

Algoritmos de compresión

18-10-19

- Comprimir se refiere a convertir datos a una representación que necesite menos espacio que la representación original.

- Existen 2 tipos de compresión:

• Sin pérdida (lossless): Puedo comprimir y descomprimir. Comprime mediante la identificación y eliminación de redundancia calculándola mediante alguna fórmula estadística.

• Con pérdida (lossy) Solo se puede comprimir.

Elimina información innecesaria. Streaming

Hay pérdida de datos. Son algoritmos muy complejos.

Especialmente usados para almacenar archivos multimedia.

Se pierde información del texto/dato original y no hay forma de recuperarla. MP3/AAC/OGG/JPEG

• Los algoritmos lossless comprimen 20% - 90% dependiendo de las características del dato original.

⇒ utilizan información estadística.

• Uno de los algoritmos de compresión Lossless más conocidos es códigos de Huffman.

Huffman Codes

- Desarrollado en 1950 por David Huffman cuando era estudiante en el MIT.


- Se utiliza en muchos formatos de compresión.

- La idea es asignar un código binario lo más corto posible a los símbolos que aparecen en los datos por comprimir.

↑ frecuencia ⇒ símbolo más corto

- Constituye un árbol binario que se utiliza para asignar dichos símbolos.

zip



The diagram shows a rectangular box representing a zip file. A horizontal line divides the box into two sections. The top section is labeled 'header can metadata' with an arrow pointing to it. The bottom section is labeled 'data' with an arrow pointing to it.

136

m: 1 1100

i = 5 00

S: 4 01

ip: 1 sample 101

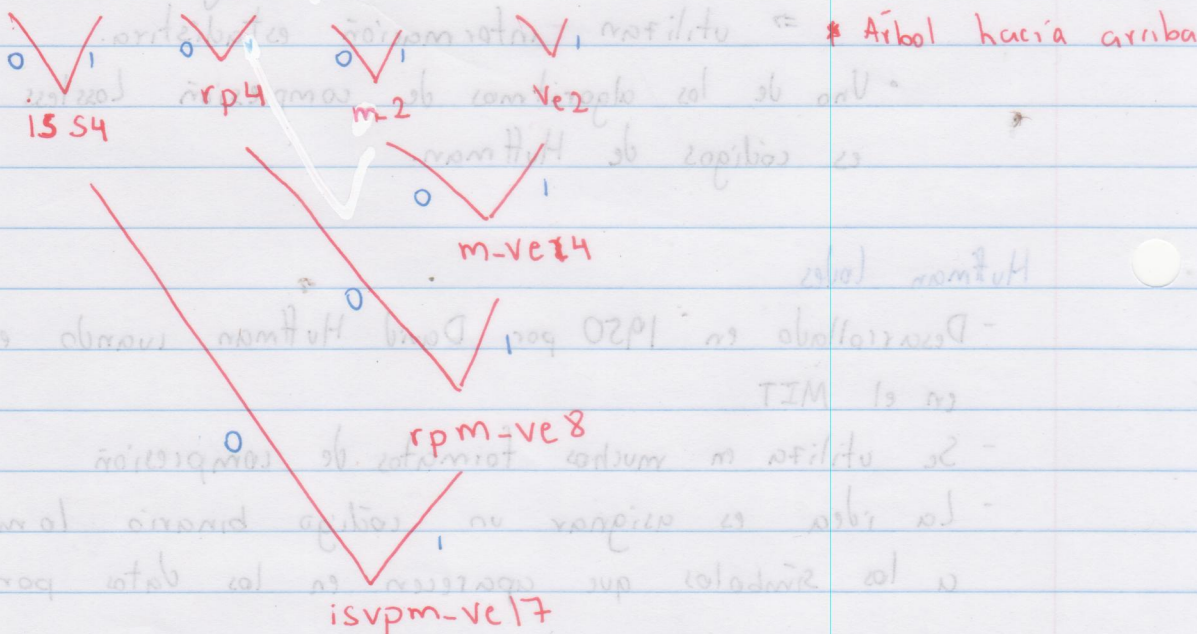
2:10pm 110

 $r = 2 \text{ m}$
$$v:1 \text{ cm}^2/\mu\text{C}$$

e: 1

Acomodar de menor a mayor

$i_5 \quad s_4 \quad r_{2,p_2} \quad m_{ab-1/b} \quad v_i e_i + z_{i+1}$



Todo a izquierda de todo a derecha

46 bit

compression 74% 