Compresión de datos sin pérdida
Práctico 4

Ejercicio 1 *Parsing LZ78*
Obtener el parsing LZ78 de la secuencia binaria $x^n = 00000110101000000110101$. Mostrar también la codificación de x^n como secuencia de pares (i, b), donde i es un índice de frase y b un símbolo.

Ejercicio 2 *Secuencia constante*
Consideramos la secuencia constante $x^n = 1111\ldots$.

1. Obtener el parsing LZ77 y LZ78 de la secuencia.
2. Mostrar que la cantidad de bits de codificación por símbolo de entrada tiende a cero con n, tanto en LZ77 como en LZ78.

Ejercicio 3 *LZ78*

1. Continuar el parsing LZ78 de la secuencia $0, 00, 001, 000000110101111$.
2. Determinar una secuencia, de un largo dado, tal que el número de frases en el parsing LZ78 se maximiza.
3. Determinar todas la secuencias, de un largo dado, para las cuales el número de frases en el parsing LZ78 se minimiza.

Ejercicio 4 *Largo de punteros en LZ77*
Consideramos una versión de LZ77 en la cual los tokens son de la forma (F, P, L) o (F, C), donde F es un bit de flag que indica el tipo de token, C es un carácter literal, P es un puntero al inicio de una concordancia de un patrón dentro de la ventana y L es el largo de la concordancia. Según convenga, una concordancia corta se puede representar con un único token de la forma (F, P, L) o como una secuencia de tokens de la forma (F, C). Asumimos que el ancho de la ventana es W y el largo de las concordancias se limita a un máximo de M.

1. ¿Cuántos bits se necesitan para representar P y para representar L?.
2. Asumir que los literales se representan con 8 bits. ¿Cuál es el largo de concordancia mínimo, en función de W y M, que debería representarse con un token de la forma (F, P, L) en lugar de una secuencia de tokens (F, C)?
3. Calcular el largo de concordancia mínimo determinado en el punto anterior para $W = 4096$ y $M = 256$.

Ejercicio 5
Sea $x = k2^k$, $k \geq 1$.

1. Demuestre que $\log k \leq \log \log x$ (sugerencia: escriba k en función de $\log x$ y $\log k$ y tome logaritmos).
2. Demuestre que $k \geq \log x - \log \log x$.
3. Demuestre que $k \leq \log x - \log \log x + o(1)$.
   (Sugerencia: escriba $2^k = \frac{x}{k}$, aplique la cota del inciso anterior, y tome logaritmos.)

Ejercicio 6 *LZ78: cota superior sobre $c(n)$*

Complete la demostración de la cota

$$c(n) \leq \frac{n}{(1 - \epsilon_n) \log n}$$

vista en las diapositivas. Puede utilizar todos los pasos de la diapositiva, y se sugiere utilizar el resultado del ejercicio anterior para completar la demostración. En particular, demuestre que $\epsilon_n = O(\frac{\log \log n}{\log n})$.