

Aula 03



Métodos de compressão básicos

RLE

Remoção de espaços com bitmap

7 bits/caracter

MTF - *Move to Front*

BWT - *Burrows-Wheeler Transform*

Remoção de espaços

Entrada: “o rato roeu a rolha do remédio ruim do rei da Rússia”

Saída: “oratoroeuarolhadoremédioruimndoreidaRússia”

Bitmap: 01000010 00010100 00010010 00000010 00010010 001001

Redução de bits/caracter

7 bits/caracter

Código Baudot: 5 b/c

Move to Front Transform

Transforma corridas (*runs*) em
sequências de zeros

Ex.: Codificação

entrada = bananaaana

alfabeto inicial = {abn} →

0	1	2
a	b	n

- 1) Saída: 1, alfabeto {ban}
- 2) Saída: 1, alfabeto {abn}
- 3) Saída: 2, alfabeto {nab}
- 4) Saída: 1, alfabeto {anb}
- 5) Saída: 0, alfabeto {anb}
- 6) Saída: 0, alfabeto {anb}
- 7) Saída: 0, alfabeto {anb}
- 8) Saída: 0, alfabeto {anb}
- 9) Saída: 1, alfabeto {nab}
- 10) Saída: 1, alfabeto {anb}
- 11) Saída final: 1,1,2,1,0,0,0,0,1,1, alfabeto {ban}

Move to Front Transform

Transforma corridas (*runs*) em
sequências de zeros

Ex.: Decodificação

entrada = 1,1,2,1,0,0,0,0,1,1

alfabeto inicial = {abn} →

0	1	2
a	b	n

- 1) Saída: b, alfabeto {ban}
- 2) Saída: a, alfabeto {abn}
- 3) Saída: n, alfabeto {nab}
- 4) Saída: a, alfabeto {anb}
- 5) Saída: a, alfabeto {anb}
- 6) Saída: a, alfabeto {anb}
- 7) Saída: a, alfabeto {anb}
- 8) Saída: a, alfabeto {anb}
- 9) Saída: n, alfabeto {nab}
- 10) Saída: a, alfabeto {anb}
- 11) Saída final: banaaaaaana

Burrows-Wheeler Transform

Rearranja strings em corridas (*runs*)

Método: ordena todos os shifts circulares de um texto (S) em ordem lexicográfica, extraíndo a última coluna e o índice (I) da string original do conjunto ordenado de permutações de (S)

Ex.: Codificação

1) S = ^banana|

2) shifts:

^banana|

|^banana

a|^banan

na|^bana

ana|^ban

nana|^ba

anana|^b

banana|^

3) shifts:

anana|^b

ana|^ban

a|^banan

banana|^

nana|^ba

na|^bana

^banana|

|^banana

4) saída (última coluna e índice)

bnn^aa|a; índice = 7

Burrows-Wheeler Transform

Rearranja strings em corridas (*runs*)

Método: ordena todos os shifts circulares de um texto (S) em ordem lexicográfica, extraíndo a última coluna e o índice (I) da string original do conjunto ordenado de permutações de (S)

Ex.: Decodificação

entr.	Sort	add 2	sort	add 3	sort	4567	add8	sort
b	a	ba	an	ban	ana	...	banana ^	anana ^b
n	a	na	an	nan	ana		nana ^ba	ana ^ban
n	a	na	a	na	a ^		na ^bana	a ^banan
^	b	^b	ba	^ba	ban		^banana	banana ^
a	n	an	na	ana	nan		anana ^b	nana ^ba
a	n	an	na	ana	na		ana ^ban	na ^bana
	^	^	^b	^b	^ba		^banana	^banana 7
a		a	^	a ^	^b		a ^banan	^banana