Assignment 2

Posted: Oct. 2
Due: Oct. 10

1. Consider the following relational schemas:

   Suppliers(*sid*: integer, *sname*: string, *address*: string)
   Parts(*pid*: integer, *pname*: string, *color*: string)
   Catalog(*sid*: integer, *pid*: integer, *cost*: real)

   Write the following queries in relational algebra:

   a. Find pairs of sids such that the first supplier charges more for pid P123 than the second supplier
   b. Find the sids of the suppliers that supplied the most expensive parts
   c. Find the name for each supplier who has supplied two parts with different color
   d. Among all the suppliers who supply pid P123, find the sids who charge the most for this part

2. Refer to the Sailors, Boats and Reserves relation schemas defined as follows:

   Sailors(*sid*: integer, *sname*: string, *rating*: integer, *age*: real)
   Boats(*bid*: integer, *bname*: string, *color*: string)
   Reserves(*sid*: integer, *bid*: integer, *day*: date)

   Specify the following queries in relational algebra:

   a. Retrieve the bids that are reserved by exactly one sailor (possibly for multiple times)
   b. Find the sailors' ages that have reserved all the boats
   c. Find the ages of the sailors that have reserved all the boats

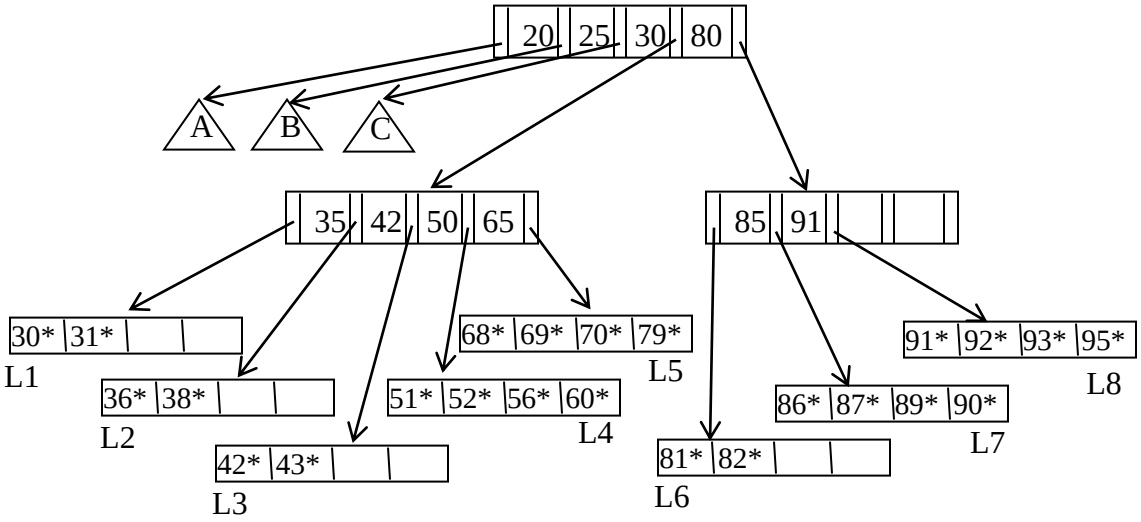
3. Consider the instance of the Students relation shown in the following figure, sorted by age. The primary key is sid, and the candidate key is login. For the purpose of this question, assume that these tuples are stored in a sorted file in the order shown. The first tuple is on page 1, the second tuple is also on page 1, and so on. Each page can store up to three data records, so the fourth tuple is on page 2. If a sparse index is used, different data entries must not duplicate page ids.

sid	name	login	age	gpa
222	Jones	Jones@cs	18	3.0
111	Madayan	madayan@music	11	3.2
444	Smith	smith@ee	18	3.2
333	Smith	smith@math	12	3.2
777	Guldu	guldu@music	19	3.8
555	Linda	Linda@cs	17	3.8

Explain what the data entries in each of the following indexes contain. If such an index cannot be constructed, say so and explain why.

a. A secondary, unclustered sparse index on name using alternative 2.
b. A secondary, unclustered dense index on name using alternative 3.
c. A secondary, clustered, sparse index on gpa using alternative 2.
d. A primary, unclustered dense index on sid using alternative 2.
e. A primary, clustered dense index on sid using alternative 2.
f. A secondary, clustered, sparse index on age using alternative 2.
g. A primary, unclustered, dense index on login using alternative 2.
h. A secondary, unclustered, dense index on age using alternative 3.

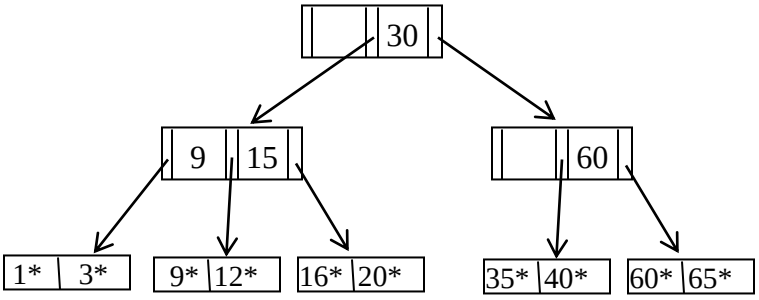
4. Consider the following B+-tree with an order of 2.



Answer the following questions:

a. Name all the tree nodes that must be fetched to answer the following query: "Get all records with search key value less than 52."
b. Insert a record with search key value 88 into the tree.
c. Delete the record with search key value 81 from the (original) tree.
d. Name a search key value such that inserting it into the (original) tree would cause an increase in the height of the tree.
e. Note that subtrees A, B and C are not fully specified. Nonetheless, what can you infer about the contents and the shape of these trees?

5. Consider the following B+-tree with an order of 1.



Answer the following questions:

a. Insert search key values 32, 21 and 75. Show how the tree expands. (You should redraw the tree if overflow occurs, otherwise simply change the values in appropriate nodes.)
b. Delete search key values 65, 60, 35, 20 and 16 (from the original tree) in that order. Show how the tree shrinks. (You should redraw the tree if underflow occurs, otherwise simply change the values in appropriate nodes.)