TEAM SOBER BITS

THEME: OPEN INNOVATION

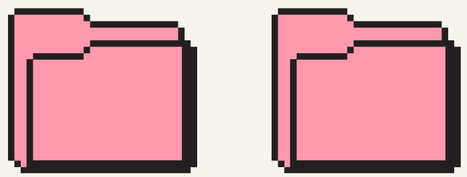
1.Vidhi Bhamare
2.Aditi Hinge
3.Shreya Watwe

SY COMP

FILE COMPRESSION

AIM:

The aim of this project is to create a prototype that can compress large files so that it is easy to transfer them.



DATA STRUCTURES USED

Binary Tree

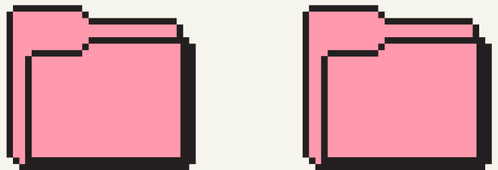
Array List

Cautious

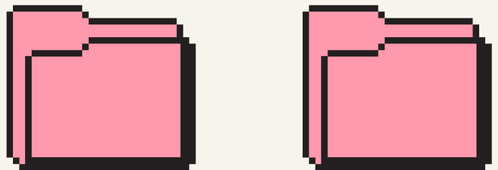
Adventurous

HOW DOES IT WORK?

1. Count the frequency of each character in the input text.
2. Create a binary tree where each leaf node represents a character and the path from the root to the leaf node represents the code for that character.
3. Merge the two nodes with the lowest frequency into a single node, until all the nodes are merged into a single root node.
4. Traverse the binary tree from the root to the leaf node for each character to generate its code.



5.Replace each character in the input text with its corresponding Huffman code to generate the compressed output.
6.To decode the compressed data, reconstruct the original Huffman tree and read the compressed data bit by bit until a leaf node is reached. Output the corresponding character and continue reading the next bit of compressed data.

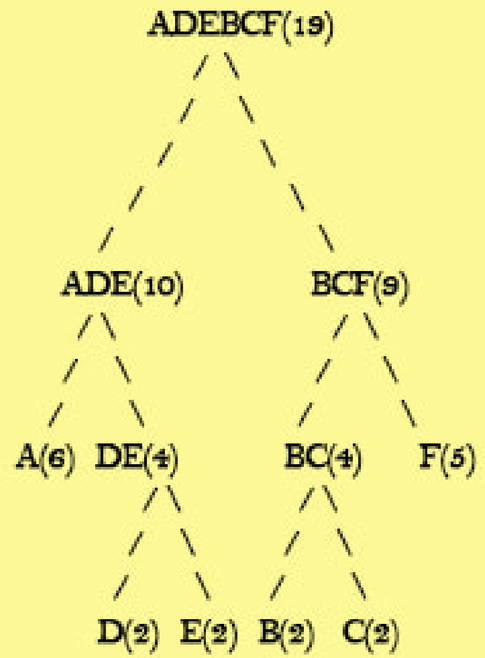


Huffman coding and decoding

Input

"AAAAAABBCCDDEEFFFFFF"

=>



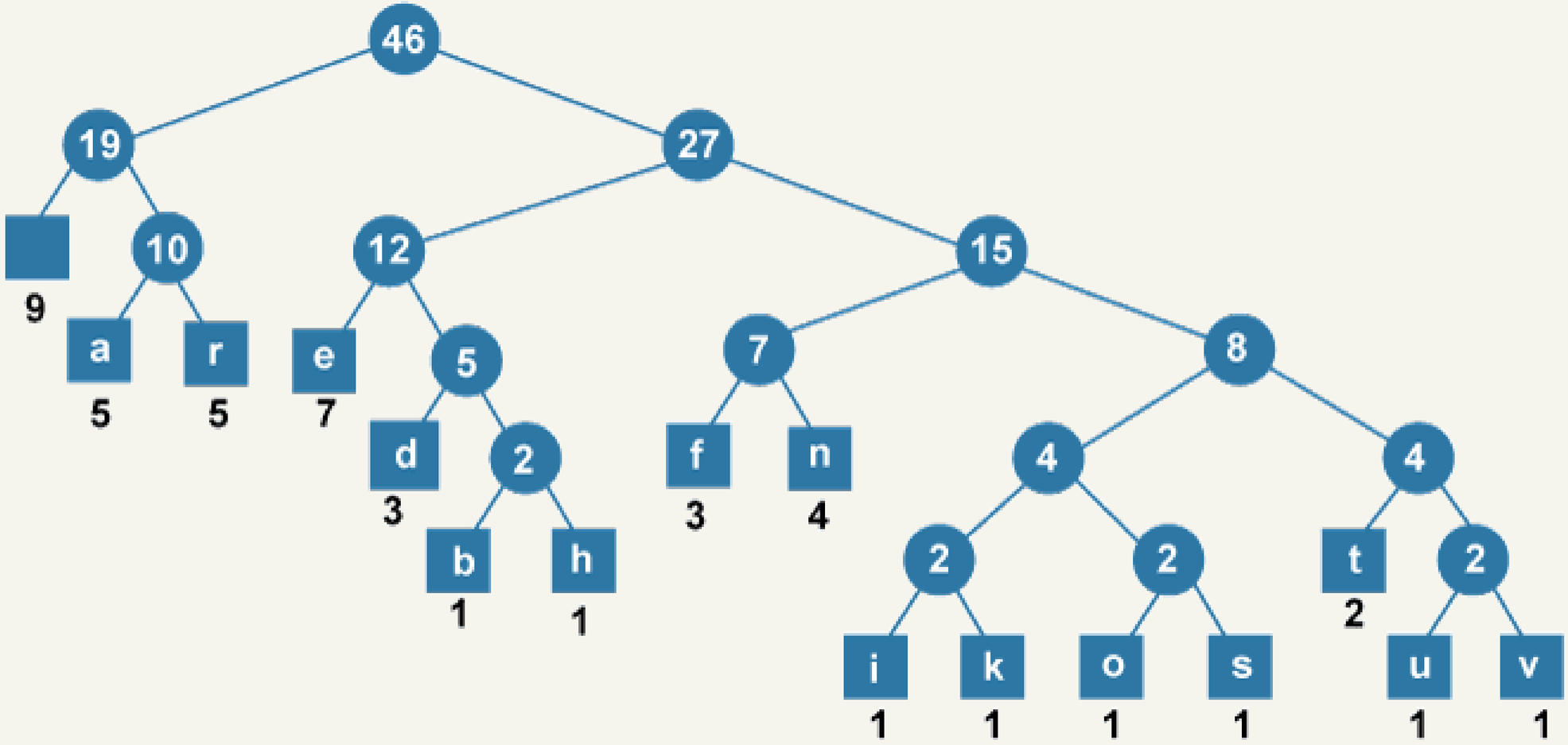
Build Frequency tree

=>

Output

A = 11
B = 000
C = 011
D = 001
E = 010
F = 10

Solution: Greedy


46
19
27
10
12
15
a
r
e
5
7
8
d
2
f
n
4
4
b
h
2
2
t
2
i
k
o
s
u
v
9
5
5
7
3
1
1
3
4
1
1
1
1
2
1
1

REAL LIFE APPLICATION

File Compression

compress files such as text, images, audio, and video files.

Network Communication

the amount of data that needs to be transferred is reduced, which can improve network performance.

Text Messaging

text messaging applications to compress messages and reduce the data usage, particularly important in countries where mobile data is expensive or limited.

Data Storage

particularly important in large-scale data storage applications, such as data centers, where the cost of storage is a major concern.

FUTURE SCOPE

1.Currently, this algorithm is implemented to compress data. It can be further used to decompress the compressed file.

2.It can be integrated with frontend to provide a good UI.

3.This can be also used to compress image and audio files.