

Matricola: 4811831

Nome: Haider Ali

Codifica di Huffman

Ogni carattere ASCII occupa **8 bit**. Se un file in input che consiste di **10^5 caratteri**, è necessario un totale di **$8 * 10^5 = 800000$ bit** per inviare questa stringa.

Usando la tecnica di codifica di Huffman, possiamo comprimere la stringa a una dimensione più piccola. La codifica di Huffman crea un albero utilizzando le frequenze del carattere e quindi genera il codice per ogni carattere.

Huffman Coding previene qualsiasi ambiguità nel processo di decodifica utilizzando due **proprietà**:

- **Decifrabilità univoca**
- **Istantaneità**

Entropia come limite insuperabile: La lunghezza attesa $L(C, X)$ di una codifica univocamente decifrabile C non può essere minore dell'entropia $H(X)$.

Algoritmo:

Input: X , insieme di $|X|$ simboli con $p(x)$ probabilità del simbolo x .

Output: T , albero binario in cui ognuna delle $|X|$ foglie contiene una codifica binaria di uno degli x .

```
Q ← X
for i = 1 → (|X| - 1)
    x ← left child of z ← EXTRACT-MIN(Q)
    y ← left child of z ← EXTRACT-MIN(Q)
    p(z) ← p(x) + p(y)
    INSERT(Q, z)
return EXTRACT-MIN(Q)
```