

Aula 07



Codificação aritmética

Codificação por entropia

Codifica toda a mensagem em um único número

Saída é um número n tal que $0.0 \leq n < 1.0$

Demorou-se a encontrar métodos práticos para codificação em sistemas com registradores de tamanho fixo

Codificação aritmética

Ex.: mensagem \rightarrow ABBCAB

Usando codificação binária de bloco de tamanho fixo

A=00, B=01, C=10 (11 não-usado)

Mensagem codificada = 000101100001 (12 bits)

Usando codificação aritmética (com iguais probabilidades para cada caractere)

Representar a mensagem como a parte fracionária em um número em base 3 $\rightarrow (0.011201)_3$, depois converter para base 2 $\rightarrow (0.0010110010)_2$

Mensagem codificada = 0010110010 (10 bits)

Decodificação: converter para base 3, arredondar para 6 dígitos (tamanho da mensagem precisa ser conhecido pelo decodificador)

Codificação aritmética para símbolos com probabilidades diferentes

Para codificação ótima, utiliza-se modelo com probabilidades

Algoritmo:

Intervalo inicial = $[0, 1)$

Enquanto houver caracteres na entrada

Intervalo = intervalo correspondente ao próximo caractere

Código final = valor fracionário dentro do último intervalo (mínimo de dígitos suficientes para que o valor fique dentro do intervalo)

Codificação aritmética para símbolos com probabilidades diferentes

Probabilidades

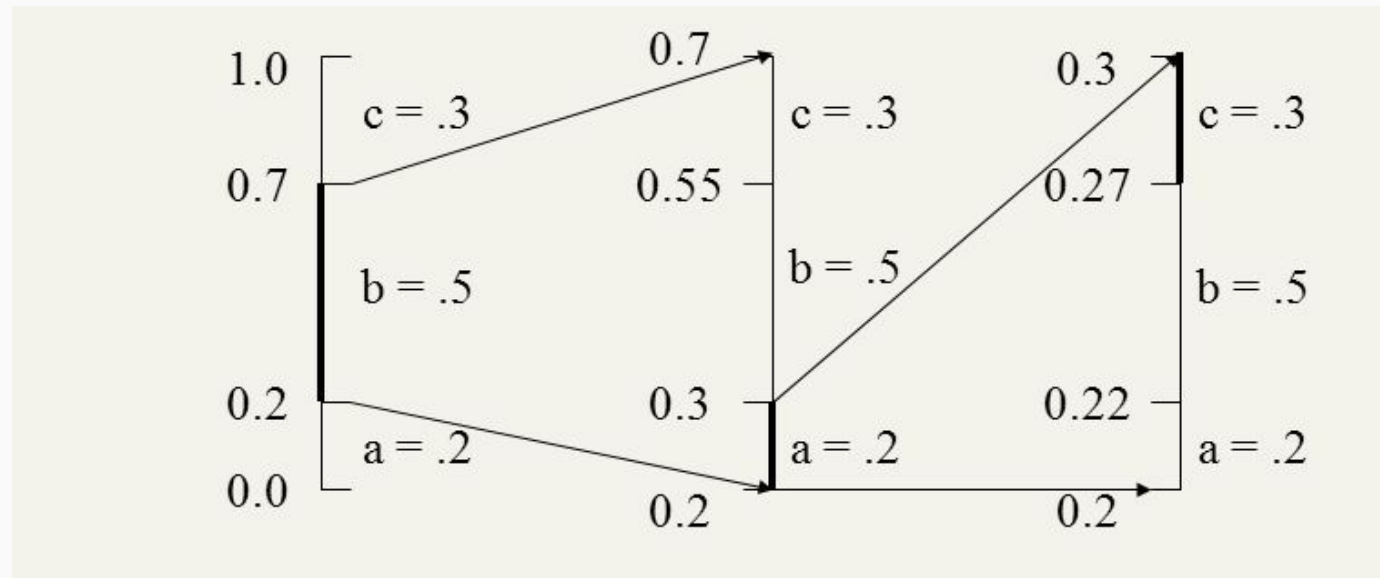
Iniciais:

$$P(a) = 0.2$$

$$P(b) = 0.5$$

$$P(c) = 0.3$$

Mensagem: **bac**

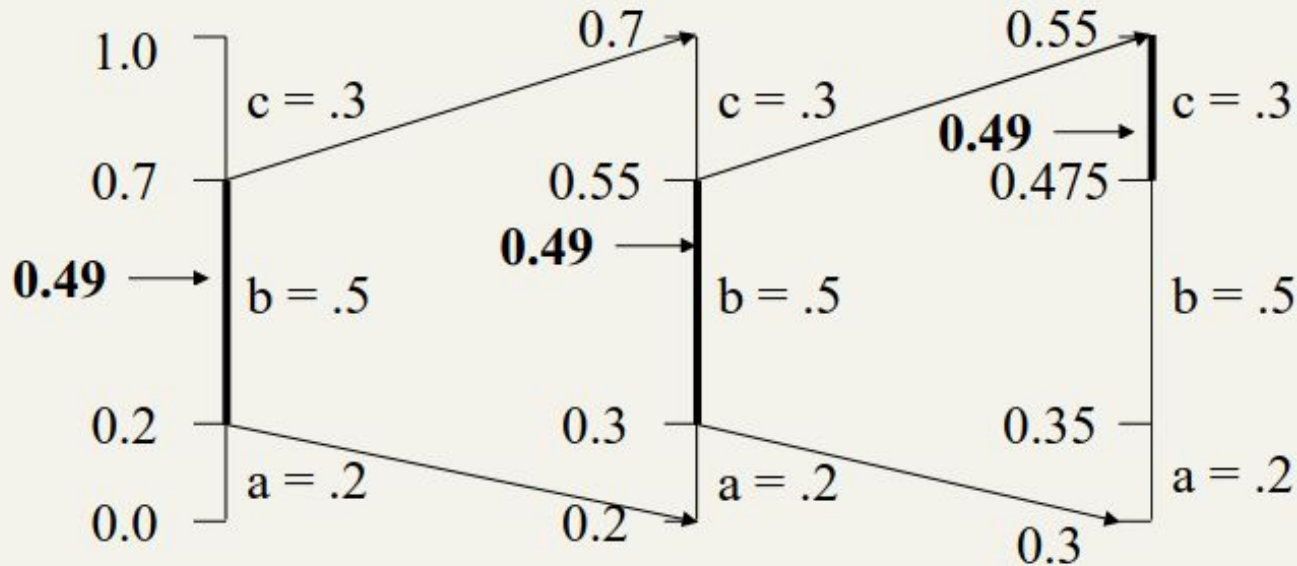


Intervalo final: [0.27, 0.3)

Decodificação aritmética para símbolos com probabilidades diferentes

Mensagem de tamanho 3

Código: 0.49



Codificação aritmética vs Huffman

Codificação aritmética	Huffman
Não é método estatístico	Método estatístico
Resultado ótimo	Resultado próximo do ótimo
Sem correspondência 1 para 1 entre símbolo original e código final (mensagem inteira tem um código único)	Cada símbolo original tem um código final