Connie Ma
56047129 // p6z8
CPSC 221 202
2.4.17

THEORY ASSIGNMENT #3

PART ① – Hashing

① Want to hash a 11-digit account number, which of the following would be best? (Use open addressing w/ 1000 cells, 250 keys, linear probing to resolve collisions.)

$h_1(k) = \text{floor}(\sqrt{k}) \% 1000$

$h_2(k) = (abcd) \% 1000$ where $abcd = i_2 i_4 i_6 i_8$ of 11-digit account

$h_3(k) = [(\#0\text{s in } k) + (\#1\text{s in } k) + \ldots + (\#9\text{s in } k)] \% 1000$

(max number of numbers you can have is 11, so everything will hash to 11 % 1000) → First of all h_3 is terrible because it will hash every account # to the same value, which coupled with linear probing will suffer from primary clustering. Unless

(if the keyspace was smaller then this wouldn't be so bad) → h_1 utilizes the sqrt function which is so much more computationally expensive than selecting the 2^{nd}, 4^{th}, 6^{th} & 8^{th} digits. Like wikipedia says the time complexity of computing number w/ n digits is equivalent to that of multiply 2 n-digit numbers.

→ So by default h_3, h_1 suck & $\boxed{h_2}$ is the best out of the choices.

↳ But just to confirm, you can already tell w/ h_2 keys will hash to more unique hash values by virtue of probability → 9×9×9 diff possible combinations, than h_3 provided account # generation is random. ← 729 > 250

② Using hash table m = 11 and h(k) = k % 11, show the hash table that results from inserting:

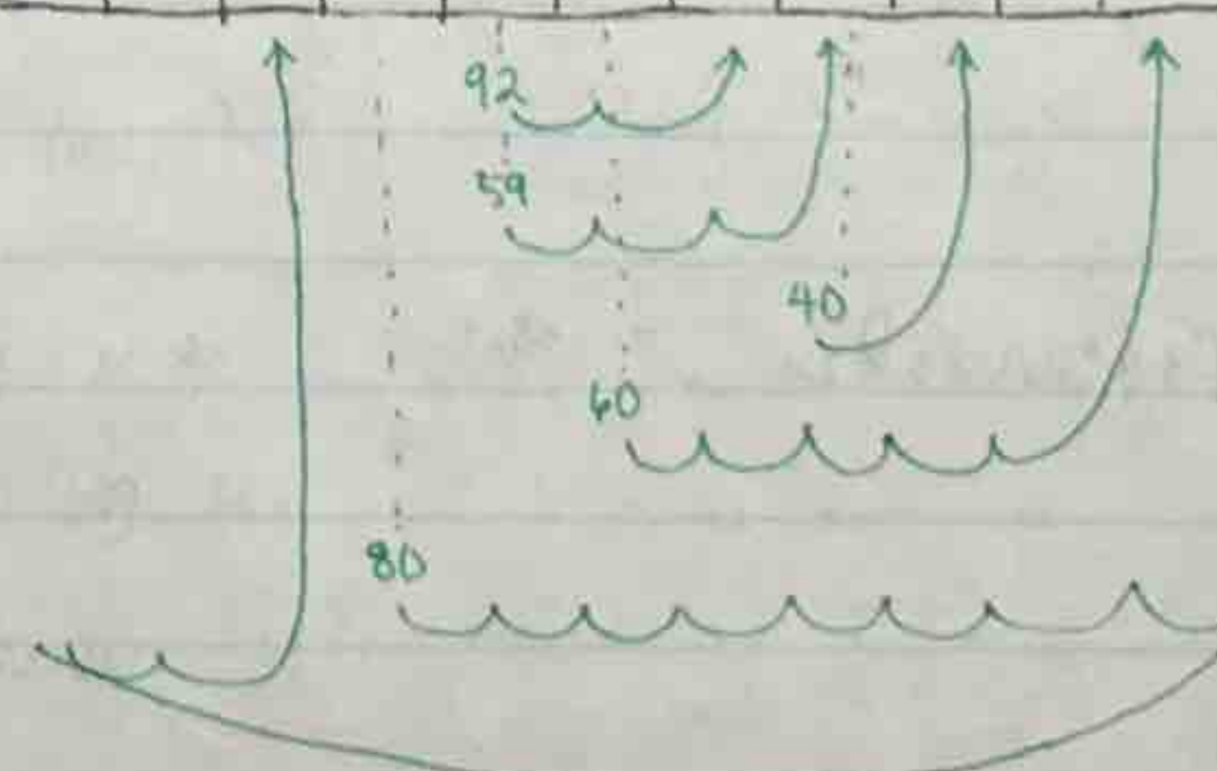
26, 42, 5, 44, 92, 59, 40, 36, 12, 60, 80

using a) linear probing / b) quadratic probing / c) double hashing w/ $h_2(x) = 2x \% 11$ if non-zero, $h_2(x) = 1$ if expression in prev line is zero for resolving collisions. Write failing cases if any after m tries.

26 → 4	36 → 3
42 → 9	12 → 1
5 → 5	60 → 5
44 → 0	80 → 3
92 → 4	
59 → 4	
40 → 7	

a) linear probing

0	1	2	3	4	5	6	7	8	9	10
44	12	80	36	26	5	92	59	40	42	60



Connie Ma
56047129 // p6z8
CPSC 221 202
2.4.17

THEORY ASSIGNMENT #3

pg 2

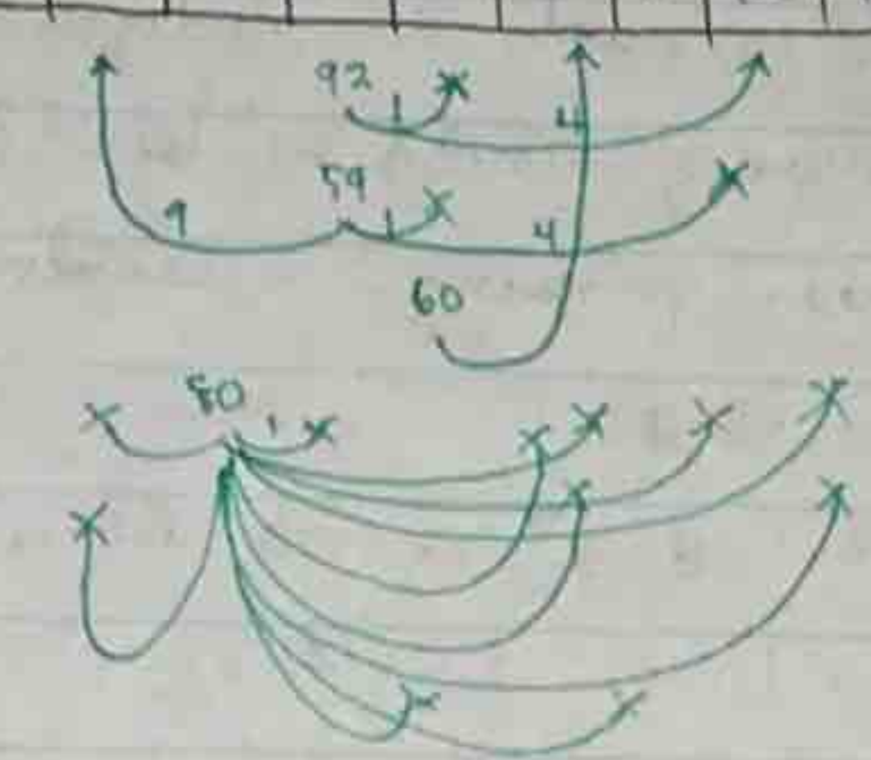
PART 1 [Q2 cont'd]

b) quadratic probing

26 42 5 44 92 59 80 36 12 60 80
4 / 9 / 5 / 0 / 4 / 4 / 7 / 3 / 1 / 5 / 3

0	1	2	3	4	5	6	7	8	9	10
44	12	59	36	26	5	60	40	92	42	

insert failed: 80



failed 11 times

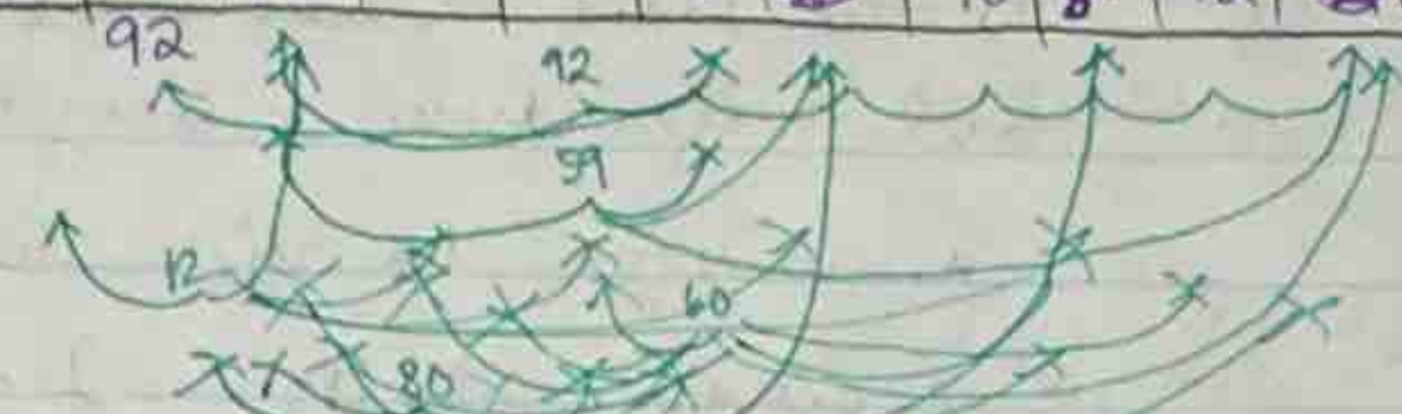
c) double hashing

w/ $h_2(x) = (2x) \% 11$ if nonzero
$h_2(x) = 1$ if prev line 0

$(92*2) \% 11 = 8$
$(59*2) \% 11 = 8$
$(60*2) \% 11 = 10$
$(80*2) \% 11 = 6$

0	1	2	3	4	5	6	7	8	9	10
44	~~92~~	12	36	26	5	59	40	80	42	~~59~~ 60

92

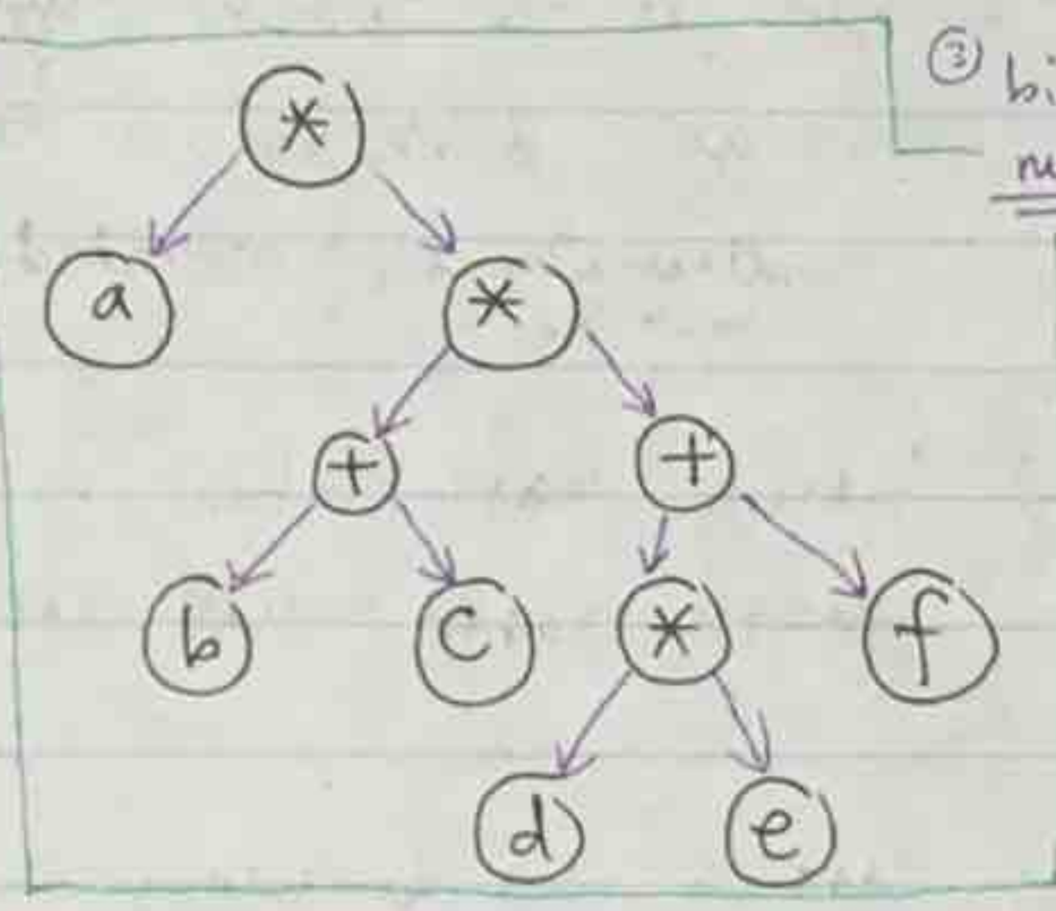


PART 2 — Tree Traversals

3) If pre-order visitation of nodes in a binT is → contains 1-char var & binary operators
a+bc+*def then:

Given:
1) - visit
-left
pre: -right
2) 11 nodes
3) binary operators must have L&R child

a) Draw expression tree:



4) in: -left
- visit
- right

b) What is in-order sequence?

→ $\boxed{a*b+c*d*e+f}$

PART 3 — Pigeonhole Principle

4) How many unique numbers from set {1, 3, 5, 7, 8, 9, 10, 11, 13, 15} must be selected to guarantee that at least one pair adds to 16?

☆ 8 & 10 never have a corresponding pair.

i. set of Possible pairs: {(1,15), (3,13), (5,11), (7,9), (8), (10)} = pigeonholes
(8), (10) ← are their own pigeon holes

ii. |pigeonholes| = 6 ∴ you just need 6+1 pigeons to guarantee a match ☆

= $\boxed{7}$ unique numbers (6 for each possible pigeonhole, 7th to make a pair)

Connie Ma
56047129 // pb78
CPSC 221 202
2.4.17

THEORY ASSIGNMENT #3

PART 3 (cont'd)

5 How many integers from 1 to 100 must you pick in order to be sure that you get one that is divisible by 5?

$$S = \left\{ \begin{array}{l} \text{set of numbers} \\ \text{b/t } 1\text{-}100 \\ \% 5 = 0 \end{array} \right\} = \{5, 10, 15, 20, \ldots, 95, 100\}$$

$\rightarrow |S| = 20$

(☆ this is assuming we don't put back the number we pick)

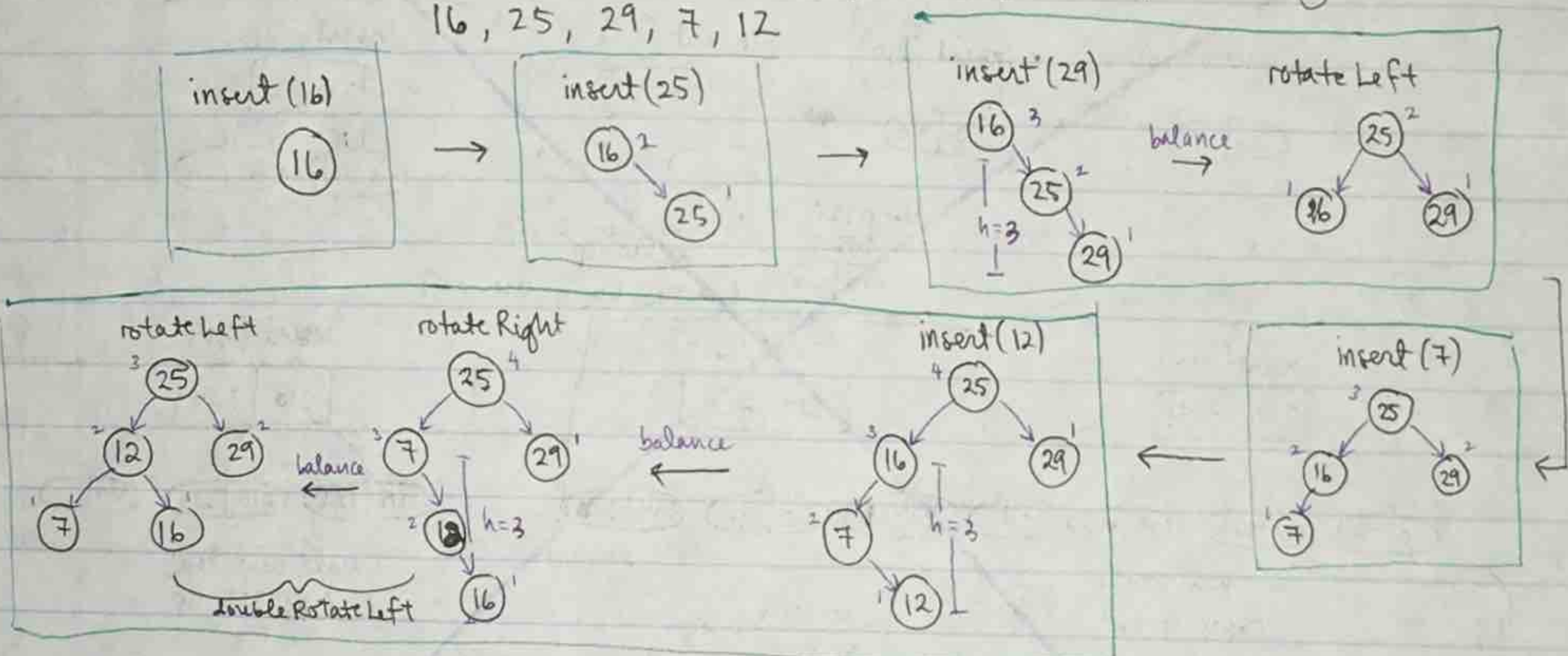
To guarantee we get one number from S, we have to pick $\underbrace{\underbrace{|\{1, 2, \ldots, 100\}| - |S|}_{\text{pigeonholes}} + 1}_{\text{pigeons}}$ numbers (by Pigeonhole principle)

$= 100 - 20 + 1 = \boxed{81}$

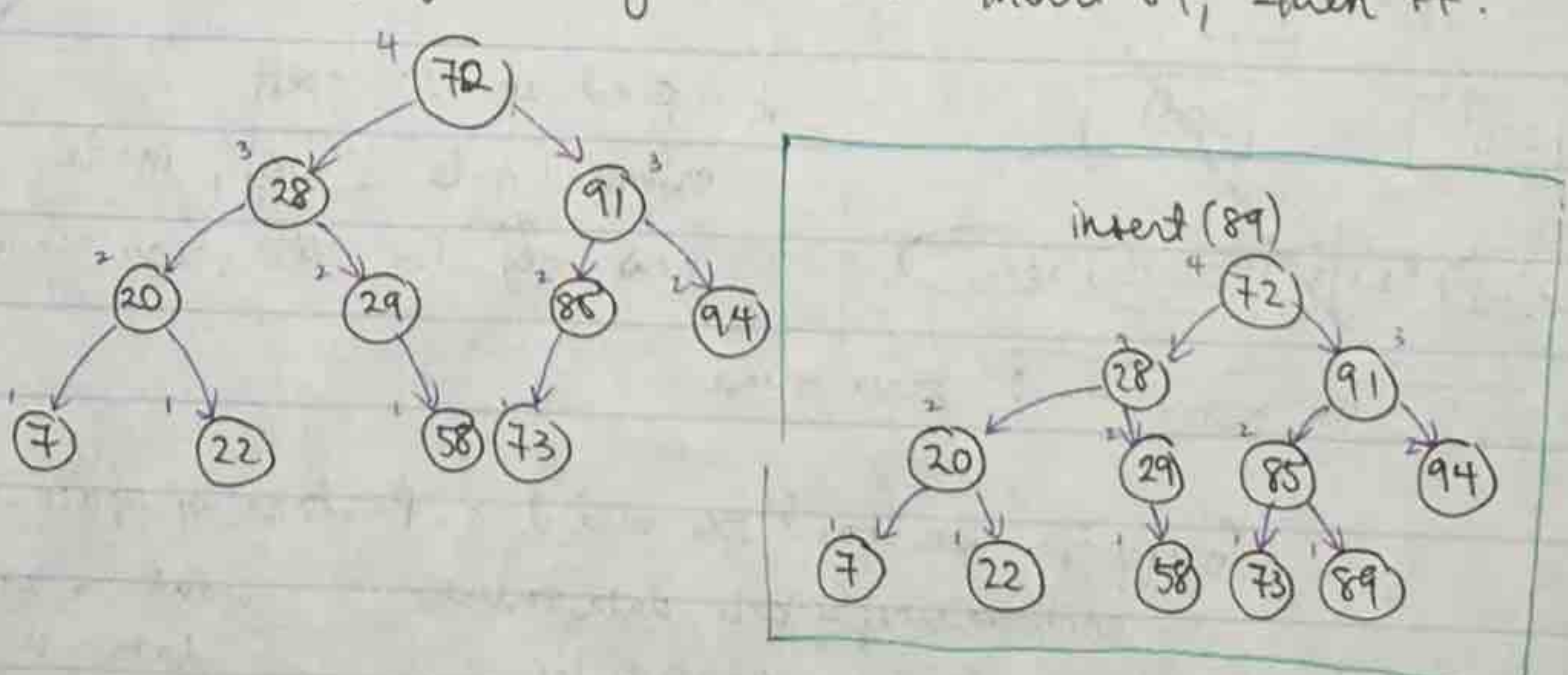
PART 4 — AVL Trees

6 Draw an AVL tree after inserting each of the following:

16, 25, 29, 7, 12



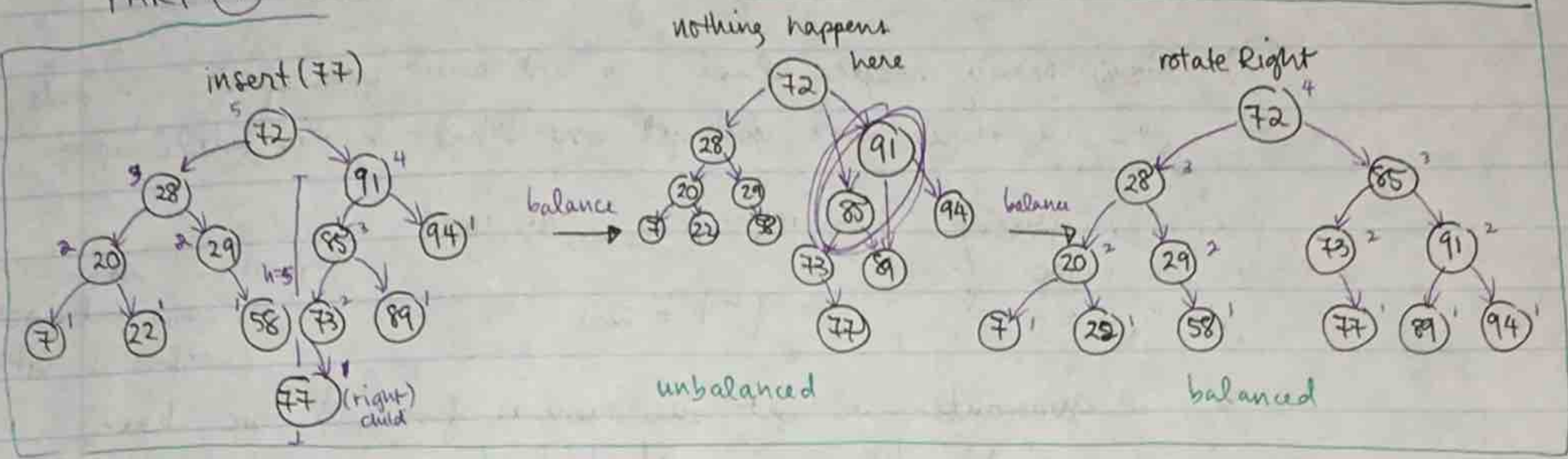
7 Repeat same process w/ following AVL tree. Insert 89, then 77.



Connie Ma
56047129 // p6z8
CPSC 221 202
2.4.17

THEORY ASSIGNMENT #3

PART 4 [Q7 cont'd]



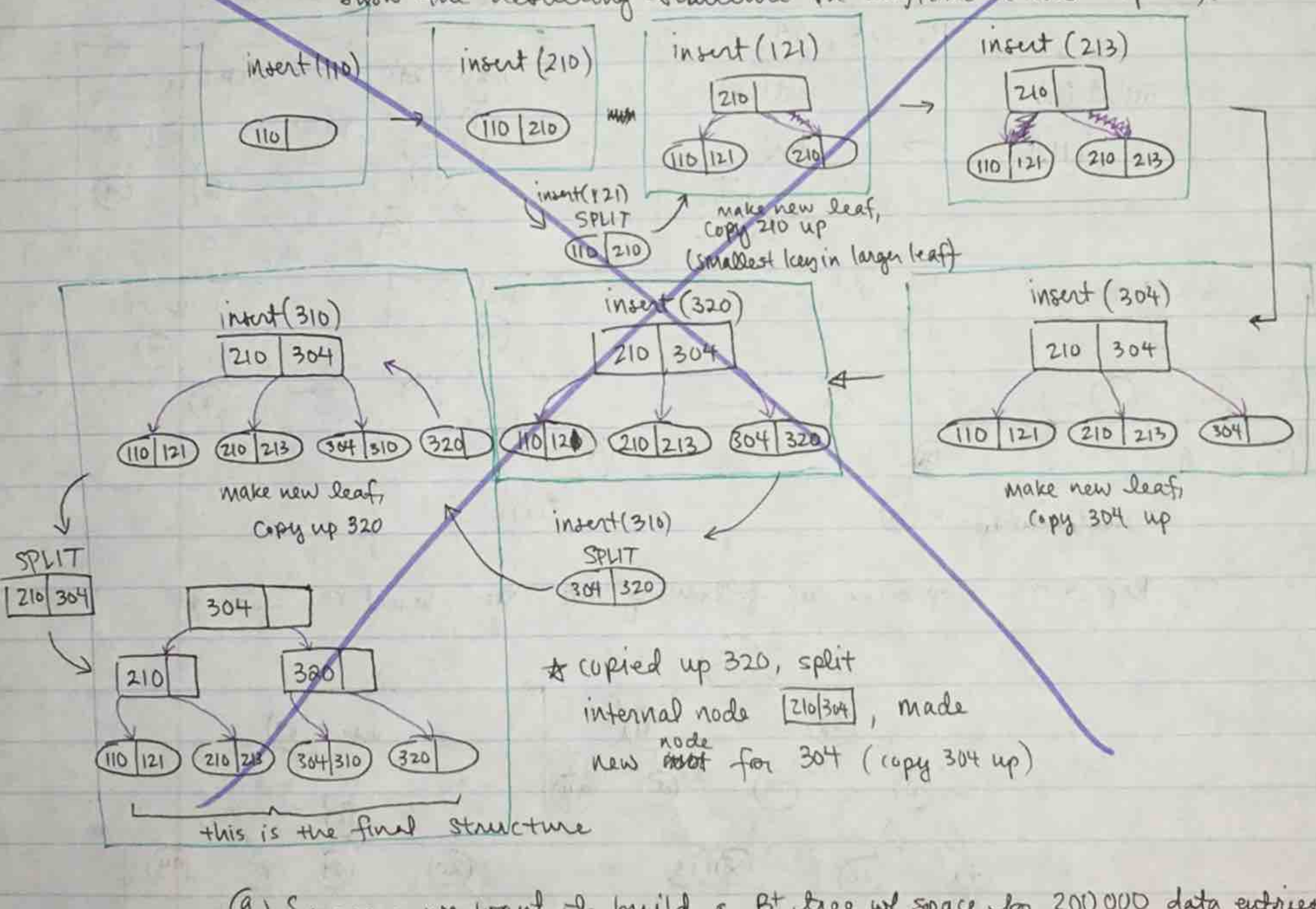
PART 5 - B^+ Trees

8) Suppose we have an empty B^+ tree ~~root~~ node that ~~can hold up to~~ 2 keys. Wish to enter the following keys

110, 210, 121, 221, 213, 304, 320, 310   course codes? uwu

OOPS FORGOT THIS ONE HAHA

Show the resulting structure (& anytime a node splits).



9) Suppose we want to build a B^+ tree w/ space for 200,000 data entries

- each entry = key, data value
- each page = 4096 bytes
- each page has 3 × 8 byte ptrs (parent, left sib, right sib) & additional key/data
- want to use even number of data entries in leaf pages
- internally also want even number

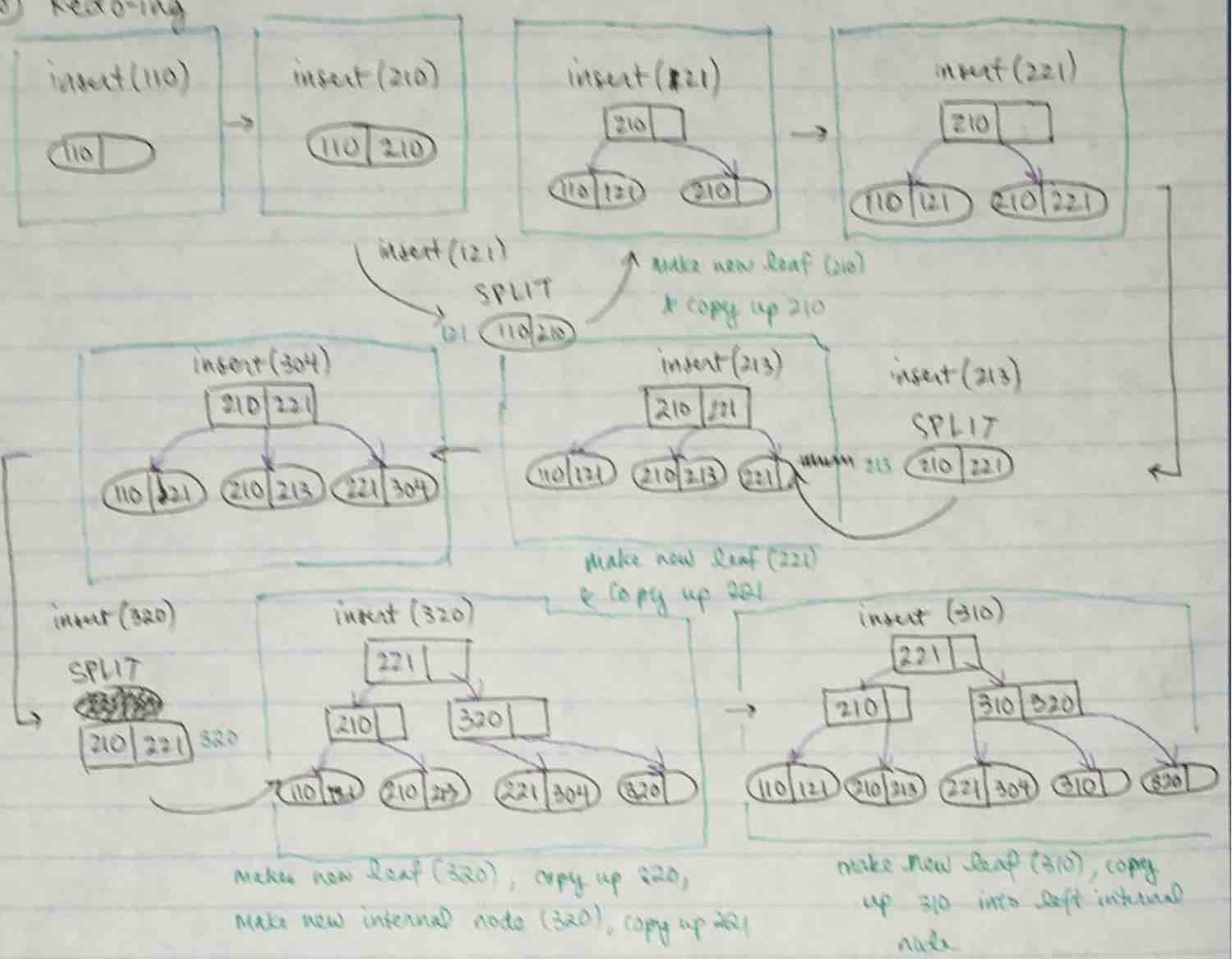
Connie Ma
56047129 / p6x8
CPSC 221 202
2.4.17

THEORY ASSIGNMENT #3

P5

PART 5 (Cont'd)

⑧ Redo-ing



⑨ a) ① each data entry: (k|v) — 56 bytes (key), 8 bytes (v)

② page: (8 bytes + 8 bytes + 8 bytes) + n (56+8) (data entries) = 4096 bytes
(parent, left, right)

$n \Rightarrow \frac{4096-24}{64} = 63.6 \Rightarrow \lfloor 63.6 \rfloor = 62$ (even)

we need:

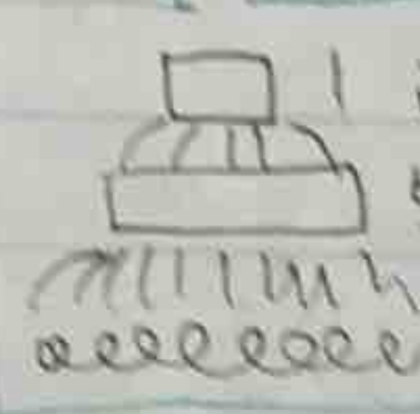
③ 200000 entries $/62 = 3225.8 \Rightarrow \lceil 3225.8 \rceil =$ $\boxed{3226 \text{ leaf pages}}$

b) ① Internal page: [L|R|P|n keys], n+1 chid $(3\times8) + n56 + (n+1)(8) = 4096$ bytes

② n max = 62 (see above) ∴ each internal page → 63 leaf pages

③ $\lceil 3226/63 \rceil =$ $\boxed{52 \text{ internal pages @ leaf+1 level}}$

⇒ ~~[illegible]~~



1 internal page (root) lvl
52 internal pages (lvl 2)
3226 leaf pages ∴