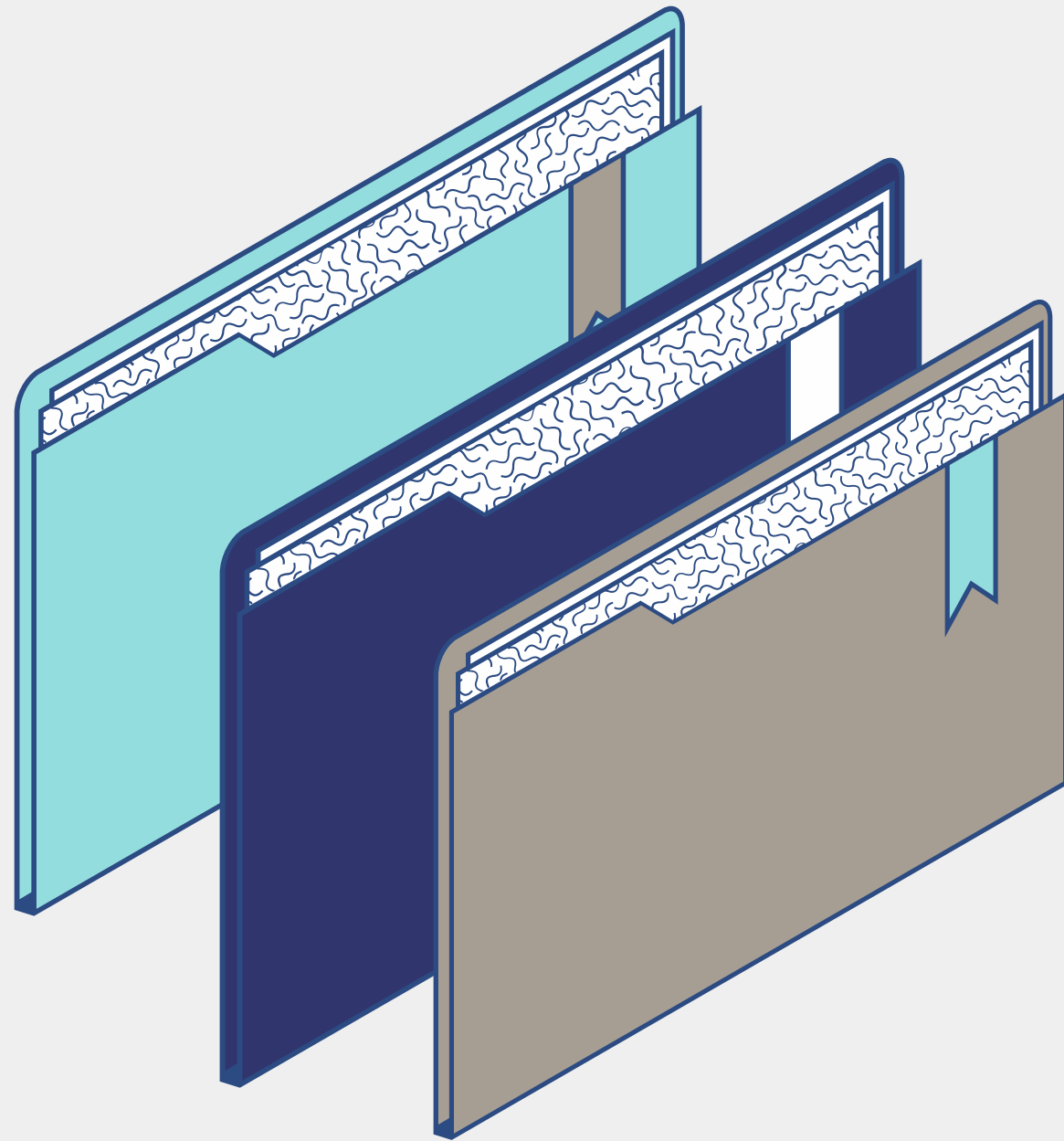




GROUP TURON

HUFFMAN CODING

CHRISBELL JADE S. TINDUGAN

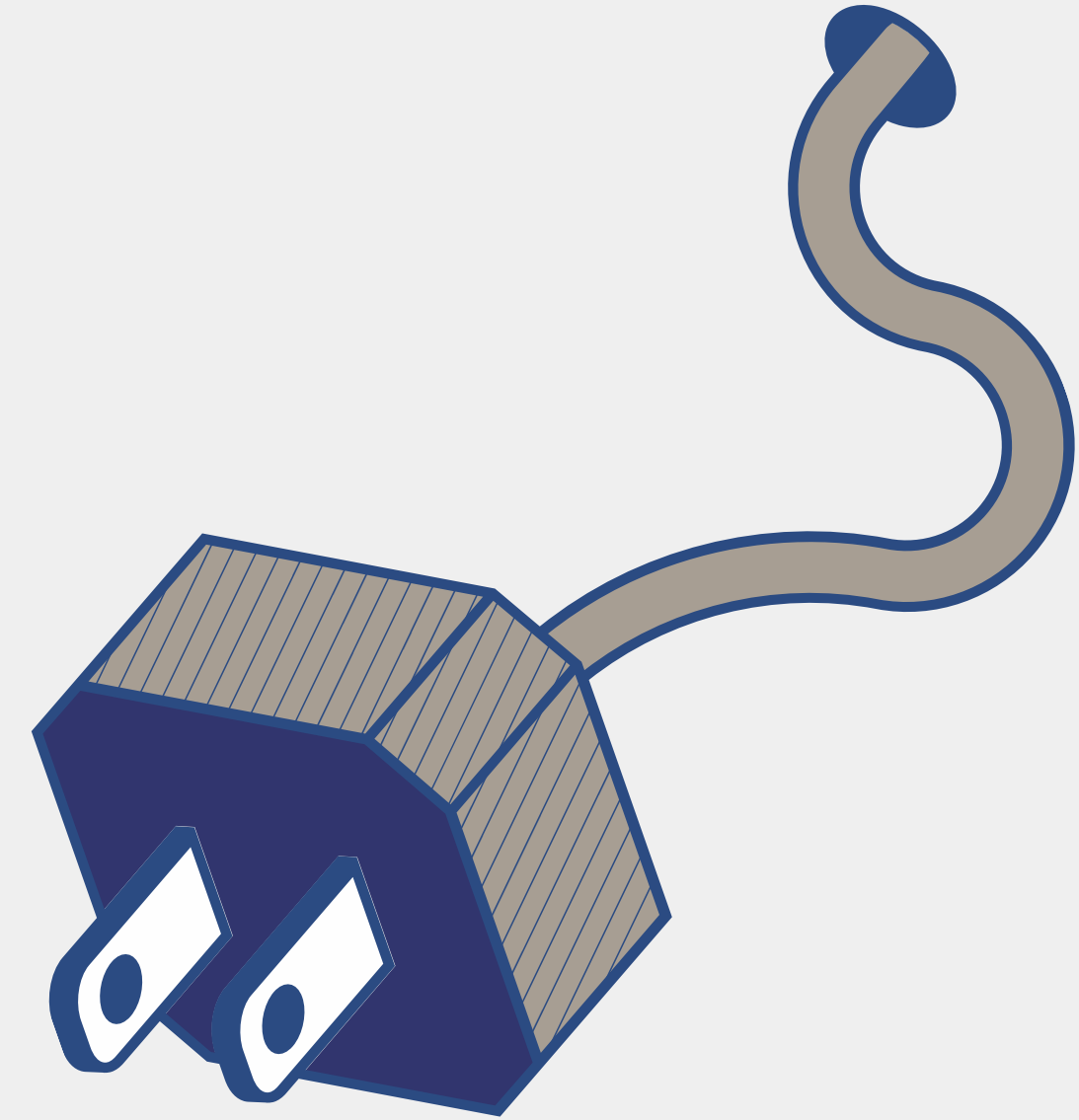


OBJECTIVES

- What is Huffman Coding?
- How does Huffman Coding work?
- Huffman Coding Complexity
- Huffman Coding Applications

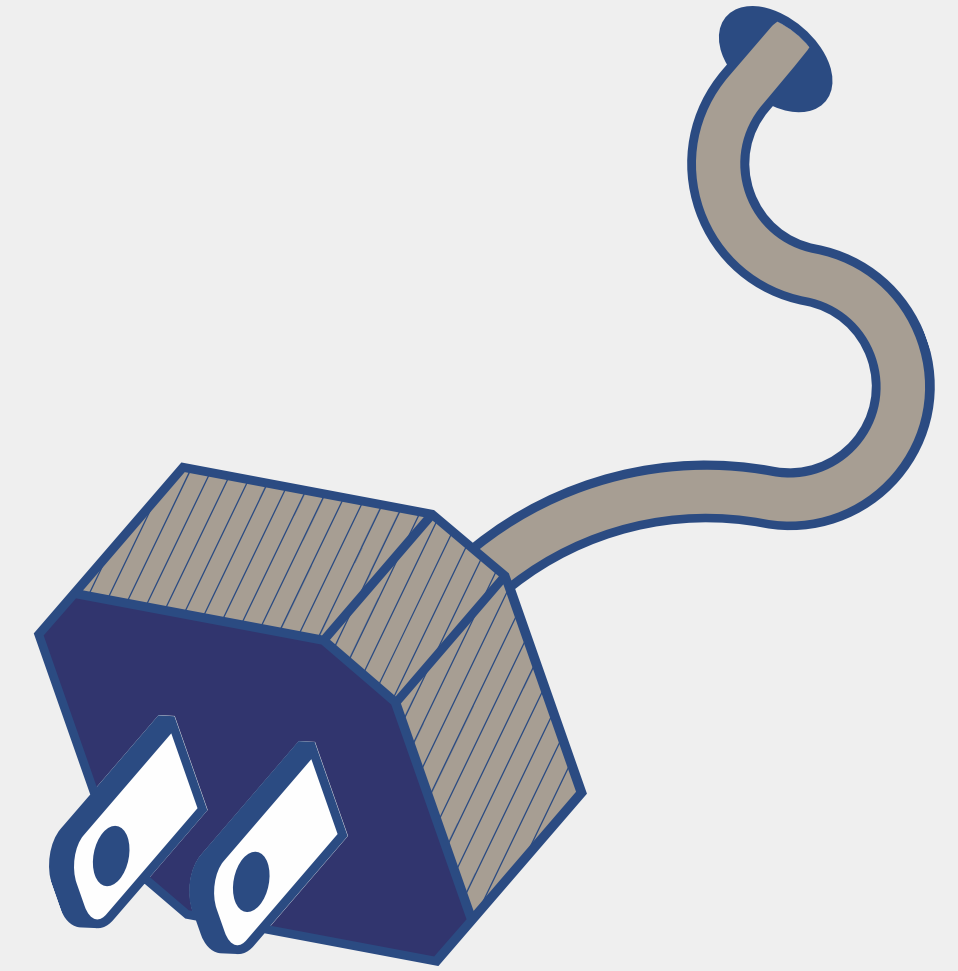
HUFFMAN CODING

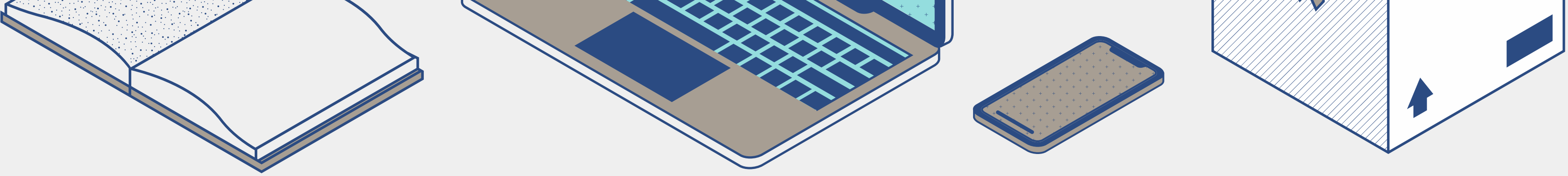
Huffman Coding is a technique of compressing data to reduce its size without losing any of the details. It was first developed by David Huffman.



HUFFMAN CODING

The idea is to assign variable-length codes to input characters, lengths of the assigned codes are based on the frequencies of corresponding characters. The variable-length codes assigned to input characters are Prefix Codes.





HOW DOES HUFFMAN CODING WORKS?

Huffman coding first creates a tree using the frequencies of the character and then generates code for each character. Once the data is encoded, it has to be decoded. Decoding is done using the same tree.

HOW DOES HUFFMAN CODING WORKS?

1

2

3

4

STEP

Calculate the frequency of each character in the string.

STEP

Sort the characters in increasing order of the frequency. These are stored in a priority queue Q .

STEP

Make each unique character as a leaf node. Create an empty node z .

STEP

Assign the minimum frequency to the left child of z and assign the second minimum frequency to the right child of z .

Set the value of the z as the sum of the above two minimum frequencies.

HOW DOES HUFFMAN CODING WORKS?

5

6

7

8

STEP

Remove these two minimum frequencies from Q and add the sum into the list of frequencies.

STEP

Insert node z into the tree.

STEP

Repeat steps 3 to 5 for all the characters.

STEP

For each non-leaf node, assign 0 to the left edge and 1 to the right edge.

REMINDER

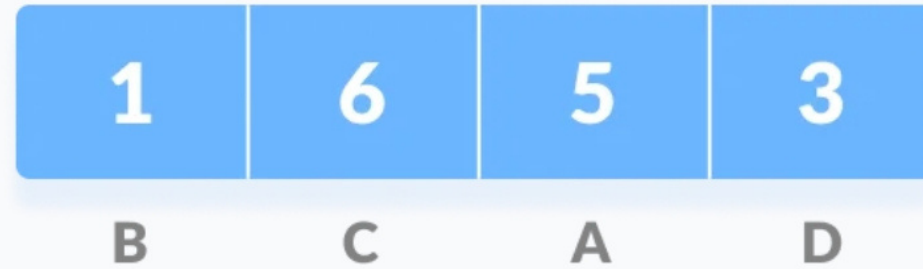
For sending the string over a network, we have to send the tree as well as the above compressed-code.

EXAMPLE

B	C	A	A	D	D	D	C	C	A	C	A	C	A	C
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Initial string

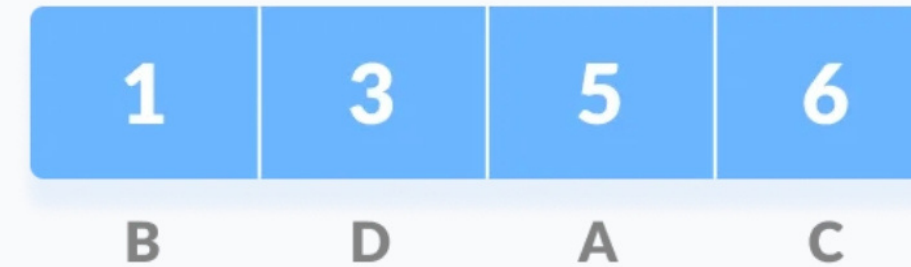
Step 1



Frequency of string



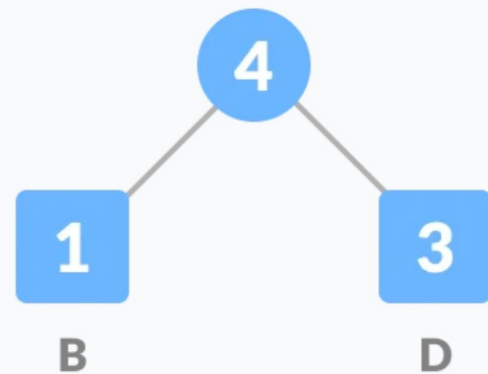
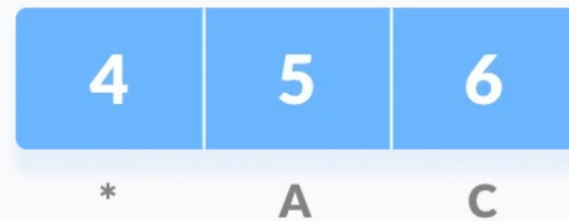
Step 2



Characters sorted according to the frequency



Step 4

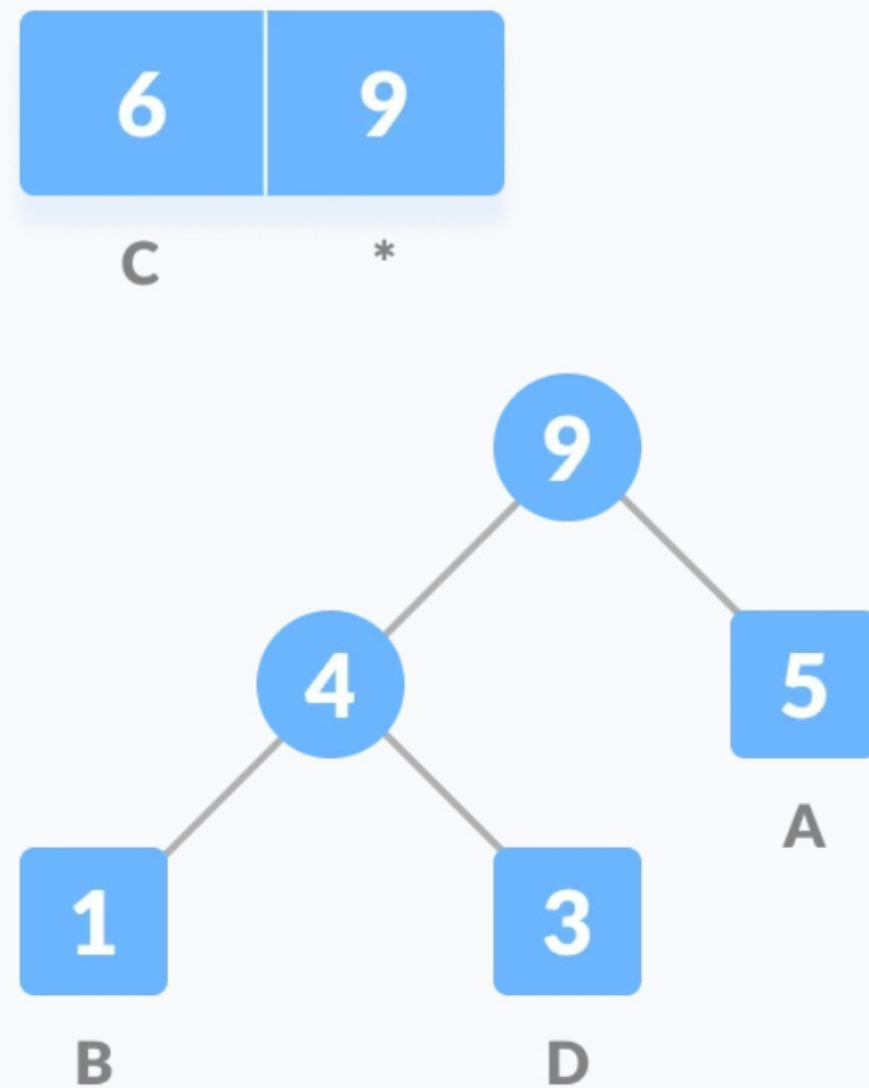


Getting the sum of the least numbers



Follow Step 3

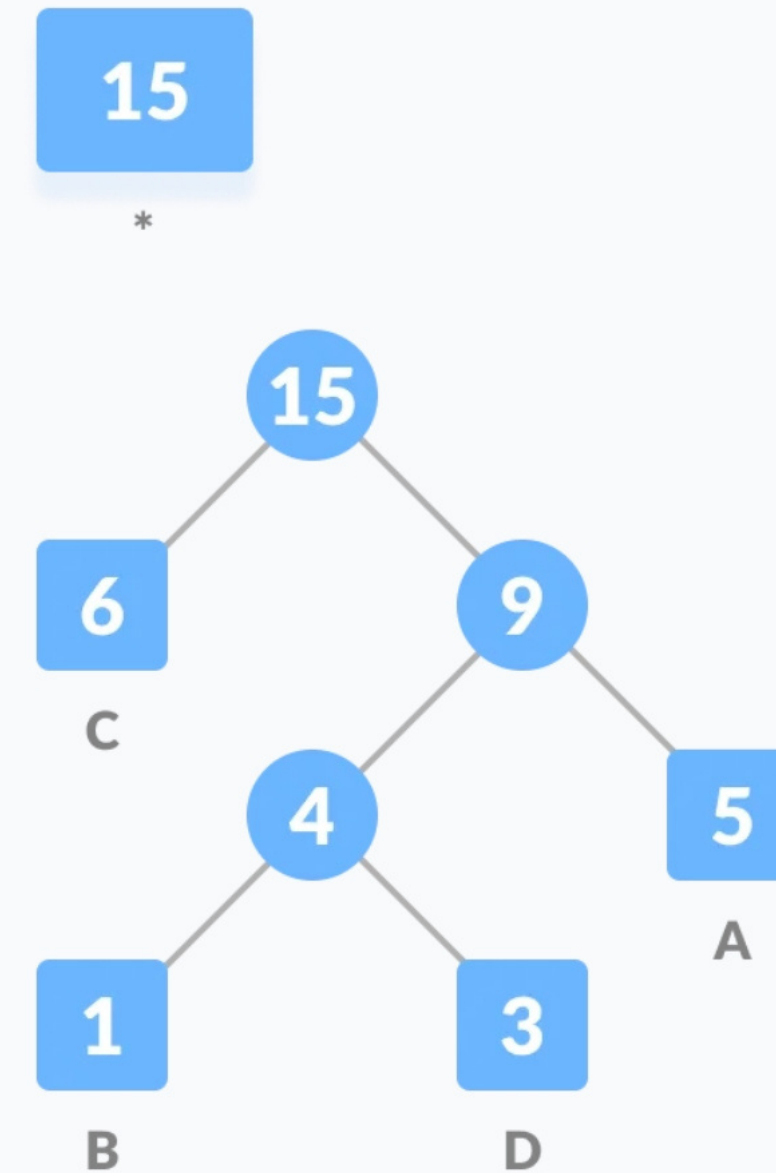
Step 5-6



Repeat steps 3 to 5 for all the characters.

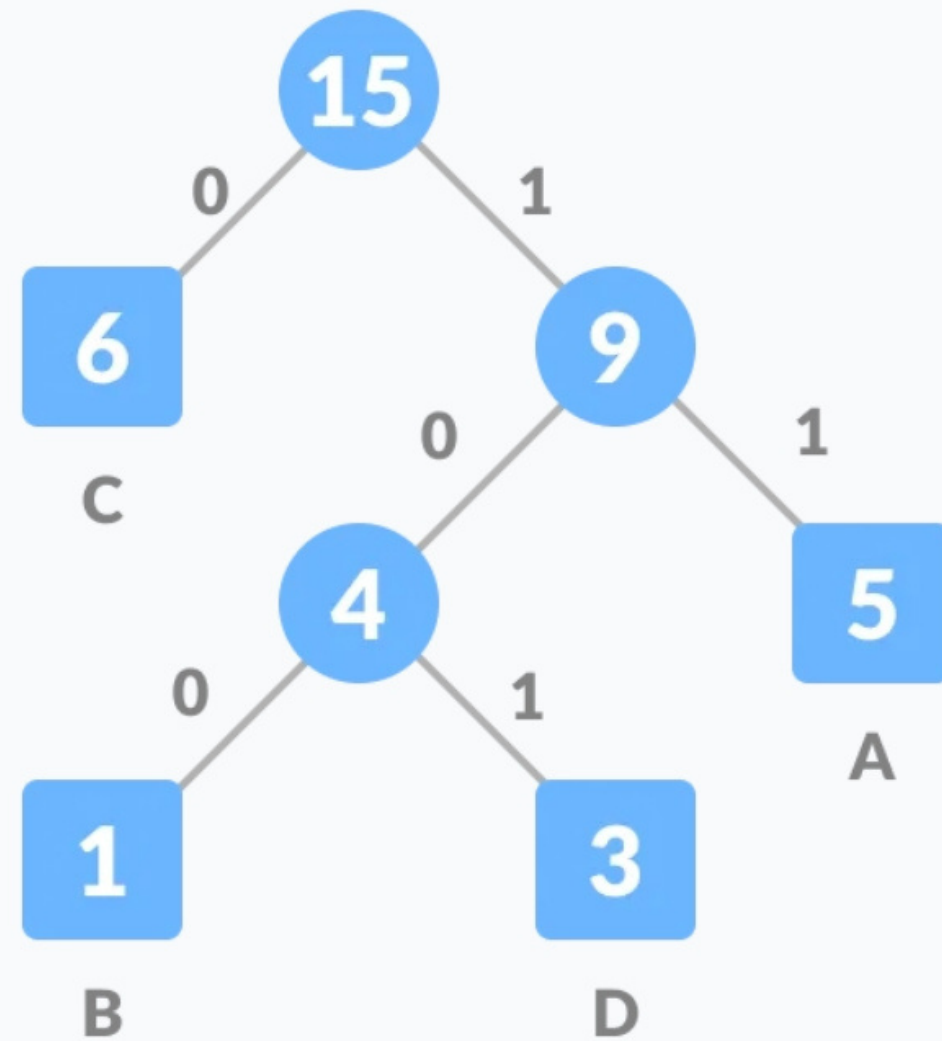


Step 7



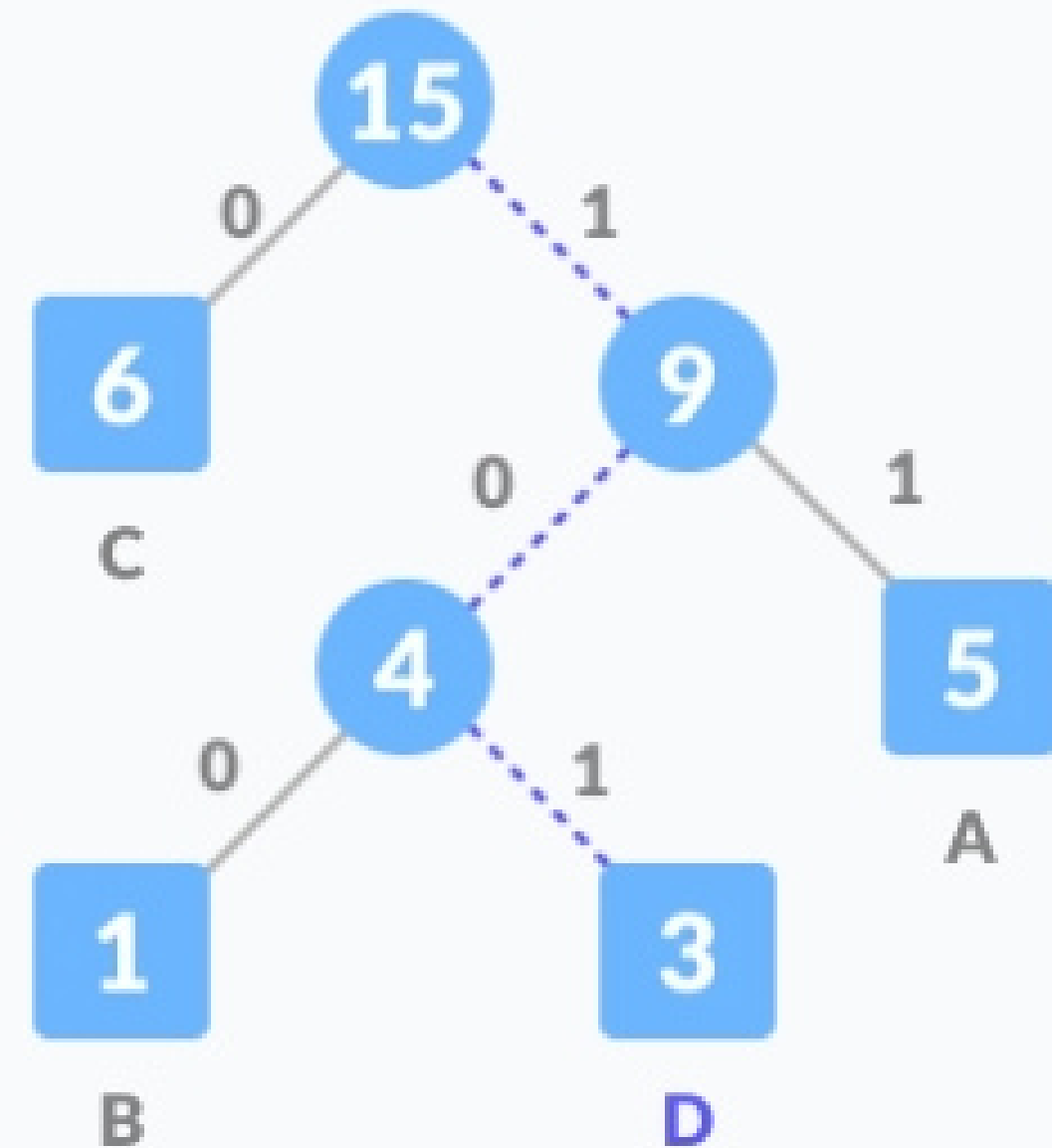
Repeat steps 3 to 5 for all the characters.

Step 8



Assign 0 to the left edge and 1
to the right edge

DECODING



Decoding



Character	Frequency	Code	Size
A	5	11	$5 \times 2 = 10$
B	1	100	$1 \times 3 = 3$
C	6	0	$6 \times 1 = 6$
D	3	101	$3 \times 3 = 9$
$4 \times 8 = 32$ bits	15 bits		28 bits



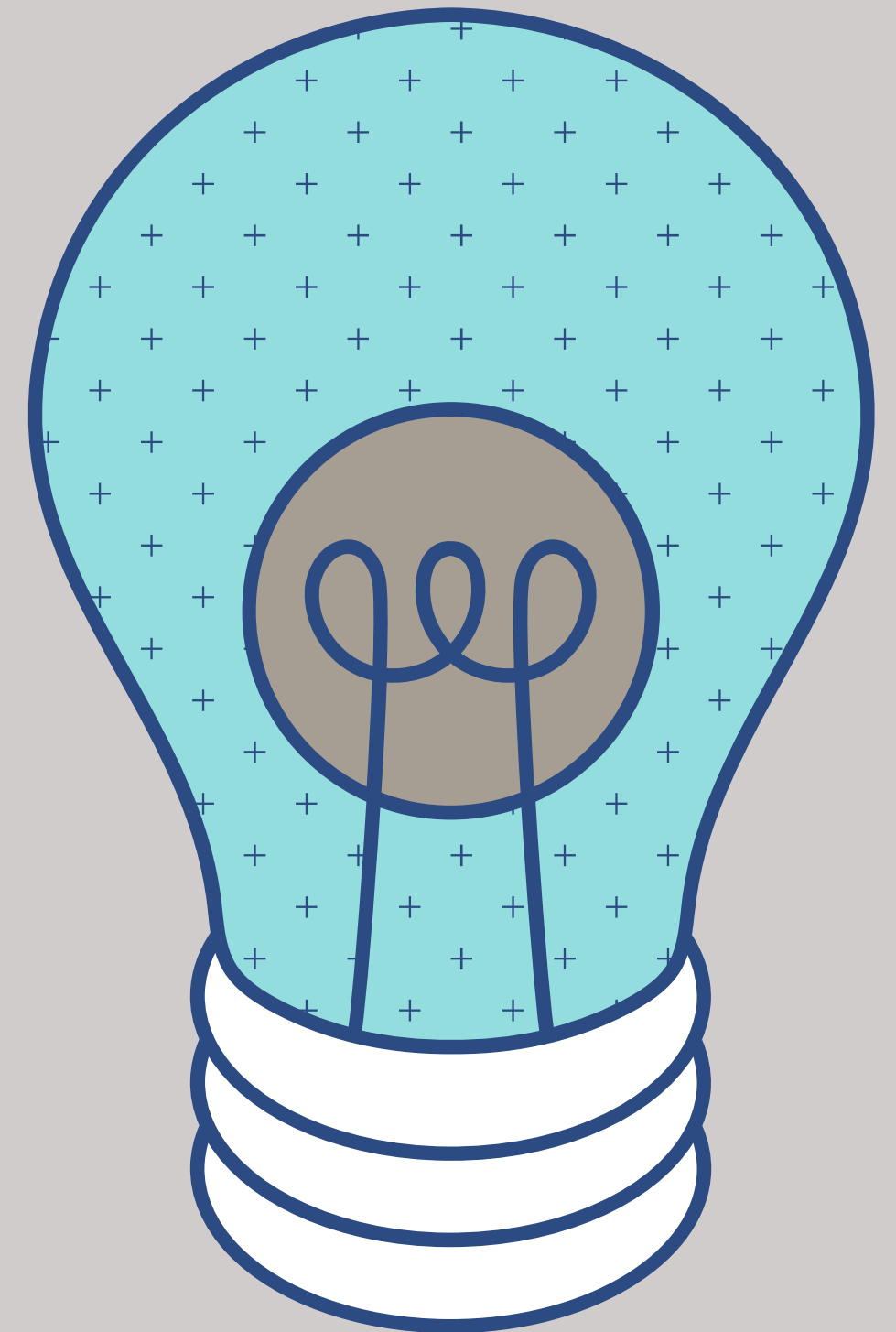
HUFFMAN CODING COMPLEXITY

The time complexity for encoding each unique character based on its frequency is $O(n \log n)$.

Extracting minimum frequency from the priority queue takes place $2^{*}(n-1)$ times and its complexity is $O(\log n)$. Thus the overall complexity is $O(n \log n)$.

HUFFMAN CODING APPLICATION

- Huffman coding is used in conventional compression formats like GZIP, BZIP2, PKZIP, etc.
- For text and fax transmissions.



Do you have any questions?

END OF REPORT. THANK YOU FOR
LISTENING

