Encoding LZW

codes 0-255 - individual characters bytes
example A-65 ; B-66 ; C-67

256 - 4095 - substrings.

(E1) Encoding for input ABABBABBB

Table	
A	65
B	66
AB	256
BA	257
ABB	258
BAB	259
BB	260

Step 1
Step 2
Step 3
Step 4
Step 5

Step 1: A → is in table ⇒ move next
AB → is not in table ⇒ 1) add in table from position 256 (increment position for next substrings)
2) print prefix A code ⇒ (65)

Step 2: Remove printed char: ~~A~~BABBABBB

B → is in table ⇒ move next
BA → is not in table ⇒ 1) add in table BA
2) print prefix B code ⇒ (66)

Step 3: ~~AB~~ABBABBB

A → is in table ⇒ move next
AB → is in table ⇒ move next
ABB → is not in table ⇒ 1) add in table ABB
2) print prefix AB ⇒ (256)

Step 4 ~~ABA~~BABBB

B → is in table
BA → is in table
BAB → is not in table ⇒ 1) add in table BAB
2) print prefix BA ⇒ (257)

Step 5 ~~ABABB~~BBB

B → is in table
BB → is not in table ⇒ 1) add in table BB
2) print prefix B (66)

Step 6 ~~ABABBB~~BB

B → is in table
BB → is in table ⟶ print BB code (260)

Output codes : 65 66 256 257 66 260

E2. Encoding for input: CAB*AB*CA

TABLE
A - 65
B - 66
C - 67
* - 42
CA - 256
AB - 257
B* - 258
*A - 259
AB* - 260
*C - 261

Step 1: C → in the table → move next
CA → not in table → 1) add CA
2) print C (67)

Step 2: ~~C~~AB*AB*CA

A → in table → move next
AB → not in table → 1. add AB
2. print A (65)

Step 3: ~~CA~~B*AB*CA

B - in table
B* - not in table ⇒ 1. add B*
2. print B code (66)

Step 4: ~~CAB~~*AB*CA

* - in table
*A - not in table ⇒ 1. add *A
2. print * code (42)

Step 5: ~~CAB*~~AB*CA

A → in table
AB - in table
AB* - not in table ⇒ 1. add AB*
2. print AB (257)

Step 6: ~~CAB*AB*~~CA

* - in table
*C - not in table ⇒ 1. add *C
2. print * (42)

Step 7: ~~CAB*AB*C~~A

C - in table ⇒ move next
CA - in table ⟶ print (256)

Output code : 67 65 66 42 257 42 256

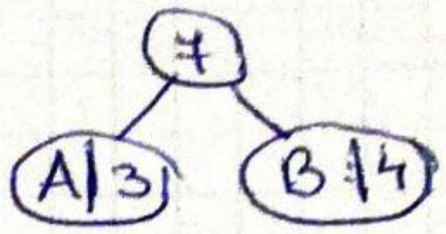
Es. Huffman codes for AAA BBBB CC DDCCC DDDDD

Step 1. a) Freq table

A	3
B	4
C	5
D	7

b) Extract 2 min freq nodes : A:3 and B:4.

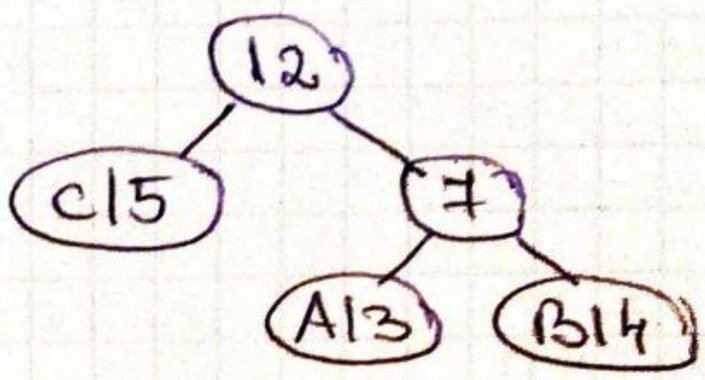
Create new node :
(sum(freq(A) + freq(B))



Step 2 a) Freq table

C	5
new node	7
D	7

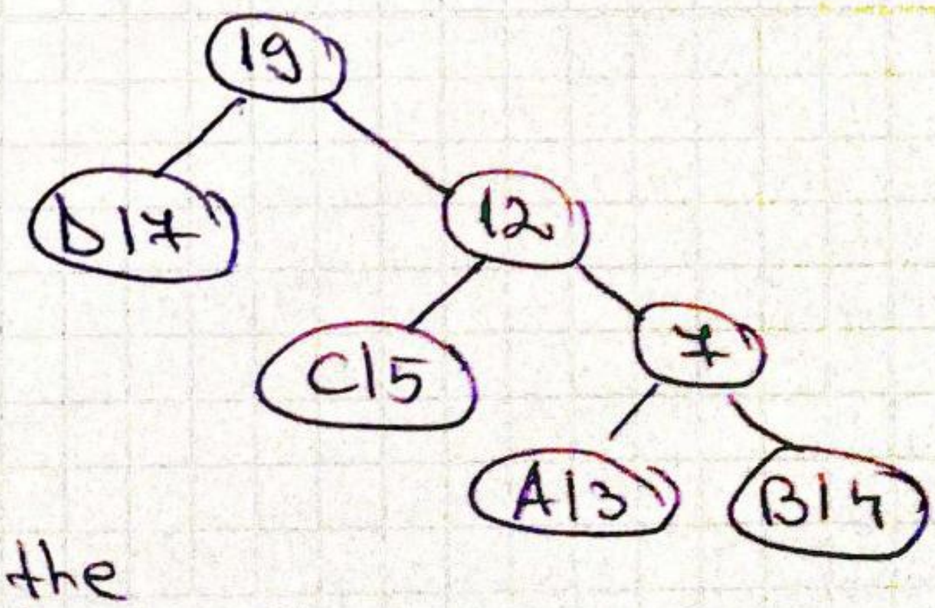
b) Extract 2 min freq nodes : C:5 and new node 7



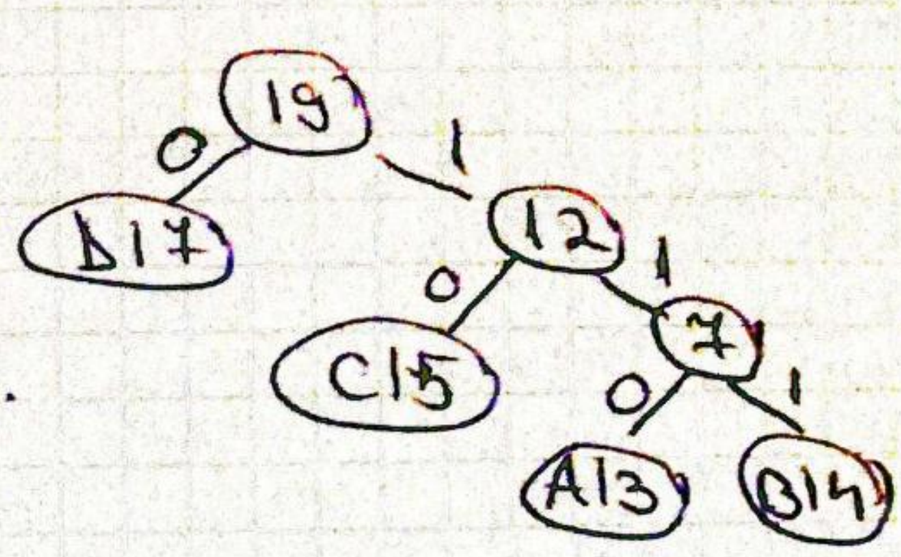
Step 3 : a) Freq table

D	7
new node	12

b) new node



Step 4 :
- when moving to the right child write 1
- when moving to the left child write 0



Step 5 : Write codes from root → leaf.

A	1 1 0
B	1 1 1
C	1 0
D	0