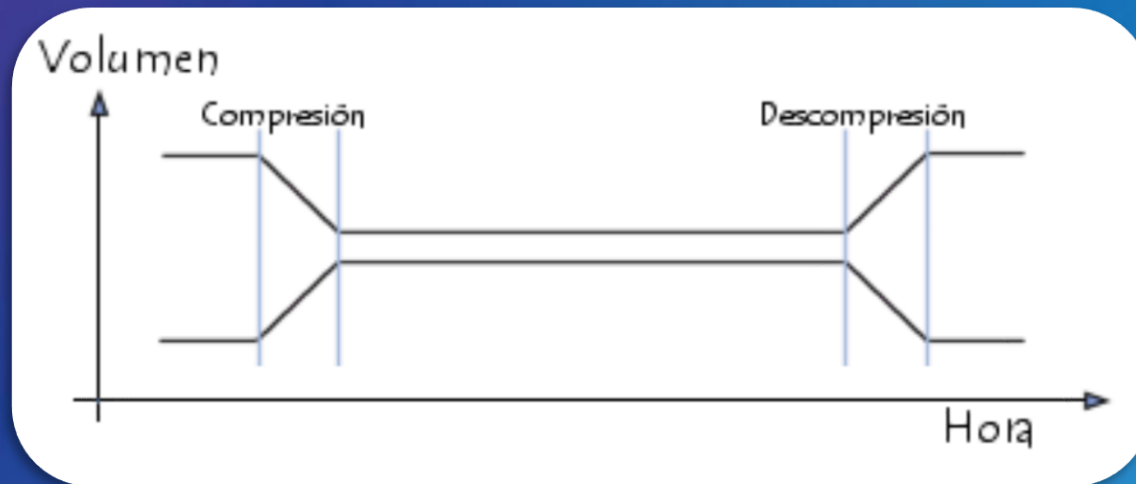




Universidad
Rafael Landívar
Tradición Jesuita en Guatemala

Compresión de datos

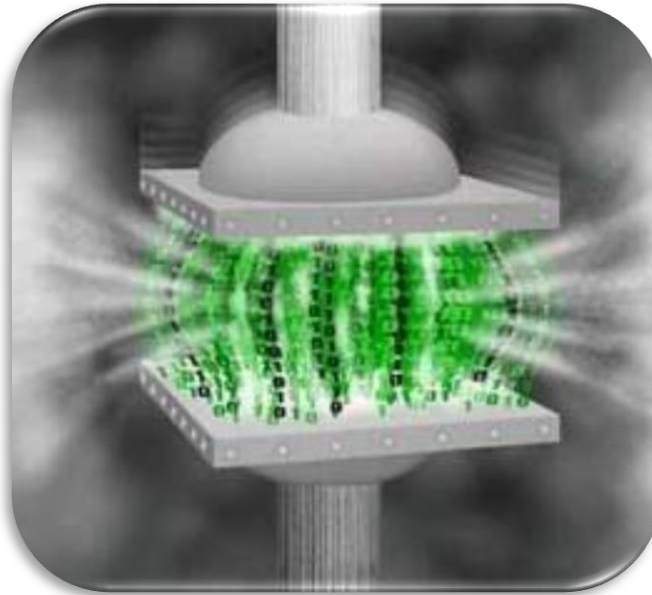


Facultad de ingeniería

Compresión de datos

En ciencias de la computación la compresión de datos es la reducción del volumen de datos tratables para representar una determinada información, empleando una menor cantidad de espacio.

Al proceso de compactar los datos se le denomina **compresión** y al proceso inverso **descompresión**.



Compresión de datos

El espacio que ocupa información sin compresión es el producto entre la frecuencia del muestreo y la resolución (tamaño de palabra), aumentando así el tamaño del archivo a gestionar.

En temas de transmisión o almacenamiento de información el tamaño del archivo a transmitir impacta en el rendimiento del mismo, por esto se utiliza la compresión para facilitar el proceso y modificar la resolución en un número de inferior bits (tecnología digital).

Compresión de datos

La compresión de datos se basa fundamentalmente en buscar repeticiones en series de datos para después almacenar solo el dato junto al número de veces que se repite. Así, por ejemplo, si en un fichero aparece una secuencia como "AAAAAA", ocupando 6 bytes se podría almacenar simplemente "6A" que ocupa solo 2 bytes, en algoritmo llamado RLE.

Aunque en realidad los procesos son mucho mas complejos, todos son basados en esta lógica de patrones y frecuencias.



A la hora de hablar de compresión hay que tener presentes dos conceptos que forman parte de la teoría de la información:

Redundancia: Es la repetición de información, entre mas repeticiones existan de una palabra mas predecible se vuelve.

Entropía: Se usa para describir la información impredecible e inesperada en el contexto que aparece. Entre mayor sea la aparición de palabras con menor redundancia más difícil se vuelve decodificar un mensaje, mientras que, si el grado de redundancia aumenta, existe más probabilidad de que un mensaje se decodifique fácilmente

Tipos de compresión

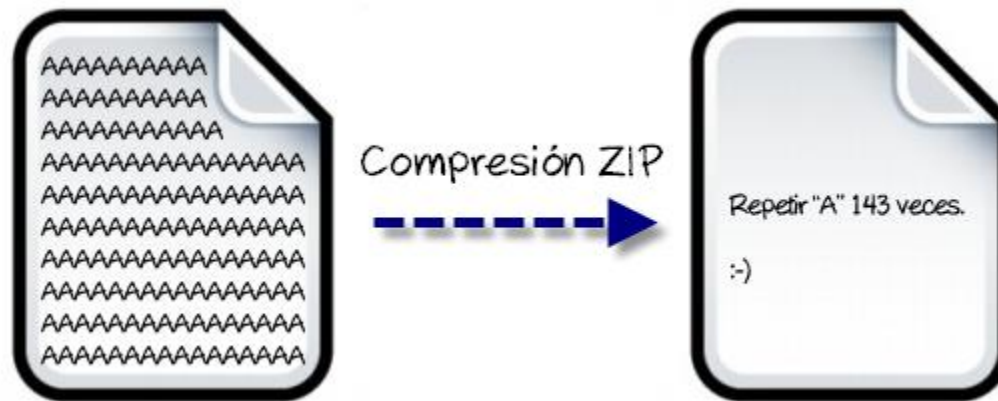
El objetivo de la compresión es reducir el peso binario, a la hora de realizar una compresión de la información tenemos ante nosotros dos posibilidades:

Optar por una compresión de datos con pérdida de calidad o bien no perder nada de calidad pero que también conlleva una menor reducción del tamaño.



Compresión sin pérdida

En este tipo de compresión, los datos antes y después del proceso son los mismos, sin sufrir ningún tipo de pérdida. En este proceso, una mayor compresión requiere de un mayor tiempo de proceso. Es utilizado principalmente para archivos de texto donde la pérdida de cualquier tipo de dato puede cambiar el significado del mismo.



Compresión con pérdida

En este tipo de sistema de compresión, el proceso puede eliminar datos para reducir su tamaño, lo que conlleva una reducción en la calidad de los datos. Una vez realizada la compresión, no se puede volver a la versión original aunque si que se puede lograr un acercamiento muy próximo. Es el sistema que se suele utilizar para la compresión de elementos multimedia como imágenes, vídeos o música.



Original



JPEG 30%



JPEG 1%

Ejercicio

- En grupos genere un tipo compresión de datos simple sin pérdida, que permita comprimir y descomprimir un mensaje de forma optima, el algoritmo debe funcionar para las siguientes frases.
- Considere el conjunto universo todas la letras utilizadas en la frase
- **LA CASA**
- **EL LEON NO VE**
- Asumiendo que cada una de las letras tiene un peso de 1 byte (8 bits), encuentre una forma de reducir el tamaño del stream de datos para poder transmitirlo.