

Proyecto 2: Códigos de Huffman

Fabián Alonso Gómez Quesada*, Wilberth Daniel Gutiérrez Montero[†]

Nagel Eduardo Mejía Segura[‡], Oscar Mario González Cambrón[§]

*^{†‡§} Escuela de Ingeniería en Electrónica, Instituto Tecnológico de Costa Rica.

Emails: fabi.goque@estudiantec.cr, wil.gutierrez@estudiantec.cr,

nagelmese@estudiantec.cr, oscargonzalez@estudiantec.cr

Resumen—

I. INTRODUCCIÓN

II. DESARROLLO DEL PROYECTO

A. Implementación del código de Huffman

¿Cómo se implementa el algoritmo de generación de códigos Huffman?

B. Resultados de archivos

1) *Archivo de texto abecedario*: Primero, en cuanto al archivo **solo_abc_cien.txt** los datos están codificados según el código ASCII, sin embargo, está bajo el estándar UTF-8, por lo que cada carácter es de 8 bits. Como se repiten todas las letras 100 veces, cada símbolo es equiprobable, lo que se podría esperar tras correr el código, es que la mayoría de símbolos tengan la misma longitud, dicho de otro modo la varianza sería pequeña.

Al correr el código con dicho archivo, se obtiene la Tabla I donde se observa que la mayoría de códigos tienen 5 bits, con excepción de 6 símbolos que contemplan 4 bits.

III. CONCLUSIÓN

Tabla I
CÓDIGOS HUFFMAN PARA CARACTERES ASCII (A-Z)

Char (decimal)	Símbolo	Código Huffman
65	A	0001
66	B	0000
67	C	0011
68	D	0010
69	E	10101
70	F	10100
71	G	10111
72	H	10110
73	I	10001
74	J	10000
75	K	10011
76	L	10010
77	M	11101
78	N	11100
79	O	11111
80	P	11110
81	Q	11001
82	R	11000
83	S	11011
84	T	11010
85	U	01101
86	V	01100
87	W	01111
88	X	01110
89	Y	0101
90	Z	0100