German University in Cairo
Media Engineering and Technology
Dr. Wael Abouelsaadat

Data Bases II, Spring 2017
Practice Assignment 1

Exercise 1-1 True or False

a) When a sparse index is used, each record must have an entry in the index.
b) Secondary indices are always dense.
c) Considering 2 levels of indices, any of them can be either sparse or dense.
d) You can only use a sparse index if the data file is sorted by the search key, while a dense index can be used for any search key.
e) In multilevel indexes, a primary index created for its first level is classified as the first level of multilevel index.
f) A type of index which is based on any candidate key or a non-key field which may have duplicate values is classified as a primary index.

Solution:

a) False, each block must have an entry in the index.
b) True, because these indexes are for columns other than the one the table is sorted with.
c) False, the first level has to be dense.
d) True, because if the data file is not sorted by the search key then we have to use a dense index.
e) False, a primary index created for its first level is classified as the second level of a multilevel index.
f) False, a type of index which is based on any candidate key or a non-key field which may have duplicate values is classified as a secondary index.

Exercise 1-2

Suppose blocks hold either three records, or ten key-pointer pairs. As a function of n, the number of records, and how many blocks do we need to hold a data file and:

a) A dense index
b) A sparse index

Solution:

a) For dense index we need a key-pointer pair for each record, and so will need $\frac{n}{10}$ blocks. For the data, we will need $\frac{n}{3}$ blocks, and so the total number of blocks is $\frac{13\,n}{30}$.
b) For sparse index we need a key-pointer pair for each of the data block, and so will need $\frac{n}{30}$ blocks. For the data, we will need $\frac{n}{3}$ blocks, and so the total number of blocks is $\frac{11\,n}{30}$.

Exercise 1-3

Repeat the exercise above if we use as many levels of index as is appropriate, until the final level of index has only one block.

Solution:

a) For dense, we will need $\log_{10} n$ levels. Total number of blocks is:

$$\begin{aligned} & 1 + 10 + 10^2 + 10^3 + \cdots + n \\ &= 1 + 10 + 10^2 + 10^3 + \cdots + 10^{\log_{10} n} \\ &= \sum_{i=0}^{\log_{10} n} 10^i \\ &= \frac{1 - 10^{\log_{10} n}}{1 - 10} \\ &= \frac{1 - n}{-9} \\ &= \frac{n - 1}{9} \end{aligned}$$

Alongside the $\frac{n}{3}$ blocks for the table, the total is

$$\frac{4n - 1}{9}$$

b) For sparse,

$$
\begin{aligned}
&1 + 10 + 10^2 + 10^3 + \cdots + \frac{n}{3} \\
&= 1 + 10 + 10^2 + 10^3 + \cdots + 10^{\log_{10} n} \\
&= \sum_{i=0}^{\log_{10} n/3} 10^i \\
&= \frac{1 - 10^{\log_{10} n/3}}{1 - 10} \\
&= \frac{1 - \frac{n}{3}}{-9} \\
&= \frac{n - 3}{27}
\end{aligned}
$$

Alongside the $\frac{n}{3}$ blocks for the table, the total is

$$\frac{10n - 3}{27}$$

Exercise 1-4

Assume that a page can hold up to two rows. If we have a Table D with 16 rows with the following keys:
D = {50, 80,100, 100, 100, 150, 150, 200, 200, 300, 410, 410, 410, 420, 430,450}
Build a 3-level index on D – one of which must be dense. Clearly explain how you handle duplicates.