



**Disciplina: Natureza da Informação (BCM0504) – 2018-Q3**

Prof. Alexandre Donizeti Alves

**GABARITO [Lista de Exercícios 04 – Compressão]**

1. Considere a mensagem “122121213”. Assumindo que cada caractere é um ASCII de 8 bits, como ela seria transmitida nos seguintes casos?

a) Codificação run-length.

**1[1] 2[2] 1[1] 2[1] 1[1] 2[1] 1[1] 3[1]**

b) Compressão LZW. Mostre os passos realizados.

Entrada	Dicionário	Transmissão
1	--	256 (START)
2	258 (12)	49 (1)
2	259 (22)	50 (2)
1	260 (21)	50 (2)
2	--	--
1	261 (121)	258 (12)
2	--	--
1	--	--
3	262 (1213)	261 (121)
		51 (3)
		257 (STOP)

c) Qual algoritmo foi mais efetivo na compressão?

**Assumindo que o run-length utiliza 8 bits por caractere da string gerada [SEM CONSIDERAR OS COLCHETES]: run-length utiliza 128 (8x8x2) bits e LZW utiliza 72 (8x9) bits. Portanto, o LZW foi mais eficiente em termos de quantidade de bits.**

2. Mostre os passos da compressão LZW para as sequências:

a) AABCBBABC (**9 x 8 = 72 bits**)

Entrada	Dicionário	Transmissão
A	--	256 (START)
A	258 (AA)	65 (A)
B	259 (AB)	65 (A)
C	260 (BC)	66 (B)
B	261 (CB)	67 (C)
B	262 (BB)	66 (B)
A	263 (BA)	66 (B)
B	--	--
C	264 (ABC)	259 (AB)
		67 (C)
		257 (STOP)

b) ABRACADABRA (11 x 8 = 88 bits)

Entrada	Dicionário	Transmissão
A	--	256 (START)
B	258 (AB)	65 (A)
R	259 (BR)	66 (B)
A	260 (RA)	82 (R)
C	261 (AC)	65 (A)
A	262 (CA)	67 (C)
D	263 (AD)	65 (A)
A	264 (DA)	68 (D)
B	--	--
R	265 (ABR)	258 (AB)
A	--	--
		260 (RA)
		257 (STOP)

3. Que compressão é conseguida em cada caso do exercício anterior?

a)  $1 - (10 \times 9) / (9 \times 8) = 1 - (90 / 72) = 1 - 1,25 = -0,25$

Como a taxa de compressão foi negativa, isso significa que o LZW acabou expandindo a mensagem ao invés de comprimir (aumentou em 25%)

b)  $1 - (11 \times 9) / (11 \times 8) = 1 - (99 / 88) = 1 - 1,125 = -0,125$

Como a taxa de compressão foi negativa, isso significa que o LZW acabou expandindo a mensagem ao invés de comprimir (aumentou em 12,5%)

4. Consideremos a seguinte variante versão portuguesa da famosa frase do político JFK:

***A pergunta de cada um de nós não deve ser o que o país pode fazer por nós; mas sim o que cada um de nós pode fazer pelo país***

a) Quantos bytes ocupa a frase?

**Considerando que cada caractere ocupa 1 byte, a frase ocupa 124 bytes.**

b) Que frequência tem cada uma das palavras?

A	1
espaço	29
pergunta	1
de	3
cada	2
um	2
nós	3
não	1
deve	1
ser	1
o	3
que	2
país	2
pode	2
fazer	2
por	1
;	1
mas	1
sim	1
pelo	1

c) Monte um dicionário estático para essa frase com base na resposta da questão b.

Dicionário	Binário	Representação
0	00000	A
1	00001	espaço
2	00010	pergunta
3	00011	de
4	00100	cada
5	00101	um
6	00110	nós
7	00111	não
8	01000	deve
9	01001	ser
10	01010	o
11	01011	que
12	01100	país
13	01101	pode
14	01110	fazer
15	01111	por
16	10000	;
17	10001	mas
18	10010	sim
19	10011	pelo

d) Quantos bytes ocupa o dicionário?

**O dicionário ocupa 5 bits por entrada e, como possui 20 entradas, ocupa 100 bits.**

5. Como as sequências a seguir seriam codificadas com run-length?

a) AAABBBBBBYYYYPPPPPPPPPTKKKKKKKKK (31 x 8 = 248 bits)

A[3] B[6] Y[4] P[9] T[1] K[8]

b) 111112223333312222221111111333333333 (36 x 8 = 288 bits)

1[5] 2[3] 3[5] 1[1] 2[6] 1[7] 3[9]

6. Qual porcentagem de compressão foi conseguida em cada caso no exercício anterior?

**Assumindo que run-length utiliza 8 bits por caractere da string gerada [SEM CONSIDERAR OS COLCHETES]:**

$$\text{a) } 1 - (6 \times 8 \times 2) / (31 \times 8) = 1 - (96 / 248) = 1 - 0,3871 = 0,6129$$

Como a taxa de compressão foi positiva, isso significa que o run-length diminuiu a mensagem em 61,29%.

$$\text{b) } 1 - (7 \times 8 \times 2) / (36 \times 8) = 1 - (112 / 288) = 1 - 0,3889 = 0,6111$$

Como a taxa de compressão foi positiva, isso significa que o run-length diminuiu a mensagem em 61,11%.

7. A mensagem abaixo representa um código construído pelo método LZW utilizando como dicionário inicial o código ASCII de 8 bits.

256 65 66 65 258 67 67 258 66 265 67 257

a) Sabendo-se que as letras A, B e C são representadas pelos códigos 65, 66 e 67, respectivamente, e que os códigos de início e de final de mensagem são dados pelos códigos 256 e 257, respectivamente, qual é a mensagem original?

Transmissão	Dicionário	Saída
256 (START)	--	--
65 (A)	--	A
66 (B)	258 (AB)	B
65 (A)	259 (BA)	A
258 (AB)	260 (AA)	AB
67 (C)	261 (ABC)	C
67 (C)	262 (CC)	C
258 (AB)	263 (CA)	AB
66 (B)	264 (ABB)	B
265 (BB)	265 (BB)	BB
67 (C)	266 (BBC)	C
257 (STOP)		

**Mensagem original = ABAABCCABBBBC**

b) Como cada pacote emitido pelo LZW consiste de uma sequência de 9 bits, qual foi a taxa de compressão obtida em relação ao texto puro em ASCII de 8 bits?

**A mensagem ABAABCCABBBBC consiste de 13 pacotes de 8 bits. Como o código obtido pelo LZW consiste de 12 pacotes de 9 bits, a taxa de compressão obtida foi de:**

$$1 - (12 \times 9) / (13 \times 8) = 1 - (108 / 104) = 1 - 1,0385 = -0,0385$$

**Como a taxa de compressão foi negativa, isso significa que o LZW acabou expandindo a mensagem ao invés de comprimir, ou seja, aumentou em 3,85%.**