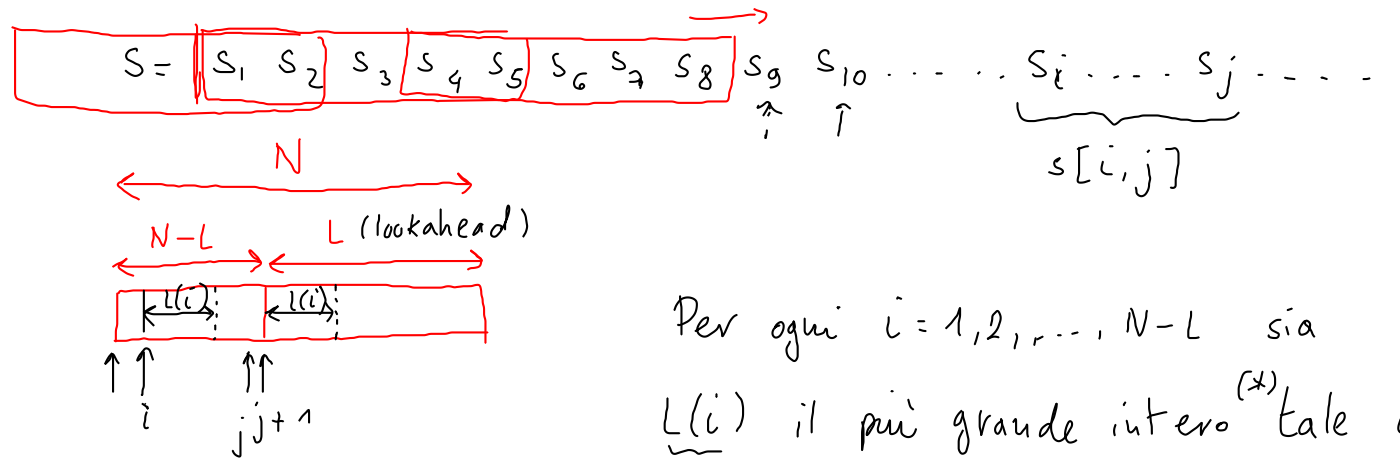


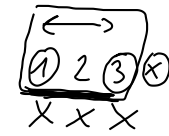
CODICE DI ZIV-LEMPER (LZ77)

→ .zip, .gzip



Finestra

$$L(i) = 3$$



$i=1$

1

3

$$3 = 1 + 3 - 1$$

$$3 = 3 - 1 + 1$$

Per ogni $i = 1, 2, \dots, N-L$ sia

$L(i)$ il più grande intero ^(*) tale che:

$$(*) < L$$

$$s[i, i+L(i)-1] = s[j+1, j+L(i)]$$

Sia

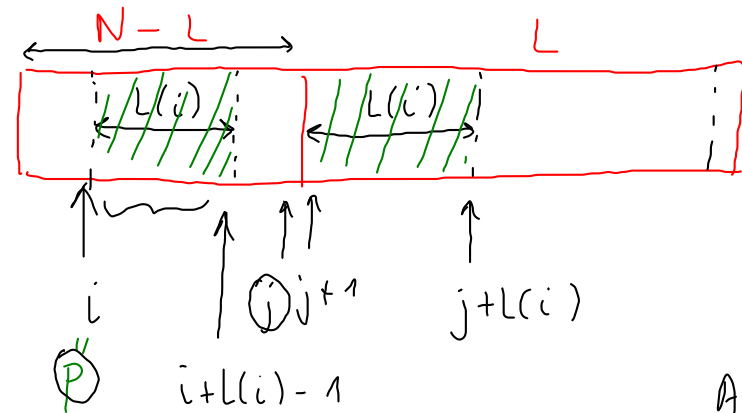
$$L(p) = \max_{1 \leq i \leq j} L(i)$$

Allora $s[j+1, j+L(p)]$ è estensione riproducibile

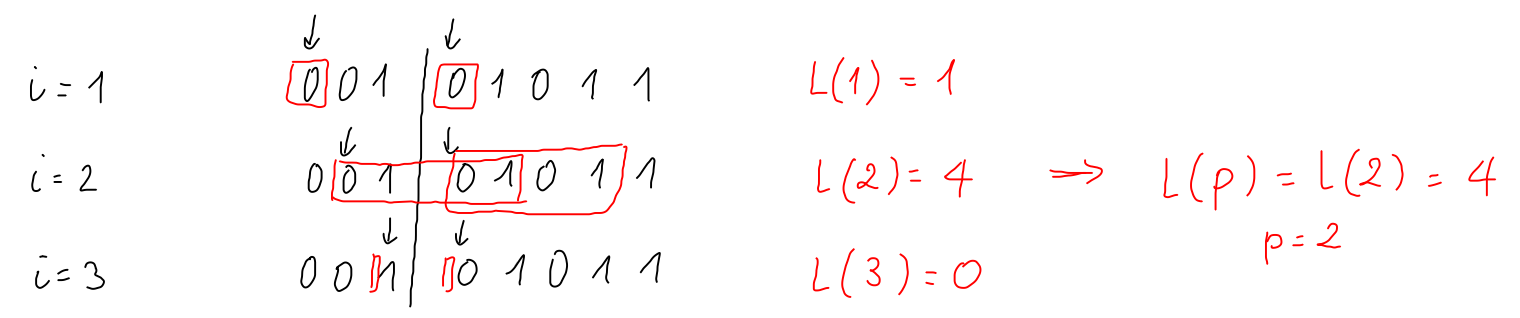
$$\left\{ \begin{array}{l} p \\ L(p) \end{array} \right.$$

è il puntatore della riproduzione

è la lunghezza della riproduzione



Esempio. Finestra di lunghezza $N=8$, $s = \overbrace{001}^3 | 01011$ ($j=3$)



La stringa $s_1 s_2 s_3 s_4 s_5 s_6 s_7 s_8 s_9 \dots$ è segmentata in messaggi

$\leftarrow \xrightarrow{m_1} \quad \leftarrow \xrightarrow{m_2} \quad \leftarrow \xrightarrow{m_3}$

Ciascun messaggio m_i è codificato in una parola di codice w_i a lunghezza costante (L_c).

$$w_i = w_{i1} w_{i2} w_{i3}$$

La lunghezza di ogni w_i è

$$L_c = \lceil \log_k(N-L) \rceil + \lceil \log_k L \rceil + 1$$

- w_{i1} : posizione del puntatore $\rightarrow p-1$ in base $K \rightarrow \lceil \log_k(N-L) \rceil$ N-L possibili valori
- w_{i2} : lunghezza della riproduzione $\rightarrow L(p)$ in base $K \rightarrow \lceil \log_k L \rceil$ L possibili valori
- w_{i3} : primo simbolo dopo la riproduzione (in base K) $\rightarrow 1$

Algoritmo Ziv-Lempel (LZ77)

Chiamiamo F_i la finestra al passo i -esimo dell'algoritmo.

1. Inizializza il contenuto della finestra F_1 di lunghezza N con $N-L$ simboli '0' e i primi L simboli di s ($s[1, L]$)

$$(F_1 = 0^{N-L} * s[1, L])$$

2. Noto F_i ($i > 1$), sia $m_i = [N-L + 1, N-L + L(p_i) + 1]$

3. Codifica m_i in $w_i = w_{i1} w_{i2} w_{i3}$ secondo lo schema già discusso (estensione riproducibile)

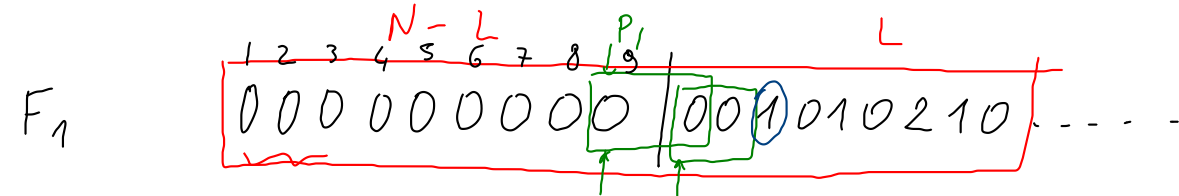
4. Aggiorna la finestra eliminando i primi $L(p_i) + 1$ simboli

e facendone subentrare altrettanti dalla sorgente, per ottenere F_{i+1} .

Esempio .

$s = \overbrace{001010210}^g \overbrace{210212021021200}^L$ con $N=18, L=9$ ($K=3$) $\mathcal{A}=\{0,1,2\}$

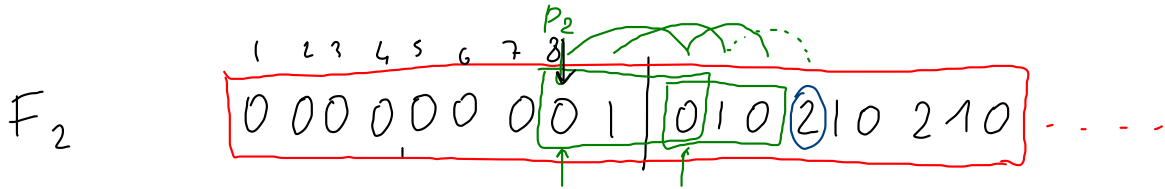
$$\Rightarrow L_c = \lceil \log_k(N-L) \rceil + \lceil \log_k L \rceil + 1 = \lceil \log_3 9 \rceil + \lceil \log_3 9 \rceil + 1 = 2 + 2 + 1 = 5.$$



$$p_1 = 9 \quad L(p_1) = 2 \quad '1'$$

$$p_1 - 1 = 8 \quad \downarrow \quad \downarrow \quad \downarrow$$

$$w_{11} = (22)_3 \quad w_{12} = (02)_3 \quad w_{13} = 1 \quad \rightarrow w_1 = 22021$$

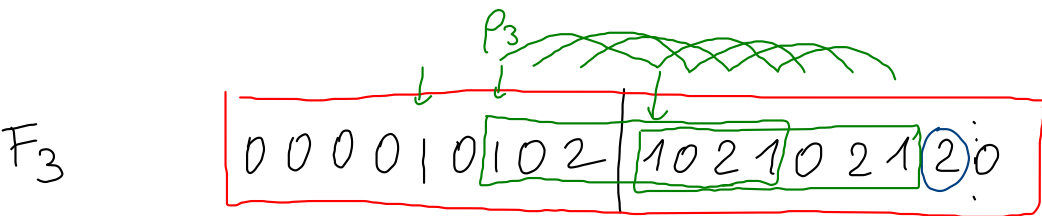


Faccio avanzare la finestra di $L(p_1)+1$ simboli

$$p_2 = 8 \quad L(p_2) = 3 \quad '2'$$

$$p_2 - 1 = 7 \quad \downarrow \quad \downarrow$$

$$w_{21} = (21)_3 \quad w_{22} = (10)_3 \quad w_{23} = 2 \quad \rightarrow w_2 = 21102$$

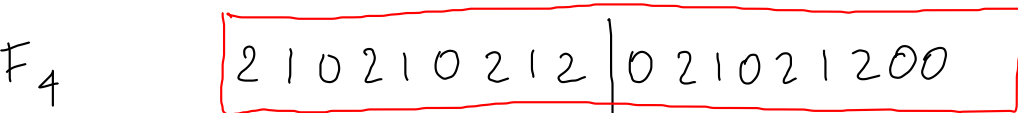


Faccio avanzare la finestra di $L(p_2)+1$ simboli

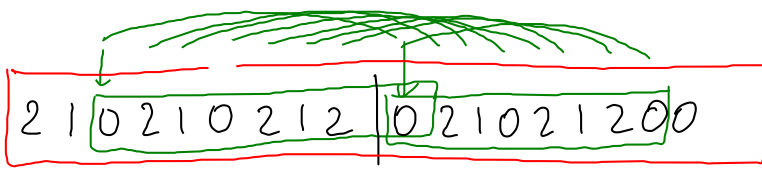
$$p_3 = 7 \quad L(p_3) = 7 \quad '2'$$

$$p_3 - 1 = 6$$

$$w_{31} = (20)_3 \quad w_{32} = (21)_3 \quad w_{33} = 2 \quad \rightarrow w_3 = 20212$$



F₄



$$p_4 = 3$$

$$p_4 - 1 = 2$$

$$L(p_4) = 8$$

'0'

$$w_{41} = (02)_3$$

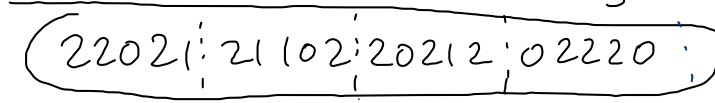
$$w_{42} = (22)_3$$

$$w_{43} = 0$$

$$w_4 = 02220$$

5 5 5 5

⇒ Ho codificato s complessivamente in



N
L
K

→ calcolo

$$\lceil \log_K (N-L) \rceil +$$

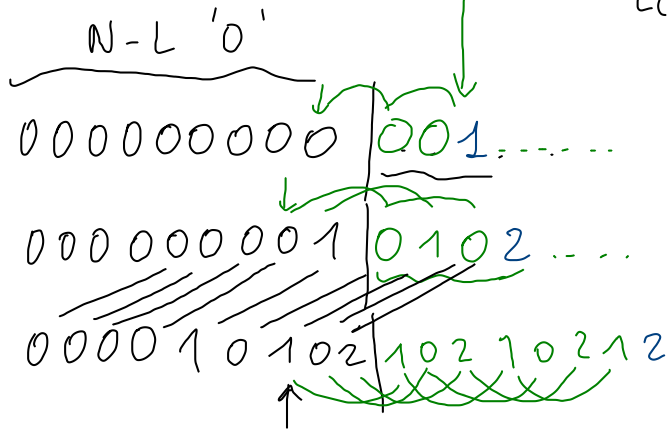
$$\lceil \log_K L \rceil +$$

$$1$$

$$= L_C$$

Decodifica: $w_1 = 2210210 \rightarrow p_{i-1} = (22)_3 = 8 \rightarrow p_i = 9$

$$L(p_i) = (02)_3 = 2$$



$$w_2 = 21102 \rightarrow p_2 - 1 = (21)_3 = 7 \quad p_2 = 8$$

$$\rightarrow L(p_2) = (10)_3 = 3 +$$

'2'

$$w_3 = 20212 \rightarrow p_3 - 1 = (20)_3 = 6 \quad p_3 = 7$$

$$L(p_3) = (21)_3 = 7$$

$$w_4 = 02220 \rightarrow p_4 - 1 = 2 \quad p_4 = 3$$

$$L(p_4) = (22)_3 = 8$$

'0'

