

# Exercici AvCont-4(b): Compresió LZ-77: TEXT

## Tecnologies Multimedia

Igor Dzinka, Vicent Roig

### Ejercicio:

*Después de haber implementado en Java un compresor/descompresor de datos binarios mediante el algoritmo LZ-77 que cumpla las especificaciones indicadas en la parte (a) del ejercicio, **analizaremos ahora la capacidad de LZ-77 para comprimir texto.***

Haced las modificaciones necesarias en el programa anterior para que sea capaz de leer ficheros de texto, generar el código ASCII (de 8 bits) de cada carácter del texto y devolverlo como una String binaria procesable por vuestros métodos de compresion /descompresión en LZ-77. Conseguid también que el programa calcule el tiempo invertido en comprimir los datos. Comprobad el correcto funcionamiento del programa; es decir, que es capaz de comprimir texto en LZ-77 y de recuperarlo tras la descompresión.

Comprimid el archivo “hamlet\_short.txt” con distintos valores de Mdes y Ment entre 4 y 4096 y analizad el factor de compresión y el tiempo invertido para conseguirlo.

¿Cuál es el mejor factor de compresión que obtenéis y con qué valores de Mdes y Ment?

mDest	mEnt	Tiempo(segundos)	Ratio de compresión
2048	32	0.33.	1.0937376
2048	64	0.62	1.1018728
2048	128	1.32	1.0901445
2048	256	2.63	1.0443102
2048	512	4.66	1.0171653
4096	32	0.15	1.1331785
4096	64	0.43	1.1917828
4096	128	1.00	1.2019705
4096	256	2.10	1.1702638
4096	512	4.03	1.1512629
4096	1024	6.96	1.0935935

El mejor factor de compresión que conseguimos en el fichero Hamlet\_short.txt es de 1.201:1 y lo conseguimos con valor de entrada deslizante de 4096 y el de ventana de entrada de 128. Para valores de ventana deslizante menores a 2048 se consiguen unos ratios de compresión muy próximos o inferiores a 1.

## Exercici AvCont-4(b): Compresió LZ-77: TEXT

### Tecnologies Multimedia

Igor Dzinka, Vicent Roig

¿Cómo varia (cómo escala) el tiempo de cálculo necesario al aumentar Mdes y Ment?

Empíricamente hemos podido comprobar que cuanto más grande es el valor de mDest, más rápido se ejecuta la compresión y cuanto más pequeño el valor de mEnt también se ejecuta más rápido.

¿Qué combinación de Mdes y Ment elegiríais?

Como se refleja en la tabla para los valores de mDest = 4096 y mEnt = 128 se consigue el mejor ratio de compresión (1.2019705). Sin embargo también vemos que para valores de mDest = 4096 y mEnt = 64, se consigue un ratio de compresión muy similar (1.1917828) pero tardando menos de la mitad de tiempo que en el caso anterior (0.43 segundos contra 1.00). Por lo tanto si nos interesa **el mayor ratio de compresión**, elegiremos **mDest = 4096 y mEnt = 128**. En caso de que nos interese **un buen ratio con tiempo de ejecución más óptimo**, elegiremos **mDest = 4096 y mEnt = 64**.

Comprimid ahora el archivo "quijote\_short.txt" y analizad para qué combinación de Mdes y Ment se obtiene el mejor factor de compresión. ¿Es el mismo que en el caso anterior? Proponed varias razones que expliquen esta diferencia.

mDest	mEnt	Tiempo(segundos)	Ratio de compresión
2048	32	0.17	1.129906
2048	64	0.52	1.1119272
2048	128	1.19	1.0672235
2048	256	2.76	1.0218723
2048	512	5.16	0.9850602
4096	32	0.18	1.110417
4096	64	0.59	1.1047156
4096	128	1.53	1.0789126
4096	256	3.14	1.0496855
4096	512	5.93	1.0286864
4096	1024	9.61	1.0055344

Como podemos ver se consiguen unos ratios de compresión inferiores de media que con el fichero hamlet\_short.txt. El mejor ratio se consigue con mDest = 2048 y mEnt = 32.

Analizando el texto, nos damos cuenta que en el texto de quijote\_short hay cierta cantidad de repeticiones de patrones de entre 1 y 4 letras pero muy pocas de mas tamaño. Para detectar estos patrones no nos hace falta una ventana de entrada de gran tamaño, con una de 32-64 bits hay mas que suficiente y además se consigue un rendimiento mayor.

# Exercici AvCont-4(b): Compensió LZ-77: TEXT

Tecnologies Multimedia

Igor Dzinka, Vicent Roig

---

En cambio en el fichero Hamlet\_short podemos encontrar patrones mas largos, de entre 6 y 10 letras, y por lo tanto una ventana de entrada de entre 64 y 128 caracteres optimiza mucho mas la compresión.