

Códigos de condição prefixa

Fonte S: emite os símbolos s_1 a s_k , com probabilidades de ocorrência p_1 a p_k . A cada símbolo corresponde uma palavra de código C_k , em que cada elemento de m_k de C_k é 0 ou 1 e n_k é o número de bits da palavra. Ou seja o código C_k é $m_{1k} m_{2k} m_{3k} \dots m_{n_k k}$. Um prefixo de C_k é qualquer seqüência $m_{1k} m_{2k}$ de comprimento inferior a n_k . Um código é de condição prefixa se para qualquer par $C_i C_j$, C_i não é prefixo de C_j . Um código de condição prefixa é sempre de decodificação única.

Exemplos

Símbolo	Probabilidade	Código I	Código II	Código III
s0	0.5	0	0	0
s1	0.25	1	10	01
s2	0.125	00	110	011
s3	0.125	11	111	0111

Código I: não é de condição prefixa: s0 é prefixo de s2 ...

Código II: é de condição prefixa

Código III: não é de condição prefixa, mas é de decodificação única

Tipos de códigos

Decodificação única – só há uma forma de dividir uma seqüência não separada de códigos em códigos separados.

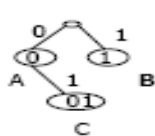
De condição prefixa – nenhum código é prefixo de outra palavra de código. É um código de decodificação única, em que quando se analisa uma seqüência não é preciso avançar sobre a seqüência para decidir qual o código corrente.

Ótimo – O código tem o menor comprimento médio possível ($=H(S)$). Não existem códigos não utilizados.

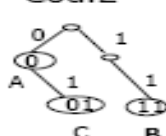
Representação em árvore

Símbolo	Cod1	Cod2	Cod3	Cod4	Notas
A	0	0	00	00	Cod.1 não é de decodificação única Cod. 2 precisa de lookahead Cod. 3 e Cod 4 são de condição prefixa Cod. 3 não é ótimo
B	1	11	01	01	
C	01	01	10	1	

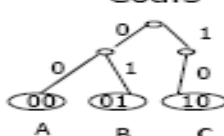
Cod.1



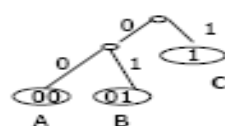
Cod.2



Cod.3

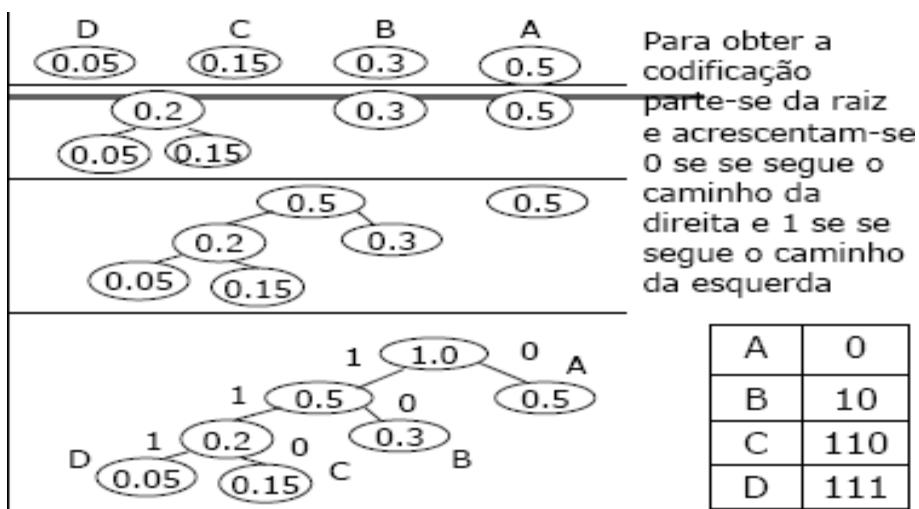
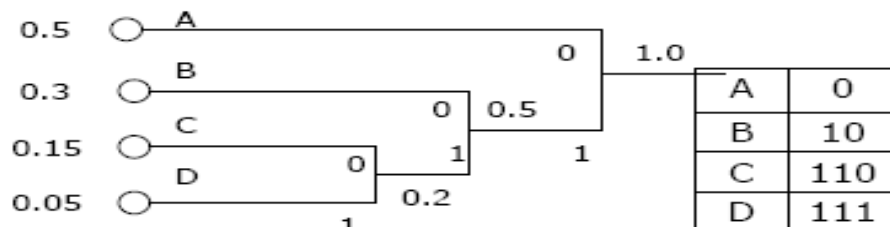


Cod.4



Codificação de Huffman

- Alfabeto $S = \{x_1, x_2, \dots, x_n\}$
- Probabilidades $P = \{p_1, p_2, \dots, p_n\}$
- p_i é ordenado ($p_1 \geq p_2 \geq p_3 \dots \geq p_n$)



Exercício 1. Seja alfabeto $A = \{1, 2, 3, 4, 5, 6, 7, 8\}$ em que $p(x_i) = 1/8$, desenhar a árvore de Huffman e estabelecer o código apropriado para cada símbolo.

Exercício 2. Na Tabela 1 estão listados os símbolos de uma fonte discreta sem memória e as respectivas probabilidades de ocorrência. Desenhar a árvore de Huffman e estabelecer o código apropriado para cada símbolo.

m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9
0.200	0.150	0.130	0.120	0.100	0.09	0.08	0.07	0.06

Tabela 1: Estatísticas dos símbolos de uma fonte discreta sem memória

Exercício 3. Uma palavra foi codificada usando o código de Huffman, tendo-se obtido a sequência binária

1 0 1 1 1 0 1 1 0 1 0 1 1 1 0 0 1 1 1 0 1 0 0

O alfabeto original era constituído pelas letras A, B, C, D, E, I, L, R e T e a letra I foi codificada como "00". Supondo que estas letras ocorriam com as probabilidades:

$P(A) = 0,26$ $P(D) = 0,01$ $P(L) = 0,01$

$P(B) = 0,09$ $P(E) = 0,07$ $P(R) = 0,23$

$P(C) = 0,08$ $P(I) = 0,22$ $P(T) = 0,03$

qual terá sido a palavra codificada?