**Esercizio sugli algoritmi di compressione**

L’esercizio richiedeva di comprimere la stringa **abcddaabbccdcdaaefaba** con l’algoritmo LZW e uno a scelta tra gli algoritmi Shannon-Fano e Huffman. Per l’esercizio ho deciso di utilizzare la codifica di Huffman in quanto è stato osservato che raggiunge sempre la soluzione ottima in termini di compressione.

Per ognuno degli algoritmi viene presentata la soluzione e lo spazio occupato dalla codifica in termini di bit.

**LZW**

La tabella risultante dall’applicazione di LZW sulla stringa fornita è la seguente:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **w** | **k** | **Output** | **Codice** | **Simbolo** |
| NULL | a |  |  |  |
| a | b | a | 256 | ab |
| b | c | b | 257 | bc |
| c | d | c | 258 | cd |
| d | d | d | 259 | dd |
| d | a | d | 260 | da |
| a | a | a | 261 | aa |
| a | b |  |  |  |
| ab | b | 256 | 262 | abb |
| b | c |  |  |  |
| bc | c | 257 | 263 | bcc |
| c | d |  |  |  |
| cd | c | 258 | 264 | cdc |
| c | d |  |  |  |
| cd | a | 258 | 265 | cda |
| a | a |  |  |  |
| aa | e | 261 | 266 | aae |
| e | f | e | 267 | ef |
| f | a | f | 268 | fa |
| a | b |  |  |  |
| ab | a | 256 | 269 | aba |
| a | EOF | a |  |  |

Tabella Tabella della codifica LZW

L’output della codifica è stato evidenziato. Per la codifica ogni carattere trasmesso deve essere di 16 bit, anche i caratteri ASCII. Questo per unificare la dimensione dei caratteri inviati. Ne consegue che devono essere trasmessi 15\*16 bit = **240 bit**.

**HUFFMAN**

L’albero risultate dall’applicazione della codifica di Huffman alla stringa fornita è il seguente:

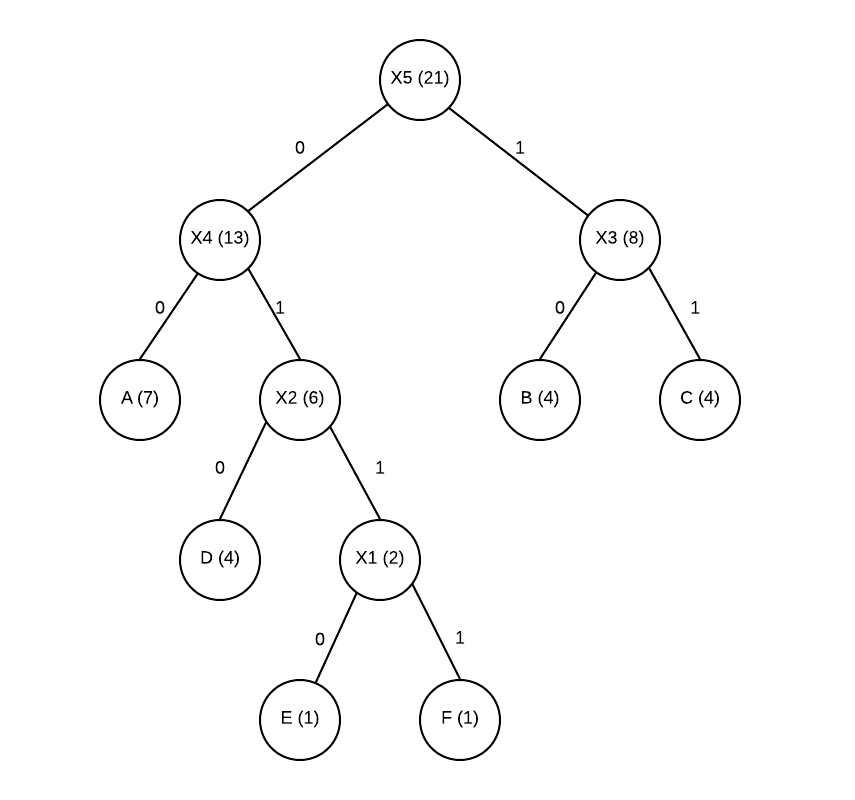


Figura Albero della codifica di Huffman

La tabella contenente i dati relativi al processo di codifica di Huffman è la seguente:

|  |  |  |  |
| --- | --- | --- | --- |
| **Simbolo** | **#** | **Codice** | **# di bit** |
| A | 7 | 00 | 14 |
| B | 4 | 10 | 8 |
| C | 4 | 11 | 8 |
| D | 4 | 010 | 12 |
| E | 1 | 0110 | 4 |
| F | 1 | 0111 | 4 |

Tabella Dati della codifica di Huffman

Le colonne che necessitano di essere trasmesse sono quelle evidenziate. Devono essere trasmetti 8 bit per ogni carattere e per ogni numero, in modo da unificare la rappresentazione dei simboli inviati. Quindi vengono trasmessi 8\*12 bit = **96 bit**.

**Comparazione dei risultati ottenuti**

Si può notare che l’algoritmo di Huffman ha ottenuto risultati migliori nonostante debba trasmettere anche la tabella per permettere la decodifica della stringa. Infatti, nonostante LZW cerchi di eliminare l’overhead della tabella permettendo la costruzione dinamica della stessa in fase di decompressione, sembra aver ottenuto un risultato peggiore di Huffman. Huffman ottiene sempre l’ottimo in termini di compressione per ogni lunghezza di stringa, mentre LZW necessita di dati di almeno 100 kb per ottenere dei risultati efficienti e comparabili a quelli di Huffman.