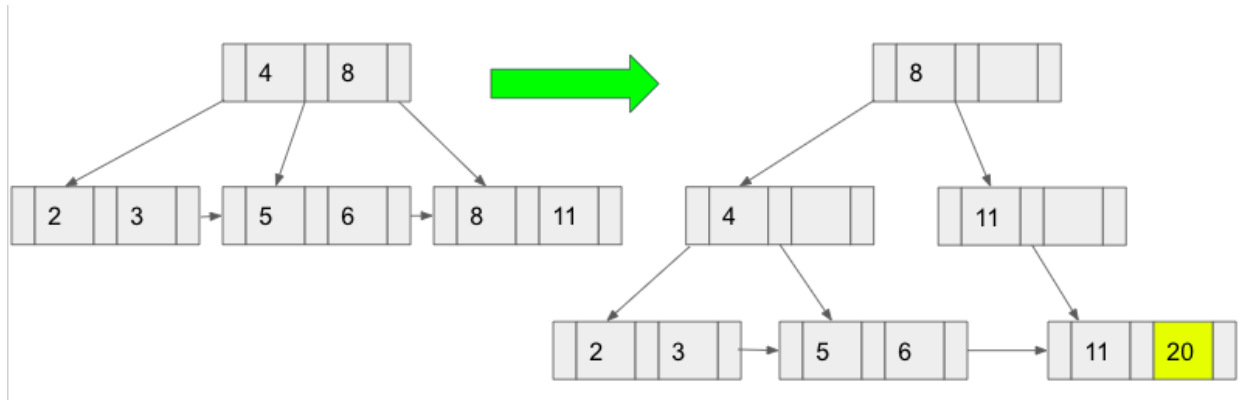


Gretel Rajamoney  
[rajamong@oregonstate.edu](mailto:rajamong@oregonstate.edu)  
 933188305

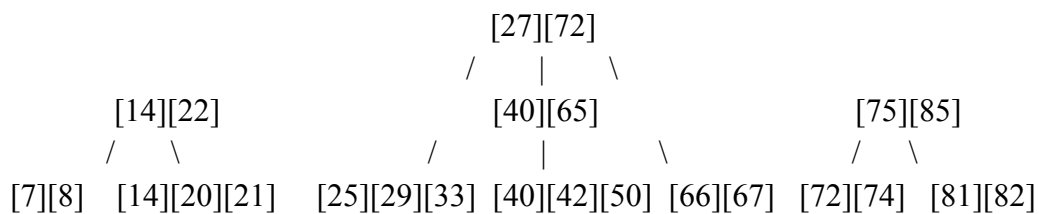
Kaavya Subramanian  
[subramka@oregonstate.edu](mailto:subramka@oregonstate.edu)  
 933291513

### Question 1: Tree Indexing

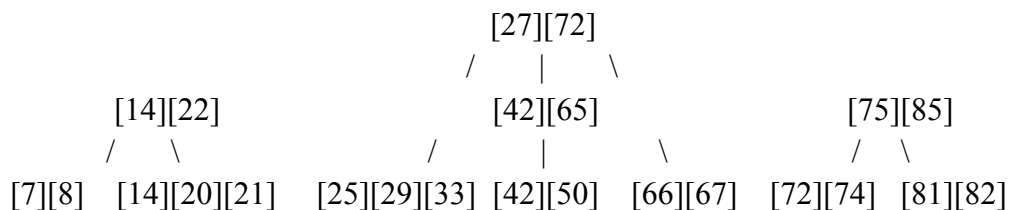
- a. The block size is 30 bytes, the data values are 8 bytes, and the record pointer is 4 bytes. So  $d$  is equal to  $(2d*8) + (2d + 1)* 4 \leq 30 \rightarrow$  each key can store between  $[1,2]$  values. The tree on the left increases in height if we add the value 20:



- b. Since the block size is 30 bytes, the data value takes in 4 bytes, and the record pointer also takes in at most 4 bytes:  $[30 / (4 + 4)] = [30 / 8] = 3$  bytes



After deleting 40 and readjusting, we are left with the following structure.



### Question 2: Indexing

- a. The first step we must take in order to consider this SQL query is to find age where age is equal to 20. This means that we must parse through all of the various ages and select only those who are 20. Since there are many people who can be 20, age is classified as a primary key. Now that we have narrowed our possibilities down to only those who are 20, we can then parse through all of the salaries and find those within our selected 20 year olds who have a salary that is greater than 20,000.

1. Find all who have age = 20
2. Search through all people with age = 20, for a salary > \$20,000

Age (Primary Key)	Salary (Secondary Key)
20	\$50,000
20	\$13,000
20	\$20,000
44	\$75,000
55	\$11,000

Since a clustered index is generally used to sort the data based on that index and store the sorted data. In B+ trees, the values that are stored are its leaves only, meanwhile the upper levels are used as keys for searching as well as branching. Another important feature is that a table can only have one clustering index due to the fact that we only have one way of storing the entirety of the data in a sorted order. These indices are generally used as a primary key since it identifies a row within the database table. Since the B+ tree index utilizes age as the primary key and salary as a hashing value it is considered clustered.

### Question 3: Sorting on External Storage

- a. Source code submitted separately through Canvas!