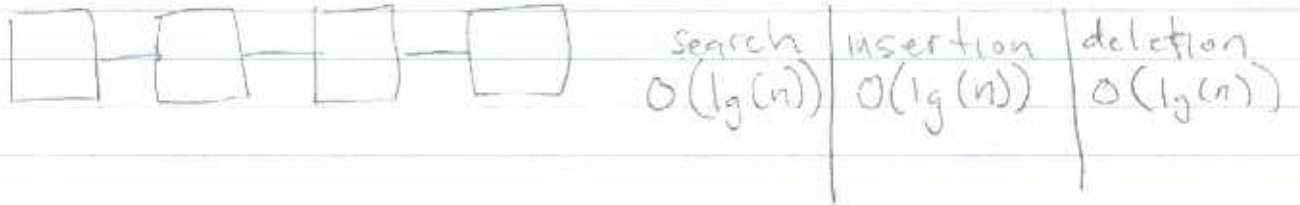
with this implementation, you have 2 linked lists and if any page in the empty LL gets filled, then move it to the filled link list, which makes insertion equal $O(1)$ instead of $O(n)$

2) Directory-based Heap file



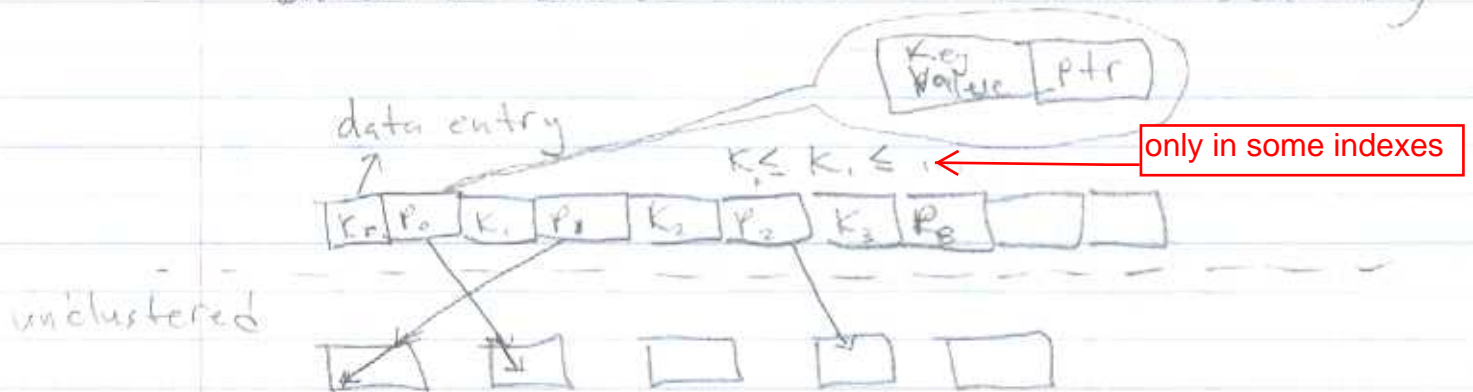
Same as DLL Heap File implementation

Sorted File - tuples are organized by order of the values of attributes



II Indexing:

Index: are data structures for efficient searching



Search key: attributes used to build indexes

I. Primary index vs Secondary Index

Primary index is the index on the primary key of the table.
(e.g. index on employee id)

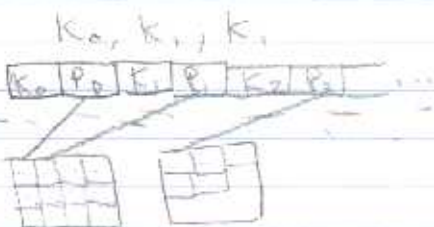
II Clustered Index vs. unclustered index

pointers don't cross cross

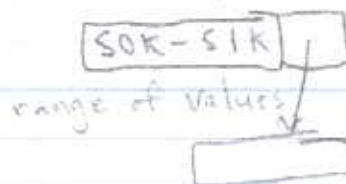
do cross cross

different ordered k when compared to tuples ordering

clustered



SPARSE VS DENSE



Clustered index: one that offers an ordered list of data entries (search key value + pointer) and the order coincides with that of the raw data (data file).

Dense index: in which you can find a data entry (search key value + pointer) for every search key value in the data file. On the contrary, you can find a corresponding data entry for only some of the tuples in a sparse index.