

CARNEGIE MELLON UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE
15-445/645 – DATABASE SYSTEMS (FALL 2017)
PROF. ANDY PAVLO

Homework 3 (by Allison Wang) – Solutions
Due: **Wednesday Oct 04, 2017 @ 11:59pm**

IMPORTANT:

- **Upload this PDF** with your answers to **Gradescope by 11:59pm on Wednesday Oct 04, 2017.**
- **Plagiarism:** Homework may be discussed with other students, but all homework is to be completed **individually.**
- **You have to use this PDF for all of your answers.**

For your information:

- Graded out of **100** points; **4** questions total
- Rough time estimate: \approx 1-4 hours (0.5-1 hours for each question)

Revision : 2017/10/22 09:35

Question	Points	Score
Extendible Hashing	30	
Linear Hashing	20	
B+ Tree	38	
Skip List and Radix Tree	12	
Total:	100	

Question 1: Extendible Hashing.....[30 points]

Graded by:

(a) Consider an extendible hashing structure that

1. Each bucket can hold up to 2 records
2. Is initially empty (only one empty bucket)

Consider the result after inserting key 8, 16, 4, 3, 11, 12 in order, using the lowest-bits for the hash function. That is, records in a bucket of local depth d agree on their **rightmost** d bits. For example, key 4 (0100) and key 12 (1100) agree on their rightmost 3 bits (100).

i. **[5 points]** What is the global depth of the resulting directory?

- 0 1 2 3 4 5 6 7

ii. **[5 points]** What is the local depth of the bucket contains key 3?

- 0 1 2 3 4 5 6 7

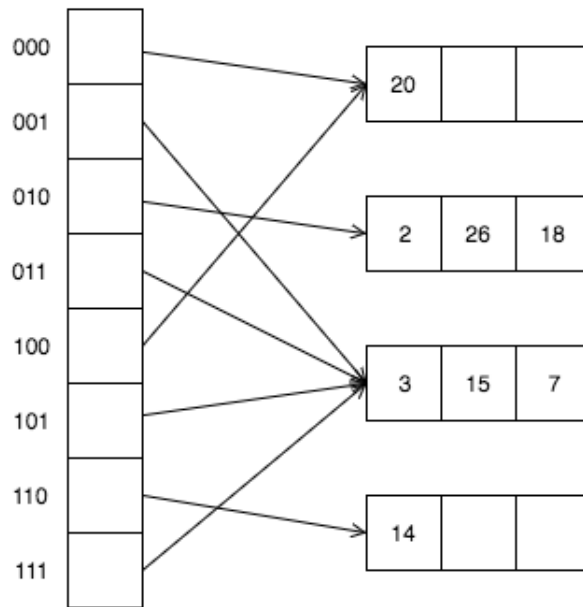


Figure 1: extendible hashing

(b) Answer the following questions about Figure 1. Suppose we insert keys 30, 21, 12, 42, 31 in order.

i. **[5 points]** Which key will cause the first split?

- 30 21 12 42 31 None of the above

ii. **[5 points]** Which key will first cause the directory to double in size?

- 30 21 12 42 31 None of the above

(c) Start from the hash table of Figure 1. Consider the result after deleting keys 18, 14 in order.

i. **[5 points]** How many buckets will remain?

- 0 1 2 3 4 5 None of the above

- ii. [5 points] What is the local depth of the bucket containing key 2?
 0 1 2 3 4 5 None of the above

Solution: Either (i)3 (ii)2, or (i)2 (ii)1 is correct.

Question 2: Linear Hashing.....[20 points]**Graded by:**

- (a) [5 points] Linear hashing always splits the bucket that overflows.
 True **False**
- (b) Answer the following questions for the hash table of Figure 2. Assume that a bucket split occurs whenever an overflow page is created. $h_0(x)$ takes the rightmost 2 bits of key x as the hash value, and $h_1(x)$ takes the rightmost 3 bits of key x as the hash value.

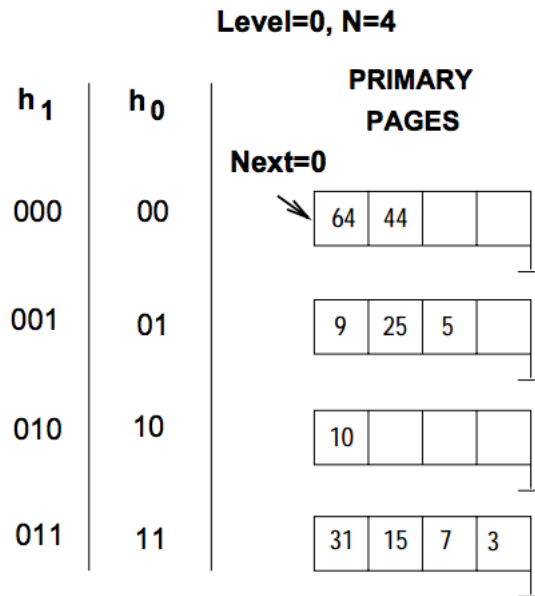


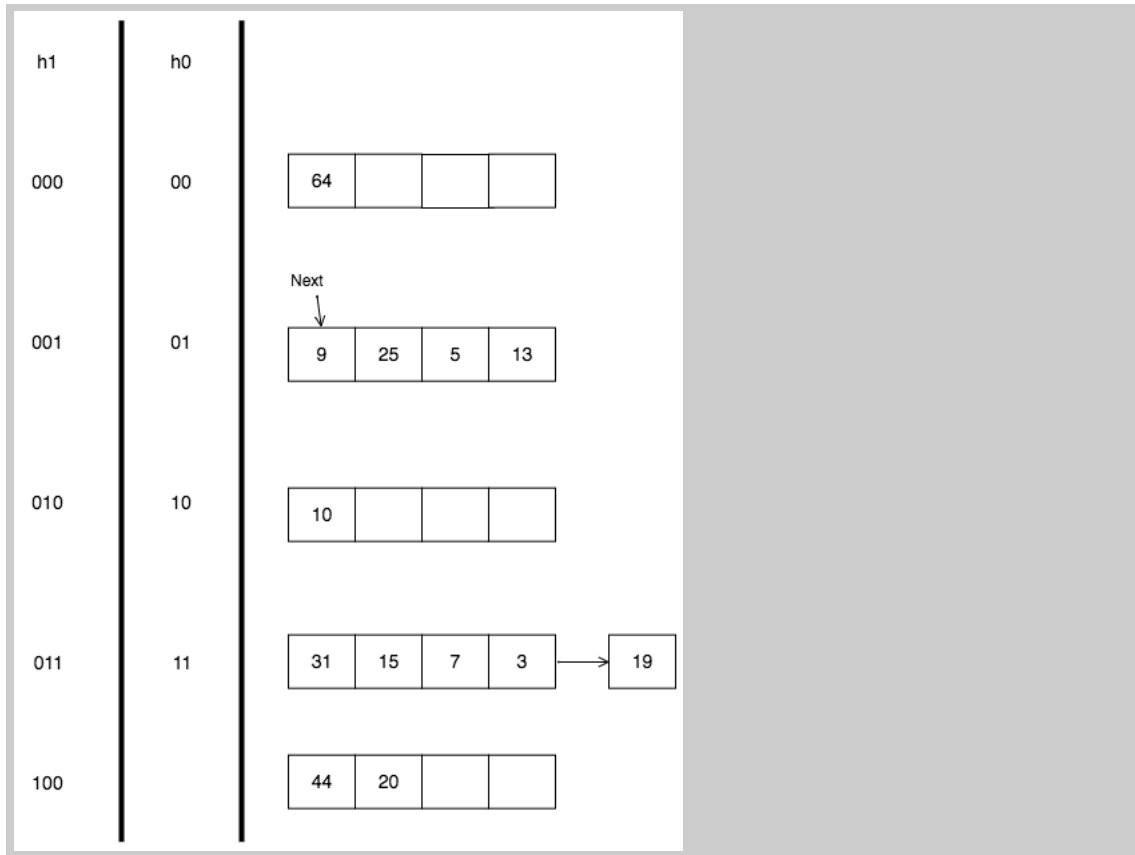
Figure 2: Linear Hashing

- (c) [5 points] What is the largest key less than 25 whose insertion will cause a split?

Solution: 23

- (d) [10 points] Starting from the hash table of Figure 2, plot the final hash table, after inserting 13, 19, 20. Remember to indicate the new hash function (if any), and to update the “Next” pointer, if needed. You may use the draw.io template, at:
<http://cmudb.io/fall2017-hw3-q2>

Solution: see below



Question 3: B+ Tree [38 points]

Graded by:

- (a) [5 points] Consider the most sparse B+ tree of order $d = 1$ containing the keys 1 through 15 inclusive. How many nodes does the B+ tree have?

Solution: This question is removed. We consider all solutions to be correct.

- (b) [5 points] Consider the B+ tree in Figure 3, what is the minimum number of pointers to be followed to satisfy the query: Get all records with key greater than 11 and less than 20?

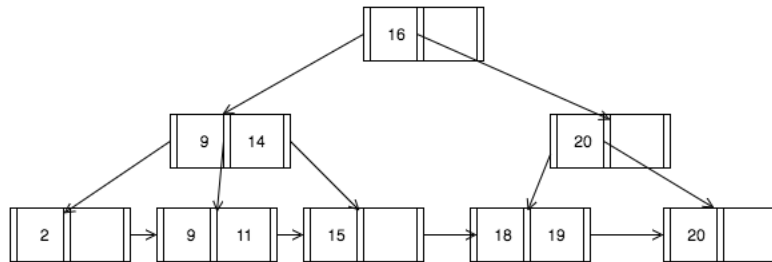


Figure 3: B+ Tree

Solution: 5

- (c) Consider the B+ tree in Figure 4 of order $d = 2$ and height $h = 2$ levels. Please make the following assumptions:
 - With respect to “ \geq ”, follow the convention used in the textbook, and in Figure 4, that is, the left pointer is for $<$, the right one for \geq .
 - In case of underflow, if you can borrow from both siblings, choose the one on the right.

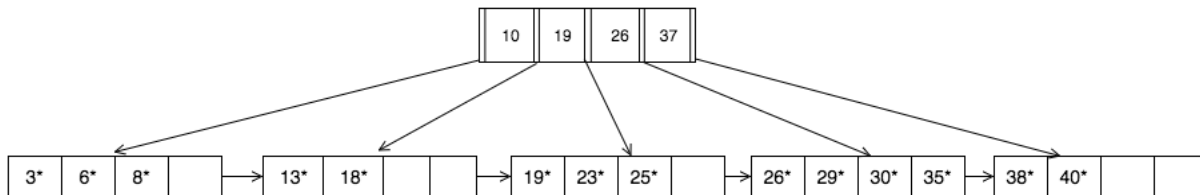


Figure 4: B+ Tree of order $d = 2$.

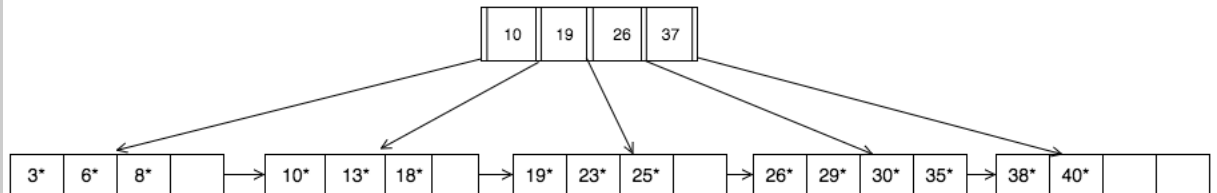
For all questions below, use the standard B+ tree algorithm given in the slides and the textbook (on insertions: 2-to-1 split, no deferred splits; on deletions: no underflowing pages). For your drawing convenience, you may use the draw.io template, at:

<http://cmudb.io/fall12017-hw3-q3>

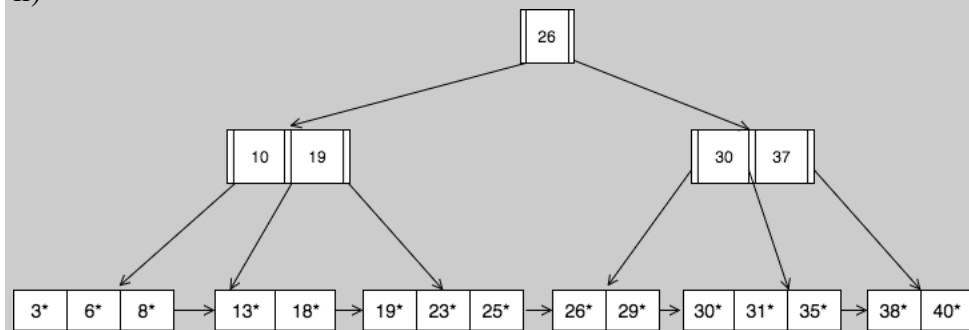
In all cases, start from the B+ tree of Figure 4.

- i. [4 points] Start from the original B+ tree; insert 10*.
- ii. [8 points] Start from the original B+ tree; insert 31*.
- iii. [4 points] Start from the original B+ tree; delete 19*.
- iv. [6 points] Start from the original B+ tree; delete 40*.
- v. [6 points] Start from the original B+ tree; delete 18*.

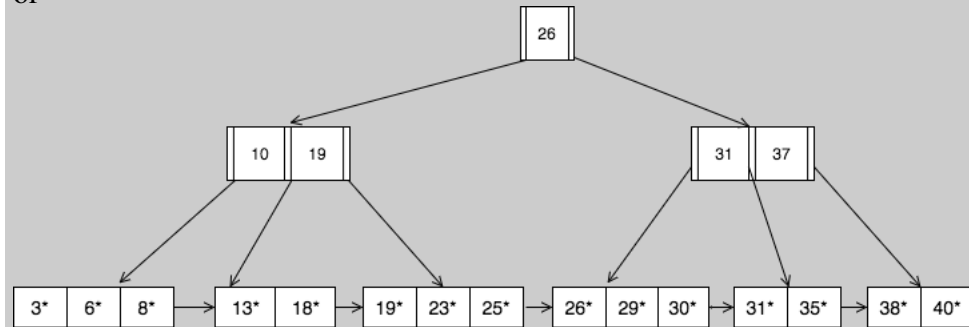
Solution: i)



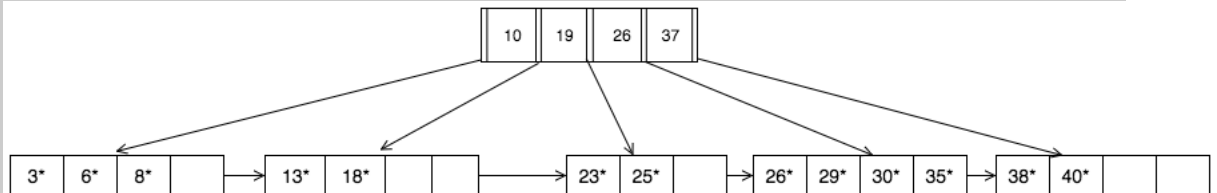
ii)



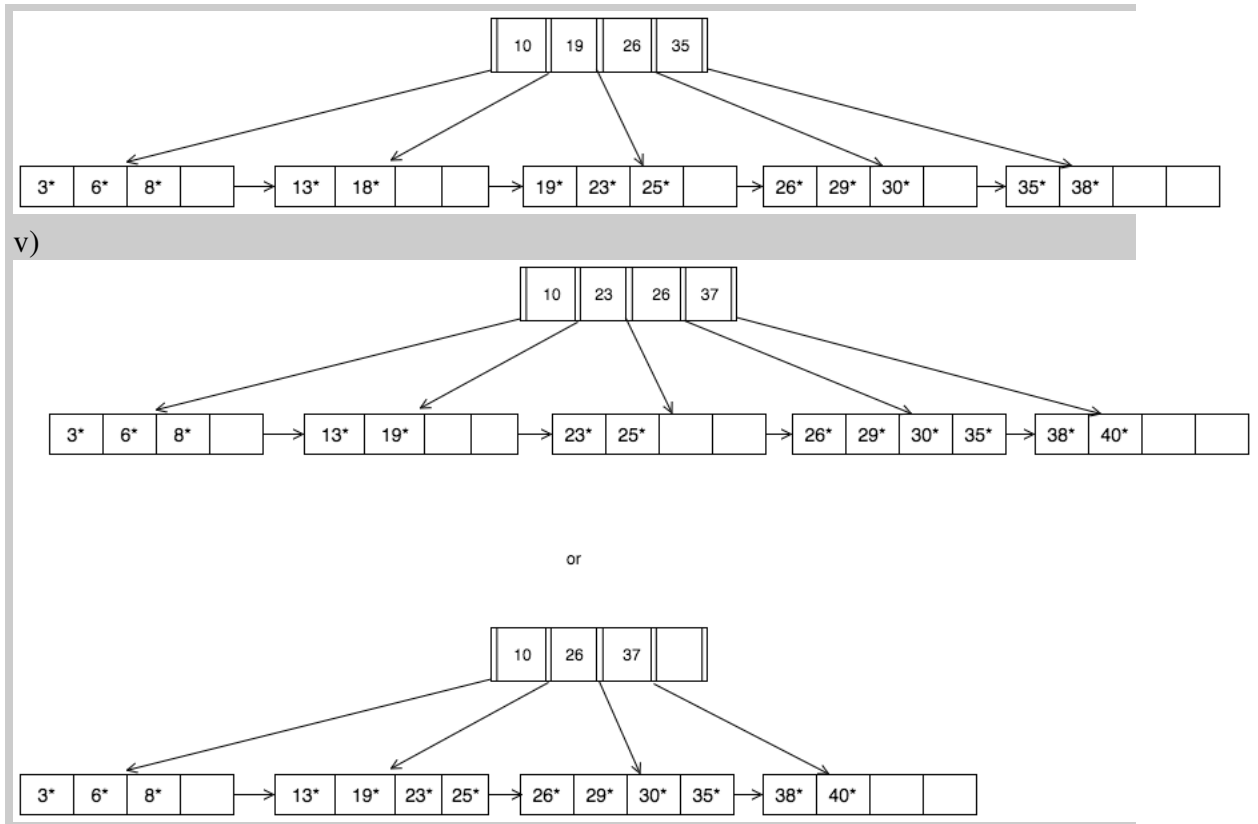
or



iii)



iv)



Question 4: Skip List and Radix Tree.....[12 points]
Graded by:

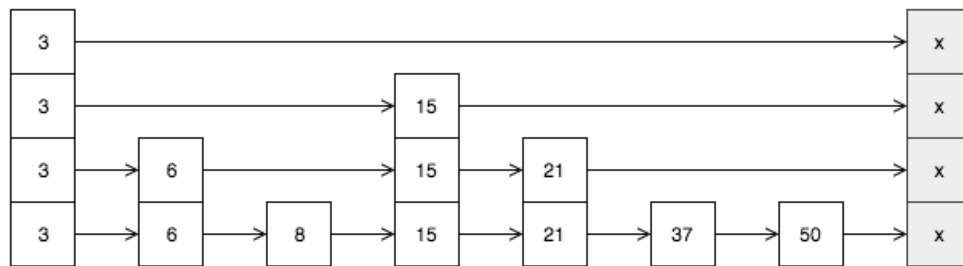


Figure 5: Skip List

(a) [6 points] Consider the skip list in Figure 5. Suppose we want to insert key 35. Which of the following node(s) might directly point to the new node? Select all that apply.

- 3
 6
 8
 15
 21
 37
 50
 Nil
 None of the above

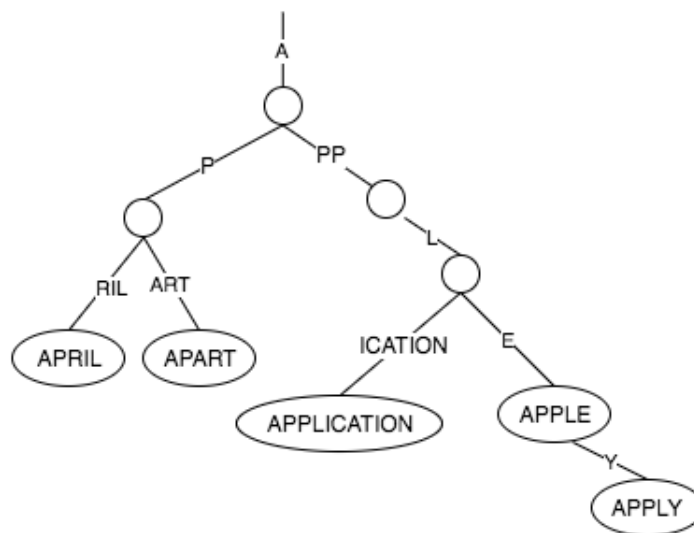


Figure 6: Radix Tree

(b) [6 points] Consider the radix tree in Figure 6. Is it a valid radix tree? If yes, draw the tree after inserting the new word "approve". Else, draw the valid radix tree with existing words. You may use the draw.io template at:

<http://cmudb.io/fall2017-hw3-q4>

Solution: Not valid

