

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA DE CAMPINAS

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

"CONSTRUÇÃO DE ALGORITMOS DE TESTE  
EFICIENTES UTILIZANDO A TEORIA DA  
INFORMAÇÃO"

JAIME MARIZ DE FARIA Jr.

ORIENTADOR: PROF. Dr. DALTON SOARES ARANTES

CO-ORIENTADOR: PROF. DR. CARLOS R. P. HARTMANN

Tese apresentada à Faculdade de Engenharia de Campinas FEC/UNICAMP, como parte dos requisitos exigidos para obtenção do título de MESTRE EM CIÊNCIA.

AGOSTO 1979

UNICAMP  
BIBLIOTECA CENTRAL

### AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Dalton Soares Arantes, pela atenção e apoio demonstrados;

Ao meu co-orientador, Prof. Dr. Carlos R.P. Hartmann, pela sua incansável dedicação, valiosas contribuições, interesse e amizade dispensadas;

A Universidade Federal do Rio Grande do Norte (UFRN) por ter me proporcionado essa oportunidade;

A Olga Regina S. de Souza e Maria Julia D. Fray pelos excelentes trabalhos de datilografia;

A todos os colegas e amigos que, de uma forma ou de outra, contribuíram para que eu aqui chegasse.

DEDICATÓRIA

À MINHA MÃE,  
MARIA JULIÊTA D. DE FARIA;

À MINHA ESPOSA,  
MARÍLIA V. BEZERRA DE FARIA;

AO MEU PAI,  
JAIME MARIZ DE FARIA.

# ÍNDICE

## SUMÁRIO:

### CAPÍTULO I - INTRODUÇÃO

#### INTRODUÇÃO

### CAPÍTULO II - PRELIMINARES

II.1 - NOTAÇÃO PARA A TEORIA DE PROBABILIDADE UTILIZADA

II.2 - DEFINIÇÃO DE AUTO-INFORMAÇÃO E ENTROPIA

II.3 - CODIFICAÇÃO DE FONTE

II.4 - CÓDIGOS DE HUFFMAN PARA FONTES (CÓDIGOS ÓTIMOS)

II.5 - ALGORITMO DE TESTE

II.6 - ALGORITMO DE TESTE ÓTIMO DE PRIMEIRA ORDEM

### CAPÍTULO III - NOVO LIMITE SUPERIOR

III.1 - INTRODUÇÃO

III.2 - DETERMINAÇÃO DE NOVO LIMITE SUPERIOR

III.3 - PROPRIEDADES DO NOVO LIMITE SUPERIOR

III.4 - COMPARAÇÕES DO NOVO LIMITE SUPERIOR COM O DE MASSEY

### CAPÍTULO IV - CONSTRUÇÃO DE ALGORITMOS EFICIENTES

IV.1 - DEFINIÇÃO

IV.2 - ESTUDO DA COMPLEXIDADE DE CONSTRUÇÃO DO NOVO ALGORITMO

### CAPÍTULO V - CONCLUSÕES

V.1 - COMENTÁRIOS FINAIS

V.2 - SUGESTÃO PARA CONSTRUÇÃO DE ALGORITMOS EFICIENTES

SUGESTÕES PARA PRÓXIMAS PESQUISAS

### APÊNDICE A

### REFERÊNCIAS BIBLIOGRÁFICAS

## ABSTRACT

Recently J. Massey has studied the complexity of certain algorithms as a source coding problem. He has applied some results from Information Theory to this problem and found upper and lower bounds for the measure of the efficiency of these algorithms. These bounds can be easily calculated and also can be used to design, in some cases, efficient algorithms.

In this work we generalize the Massey's procedure. Again, using some results from Information Theory, we obtained new upper bounds on the efficiency of these algorithms which are, in general, more tight than Massey's. With these new upper bound we can construct more efficient algorithms. We have studied the complexity required in the design of these new algorithms and also suggest a method for the design of more efficient ones.

## SUMÁRIO

Recentemente J. Massey abordou o estudo da complexidade de certos tipos de algoritmos como um problema de codificação de fontes. Ele aplicou alguns resultados da Teoria da Informação a esse problema e obteve limites inferiores e superiores para a medida da eficiência de certos tipos de algoritmos. Esses limites são facilmente calculados e possibilitam a construção, em alguns casos, de algoritmos eficientes.

Nesse trabalho nos generalizamos o procedimento introduzido por Massey. Novamente utilizando alguns resultados da Teoria da Informação, estabelecemos novos limites superiores que são menores ou iguais aos por ele estabelecidos. Com a ajuda desses novos limites pudemos então construir algoritmos mais eficientes. Estudamos também a complexidade para essa nova construção, e sugerimos, finalmente, um método para construção de algoritmos ainda mais eficientes.

## CAPÍTULO I

### INTRODUÇÃO

## INTRODUÇÃO:

Começaremos nosso trabalho com um exemplo simples e característico. Sempre que possível, ao longo de seu desenvolvimento, voltaremos a fazer citações e comparações com esse mesmo exemplo de modo que ele nos acompanhe e ajude a focalizar o problema, bem como auxilie no desenvolvimento de nosso raciocínio.

Exemplo 1.1: Uma fábrica produz seis tipos diferentes de resistores ( $u_1, u_2, u_3, u_4, u_5$  e  $u_6$ ), cujas probabilidades de ocorrência estão indicadas na tabela abaixo:

| $u$      | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ |
|----------|-------|-------|-------|-------|-------|-------|
| $P_U(u)$ | 0,10  | 0,10  | 0,30  | 0,20  | 0,20  | 0,10  |

FIG.I.1 - Probabilidades de Ocorrência dos Resistores

Para identificação dos tipos de resistores são dados cinco testes ( $T_1, T_2, T_3, T_4$  e  $T_5$ ). Estes, quando aplicados, produzem resultados binários (0 e 1) conforme a seguinte tabela de decisão:

| $T \backslash u$ | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ |
|------------------|-------|-------|-------|-------|-------|-------|
| $T_1$            | 0     | 0     | 0     | 1     | 1     | 1     |
| $T_2$            | 1     | 0     | 0     | 1     | 1     | 1     |
| $T_3$            | 0     | 1     | 0     | 0     | 0     | 1     |
| $T_4$            | 0     | 1     | 0     | 1     | 0     | 1     |
| $T_5$            | 0     | 1     | 1     | 0     | 1     | 1     |

FIG.I.2 - Tabela de Decisão

Estamos a procura de um algoritmo que unicamente identifique todos os tipos de resistores. Neste caso sabemos que tal algoritmo existe já que todas as colunas da Fig. I.2 são distintas.

Então o diagrama de bloco de um desses algoritmos teria o aspecto da Fig. I.3:

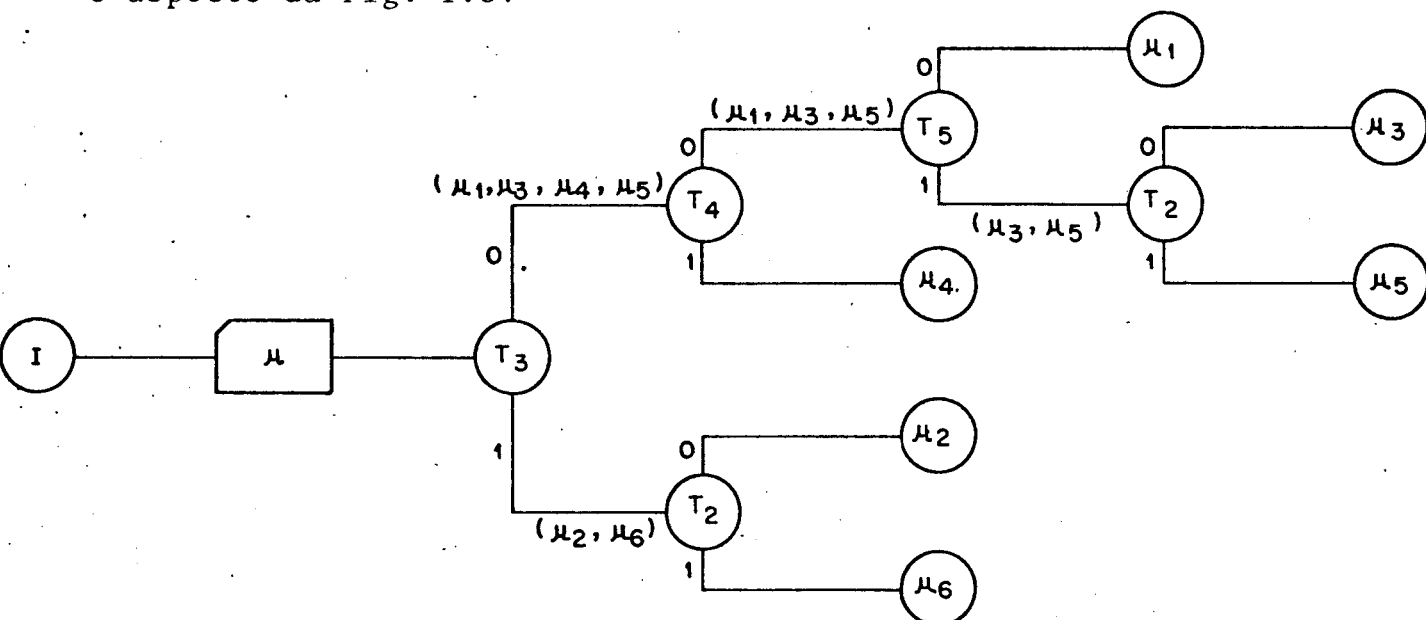


Fig. I.3 - Diagrama de blocos de um algoritmo que unicamente identifica todos os  $\mu_i$ 's.

Podemos associar ao algoritmo acima a seguinte árvore binária:

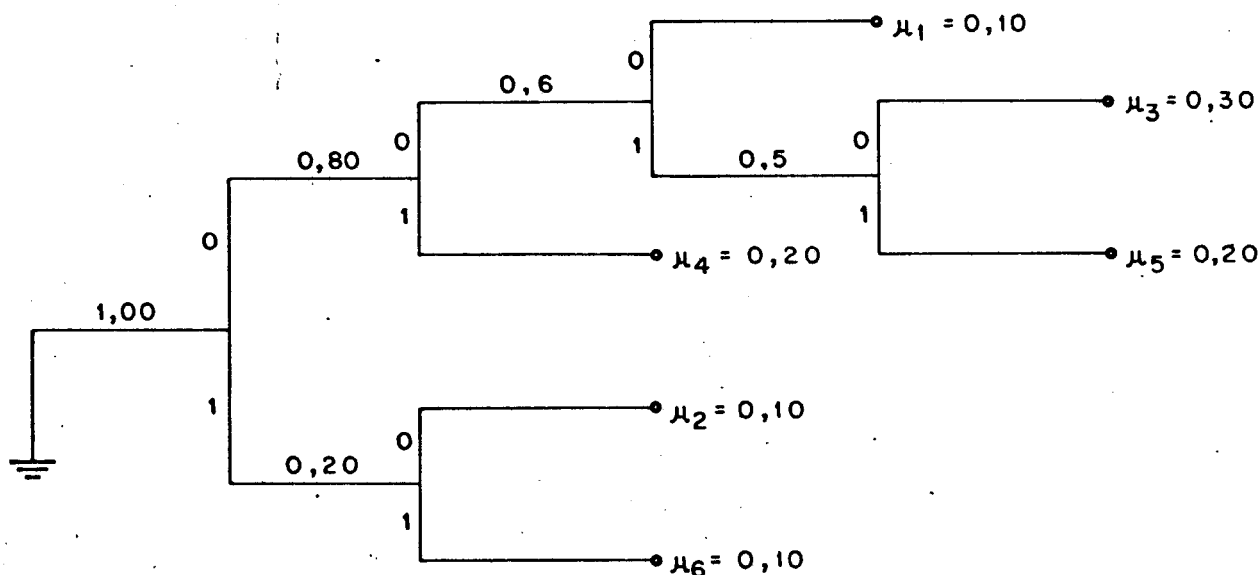


Fig. I.4 - Árvore binária associada ao diagrama de bloco do algoritmo da Fig. I.3.

Correspondente à esta árvore binária temos um código binário:

| $u$   | $P_U(u_i)$ | CÓDIGO 2 |   |   |   |
|-------|------------|----------|---|---|---|
| $u_1$ | 0,10       | 0        | 0 | 0 |   |
| $u_2$ | 0,10       | 1        | 0 |   |   |
| $u_3$ | 0,30       | 0        | 0 | 1 | 0 |
| $u_4$ | 0,20       | 0        | 1 |   |   |
| $u_5$ | 0,20       | 0        | 0 | 1 | 1 |
| $u_6$ | 0,10       | 1        | 1 |   |   |

FIG.I.5 - Código Binário Correspondente à Árvore da Fig.I.4

Podemos calcular então o número médio de testes,  $\bar{W}$ , usados na identificação dos diferentes tipos de resistores, Seja  $w_i$  o número de testes necessários para identificar o resistor  $u_i$ . Então,

$$\bar{W} = \sum_{i=1}^6 w_i P_U(u_i) = 2 (0,10 + 0,20 + 0,10) + 3 \times 0,10 + 4 (0,30 + 0,20) =$$

$$\bar{W} = 3,10 \text{ testes/identificação}$$

OBSERVAÇÃO: Daqui para frente representaremos "testes/identificação" apenas por "t/i".

Os engenheiros estão sempre preocupados com a otimização de problemas desse tipo, isto é, eles estão sempre à procura de algoritmos eficientes para unicamente distinguir as variáveis dadas, usando um conjunto de testes conhecido. No nosso caso o critério de eficiência será medido através do seu número médio de testes por identificação.

Assim sendo, diante do resultado acima obtido levanta-se, principalmente, duas perguntas:

1ª) Quão eficiente é esse algoritmo?

2ª) Pode-se construir um outro ainda mais eficiente?

Uma maneira de responder à primeira pergunta seria encontrar limite inferior para o  $\bar{W}$  e ver quão próximo desse limite está o valor obtido. Com relação à segunda pergunta, uma maneira de se investigar seria a partir da determinação de um limite superior para o  $\bar{W}$ . Aí então se tentaria construir um algoritmo de modo a minimizar esse limite superior.

Notamos que qualquer código associado a uma árvore correspondente a um algoritmo de identificação, como visto acima, apresenta a seguinte propriedade: nenhuma palavra código é prefixo de qualquer outra [4]. Para esse tipo de código a Teoria da Informação nos fornece ferramentas que nos permitem, facilmente, determinar esses limites que precisamos para  $\bar{W}$ .

No Capítulo II daremos noções da Teoria da Informação, idéias e definições necessárias à compreensão do nosso trabalho, como também determinaremos os tais limites para  $\bar{W}$  encontrados por Massey, [4]. No Capítulo III apresentaremos um novo limite superior e o estudo de algumas de suas propriedades. No Capítulo IV usaremos esse novo limite superior para a construção de algoritmos eficientes, além de analisarmos o problema da complexidade de tal construção. Finalmente, no Capítulo V, apresentaremos as conclusões dessa investigação.

## CAPÍTULO II

### PRELIMINARES

## PRELIMINARES:

Nesse Capítulo apresentaremos noções de alguns assuntos básicos para o desenvolvimento do nosso trabalho.

### II.1 - NOTAÇÃO PARA A TEORIA DE PROBABILIDADE UTILIZADA [ 1,3,4 ] :

Definição 2.1 : O espaço de probabilidade discreta,  $U$ , consiste de um espaço amostral  $S$  de eventos elementares e uma medida de probabilidade,  $P_U(u)$ , associada a cada evento elementar  $u$ , satisfazendo as seguintes propriedades:

$$1) P_U(u) \geq 0$$

$$2) \sum_{u \in S} P_U(u) = 1,$$

$$\text{onde } S = \{u_1, \dots, u_K\}$$

Experimento com um resultado:

$$u \in \{u_1, \dots, u_K\}$$

| $u$     | $P_U(u)$   |
|---------|------------|
| $u_1$   | $P_U(u_1)$ |
| $\cdot$ |            |
| $\cdot$ |            |
| $\cdot$ |            |
| $u_K$   | $P_U(u_K)$ |

Experimento com dois resultados:

$$x \in \{a_1, \dots, a_I\}$$

$$y \in \{b_1, \dots, b_J\}$$

$$(x,y) \in \{(a_i, b_j) \mid i = 1, \dots, I \text{ e } j = 1, \dots, J\}$$

| $(x,y)$     | $P_{XY} (x,y)$     |
|-------------|--------------------|
| $(a_1,b_1)$ | $P_{XY} (a_1,b_1)$ |
| $(a_1,b_2)$ | $P_{XY} (a_1,b_2)$ |
| $\vdots$    | $\vdots$           |
| $\vdots$    | $\vdots$           |
| $\vdots$    | $\vdots$           |
| $(a_I,b_J)$ | $P_{XY} (a_I,b_J)$ |

Sabemos que

$$P_X(a_i) = \sum_{j=1}^J P_{XY}(a_i, b_j) \text{ e } P_Y(b_j) = \sum_{i=1}^I P_{XY}(a_i, b_j)$$

Definição 2.2 : Se  $P_X(a_i) > 0$ , a probabilidade condicionada de que o resultado  $y$  seja  $b_j$  dado que o resultado  $x$  é  $a_i$  é definido como:

$$P_{Y|X}(b_j|a_i) = \frac{P_{XY}(a_i, b_j)}{P_X(a_i)}$$

Definição 2.3: A variável aleatória  $X$  é uma função do espaço amostral  $S$  para o alfabeto  $\{x_1, \dots, x_I\}$  onde  $S = \{u_1, \dots, u_K\}$ .

Definição 2.4:  $\{X = x_i\} = \{u \mid u \in S \text{ e } X(u) = x_i\}$

e

$$P_U(X = x_i) = P_U(\{u \mid u \in S \text{ e } X(u) = x_i\})$$

Definição 2.5: A distribuição de probabilidades  $P_X$  para  $X$  é um mapeamento de  $\{x_1, \dots, x_I\}$  para os números reais, tal que:

$$P_X(x) = P_U(X=x) = P_U(\{u \mid u \in S \text{ e } X(u) = x\})$$

## II.2 - DEFINIÇÃO DE AUTO INFORMAÇÃO E ENTROPIA:

Definição 2.6: A Auto-Informação do evento  $x$  é definida como

$$I(x) = -\log_a P_X(x)$$

isto é, seja o evento  $x$  que tem uma probabilidade de ocorrência  $P_X(x)$ . Se soubermos que tal evento ocorreu teremos então recebido  $I(x)$  unidades de informação.

Essa unidade de informação recebe nomes distintos dependendo da base "a" do logaritmo da definição. Se

$a = 2 \Rightarrow I(x)$  será dado em "bits"

$a = e \Rightarrow I(x)$  será dado em "nats"

$a = 10 \Rightarrow I(x)$  será dado em "Hartley"

onde  $1 \text{ Hartley} = 3.32 \text{ bits}$

$1 \text{ nat} = 1.44 \text{ bits}$

OBSERVAÇÃO: Daqui para frente não mais especificaremos o valor de "a" representando a base do logaritmo. Isto é, teremos  $\log_a x = \log x$ .

Estudando-se a faixa de variação do  $I(x)$  podemos concluir que:

$$0 \leq I(x) \leq \infty$$

$I(x) = 0 \Rightarrow P_X(x) = 1$ , isto é, a ocorrência do evento  $x$  não nos trouxe nenhuma informação, pois já tínhamos certeza de tal ocorrência.

$I(x) = \infty \Rightarrow P_X(x) = 0$ , ou seja, a ocorrência do evento  $x$  nos trouxe uma quantidade infinita de informação, já que não a esperávamos de modo algum.

Definição 2.7 : A ENTROPIA (Auto-Informação média ou Incerteza Média) da variável aleatória  $X$  que tem valores ....  $\{x_1, \dots, x_I\}$  é definida como:

$$H(X) = \overline{I(x)} \triangleq - \sum_{i=1}^I P_X(x_i) \log P_X(x_i)$$

TEOREMA 2.1[4]: Se  $X$  toma valores sobre o alfabeto  $\{x_1, \dots, x_I\}$ , então:

$$0 \leq H(X) \leq \log I,$$

com igualdade do lado esquerdo se, e somente se, existe um "i" tal que  $P_X(x_i) = 1$ , e igualdade do lado direito se, e somente se,  $P_X(x_i) = 1/I$ , para  $i = 1, \dots, I$ .

Um exemplo simples e bastante frequente de Entropia é quando a variável aleatória  $X$  assume apenas dois possíveis valores,  $x_1$  e  $x_2$  (caso binário). Assim temos  $P_X(x_1) = p$  e  $P_X(x_2) = 1-p$  e a expressão de sua entropia será:

$$H(X) = -p \log_2 p - (1-p) \log_2 (1-p)$$

Essa expressão aparece tão frequentemente que foi dado o nome especial para ela de "Função de Entropia Binária", e representada por  $h(p)$ . Assim:

$$h(p) = -p \log_2 p - (1-p) \log_2 (1-p),$$

cujo gráfico está representado na figura abaixo:

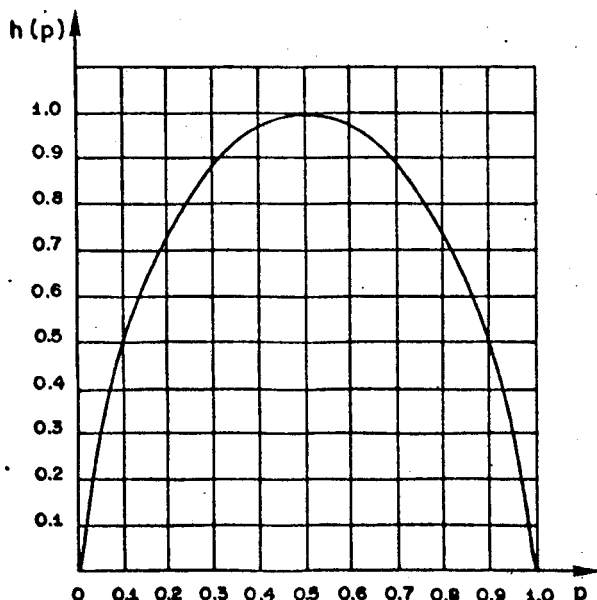


Fig. II.4

Gráfico da Entropia Binária

No apêndice A daremos uma tabela com inúmeros valores da função.

Definição 2.8: A ENTROPIA CONDICIONAL (ou Incerteza Condicional) de uma variável aleatória  $X$  dado que o evento  $Y = y_j$  ocorreu, é dada pela expressão:

$$H(X|Y = y_j) = - \sum_{i=1}^I P_{X|Y}(x_i|y_j) \log P_{X|Y}(x_i|y_j)$$

TEOREMA 2.2[4]: Se  $P_Y(y_j) \neq 0$  e  $P_{X|Y}(x_i|y_j)$  existir, então

$$0 \leq H(X|Y = y_j) \leq \log I$$

com igualdade do lado esquerdo se, e somente se, existir algum "i" tal que  $P_{X|Y}(x_i|y_j) = 1$  e, contrariamente, com igualdade do lado direito se, e somente se,  $P_{X|Y}(x_i|y_j) = 1/I$  para todo "i".

Definição 2.9: A ENTROPIA CONDICIONAL MÉDIA (Incerteza Condicional Média) da variável aleatória  $X$ , dada a variável aleatória  $Y$ , é dada pela média sobre as entropias condicionais:

$$H(X|Y) = - \sum_{i=1}^I \sum_{j=1}^J P_{XY}(x_i, y_j) \log P_{X|Y}(x_i|y_j)$$

TEOREMA 2.3,[4]:

$$0 \leq H(X|Y) \leq \log I$$

com igualdade do lado esquerdo se, e somente se, para cada  $j$  tal que  $P_Y(y_j) \neq 0$ , existe algum  $i$  tal que  $P_{X|Y}(x_i|y_j) = 1$ , e com igualdade do lado direito se, e somente se, para todo  $j$  tal que  $P_Y(y_j) \neq 0$ ,  $P_{X|Y}(x_i|y_j) = 1/I$  para todo  $i$ .

TEOREMA 2.4[4]:

$$H(X|Y) = \sum_{j=1}^J H(X|Y = y_j) P_Y(y_j)$$

## II.3 - CODIFICAÇÃO DE FONTE



Fig. II.2 - Esquema de uma Codificação de Fonte

Conceitos Fundamentais:

- (1)  $U$  é uma variável aleatória com alfabeto  $\{u_1, \dots, u_k\}$ .
- (2) Cada  $X_i$  assume letras de um alfabeto D-ário que usualmente serão  $\{0, 1, \dots, D-1\}$ .
- (3)  $W$  é uma variável aleatória, isto é, os valores da variável aleatória  $Z$  são vetores D-ários de comprimento variável.

O comprimento médio de um código de comprimento variável  $[2, 3, 4]$ ,  $\bar{W}$ , será calculado pela expressão:

$$\bar{W} = \sum w_i P_U(u_i),$$

onde  $w_i$  é o comprimento da palavra código  $Z_i = \{x_{i1} x_{i2} \dots x_{iw_i}\}$  associada a  $u_i$ . Salientamos novamente aqui que esse comprimento médio,  $\bar{W}$ , será o parâmetro que nos indicará quão eficiente um determinado código é.

Definição 2.10: Um código D-ário com Propriedade de Prefixo é aquele no qual uma palavra código não é prefixo de nenhuma outra.

OBSERVAÇÃO: Daqui para frente abreviaremos "Código com Propriedade de Prefixo" apenas por "Códigos com PP".

Exemplo 2.1 : Seja o código binário

| $u$   | Código $Z$ |
|-------|------------|
| $u_1$ | 0          |
| $u_2$ | 1 0        |
| $u_3$ | 1 1 0      |
| $u_4$ | 1 1 1      |

Esse código é um Código com PP.

Podemos ver que existe uma correspondência biunívoca entre um código D-ário com PP e uma árvore D-ária, e, inversamente, qualquer conjunto de folhas de uma árvore D-ária define um código D-ário com PP, [3,4].

Exemplo 2.2: A árvore binária do Exemplo 1.1 tem o aspecto da figura abaixo:

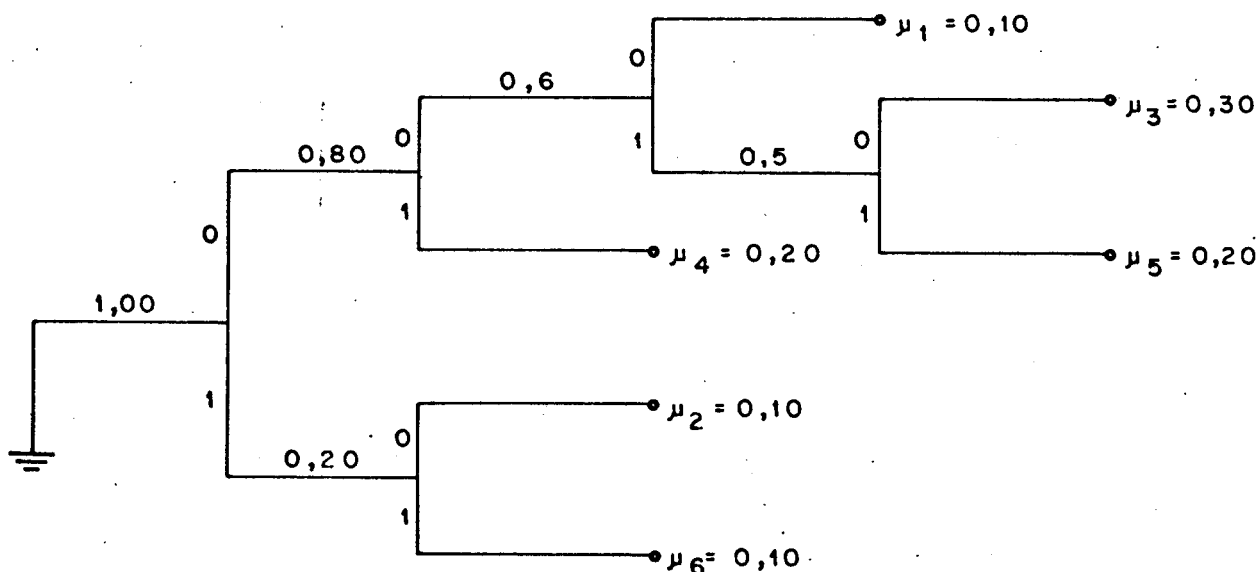


Fig. II.3 - Árvore binária correspondente ao Exemplo 1.1.

Atribuindo-se "0" ao ramo que parte subindo de um nó e "1" ao que parte descendo, temos então um código associado a essa árvore, Fig. II.4. Notemos que esse código é um código com PP:

| $P_U(u_i)$ | $u$  | CÓDIGO Z |       |       |       |
|------------|------|----------|-------|-------|-------|
|            |      | $X_1$    | $X_2$ | $X_3$ | $X_4$ |
| $u_1$      | 0.10 | 0        | 0     | 0     |       |
| $u_2$      | 0.10 | 1        | 0     |       |       |
| $u_3$      | 0.30 | 0        | 0     | 1     | 0     |
| $u_4$      | 0.20 | 0        | 1     |       |       |
| $u_5$      | 0.20 | 0        | 0     | 1     | 1     |
| $u_6$      | 0.10 | 1        | 1     |       |       |

FIG.II.4 - Código Binário com PP Associado a Árvore Binária da Fig. II.3

Chamemos a primeira coluna do código acima de  $X_1$ , a segunda de  $X_2, \dots$ , e a última de  $X_N$ , onde  $N = \max_i w_i$  (para o código da Fig.II.4 temos  $N = 4$ ). O código Z pode então ser escrito como:  $Z = X_1 \dots X_N$ , onde nem todas as palavras códigos têm necessariamente comprimento N. Observemos que  $X_i$  não é, obrigatoriamente, uma variável aleatória, pois nem todos os  $X_i$  assumem valores para todos os  $u_i$ 's. Para tornar  $X_i$  uma variável aleatória teremos que fazer um "truque".

**"Truque"**: Transforme  $Z = X_1 \dots Y_N$  em  $Z' = X'_1 \dots X'_N$ , onde todas as palavras código agora têm comprimento N, já que  $Z'$  é obtido a partir de Z completando-se com zeros as palavras código de comprimento menor que N, conforme Fig.II.5.

É importante salientar que esse "truque" transforma  $X_i$  em  $X'_i$ , que agora será uma variável aleatória. Daqui para frente sempre o aplicaremos para que tenhamos  $X'_1 \dots X'_N$  como variáveis aleatórias.

"TRUQUE"

| CÓDIGO Z |       |         |       | CÓDIGO Z' |        |         |        |
|----------|-------|---------|-------|-----------|--------|---------|--------|
| $x_1$    | $x_2$ | $\dots$ | $x_N$ | $x'_1$    | $x'_2$ | $\dots$ | $x'_N$ |
| 0        | 0     | 0       |       | 0         | 0      | 0       | 0      |
| 1        | 0     |         |       | 1         | 0      | 0       | 0      |
| 0        | 0     | 1       | 0     | 0         | 0      | 1       | 0      |
| 0        | 1     |         |       | 0         | 1      | 0       | 0      |
| 0        | 0     | 1       | 1     | 0         | 0      | 1       | 1      |
| 1        | 1     |         |       | 1         | 1      | 0       | 0      |

FIG.II.5 - Códigos Z e Z' do Exemplo 2.2

Definição 2.11: Seja  $Q_i$  o conjunto de prefixos de palavras código de comprimento maior ou igual a  $i + 1$  para o código D-ário original Z, isto é,

$$Q_i = \{x_1, \dots, x_i \mid x_1, \dots, x_i \text{ é o prefixo de uma palavra código de comprimento maior ou igual a } i + 1 \text{ do código D-ário original Z}\}$$

Exemplo 2.3 : Seja o código binário com PP do Exemplo 2.1:

|       |          | CÓDIGO Z |       |       |                |
|-------|----------|----------|-------|-------|----------------|
| $u$   | $P_U(u)$ | $x_1$    | $x_2$ | $x_3$ |                |
| $u_1$ | 1/2      | 0        |       |       | $Q_1 = \{1\}$  |
| $u_2$ | 1/4      | 1        | 0     |       |                |
| $u_3$ | 1/8      | 1        | 1     | 0     | $Q_2 = \{11\}$ |
| $u_4$ | 1/8      | 1        | 1     | 1     |                |

Definição 2.12: Definiremos  $H_{\min}$  como:

$$H_{\min} = \min_{0 \leq i < N} \min_{x_1 \dots x_i \in Q_i} H(x'_{i+1} \mid x'_1 \dots x'_i = x_1 \dots x_i)$$

onde

$$H(X'_{i+1} | X'_1 \dots X'_i = x_1 \dots x_i) = H(X'_1), \text{ para } i = 0.$$

TEOREMA 2.3,[4] : Para qualquer código com PP, seu comprimento médio está limitado por um valor superior e outro inferior dados por:

$$\frac{H(U)}{\log D} \leq \bar{W} \leq \frac{H(U)}{H_{\min}} \quad (2.1)$$

Exemplo 2.4 : Tomemos novamente o Exemplo 1.1 e calculemos seu limite superior. Antes disso, porém, precisamos dos valores das entropias em cada ponto. Do código da Fig. 1.5 temos

$$Q_1 = \{0,1\}, Q_2 = \{00\} \text{ e } Q_3 = \{001\}$$

Considerando então todos os logaritmos abaixo na base dois temos:

$$H(X'_1) = -0,20 \log 0,20 - 0,80 \log 0,80 = 0,7219 \text{ bits}$$

$$H(X'_2 | X'_1 = 0) = - \frac{0,20}{0,80} \log \frac{0,20}{0,80} - \frac{0,60}{0,80} \log \frac{0,60}{0,80} = 0,8113 \text{ bits}$$

$$H(X'_2 | X'_1 = 1) = -0,50 \log 0,50 - 0,50 \log 0,50 = 1,00 \text{ bit}$$

$$H(X'_3 | X'_1 X'_2 = 00) = - \frac{0,10}{0,60} \log \frac{0,10}{0,60} - \frac{0,50}{0,60} \log \frac{0,50}{0,60} = 0,650 \text{ bits}$$

$$H(X'_4 | X'_1 X'_2 X'_3 = 001) = - \frac{0,20}{0,50} \log \frac{0,20}{0,50} - \frac{0,30}{0,50} \log \frac{0,30}{0,50} = 0,9710 \text{ bits}$$

Então,

$$H_{\min} = \min \{ H(X'_1), H(X'_2 | X'_1 = 0), H(X'_2 | X'_1 = 1), H(X'_3 | X'_1 X'_2 = 00), \\ H(X'_4 | X'_1 X'_2 X'_3 = 001) \} =$$

$$H_{\min} = \min \{0,7219; 0,8113; 1,00; 0,65; 0,9710\}$$

$$H_{\min} = 0,650 \text{ bits}$$

Calculando-se a entropia,  $H(U)$  em bits, teremos:

$$\begin{aligned} H(U) &= -3 (0,10 \log 0,10) - 2 (0,20 \log 0,20) - 0,3 \log 0,3 = \\ &= 3 \times 0,3322 + 2 \times 0,4644 + 0,5211 \end{aligned}$$

$$H(U) = 2,4465 \text{ bits}$$

Caculemos agora o limite superior:

$$\frac{H(U)}{H_{\min}} = \frac{2,4465}{0,650} = 3,7638$$

Então, a Equação (2,1) estabelece os limites para o comprimento médio desse código:

$$2,4465 \leq \bar{W} \leq 3,7638$$

onde o valor obtido foi  $\bar{W} = 3,10 \text{ t/i}$ .

#### II.4 - CÓDIGOS DE HUFFMAN PARA FONTES (CÓDIGOS ÓTIMOS):

HUFFMAN [4,6] desenvolveu um algoritmo que nos conduz a códigos ótimos D-ários com PP. A idéia básica dessa construção é associar às palavras código de menor comprimento, variáveis aleatórias com maiores probabilidades de ocorrência.

Para uma variável aleatória  $U$  com  $S = \{u_1, \dots, u_K\}$  o método de construção de um código D-ário de Huffman, onde  $K > D$ , pode ser enunciado da seguinte maneira:

- 1) Conhecido os  $K$  nós terminais da árvore D-ária  $(u_1, \dots, u_K)$ , com suas respectivas probabilidades  $[P_U(u_1), \dots, P_U(u_K)]$ , cacule o número deles que não serão usados através da expressão:

$$r = R_{D-1} \left[ (K - D) (D - 2) \right]$$

onde  $r$  é o valor a ser calculado e  $R_A [B]$  é o resto do Algoritmo de Euclides da divisão de  $B$  por  $A$ .

OBSERVAÇÃO: Notemos que para o caso em que  $D = 2$  (árvore binária) se calcularmos esse número teremos sempre  $r = 0$ .

- 2) Una os  $D-r$  nós terminais menos prováveis (dos  $K$  existentes) juntamente com os  $r$  não usados (acima calculados) num único nó da árvore  $D$ -ária, e lhe atribua a probabilidade igual a soma de todas as probabilidades dos nós terminais unidos, considerando nula a dos  $r$  não usados. O novo conjunto de nós será formado por todos os nós não utilizados juntamente com o novo nó formado acima.

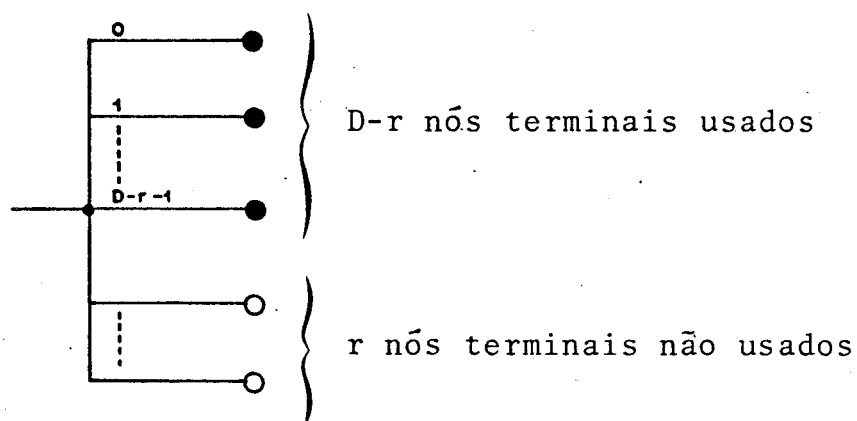


Fig. II.6 - Construção da árvore  $D$ -ária associada a um código de Huffman.

- 3) Dos novos nós (terminais ou não) tome novamente os  $D$  deles menos prováveis e una-os num único com probabilidade igual à soma de todos eles.
- 4) Repita o item anterior até tomar os  $D$  últimos nós, que se reduzirão a um único. A soma de suas probabilidades terá, obrigatoriamente, que ser "1".
- 5) Depois da árvore  $D$ -ária construída, seu código com PP correspondente é obtido atribuindo-se, em todos os nós da árvore, "0" para a primeira ramificação, "1" para a segunda, ..., e " $D - 1$ " para a última, ou outra convenção qualquer equivalente.

Notemos aqui que existem vários códigos de HUFFMAN associados à mesma árvore  $D$ -ária, dependendo da convenção adotada

para a denominação dos D ramos que partem de cada nó.

Chamando de  $\bar{W}_{HUF}$  o comprimento médio desses códigos, temos:

$$\frac{H(U)}{\log D} \leq \bar{W}_{HUF} \leq \bar{W}$$

Exemplo 2.5: Seja a variável aleatória U com as probabilidades

| $u$      | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ |
|----------|-------|-------|-------|-------|-------|-------|
| $P_U(u)$ | 0,27  | 0,23  | 0,20  | 0,15  | 0,10  | 0,05  |

Para  $D = 3$  (árvore ternária) calculemos o número de nós terminais não usados:

$$r = R_{D-1} \left[ (K - D)(D - 2) \right] = R_{3-1} \left[ (6 - 3)(3 - 2) \right]$$

$$R = R_2(3)$$

A expressão acima nos diz que r será o resto da divisão de 3 por 2. Assim  $r = 1$  e sua árvore correspondente será:

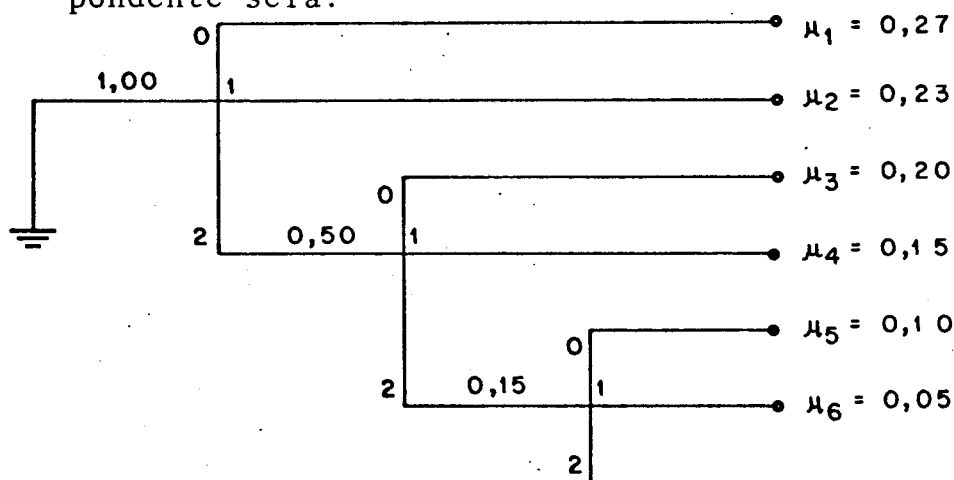


Fig. II.7 - Árvore ternária de um código de Huffman correspondente ao espaço de probabilidades acima representado.

O Código de HUFFMAN correspondente a árvore acima será então:

| $u$   | $P_U(u)$ | Código Z |
|-------|----------|----------|
| $u_1$ | 0,27     | 0        |
| $u_2$ | 0,23     | 1        |
| $u_3$ | 0,20     | 2 0      |
| $u_4$ | 0,15     | 2 1      |
| $u_5$ | 0,10     | 2 2 0    |
| $u_6$ | 0,05     | 2 2 1    |

FIG. II.8 - Código Ternário com PP Associado a árvore da Fig. II.7

Exemplo 2.6: Consideremos novamente o Exemplo 1.1. A árvore binária de um dos Códigos de HUFFMAN binário associada ao espaço de probabilidade da Fig. I.1 será:

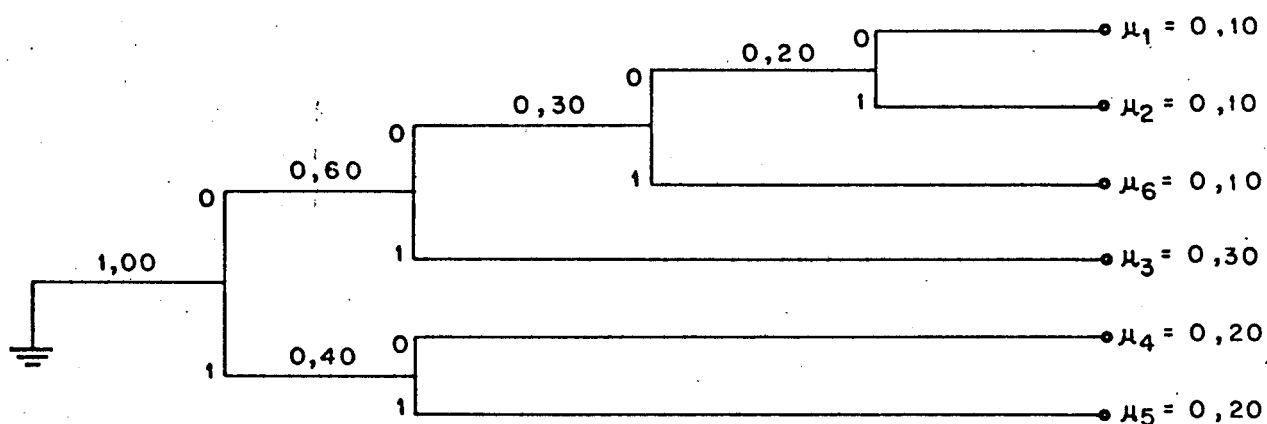


Fig. II.9 - Árvore binária de um código de Huffman correspondente ao espaço de probabilidade do Exemplo 1.1.

O código correspondente à árvore binária acima será:

| $u$   | $P_U(u)$ | Código Z |
|-------|----------|----------|
| $u_1$ | 0,10     | 0 0 0 0  |
| $u_2$ | 0,10     | 0 0 0 1  |
| $u_3$ | 0,30     | 0 1      |
| $u_4$ | 0,20     | 1 0      |
| $u_5$ | 0,20     | 1 1      |
| $u_6$ | 0,10     | 0 0 1    |

FIG.II.10 - Código Binário Associado à Árvore da FIG.II.9

Calculando-se o seu  $\bar{W}$  teremos:

$$\bar{W}_{HUF} = 2 (0,30 + 0,20 + 0,20) + (3 \times 0,10) + 4(0,10 + 0,10)$$

$$\bar{W}_{HUF} = 2,50 \text{ t/i}$$

## II.5 - ALGORITMO DE TESTE:

Definição 2.13: Um algoritmo de teste para a variável aleatória  $U$  é um algoritmo que unicamente identifica todos os  $u_i$ 's.

Daqui para frente consideraremos que o resultado dos testes podem assumir  $D$  valores diferentes. Como ressaltamos anteriormente, todo código associado a um algoritmo de teste é um código com PP. Logo, para esse código temos:

$$\bar{W}_{HUF} < \bar{W} < \frac{H(U)}{H_{\min}}$$

Entretanto, na maioria dos casos teremos

$$\bar{W} > \bar{W}_{HUF} ,$$

porque não teremos os testes necessários para a construção do Algoritmo de HUFFMAN.

Exemplo 2.7 : Consideremos um problema análogo ao Exemplo 1.1, no qual são dados cinco testes ( $T_1, \dots, T_5$ ) que deverão ser usados na identificação de seis elementos ( $u_1, \dots, u_6$ ), conforme a tabela de decisão:

| $P_U(u)$         | 0,30  | 0,22  | 0,20  | 0,18  | 0,05  | 0,05  |
|------------------|-------|-------|-------|-------|-------|-------|
| $T \backslash u$ | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ |
| $T_1$            | 1     | 0     | 1     | 0     | 1     | 0     |
| $T_2$            | 1     | 1     | 1     | 0     | 0     | 0     |
| $T_3$            | 0     | 0     | 0     | 1     | 0     | 1     |
| $T_4$            | 0     | 0     | 1     | 1     | 0     | 1     |
| $T_5$            | 1     | 0     | 0     | 1     | 0     | 0     |

FIG. II.11 - Tabela de Decisão

A árvore binária correspondente a um dos códigos de HUFFMAN associado ao espaço de probabilidades acima será:

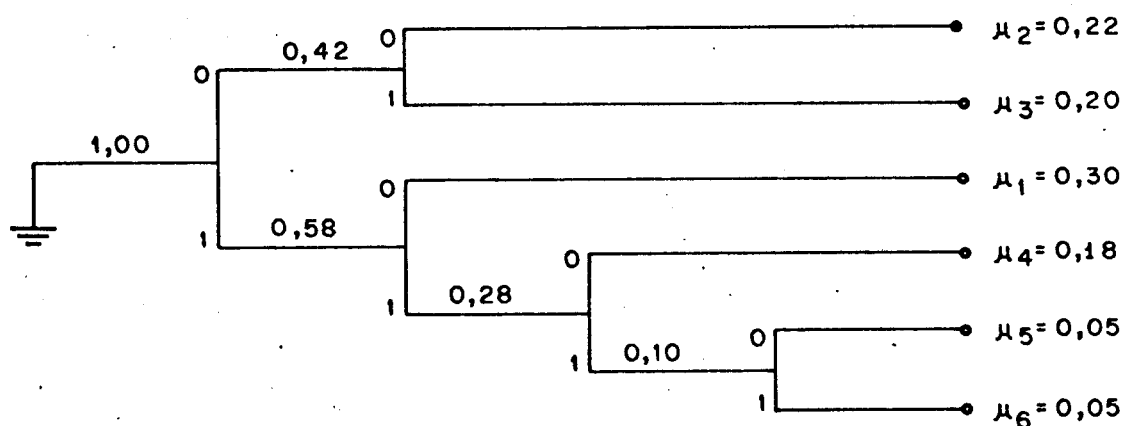


Fig. II.12 - Árvore binária de um dos códigos de Huffman associada ao espaço de probabilidade dado.

Qualquer código de HUFFMAN associado a essa árvore acima será ótimo. Entretanto, como podemos observar, a implementação de qualquer um deles é impossível já que necessitaríamos de um teste que, no primeiro passo, separasse ( $u_2$  e  $u_3$ ) do restante, e não dispomos de tal teste.

## II.6 - ALGORITMO DE TESTES ÓTIMOS DE PRIMEIRA ORDEM

Começaremos o estudo desse algoritmo novamente procurando ilustrar com o mesmo Exemplo 1.1 considerado.

Suponhamos que, ao invés de cinco, tenhamos sete testes ( $T_1, \dots, T_7$ ), cada um apresentando resultados binários usados para identificação dos resistores. Os resultados de cada teste para cada tipo de resistor são dados na tabela abaixo:

| $T \backslash u$ | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ |
|------------------|-------|-------|-------|-------|-------|-------|
| $T_1$            | 0     | 0     | 0     | 1     | 1     | 1     |
| $T_2$            | 0     | 0     | 0     | 0     | 0     | 0     |
| $T_3$            | 1     | 0     | 0     | 1     | 1     | 1     |
| $T_4$            | 0     | 1     | 0     | 0     | 0     | 1     |
| $T_5$            | 0     | 1     | 0     | 1     | 0     | 1     |
| $T_6$            | 1     | 0     | 1     | 0     | 1     | 0     |
| $T_7$            | 0     | 1     | 1     | 0     | 1     | 1     |

FIG.II.13 - Tabela de Decisão

Analisando a tabela acima podemos tirar algumas conclusões:

- 1) Será sempre possível se identificar todos os elementos se, e somente se, todas as colunas da tabela de decisão forem distintas. Isso porque caso haja duas colunas idênticas,  $u_i$  e  $u_j$ , nenhum teste irá distinguí-las.

- 2) Um teste que apresente sempre o mesmo resultado para todos os  $u_i$ 's é inútil e pode ser eliminado da tabela (teste  $T_2$  da Tabela da Fig.II.13).
- 3) Um teste é dito "essencial" se, quando retirado, faz com que duas colunas se tornem idênticas. Esse tipo é imprescindível à tabela.
- 4) Dois testes são "equivalentes" se apresentam resultados idênticos ou complementares. Assim um deles poderá ser eliminado sem prejuízo ( $T_5$  e  $T_6$  da Fig. II.13 são equivalentes).

Assim, depois de eliminados da tabela da Fig. II.13 todos os testes dispensáveis teremos então a tabela da Fig. I.1.

Daqui para frente consideraremos sempre tabelas de decisão que não tenham testes dispensáveis.

No exemplo 1.1 construímos um algoritmo de teste. Correspondente a ele construímos sua árvore binária, cujo comprimento médio do código com PP a ela associado é 3,10 t/i. Depois, nas seções II.3 e II.4 fixamos limites (superior e inferior) para o  $\bar{W}$  que, calculados para o Exemplo 1.1, encontramos:

$$\bar{W}_{HUF} = 2,50 < \bar{W} = 3,10 < \frac{H(U)}{H_{\min}} = 3,7638 \text{ bits}$$

Analisando o limite superior acima Massey [4] desenvolveu um algoritmo que o chamou de "Algoritmo de Teste Ótimo de Primeira Ordem" (daqui para frente abreviado apenas por "Algoritmo TOPO"). Basicamente ele é construído de maneira a maximizar o  $H_{\min}$ , tornando então o mínimo possível o limite superior  $H(U) / H_{\min}$ . Assim, se em cada ponto da construção de um algoritmo escolhermos um teste que maximize a entropia a esse ponto associada estaremos minimizando o limite superior do seu código correspondente.

Exemplo 2.8: Procuraremos então aplicar essa idéia de construção ao Exemplo 1.1. Para a escolha do primeiro teste a ser aplicado teremos que analisar todos os possíveis valores:

| 1º Teste | $H(X_1)$  |
|----------|-----------|
| $T_1$    | $h(0,50)$ |
| $T_2$    | $h(0,40)$ |
| $T_3$    | $h(0,20)$ |
| $T_4$    | $h(0,40)$ |
| $T_5$    | $h(0,30)$ |

FIG.II.14.- Pesquisa para a escolha do primeiro teste a ser aplicado na construção do Algoritmo TOPO.

Como a entropia do  $T_1$  foi a maior de todas, será ele o primeiro teste aplicado:

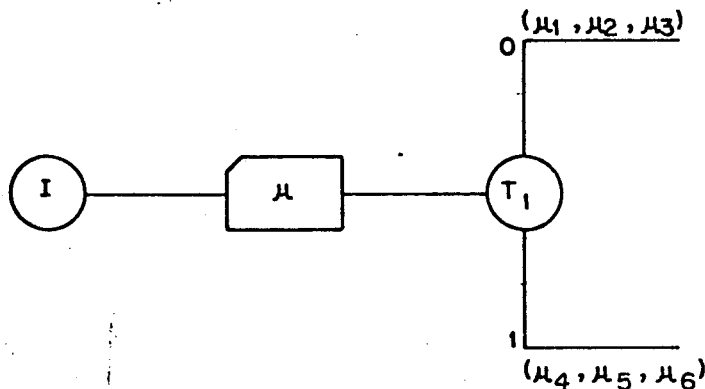


Fig. II.15 - Início da construção do Algoritmo TOPO.

Novamente calcularemos as entropias para cada resultado do primeiro teste, para que se tome os testes que maximizem-nas:

| 2ºs Testes | $H(X_2   X_1=0)$ | $H(X_2   X_1=1)$ |
|------------|------------------|------------------|
| $T_2$      | $h(0,20)$        | $h(0)$           |
| $T_3$      | $h(0,20)$        | $h(0,20)$        |
| $T_4$      | $h(0,20)$        | $h(0,40)$        |
| $T_5$      | $h(0,20)$        | $h(0,40)$        |

FIG.II.16 - Escolha dos segundos testes a serem aplicados na construção do Algoritmo TOPO.

Como na primeira coluna acima todas as entropias têm o mesmo valor, qualquer escolha nos conduzirá a resultados idênticos. Tomemos  $T_5$  arbitrariamente. Na segunda coluna temos dois valores iguais. Tomemos, novamente ao acaso, um deles,  $T_5$  por exemplo. Assim teremos:

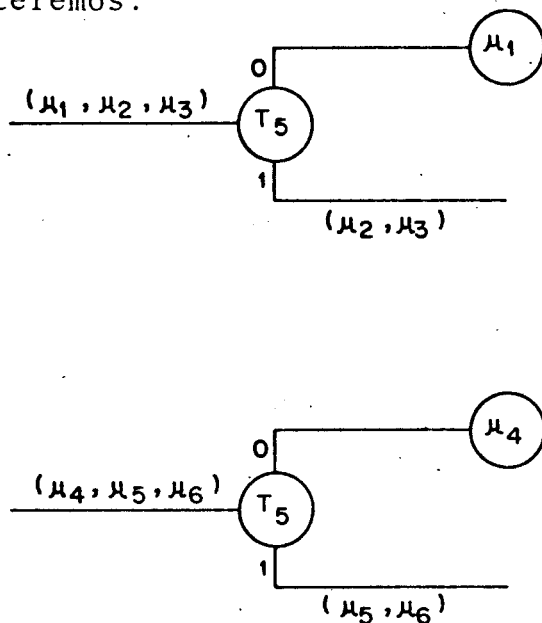


Fig. II.17 - Segundo passo da construção do Algoritmo TOPO

Procedendo analogamente para  $H(X_3|X_1X_2=01)$  e  $H(X_3|X_1X_2=11)$  teremos o diagrama de bloco do Algoritmo TOPO com o seguinte aspecto:

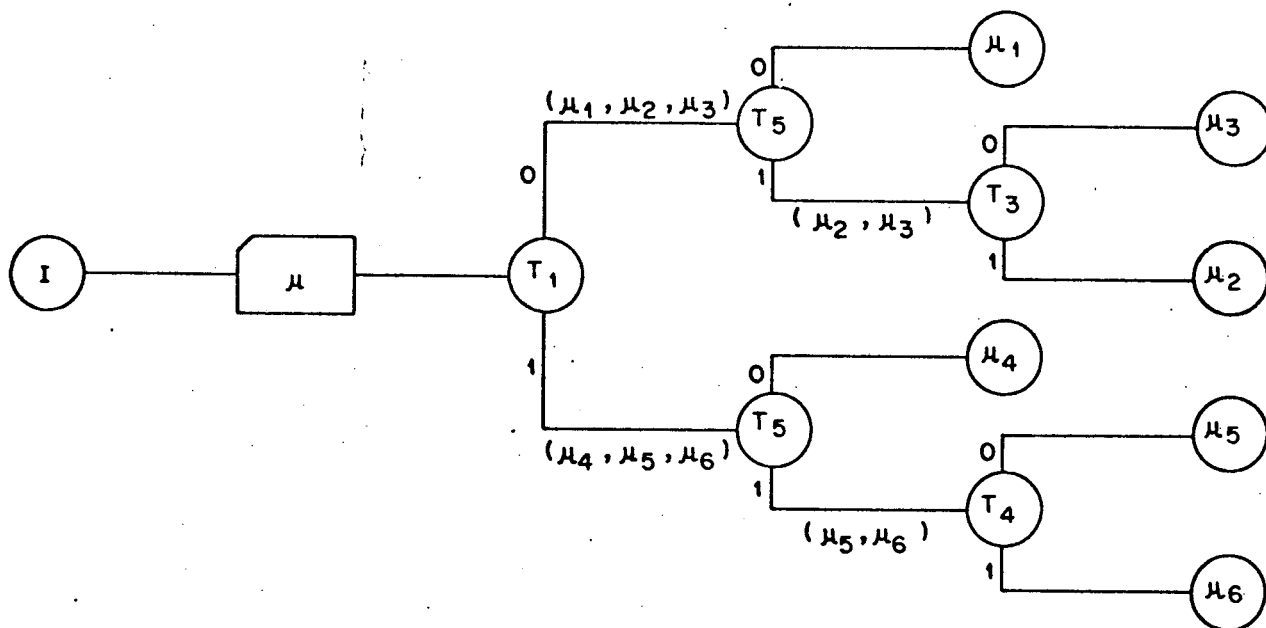


Fig. II.18 - Diagrama de bloco do Algoritmo TOPO de Massey

sua árvore binária correspondente será:

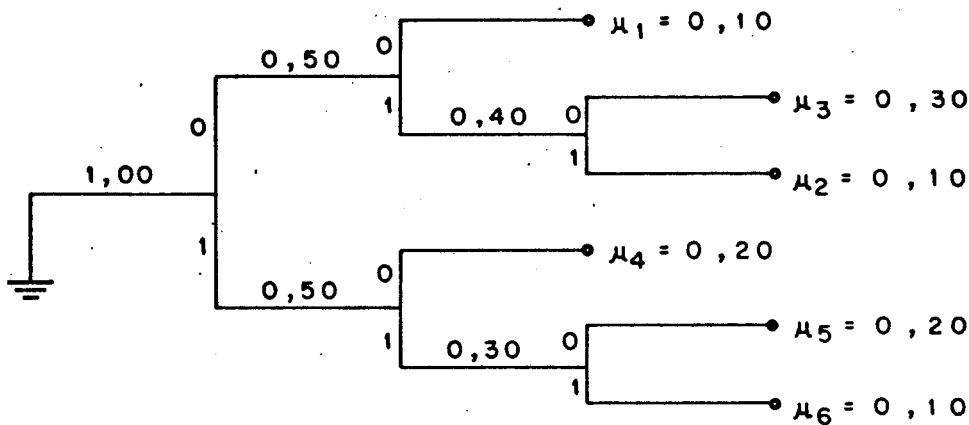


Fig. II.19 - Árvore binária associada ao Algoritmo TOPO da Fig. II.18.

Assoçado a essa árvore binária temos um código binário com PP:

| $u$   | $P_U(u)$ | Código Z |
|-------|----------|----------|
| $u_1$ | 0,10     | 0 0      |
| $u_2$ | 0,10     | 0 1 1    |
| $u_3$ | 0,30     | 0 1 0    |
| $u_4$ | 0,20     | 1 0      |
| $u_5$ | 0,20     | 1 1 0    |
| $u_6$ | 0,10     | 1 1 1    |

FIG.II.20 - Código com PP associado à árvore binária da Fig. II.19

Calculando-se seu comprimento médio correspondente teremos:

$$\bar{W} = 2(0,10 + 0,20) + 3(0,10 + 0,30 + 0,20 + 0,10)$$
$$\bar{W} = 2,70 \text{ t/i}$$

Observemos que esse valor é menor do que o  $\bar{W}$  do algoritmo do Exemplo 1.1 (3.10 t/i). Entretanto, poderia ser mais próximo do limite inferior,  $\bar{W}_{HUF} = 2.50$  t/i, conforme o comprimento médio do algoritmo de teste apresentado a seguir:

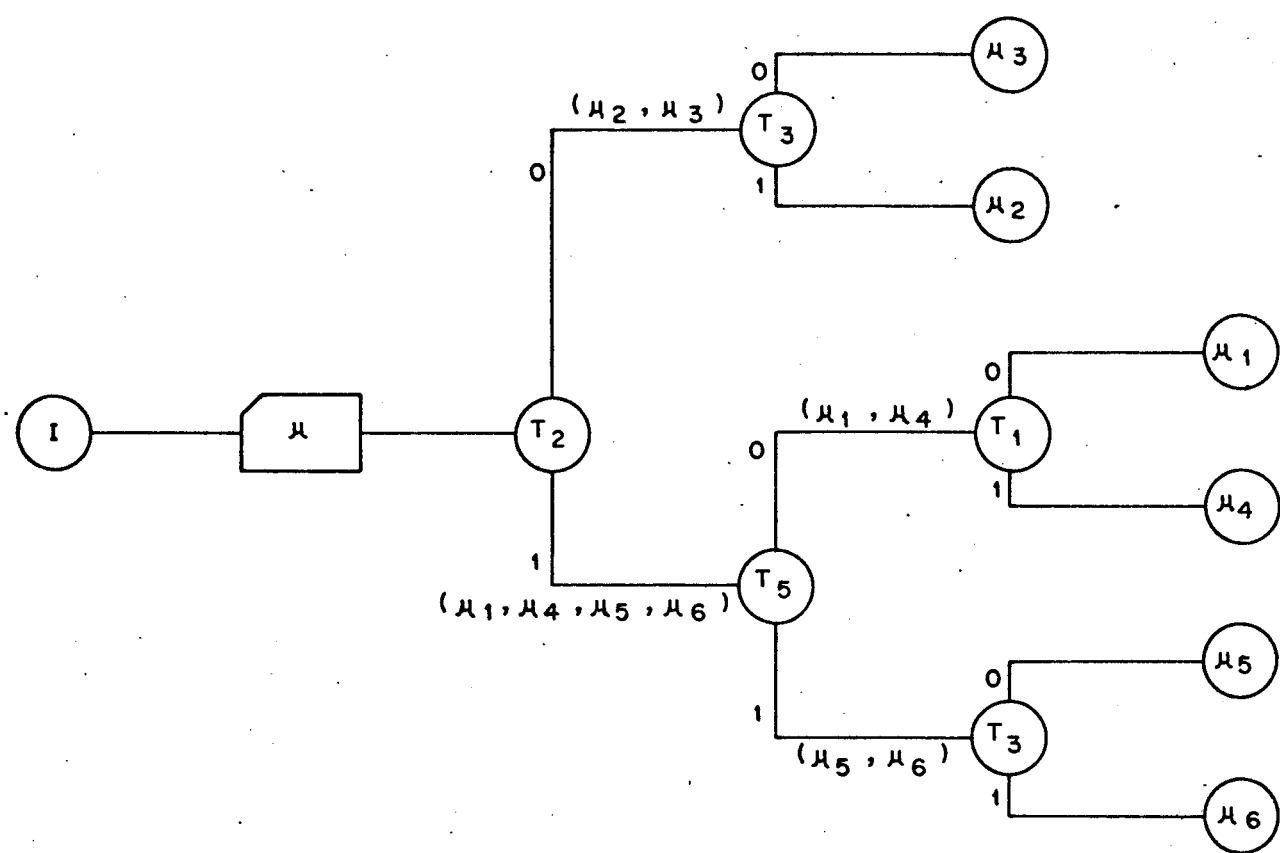


Fig. II.21 - Código com PP associado à árvore binária da Fig.II.19.

De sua árvore binária associada podemos ter o seu código binário correspondente:

| $u$   | $P_U(u)$ | Código Z |   |   |
|-------|----------|----------|---|---|
| $u_1$ | 0,10     | 1        | 0 | 0 |
| $u_2$ | 0,10     | 0        | 1 |   |
| $u_3$ | 0,30     | 0        | 0 |   |
| $u_4$ | 0,20     | 1        | 0 | 1 |
| $u_5$ | 0,20     | 1        | 1 | 0 |
| $u_6$ | 0,10     | 1        | 1 | 1 |

FIG.II.22 - Código com PP associado a árvore binária da Fig. II.22.

Calculando-se seu comprimento médio teremos:

$$\bar{W} = 2,60 \text{ t/i}$$

Logo concluimos que o Algoritmo TOPO não é, necessariamente, um algoritmo ótimo [4], já que ele apenas otimiza o próximo teste a ser aplicado.

Nosso estudo procurará retirar parte dessa limitação no intuito de construir um algoritmo mais eficiente. Para isso, no próximo capítulo apresentaremos novos limites superiores que serão menores ou iguais a  $H(U)/H_{\min}$ .

CAPÍTULO III

NOVO LIMITE SUPERIOR

### III.1 - INTRODUÇÃO

Nesse capítulo obteremos um novo limite superior que se aplica a todo e qualquer código D-ário com PP, e mostraremos que tal limite é menor ou igual àquele obtido por Massey, apresentado na seção II.6 do capítulo anterior.

Antes porém de iniciarmos o desenvolvimento para chegarmos a esse novo limite, daremos um exemplo que ilustrará o procedimento de sua obtenção:

Exemplo 3.1: Seja o Código Binário ( $D=2$ )  $Z$  com PP para a variável aleatória  $U$ , dado abaixo, e seu correspondente Código modificado  $Z'$ :

CÓDIGO  $Z$

| $u$      | $P_U(u)$ | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ |
|----------|----------|-------|-------|-------|-------|-------|
| $u_1$    | 0.05     | 0     | 0     | 1     | 0     |       |
| $u_2$    | 0.20     | 1     | 0     |       |       |       |
| $u_3$    | 0.08     | 0     | 0     | 1     | 1     | 1     |
| $u_4$    | 0.07     | 0     | 0     | 1     | 1     | 0     |
| $u_5$    | 0.15     | 1     | 1     | 1     |       |       |
| $u_6$    | 0.15     | 0     | 1     | 1     |       |       |
| $u_7$    | 0.05     | 1     | 1     | 0     | 1     |       |
| $u_8$    | 0.10     | 1     | 1     | 0     | 0     |       |
| $u_9$    | 0.05     | 0     | 1     | 0     |       |       |
| $u_{10}$ | 0.05     | 0     | 0     | 0     | 1     |       |
| $u_{11}$ | 0.05     | 0     | 0     | 0     | 0     |       |

CÓDIGO  $Z'$

| $u$      | $P_U(u)$ | $X'_1$ | $X'_2$ | $X'_3$ | $X'_4$ | $X'_5$ |
|----------|----------|--------|--------|--------|--------|--------|
| $u_1$    | 0.05     | 0      | 0      | 1      | 0      | 0      |
| $u_2$    | 0.20     | 1      | 0      | 0      | 0      | 0      |
| $u_3$    | 0.08     | 0      | 0      | 1      | 1      | 1      |
| $u_4$    | 0.07     | 0      | 0      | 1      | 1      | 0      |
| $u_5$    | 0.15     | 1      | 1      | 1      | 0      | 0      |
| $u_6$    | 0.15     | 0      | 1      | 1      | 0      | 0      |
| $u_7$    | 0.05     | 1      | 1      | 0      | 1      | 0      |
| $u_8$    | 0.10     | 1      | 1      | 0      | 0      | 0      |
| $u_9$    | 0.05     | 0      | 1      | 0      | 0      | 0      |
| $u_{10}$ | 0.05     | 0      | 0      | 0      | 1      | 0      |
| $u_{11}$ | 0.05     | 0      | 0      | 0      | 0      | 0      |

O comprimento médio do Código  $Z$  é:

$$\bar{W} = 5(0.08+0.07) + 4(0.05+0.05+0.10+0.05+0.05) + 3(0.15+0.15+0.05) + 2(0.20) =$$

$$\bar{W} = 3.40 \text{ t/i.}$$

A entropia de U será:

$$H(U) = 5(-0.05 \log 0.05) - 0.07 \log 0.07 - 0.08 \log 0.08 - 0.10 \log 0.10 - 2(0.15 \log 0.15) =$$

$$H(U) = 3.2582 \text{ bits.}$$

$$\begin{aligned} H(U) &= H(X'_1 X'_2 X'_3 X'_4 X'_5) = H(X'_1 X'_2) + H(X'_3 X'_4 \mid X'_1 X'_2) + \\ &+ H(X'_5 \mid X'_1 X'_2 X'_3 X'_4) \\ &= \left( \frac{H(X'_1 X'_2)}{\bar{W}_{X_1 X_2}} \right) \bar{W}_{X_1 X_2} + \left( \frac{H(X'_3 X'_4 \mid X'_1 X'_2)}{\bar{W}_{X_3 X_4}} \right) \bar{W}_{X_3 X_4} + \\ &+ \left( \frac{H(X'_5 \mid X'_1 X'_2 X'_3 X'_4)}{\bar{W}_{X_5}} \right) \bar{W}_{X_5} \end{aligned} \quad (3.1)$$

onde  $\bar{W}_{X_1 X_2}$  é o comprimento médio do Código Z correspondente às variáveis  $X_1$  e  $X_2$ , e assim por diante. Calculando-se os termos acima encontraremos:

$$\begin{aligned} H(X'_1 X'_2) &= 1.971 \text{ bits} & \bar{W}_{X_1 X_2} &= 2 \\ H(X'_3 X'_4 \mid X'_1 X'_2) &= 0.7377 \text{ bits} & \bar{W}_{X_3 X_4} &= 1.25 \\ H(X'_5 \mid X'_1 X'_2 X'_3 X'_4) &= 0.1495 \text{ bits} & \bar{W}_{X_5} &= 0.15. \end{aligned}$$

Substituindo-se esses valores na Expressão (3.1) ficaremos com:

$$\begin{aligned} H(U) &= \left( \frac{1.971}{2} \right) 2 + \left( \frac{0.7377}{1.25} \right) 1.25 + \left( \frac{0.1495}{0.15} \right) 0.15 \\ &= (0.9855)2 + (0.9102)1.25 + (0.9967)0.15 \end{aligned} \quad (3.2)$$

Tomando-se o mínimo termo dos acima entre parênteses, teremos:

$$H_{\min}(2,2,1) = 0.9102 \text{ bits,}$$

onde (2,2,1) indica que a entropia  $H(U)$  foi desdobrada em parcelas,

cuja primeira é a entropia das duas primeiras variáveis ( $X'_1$  e  $X'_2$ ); cuja segunda é a entropia das duas seguintes variáveis ( $X'_3$  e  $X'_4$ ), dada as duas primeiras; e por último a da última variável ( $X'_5$ ), dada as quatro iniciais. Assim, podemos escrever a Expressão (3.1) em função do seu  $H_{\min}(2,2,1)$ :

$$H(U) \geq H_{\min}(2,2,1) \cdot (\bar{W}_{X_1 X_2} + \bar{W}_{X_3 X_4} + \bar{W}_{X_5}),$$

onde o último termo entre parêntese é exatamente  $\bar{W}$ .

Então teremos:

$$H(U) \geq H_{\min}(2,2,1) \cdot \bar{W}$$

$$\bar{W} \leq \frac{H(U)}{H_{\min}(2,2,1)}.$$

Substituindo-se os valores de  $H(U)$  e  $H_{\min}(2,2,1)$  concluiremos:

$$\bar{W} \leq \frac{3.2582}{0.9102} = 3.5796.$$

Comparando-se o valor de  $\bar{W}$ , anteriormente calculado,  $\bar{W} = 3,40$ , podemos observar que ele satisfaz a desigualdade. Assim, procuraremos provar que  $H(U)/H_{\min}(L_1, \dots, L_K)$  (notação introduzida posteriormente) é realmente um limite superior para  $\bar{W}$ , e que tem a vantagem apresentada no início do capítulo.

### III.2 - DETERMINAÇÃO DO LIMITE SUPERIOR

Seja  $U$  uma variável aleatória,

| $U$      | $P_U(u)$   |
|----------|------------|
| $u_1$    | $P_U(u_1)$ |
| $u_2$    | $P_U(u_2)$ |
| $\vdots$ | $\vdots$   |
| $u_K$    | $P_U(u_K)$ |

e seja  $z_1, z_2, \dots, z_K$  um código D-ário com PP para U, de tal maneira que a Palavra Código  $z_k$  está associada à variável aleatória  $u_k$ :

| $u$      | Código Z |
|----------|----------|
| $u_1$    | $z_1$    |
| $\vdots$ | $\vdots$ |
| $u_K$    | $z_K$    |

Como fizemos anteriormente, primeiro converteremos  $Z = X_1 X_2 \dots X_N$  em  $Z' = X'_1 X'_2 \dots X'_N$ , onde  $N = \max w_k$  e  $w_k$  é o comprimento da Palavra Código  $z_k$  do Código original Z, e também onde a Palavra Código  $z'_k$  é obtida adicionando-se  $N - w_k$  zeros à Palavra Código  $z_k$ .

Exemplo 3.2: Seja um código binário Z com PP:

| CÓDIGO Z |       |       |       | CÓDIGO Z' |        |        |        |
|----------|-------|-------|-------|-----------|--------|--------|--------|
| $u$      | $X_1$ | $X_2$ | $X_3$ | $u$       | $X'_1$ | $X'_2$ | $X'_3$ |
| $u_1$    | 0     |       |       | $u_1$     | 0      | 0      | 0      |
| $u_2$    | 1     | 0     |       | $u_2$     | 1      | 0      | 0      |
| $u_3$    | 1     | 1     | 0     | $u_3$     | 1      | 1      | 0      |
| $u_4$    | 1     | 1     | 1     | $u_4$     | 1      | 1      | 1      |

Note que  $X_1 X_2 \dots X_N$  não são necessariamente variáveis aleatórias, enquanto que  $X'_1 X'_2 \dots X'_N$  sempre o serão. Como o Código Z é um código D-ário com PP, consequentemente o Código Z' também o será. Assim podemos facilmente concluir que:

$$H(U) = H(X'_1 X'_2 \dots X'_N). \quad (3.3)$$

Seja  $L_1, L_2, \dots, L_K$  um conjunto de números inteiros e positivos tal que  $N = L_1 + L_2 + \dots + L_K$ . Definamos  $s_i = L_1 + L_2 + \dots + L_i$  para  $0 \leq i < K$  e também  $s_0 = 0$ . Podemos então fazer uma partição na Eq. (3.3) acima:

$$H(X'_1 \dots X'_N) = H(X'_1 \dots X'_{s_1}) + H(X'_{s_1+1} \dots X'_{s_2} \mid X'_1 \dots X'_{s_1}) +$$

$$\begin{aligned}
 & + (X'_{s_2+1} \dots X'_{s_3} \mid X'_1 \dots X'_{s_2}) + \dots \\
 & \dots + H(X'_{s_{K-1}+1} \dots X'_N \mid X'_1 \dots X'_{s_{K-1}}), \quad (3.4)
 \end{aligned}$$

o que é uma generalização da expressão

$$H(X'_1 \dots X'_N) = H(X'_1) + H(X'_2 \mid X'_1) + \dots + H(X'_N \mid X'_1 \dots X'_{N-1}), [4].$$

Seja

$$\bar{W}_{X'_{s_i+1} \dots X'_{s_{i+1}}}$$

o comprimento médio do Código original Z referente às variáveis  $X'_{s_i+1} \dots X'_{s_{i+1}}$ ,  $0 \leq i < K$

Exemplo 3.3: Considerando-se o mesmo código do Exemplo 3.2 anterior, teremos:

$$\bar{W}_{23} = 0 P(u_1) + 1 P(u_2) + 2 P(u_3) + 2 P(u_4).$$

Por convenção

$$H(X'_{s_i+1} \dots X'_{s_{i+1}} \mid X'_1 \dots X'_{s_i}) = H(X'_1 \dots X'_{s_i})$$

para o caso  $i = 0$ . Podemos então tomar a Eq. (3.4) e obtermos:

$$\begin{aligned}
 H(U) = & \frac{H(X'_1 \dots X'_{s_1})}{\bar{W}_{X'_1 \dots X'_{s_1}}} \bar{W}_{X'_1 \dots X'_{s_1}} + \frac{H(X'_{s_1+1} \dots X'_{s_2} \mid X'_1 \dots X'_{s_1})}{\bar{W}_{X'_{s_1+1} \dots X'_{s_2}}} \bar{W}_{X'_{s_1+1} \dots X'_{s_2}} + \dots \\
 & \dots + \frac{H(X'_{s_{K-1}+1} \dots X'_N \mid X'_1 \dots X'_{s_{K-1}})}{\bar{W}_{X'_{s_{K-1}+1} \dots X'_N}} \bar{W}_{X'_{s_{K-1}+1} \dots X'_N}. \quad (3.5)
 \end{aligned}$$

Seja:

$$H_{\min}(L_1, \dots, L_K) = \min_{i=0,1,\dots,K,1} \left\{ \frac{H(X'_{s_{i+1}} \dots X'_{s_{i+1}} | X'_1 \dots X'_{s_i})}{\bar{W}_{X_{s_{i+1}}} \dots X_{s_{i+1}}} \right\}$$

Logo podemos escrever a Eq. (3.5) em função de  $H_{\min}(L_1, \dots, L_K)$ :

$$H(U) \geq H_{\min}(L_1, \dots, L_K) \left\{ \sum_{i=0}^{K-1} \bar{W}_{X_{s_{i+1}}} \dots X_{s_{i+1}} \right\} \quad (3.6)$$

Como:

$$\sum_{i=0}^{K-1} \bar{W}_{X_{s_{i+1}}} \dots X_{s_{i+1}} = \bar{W},$$

então a Eq. (3.6) fica:

$$\bar{W} \leq \frac{H(U)}{H_{\min}(L_1, \dots, L_K)} \quad (3.7)$$

### III.3 - PROPRIEDADES DO LIMITE SUPERIOR

Investigaremos a seguir algumas propriedades desse limite superior, Eq. (3.7):

Teorema 3.1: Dado um código D-ário com PP para a variável aleatória  $U$ , temos que:

$$H_{\min}(L_1, \dots, L_K) \leq H_{\min}(L'_1, \dots, L'_M)$$

para  $M \leq K$

$$e \quad L'_{j+1} = L_{t_{j-1}+1} + \dots + L_{t_j}, \quad j = 0, \dots, M-1$$

e onde  $t_{-1} = 0$ .

Prova: Seja

$$H_{\min}(L_1, \dots, L_K) = \frac{H(X'_{s_i+1} \dots X'_{s_{i+1}} \mid X'_1 \dots X'_{s_i})}{\bar{W}_{X_{s_i+1} \dots X_{s_{i+1}}}} \quad (3.8)$$

e

$$H_{\min}(L'_1, \dots, L'_M) = \frac{H(X'_{s_j+1} \dots X'_{s_{j+1}} \mid X'_1 \dots X'_{s_j})}{\bar{W}_{X_{s_j+1} \dots X_{s_{j+1}}}} .$$

A prova apresentada será por contradição. Para isso as sumiremos que:

$$H_{\min}(L_1, \dots, L_K) > H_{\min}(L'_1, \dots, L'_M) \quad (3.9)$$

De onde concluímos que:

$$H_{\min}(L_1, \dots, L_K) > 0.$$

Como:

$$L'_{j+1} = L_{t_{j-1}+1} + \dots + L_{t_j}, \quad j = 0, \dots, M-1,$$

podemos escrever:

$$\begin{aligned} H(X'_{s_j+1} \dots X'_{s_{j+1}} \mid X'_1 \dots X'_{s_j}) &= H(X'_{s_j+1} \dots X'_{s_j+L_{t_{j-1}+1}} \mid X'_1 \dots X'_{s_j}) + \\ &+ H(X'_{s_j+L_{t_{j-1}+1}+1} \dots X'_{s_j+L_{t_{j-1}+1}+L_{t_{j-1}+2}} \mid X'_1 \dots X'_{s_j+L_{t_{j-1}+1}}) + \\ &+ H(X'_{s_j+L_{t_{j-1}+1}+\dots+L_{t_{j-1}+1}} \dots X'_{s_j+L_{t_{j-1}+1}+\dots+L_{t_j}} \mid X'_1 \dots X'_{s_j+L_{t_{j-1}+1}+\dots+L_{t_{j-1}}}). \end{aligned} \quad (3.10)$$

Observemos que todas as entropias condicionais da expres são acima são diferentes de zero. Caso contrário teríamos:

$$H_{\min}(L_1, \dots, L_K) = 0.$$

A Expressão (3.10) pode ser escrita sob a forma:

$$\begin{aligned}
 H \left( X'_{s_j+1} \dots X'_{s_{j+1}} \mid X'_1 \dots X'_{s_j} \right) &= \left( \frac{H \left( X'_{s_j+1} \dots X'_{s_j+L_{t_{j-1}+1}} \mid X'_1 \dots X'_{s_j} \right)}{\bar{W}_{X_{s_j+1} \dots X_{s_j+L_{t_{j-1}+1}}}} \right) \bar{W}_{X_{s_j+1} \dots X_{s_j+L_{t_{j-1}+1}}} \\
 &+ \left( \frac{H \left( X'_{s_j+L_{t_{j-1}+1}+1} \dots X'_{s_j+L_{t_{j-1}+1}+L_{t_{j-1}+2}} \mid X'_1 \dots X'_{s_j+L_{t_{j-1}+1}} \right)}{\bar{W}_{X_{s_j+L_{t_{j-1}+1}+1} \dots X_{s_j+L_{t_{j-1}+1}+L_{t_{j-1}+2}}}} \right) \bar{W}_{X_{s_j+L_{t_{j-1}+1}+1} \dots} \\
 &\dots X_{s_j+L_{t_{j-1}+1}+L_{t_{j-1}+2}} + \\
 &+ \left( \frac{H \left( X'_{s_j+L_{t_{j-1}+1}+\dots+L_{t_{j-1}+1}} \dots X'_{s_j+L_{t_{j-1}+1}+\dots+L_{t_j}} \mid X'_1 \dots X'_{s_j+L_{t_{j-1}+1}+\dots+L_{t_{j-1}}} \right)}{\bar{W}_{X_{s_j+L_{t_{j-1}+1}+\dots+L_{t_{j-1}+1}} \dots X_{s_j+L_{t_{j-1}+1}+\dots+L_{t_j}}}} \right) \dots \\
 &\dots \bar{W}_{X_{s_j+L_{t_{j-1}+1}+\dots+L_{t_{j-1}+1}} \dots X_{s_j+L_{t_{j-1}+1}+\dots+L_{t_j}}} \quad (3.11)
 \end{aligned}$$

Chamemos de  $H_{mp1}$  o mínimo valor dos termos entre parênteses da Expressão (3.11) acima. Assim podemos escrever:

$$\begin{aligned}
 H \left( X'_{s_j+L} \dots X'_{s_{j+1}} \mid X'_1 \dots X'_{s_j} \right) &\geq H_{mp1} \left( \bar{W}_{X_{s_j+1} \dots X_{s_j}} + \right. \\
 &+ \bar{W}_{X_{s_j+L_{t_{j-1}+1}+1} \dots X_{s_j+L_{t_{j-1}+1}+L_{t_{j-1}+2}}} + \dots
 \end{aligned}$$

$$\dots + \bar{W}_{X_{s_j + L_{t_{j-1}+1} + \dots + L_{t_j-1} + 1} \dots X_{s_j + L_{t_{j-1}+1} + \dots + L_{t_j}}}$$

$$H\left(X'_{s_j+1} \dots X'_{s_{j+1}} \mid X'_1 \dots X'_{s_j}\right) \geq H_{mp_1} \cdot \bar{W}_{X_{s_j+1} \dots X_{s_{j+1}}} \quad (3.12)$$

Baseados nas Expressões (3.8), (3.9) e (3.12) podemos afirmar que:

$$H_{\min}(L_1, \dots, L_K) > H_{\min}(L'_1, \dots, L'_M) \geq H_{mp_1}$$

o que é uma contradição, já que:

$$H_{mp_1} \geq H_{\min}(L_1, \dots, L_K)$$

Então podemos concluir que:

$$H_{\min}(L_1, \dots, L_K) \leq H_{\min}(L'_1, \dots, L'_M) \quad \text{para } M \leq K$$

$$e \quad L'_j = L_{t_{j-1}+1} + \dots + L_{t_j}, \quad j = 1, \dots, M$$

$$e \text{ onde: } t_0 = 0.$$

c.q.d.

Corolário 3.1: Dado um código D-ário com PP para a variável aleatória U, temos que:

$$H_{\min}(L_1, \dots, L_N) \leq H_{\min}(L_1, \dots, L_K)$$

Como ilustração do Teorema 3.1, apresentaremos o exemplo abaixo:

Exemplo 3.4: Consideremos o mesmo Código binário Z com PP para a variável aleatória U e seu Código correspondente Z', considerados no Exemplo 3.1:

Como já determinamos, temos:

$$\bar{W} = 3.40 \text{ t/i} \quad \text{e} \quad H(U) = 3.2582 \text{ bits}$$

$$H(U) = H(X'_1 X'_2 X'_3 X'_4 X'_5) = H(X'_1 X'_2 X'_3) + H(X'_4 X'_5 | X'_1 X'_2 X'_3)$$

$$H(U) = \left( \frac{H(X'_1 X'_2 X'_3)}{\bar{W}_{X'_1 X'_2 X'_3}} \right) \bar{W}_{X'_1 X'_2 X'_3} + \left( \frac{H(X'_4 X'_5 | X'_1 X'_2 X'_3)}{\bar{W}_{X'_4 X'_5}} \right) \bar{W}_{X'_4 X'_5},$$

onde:  $H(X'_1 X'_2 X'_3) = 2.7088 \text{ bits}$   $\bar{W}_{X'_1 X'_2 X'_3} = 2.80$

$H(X'_4 X'_5 | X'_1 X'_2 X'_3) = 0.5495 \text{ bits}$   $\bar{W}_{X'_4 X'_5} = 0.60$ .

Substituindo-se os valores acima, teremos:

$$H(U) = \left( \frac{2.7088}{2.80} \right) 2.80 + \left( \frac{0.5495}{0.60} \right) 0.60$$

$$H(U) = (0.9674) 2.80 + (0.9158) 0.60.$$

Tomando-se o mínimo termo dos acima entre parênteses, te  
remos:

$$H_{\min}(3,2) = 0.9158 \text{ bits}.$$

Por outro lado, poderemos escrever a expressão:

$$H(U) = H(X'_1 X'_2 X'_3 X'_4 X'_5) = H(X'_1 X'_2) + H(X'_3 | X'_1 X'_2) + \\ + H(X'_4 | X'_1 X'_2 X'_3) + H(X'_5 | X'_1 X'_2 X'_3 X'_4)$$

$$= \left( \frac{H(X'_1 X'_2)}{\bar{W}_{X'_1 X'_2}} \right) \bar{W}_{X'_1 X'_2} + \left( \frac{H(X'_3 | X'_1 X'_2)}{\bar{W}_{X'_3}} \right) \bar{W}_{X'_3} +$$

$$+ \left( \frac{H(X'_4 | X'_1 X'_2 X'_3)}{\bar{W}_{X'_4}} \right) \bar{W}_{X'_4} + \left( \frac{H(X'_5 | X'_1 X'_2 X'_3 X'_4)}{\bar{W}_{X'_5}} \right) \bar{W}_{X'_5},$$

onde:

$$\begin{aligned}
 H(X'_1 X'_2) &= 1.971 \text{ bits} & \bar{W}_{X_1 X_2} &= 2 \\
 H(X'_3 | X'_1 X'_2) &= 0.7377 \text{ bits} & \bar{W}_{X_3} &= 0.8 \\
 H(X'_4 | X'_1 X'_2 X'_3) &= 0.400 \text{ bits} & \bar{W}_{X_4} &= 0.45 \\
 H(X'_5 | X'_1 X'_2 X'_3 X'_4) &= 0.1495 \text{ bits} & \bar{W}_{X_5} &= 0.15
 \end{aligned}$$

Substituindo-se esses valores, teremos:

$$\begin{aligned}
 H(U) &= \left( \frac{1.971}{2} \right) 2 + \left( \frac{0.7377}{0.8} \right) 0.8 + \left( \frac{0.400}{0.45} \right) 0.15 + \\
 &\quad + \left( \frac{0.1495}{0.15} \right) 0.15 \\
 &= (0.9855)2 + (0.9221)0.8 + (0.8889)0.45 + (0.9967)0.15
 \end{aligned}$$

Tomando-se aqui também o mínimo termo dos acima entre parênteses, teremos:

$$H_{\min}(2,1,1,1) = 0.8889 \text{ bits.}$$

Como podemos observar,

$$H_{\min}(2,1,1,1) = 0.8889 < H_{\min}(3,2) = 0.9158.$$

Como ilustração do Corolário 3.1 apresentaremos o exemplo abaixo:

Exemplo 3.5: Consideremos novamente o mesmo Código binário  $Z$  com PP para a variável aleatória  $U$  e seu código correspondente  $Z'$ , considerados nos Exemplos 3.1 e 3.4:

Como já determinado,

$$\bar{W} = 3.40 \text{ t/i} \quad \text{e} \quad H(U) = 3.2582 \text{ bits.}$$

$$\begin{aligned}
H(U) &= H(X'_1 X'_2 X'_3 X'_4 X'_5) = H(X'_1) + H(X'_2 | X'_1) + H(X'_3 | X'_1 X'_2) + \\
&+ H(X'_4 | X'_1 X'_2 X'_3) + H(X'_5 | X'_1 X'_2 X'_3 X'_4) \\
&= \left( \frac{H(X'_1)}{\bar{W}_{X_1}} \right) \bar{W}_{X_1} + \left( \frac{H(X'_2 | X'_1)}{\bar{W}_{X_2}} \right) \bar{W}_{X_2} + \left( \frac{H(X'_3 | X'_1 X'_2)}{\bar{W}_{X_3}} \right) \bar{W}_{X_3} \\
&+ \left( \frac{H(X'_4 | X'_1 X'_2 X'_3)}{\bar{W}_{X_4}} \right) \bar{W}_{X_4} + \left( \frac{H(X'_5 | X'_1 X'_2 X'_3 X'_4)}{\bar{W}_{X_5}} \right) \bar{W}_{X_5} .
\end{aligned}$$

Pela tabela do Código Z do Exemplo 3.1 podemos calcular:

$$\begin{aligned}
H(X'_1) &= 1 \quad \text{bit} & \bar{W}_{X_1} &= 1 \\
H(X'_2 | X'_1) &= 0.971 \quad \text{bits} & \bar{W}_{X_2} &= 1 \\
H(X'_3 | X'_1 X'_2) &= 0.7377 \quad \text{bits} & \bar{W}_{X_3} &= 0.8 \\
H(X'_4 | X'_1 X'_2 X'_3) &= 0.400 \quad \text{bits} & \bar{W}_{X_4} &= 0.45 \\
H(X'_5 | X'_1 X'_2 X'_3 X'_4) &= 0.1495 \quad \text{bits} & \bar{W}_{X_5} &= 0.15
\end{aligned}$$

Substituindo-se esses valores teremos:

$$\begin{aligned}
H(U) &= \left( \frac{1}{1} \right) 1 + \left( \frac{0.971}{1} \right) 1 + \left( \frac{0.7377}{0.8} \right) 0.8 + \\
&+ \left( \frac{0.400}{0.45} \right) 0.45 + \left( \frac{0.1495}{0.45} \right) 0.15 . \\
&= (1) + (0.971) + (0.9221) 0.8 + (0.8889) 0.45 + \\
&+ (0.9967) 0.15 .
\end{aligned}$$

Tomando-se o mínimo termo dos que estão entre parênteses,

teremos o  $H_{\min}(L_1, \dots, L_N)$ , que nesse caso será  $H_{\min}(1,1,1,1,1)$ :

$$H_{\min}(L_1, \dots, L_N) = H_{\min}(1,1,1,1,1) = 0.8889.$$

Para se encontrar o  $H_{\min}(L_1, \dots, L_K)$ , teremos:

$$\begin{aligned} H(U) &= H(X'_1 X'_2) + H(X'_3 X'_4 | X'_1 X'_2) + H(X'_5 | X'_1 X'_2 X'_3 X'_4) \\ &= \left( \frac{H(X'_1 X'_2)}{\bar{W}_{X_1 X_2}} \right) \bar{W}_{X_1 X_2} + \left( \frac{H(X'_3 X'_4 | X'_1 X'_2)}{\bar{W}_{X_3 X_4}} \right) \bar{W}_{X_3 X_4} + \\ &\quad + \left( \frac{H(X'_5 | X'_1 X'_2 X'_3 X'_4)}{\bar{W}_{X_5}} \right) \bar{W}_{X_5}, \end{aligned}$$

onde:

$$H(X'_1 X'_2) = 1.971 \text{ bits} \quad \bar{W}_{X_1 X_2} = 2$$

$$H(X'_3 X'_4 | X'_1 X'_2) = 1.1377 \text{ bits} \quad \bar{W}_{X_3 X_4} = 1.25$$

$$H(X'_5 | X'_1 X'_2 X'_3 X'_4) = 0.1495 \text{ bits} \quad \bar{W}_{X_5} = 0.15.$$

Assim,

$$\begin{aligned} H(U) &= \left( \frac{1.971}{2} \right) 2 + \left( \frac{1.1377}{1.25} \right) 1.25 + \left( \frac{0.1495}{0.15} \right) 0.15 \\ &= (0.9855)2 + (0.9102)1.25 + (0.9967)0.15. \end{aligned}$$

Tomando-se o mínimo dos termos que estão entre parênteses na expressão acima teremos o  $H_{\min}(L_1, \dots, L_K)$  que nesse caso será  $H_{\min}(2,2,1)$ :

$$H_{\min}(L_1, \dots, L_K) = H_{\min}(2,2,1) = 0.902 \text{ bits.}$$

Então, concordando com o Corolário 3.1, temos:

$$H_{\min}(1,1,1,1,1) = 0.8889 < H_{\min}(2,2,1) = 0.9102 \text{ bits.}$$

Agora mostraremos a relação que existe entre esse novo limite superior e o de Massey. Para isso revisaremos algumas definições:

#### III.4 - COMPARAÇÕES DO NOVO LIMITE SUPERIOR COM O DE MASSEY

Seja  $Q_i$  o conjunto de prefixos de comprimento "i" para o código original D-ário Z, isto é,

$$Q_i = \{ x_1 \dots x_i \mid x_1 \dots x_i \text{ é o prefixo de uma palavra código de Z de comprimento maior ou igual a } i + 1 \}.$$

Analogamente definamos  $Q'_i$  como o conjunto de prefixos de comprimento "i" para o código D-ário modificado Z', ou seja,

$$Q'_i = \{ x'_1 \dots x'_i \mid x'_1 \dots x'_i \text{ é o prefixo de uma palavra código de Z' de comprimento maior ou igual a } i + 1 \}.$$

Exemplo 3.6: Se os Códigos Z e Z' são os mesmos considerados nos Exemplos 3.2 e 3.3, teremos:

$$Q_1 = \{1\}, Q_2 = \{11\}, Q'_1 = \{1,0\} \text{ e } Q'_2 = \{00,10,11\}.$$

Assim podemos enunciar as seguintes propriedades:

Propriedade 3.1 [4]:

$$\sum_{x_1 \dots x_i \in Q_i} P_{X'_1 \dots X'_i}(x_1, \dots, x_i) = P(W > i)$$

Esta soma não é mais do que a probabilidade de que uma palavra código do código original D-ário Z tenha comprimento maior do que W.

Propriedade 3.2 [4]:

$$H(X'_{s_i+1} \mid X'_1 \dots X'_{s_i} = x_1 \dots x_{s_i}) = 0, \quad 0 \leq i \leq K-1,$$

quando  $x_1 \dots x_{s_i} \in Q'_{s_i}$  mas não a  $Q_{s_i}$ .

Neste caso não existe incerteza nenhuma porque  $X'_j = 0$  para  $j = s_i + 1, \dots, N$ .

Exemplo 3.7: Considerando-se ainda os Códigos Z e Z' dos Exemplos 3.2, 3.3 e 3.6 anteriores, temos:

$$H(X'_3 | X'_1 X'_2 = 10) = 0$$

pois dado  $X'_1 X'_2 = 10$  não existe incerteza alguma sobre  $X'_3$ .

Teorema 3.2: Dado um código D-ário com PP para a variável aleatória U, temos que:

$$H_{\min} \leq H_{\min}(L_1, \dots, L_N).$$

Prova: Seja

$$H_{\min}(L_1, \dots, L_N) = \frac{H(X'_{s_i+1} | X'_1 \dots X'_{s_i})}{\bar{W}_{X_{s_i+1}}} \quad (3.13)$$

e também 
$$H_{\min} = H(X'_{s_j+1} | X'_1 \dots X'_{s_j} = x_1 \dots x_{s_j}) .$$

A prova apresentada será novamente por contradição . Para isso assumiremos que:

$$H_{\min} > H_{\min}(L_1, \dots, L_N), \quad (3.14)$$

de onde concluímos que  $H_{\min} > 0$ .

Podemos escrever que:

$$H(X'_{s_i+1} | X'_1 \dots X'_{s_i}) = \sum_{x_1 \dots x_{s_i} \in Q'_{s_i}} H(X'_{s_i+1} | X'_1 \dots X'_{s_i} = x_1 \dots x_{s_i}) P_{X'_1 \dots X'_{s_i}}(x_1 \dots x_{s_i}) .$$

Usando-se a Propriedade 3.2 teremos:

$$H(X'_{s_i+1} \mid X'_1 \dots X'_{s_i}) = \sum_{x_1 \dots x_{s_i} \in Q_{s_i}} H(X'_{s_i+1} \mid X'_1 \dots X'_{s_i} = x_1 \dots x_{s_i}) \\ P_{X'_1 \dots X'_{s_i}}(x_1 \dots x_{s_i}). \quad (3.15)$$

Observemos que todas as entropias condicionais da expressão acima são diferentes de zero. Caso contrário teríamos:

$$H_{\min} = 0.$$

Seja:

$$H_{mp_2} = \min_{x_1 \dots x_{s_i} \in Q_{s_i}} H(X'_{s_i+1} \mid X'_1 \dots X'_{s_i} = x_1 \dots x_{s_i}).$$

Logo a Eq. (3.15) pode ser escrita como:

$$H(X'_{s_i+1} \mid X'_1 \dots X'_{s_i}) \geq H_{mp_2} \sum_{x_1 \dots x_{s_i} \in Q_{s_i}} P_{X'_1 \dots X'_{s_i}}(x_1 \dots x_{s_i}) \quad (3.16)$$

De acordo com a Propriedade 3.1:

$$\sum_{x_1 \dots x_{s_i} \in Q_{s_i}} P_{X'_1 \dots X'_{s_i}}(x_1 \dots x_{s_i}) = P(W > s_i) = \bar{W}_{X_{s_i+1}}.$$

Logo podemos escrever a Eq. (3.16) como:

$$H(X'_{s_i+1} \mid X'_1 \dots X'_{s_i}) \geq H_{mp_2} \cdot \bar{W}_{X_{s_i+1}}. \quad (3.17)$$

Assim, baseado nas Eq. (3.17), (3.13) e (3.14) podemos escrever:

$$H_{\min} > H_{\min}(L_1, \dots, L_N) > H_{mp_2},$$

o que é uma contradição, já que:

$$H_{mp_2} > H_{\min}(L_1, \dots, L_N).$$

Concluimos então que,

$$H_{\min} \leq H_{\min}(L_1, \dots, L_N).$$

c.q.d.

Corolário 3.2: Dado um código D-ário com PP para a variável aleatória  $U$ ; temos que:

$$H_{\min} \leq H_{\min}(L_1, \dots, L_K).$$

Como ilustração do Teorema 3.2 apresentaremos o exemplo abaixo:

Exemplo 3.8: Consideremos o mesmo Código binário dos Exemplos 3.1, 3.4 e 3.5. Podemos então escrever:

$$\begin{aligned} H(U) = & H(X'_1) + H(X'_2 | X'_1=0)P_{X'_1}(0) + H(X'_2 | X'_1=1)P_{X'_1} + \\ & + H(X'_3 | X'_1X'_2=00)P_{X'_1X'_2} + H(X'_3 | X'_1X'_2=01)P_{X'_1X'_2}(01) + \\ & + H(X'_3 | X'_1X'_2=11)P_{X'_1X'_2}(11) + H(X'_4 | X'_1X'_2X'_3=000)P_{X'_1X'_2X'_3}(000) + \\ & + H(X'_4 | X'_1X'_2X'_3=001)P_{X'_1X'_2X'_3}(001) + \\ & + H(X'_5 | X'_1X'_2X'_3X'_4=0011)P_{X'_1X'_2X'_3X'_4}(0011). \end{aligned}$$

Calculando-se essas entropias acima teremos:

|                                       |                                                 |
|---------------------------------------|-------------------------------------------------|
| $H(X'_1) = 1$ bit                     | $H(X'_4   X'_1X'_2X'_3=000) = 1$ bit            |
| $H(X'_2   X'_1=1) = 0.971$ bits       | $H(X'_4   X'_1X'_2X'_3=001) = 0.8113$ bits      |
| $H(X'_2   X'_1=1) = 0.971$ bits       | $H(X'_4   X'_1X'_2X'_3=110) = 0.9183$ bits      |
| $H(X'_3   X'_1X'_2=00) = 0.9183$ bits | $H(X'_4   X'_1X'_2X'_3=110) = 0.9183$ bits      |
| $H(X'_3   X'_1X'_2=01) = 0.8113$ bits | $H(X'_5   X'_1X'_2X'_3X'_4=0011) = 0.9967$ bits |
| $H(X'_3   X'_1X'_2=11) = 1$ bit       |                                                 |

$$P_{X_1}(0) = 0.50$$

$$P_{X_1}(1) = 0.50$$

$$P_{X_1X_2}(00) = 0.30$$

$$P_{X_1X_2}(01) = 0.20$$

$$P_{X_1X_2}(11) = 0.30$$

$$P_{X_1X_2X_3}(000) = 0.10$$

$$P_{X_1X_2X_3}(001) = 0.20$$

$$P_{X_1X_2X_3}(110) = 0.15$$

$$P_{X_1X_2X_3X_4}(0011) = 0.15$$

Tomando-se o mínimo valor das entropias condicionais acima teremos então:

$$H_{\min} = 0.8113 \text{ bits} \quad (3.18)$$

que é o mínimo valor de Massey. Procurando-se o  $H_{\min}(L_1, \dots, L_K)$  teremos:

$$\begin{aligned} H(U) &= H(X'_1) + H(X'_2|X'_1) + H(X'_3|X'_2X'_1) + H(X'_4|X'_1X'_2X'_3) + \\ &+ H(X'_5|X'_1X'_2X'_3X'_4) \end{aligned}$$

$$\begin{aligned} &= \left( \frac{H(X'_1)}{\bar{w}_{X_1}} \right) \bar{w}_{X_1} + \left( \frac{H(X'_2|X'_1)}{\bar{w}_{X_2}} \right) \bar{w}_{X_2} + \left( \frac{H(X'_3|X'_1X'_2)}{\bar{w}_{X_3}} \right) \bar{w}_{X_3} + \\ &+ \left( \frac{H(X'_4|X'_1X'_2X'_3)}{\bar{w}_{X_4}} \right) \bar{w}_{X_4} + \left( \frac{H(X'_5|X'_1X'_2X'_3X'_4)}{\bar{w}_{X_5}} \right) \bar{w}_{X_5} \end{aligned}$$

$$\begin{aligned}
H(U) &= \left( \frac{1}{1} \right) 1 + \left( \frac{0.971}{1} \right) 1 + \left( \frac{0.7377}{0.8} \right) 0.8 + \\
&+ \left( \frac{0.400}{0.45} \right) 0.45 + \left( \frac{0.1495}{0.15} \right) 0.15 = \\
&= (1) 1 + (0.971) 1 + (0.9221) 0.8 + \\
&+ (0.8889) 0.45 + (0.9967) 0.15 .
\end{aligned}$$

Tomando-se o mínimo valor dos termos entre parênteses acima teremos:

$$H_{\min}(L_1, \dots, L_N) = H_{\min}(1, 1, 1, 1, 1) = 0.8889 \text{ bits. (3.19)}$$

Comparando as Expressões (3.18) e (3.19) podemos dizer que:

$$H_{\min} = 0.8113 < H_{\min}(1, 1, 1, 1, 1) = 0.8889 ,$$

o que está de acordo com o Teorema 3.2.

Como ilustração do Corolário 3.2 apresentaremos o exemplo abaixo:

Exemplo 3.9: Consideremos ainda o mesmo Código binário dos Exemplos 3.1, 3.4, 3.5 e 3.8.

Para esse Código temos:

$$H_{\min} = 0.8113 \text{ bits (já calculado no Exemplo 3.8)}$$

e 
$$H_{\min}(2, 2, 1) = 0.9102 \text{ bits (já calculado no Exemplo 3.5).}$$

Como podemos observar,

$$H_{\min} = 0.8113 < H_{\min}(2, 2, 1) = 0.9102.$$

No próximo capítulo aplicaremos os resultados aqui obtidos na construção de algoritmos eficientes.

## CAPÍTULO IV

### CONSTRUÇÃO DE ALGORITMOS EFICIENTES

## CONSTRUÇÃO DE ALGORITMOS EFICIENTES

### IV.1 - DEFINIÇÃO

Os limites superiores estabelecidos no capítulo anterior nos fornece meios para construirmos algoritmos eficientes. A idéia básica usada na construção desses algoritmos será tornar o mínimo possível o limite superior. Relembremos a Definição 2.13:

"Um algoritmo de teste para a variável aleatória  $U$  é um algoritmo que unicamente identifica todos os  $u_i$ 's."

Podemos escrever a Equação (3.5) sob a forma:

$$\begin{aligned}
 H(U) = & \left( \frac{H(X'_1 \dots X'_{L_1})}{\bar{W}_{X_1 \dots X_{L_1}}} \right) \bar{W}_{X_1 \dots X_{L_1}} + \left( \frac{H(X'_{L_1+1} \dots X'_{L_1+L_2} | X'_1 \dots X'_{L_1})}{\bar{W}_{X_{L_1+1} \dots X_{L_1+L_2}}} \right) \bar{W}_{X_{L_1+1} \dots X_{L_1+L_2}} + \dots \\
 & \dots + \left( \frac{H(X'_{L_1+L_2+\dots+L_{K-1}+1} \dots X'_{L_1+L_2+\dots+L_K} | X'_1 \dots X'_{L_1+L_2+\dots+L_{K-1}})}{\bar{W}_{X_{L_1+L_2+\dots+L_{K-1}} \dots X_{L_1+L_2+\dots+L_K}}} \right) \dots \\
 & \dots \bar{W}_{X_{L_1+L_2+\dots+L_{K-1}+1} \dots X_{L_1+L_2+\dots+L_K}}
 \end{aligned} \tag{4.1}$$

Definiremos  $F(j)$  como:

$$F(j) = \frac{H(X'_{L_1+L_2+\dots+L_{j-1}+1} \dots X'_{L_1+L_2+\dots+L_j} | X'_1 \dots X'_{L_1+L_2+\dots+L_{j-1}})}{\bar{W}_{X_{L_1+L_2+\dots+L_{j-1}+1} \dots X_{L_1+L_2+\dots+L_j}}} \tag{4.2}$$

$$\text{para } j=2,3,\dots,K \text{ e seja } F(1) = \frac{H(X'_1 \dots X'_{L_1})}{\bar{W}_{X_1 \dots X_{L_1}}} \tag{4.3}$$

Definição 4.1: Um algoritmo de testes ótimo de ordem  $(L_1, L_2, \dots, L_K)$  (daqui para frente abreviado apenas por "Algoritmo TOO  $(L_1, L_2, \dots, L_K)$ ") para a variável aleatória  $U$ , é um algoritmo tal que  $F(j) \neq 0$ , para  $j=1, \dots, K$ , e que tem como filosofia básica de construção a maximização, em cada passo, do termo  $F(j)$ .

Iniciamos a construção escolhendo os testes que maximizam  $F(1)$ . Baseados nessa escolha tomaremos, entre os testes restantes, aqueles que maximizam  $F(2)$ , e assim por diante. Observe-mos aqui que tal procedimento não nos leva, necessariamente, ao algoritmo ótimo, pois a escolha ótima de  $F(j)$  depende das escolhas anteriores  $F(1), F(2), \dots, F(j-1)$ . O que faremos é, a cada passo, maximizar a incerteza média, por comprimento médio. O exemplo que segue ilustra o procedimento de construção de Algoritmos TOO  $(L_1, \dots, L_K)$ :

Exemplo 4.1: Tomemos novamente a Tabela de Decisão do Exemplo 1.1:

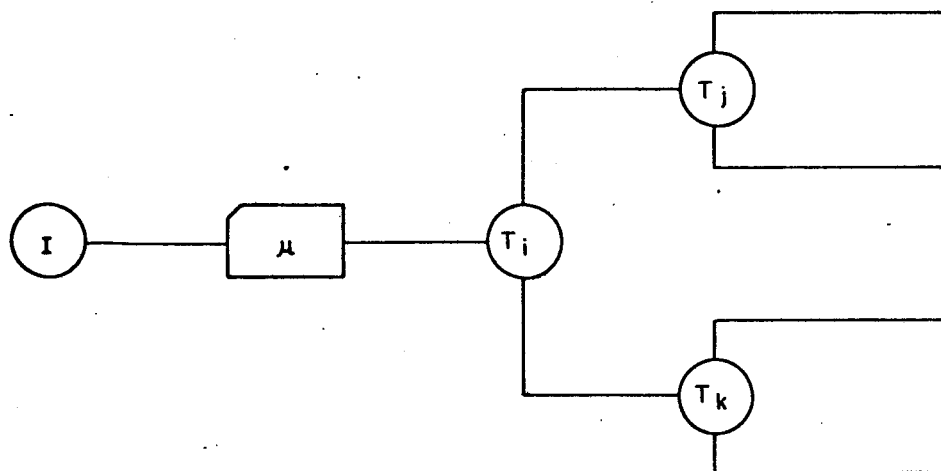
| $T \backslash u$ | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ |
|------------------|-------|-------|-------|-------|-------|-------|
| $P_U(u)$         | 0.1   | 0.1   | 0.3   | 0.2   | 0.2   | 0.1   |
| $T_1$            | 0     | 0     | 0     | 1     | 1     | 1     |
| $T_2$            | 1     | 0     | 0     | 1     | 1     | 1     |
| $T_3$            | 0     | 1     | 0     | 0     | 0     | 1     |
| $T_4$            | 0     | 1     | 0     | 1     | 0     | 1     |
| $T_5$            | 0     | 1     | 1     | 0     | 1     | 1     |

FIG.IV.1 - Tabela de Decisão do Exemplo 1.1

Façamos  $L_1$  (na Equação 4.3) igual a 2. Assim,

$$F(1) = \frac{H(X'_1, X'_2)}{W_{X_1 X_2}},$$

que será o termo a ser maximizado na construção do algoritmo. Uma pesquisa terá que ser feita para escolhermos os testes de modo a maximizar o  $F(1)$  acima. Por convenção, quando falarmos numa sequência de testes do tipo  $T_i T_j T_k$ , significará, respectivamente, que serão usados nas posições abaixo indicadas:



Como podemos observar na tabela da Fig.IV.1 nenhum teste aplicado separa, no primeiro passo, um dos  $u_i$ 's. Assim sendo, todas as palavras código associadas a qualquer algoritmo que unicamente identifique todos os  $u_i$ 's terão, no mínimo, comprimento dois. Então o  $\bar{W}_{X_1 X_2}$  será igual a dois, e maximizar  $F(1)$  da Equação (4.3) será maximizar  $H(X_1 X_2)$ . Este termo, por sua vez, será dado por:

$$H(X_1 X_2) = H(X_1) + H(X_2 | X_1 = 0) P_{X_1}(0) + H(X_2 | X_1 = 1) P_{X_1}(1)$$

Notemos que, para um determinado teste inicial, maximizar  $H(X_1 X_2)$  será equivalente a maximizar  $H(X_2 | X_1 = 0)$  e .....  $H(X_2 | X_1 = 1)$ . Passaremos então a investigar os testes que maximizam os termos acima, dado cada teste inicial em separado:

Teste inicial  $T_1$ :  $H(X_1) = h(0,50) = 1$  bit,  $P_{X_1}(0) = 0,50$  e  
 $P_{X_1}(1) = 0,50$

| 2 <sup>os</sup> TESTES | $H(X'_2 X'_1=0)$ | $H(X'_2 X'_1=1)$ |
|------------------------|------------------|------------------|
| $T_2$                  | $h(0,20)$        | $h(0)$           |
| $T_3$                  | $h(0,20)$        | $h(0,20)$        |
| $T_4$                  | $h(0,20)$        | $h(0,40)$        |
| $T_5$                  | $h(0,20)$        | $h(0,40)$        |

Valor máximo de  $H(X'_1X'_2)$  dado  $T_1$  como teste inicial:

$$\begin{aligned}
 H(X'X') &= h(0,50) + ((h(0,20) \times 0,50) + (h(0,40) \times 0,50)) \\
 &= 1 + (0,7219 \times 0,50) + (0,9710 \times 0,50) = 1,8465 \text{ bits} .
 \end{aligned}$$

Teste inicial  $T_2$ :  $H(X'_1) = h(0,40) = 0,9710$  bits,  $P_{X_1}(0) = 0,40$  e  
 $P_{X_1}(1) = 0,60$

| 2 <sup>os</sup> TESTES | $H(X'_2 X'_1=0)$ | $H(X'_2 X'_1=1)$ |
|------------------------|------------------|------------------|
| $T_1$                  | $h(0)$           | $h(0,1667)$      |
| $T_3$                  | $h(0,25)$        | $h(0,1667)$      |
| $T_4$                  | $h(0,25)$        | $h(0,50)$        |
| $T_5$                  | $h(0)$           | $h(0,50)$        |

Valor máximo de  $H(X'_1X'_2)$  dado  $T_2$  como teste inicial:

$$\begin{aligned}
 H(X'_1X'_2) &= h(0,40) + (h(0,25) \times 0,40) + (h(0,50) \times 0,60) = \\
 &= 0,9710 + (0,8113 \times 0,40) + (1 \times 0,60) = 1,8955 \text{ bits}
 \end{aligned}$$

Teste inicial  $T_3$ :  $H(X'_1) = h(0,20) = 0,7219$  bits,  $P_{X_1}(0) = 0,80$  e  
 $P_{X_1}(1) = 0,20$

| 2 <sup>os</sup> TESTES | $H(X'_2 X'_1=0)$ | $H(X'_2 X'_1=1)$ |
|------------------------|------------------|------------------|
| $T_1$                  | $h(0,50)$        | $h(0,50)$        |
| $T_2$                  | $h(0,375)$       | $h(0,50)$        |
| $T_4$                  | $h(0,25)$        | $h(0)$           |
| $T_5$                  | $h(0,375)$       | $h(0)$           |

Valor máximo de  $H(X'_1X'_2)$  dado  $T_3$  como teste inicial:

$$\begin{aligned}
 h(X'_1X'_2) &= h(0,20) + (h(0,50) \times 0,80) + (h(0,50) \times 0,20) = \\
 &= 0,7219 + (1 \times 0,80) + (1 \times 0,20) = 1,7219 \text{ bits.}
 \end{aligned}$$

Teste inicial  $T_4$ :  $H(X') = h(0,40) = 0,9710$  bits,  $P_{X_1}(0) = 0,60$  e  
 $P_{X_1}(1) = 0,40$

| 2 <sup>os</sup> TESTES | $H(X'_2 X'_1=0)$ | $H(X'_2 X'_1=1)$ |
|------------------------|------------------|------------------|
| $T_1$                  | $h(0,333)$       | $h(0,25)$        |
| $T_2$                  | $h(0,50)$        | $h(0,25)$        |
| $T_3$                  | $h(0)$           | $h(0,50)$        |
| $T_5$                  | $h(0,1667)$      | $h(0,50)$        |

Valor máximo de  $H(X'_1X'_2)$ , dado  $T_4$  como teste inicial:

$$\begin{aligned}
 H(X'_1X'_2) &= h(0,40) + (h(0,50) \times 0,60) + (h(0,50) \times 0,40) \\
 &= 0,9710 + (1 \times 0,60) + (1 \times 0,40) = 1,9710 \text{ bits}
 \end{aligned}$$

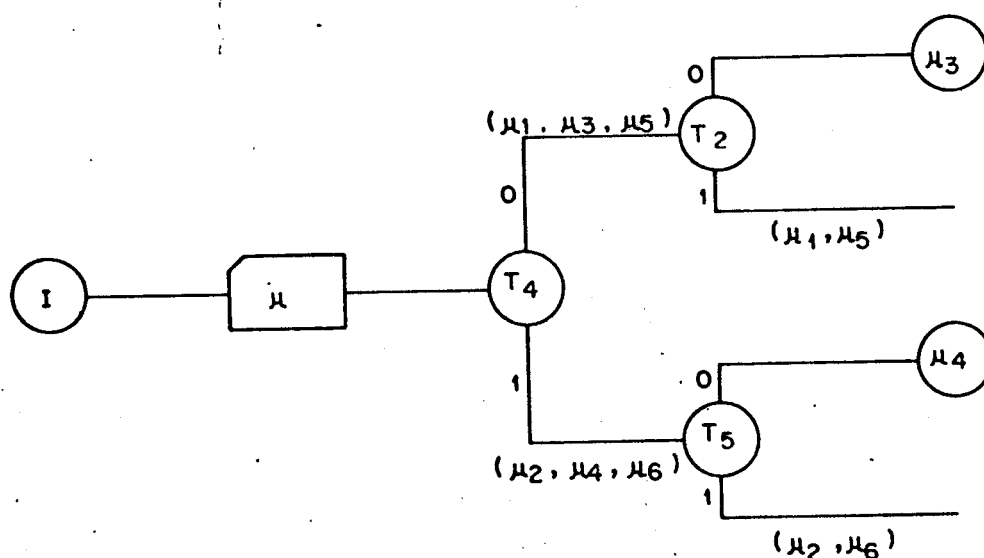
Teste inicial  $T_5$  :  $H(X'_1) = h(0,30) = 0,8813$  bits,  $P_{X_1}(0)=0,30$  e  
 $P_{X_1}(1) = 0,70$

| 2 <sup>os</sup> TESTES | $H(X'_2 \mid X'_1=0)$ | $H(X'_2 \mid X'_1=1)$ |
|------------------------|-----------------------|-----------------------|
| $T_1$                  | $h(0,333)$            | $h(0,4286)$           |
| $T_2$                  | $h(0)$                | $h(0,4286)$           |
| $T_3$                  | $h(0)$                | $h(0,2857)$           |
| $T_4$                  | $h(0,333)$            | $h(0,2857)$           |

Valor máximo de  $H(X'_1 X'_2)$ , dado  $T_5$  como teste inicial:

$$\begin{aligned}
 H(X'_1 X'_2) &= h(0,30 + (h(0,3333) \times 0,30) + (h(0,4286) \times 0,70)) \\
 &= 0,8813 + (0,9183 \times 0,30) + (0,9852 \times 0,70) = 1,8464 \text{ bits.}
 \end{aligned}$$

Assim, o valor máximo encontrado para  $H(X'_1 X'_2)$  foi 1,9710 bits correspondente aos testes  $T_4 T_2 T_5$  ou  $T_4 T_2 T_3$ . Uma dessas se quências de teste é a melhor maneira de se iniciar nosso algoritmo. Tomando-se uma delas teremos:



Notemos que para terminar o algoritmo precisamos ainda de dois testes: um para distinguir  $u_1$  de  $u_5$ , e outro  $u_2$  de  $u_6$ . Aqui não mais será necessário pesquisa para ver qual o teste ótimo para esse ponto, pois todos que fizerem a separação desejada terá o mesmo valor de entropia. Assim, tomando-se quaisquer dois testes que façam as referidas distinções teremos:

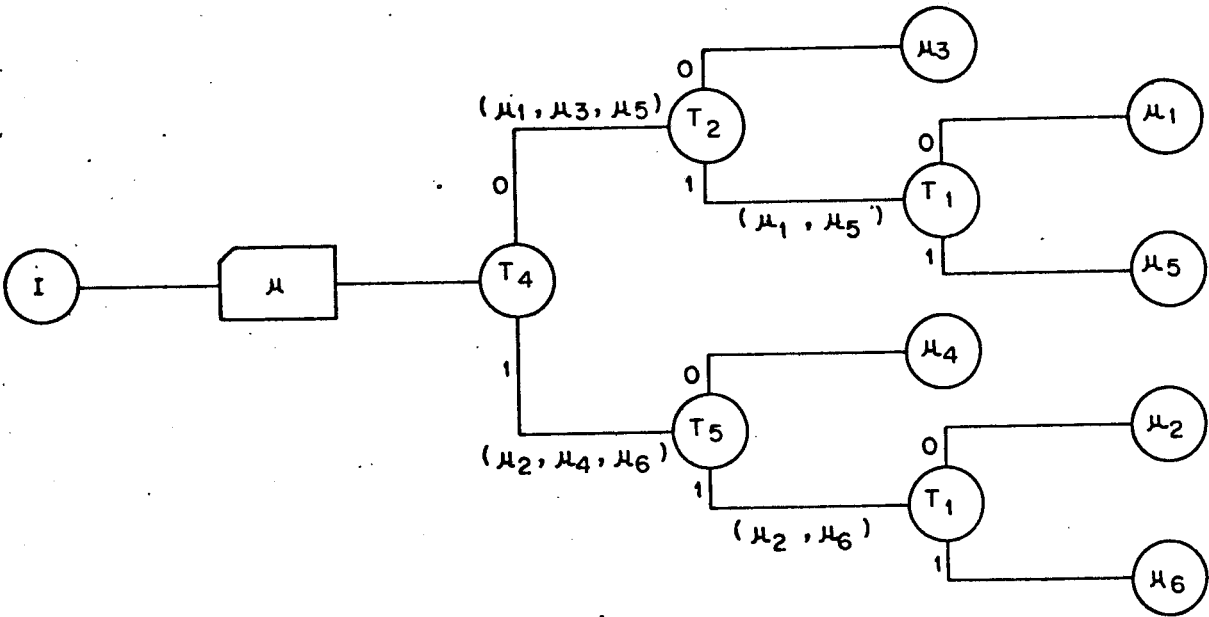


Fig. IV.2 - Algoritmo Ótimo de Ordem (2,1) associado à Tabela de Decisão do Exemplo 1.1.

De sua árvore binária associada temos um Código com PP:

| $u$   | $P_U(u)$ | CÓDIGO Z |   |   |
|-------|----------|----------|---|---|
| $u_1$ | 0,10     | 0        | 1 | 0 |
| $u_2$ | 0,10     | 1        | 1 | 0 |
| $u_3$ | 0,30     | 0        | 0 |   |
| $u_4$ | 0,20     | 1        | 0 |   |
| $u_5$ | 0,20     | 0        | 1 | 1 |
| $u_6$ | 0,10     | 1        | 1 | 1 |

Calculando-se seu comprimento médio, teremos:

$$\begin{aligned}\bar{W} &= 2(0,30 + 0,20) + 3(0,10 + 0,10 + 0,20 + 0,10) = \\ &= 1,00 + 1,50\end{aligned}$$

$$\bar{W} = \underline{2,50 \text{ t/i}}$$

Como podemos ver, esse valor é menor ainda (8%) do que o comprimento médio do Algoritmo TOPO de Massey. ( $\bar{W} = 2,70 \text{ t/i}$ ), do Exemplo 2.8, aplicado a essa mesma tabela de decisão.

Este exemplo nos conduziu, coincidentemente, ao valor ótimo, já que seu comprimento médio foi igual ao do Código de HUFFMAN, calculado no Exemplo 2.6. Apesar disso nem sempre acontecer, aqui temos uma ilustração de como esse novo método de construção nos conduz a algoritmos mais eficientes.

O teorema que se segue nos fornece uma comparação entre o comprimento médio do Algoritmo TOO (1,...1) e o Algoritmo TOPO de Massey, apresentado na Seção II.6. Para tanto seja ...  $\bar{W}_{M_1}$  o comprimento médio do algoritmo TOPO e  $\bar{W}_{M_2}$  do Algoritmo TOO.

Teorema 4.1 :

$$\bar{W}_{M_1} = \bar{W}_{M_2}$$

Prova: Sendo que

$$H(X'_i | X'_1 \dots X'_{i-1}) = \sum_{d_1=0}^{D-1} \sum_{d_2=0}^{D-1} \dots \sum_{d_{i-1}=0}^{D-1} H(X'_i | X'_1 \dots X'_{i-1} = d_1 \dots d_{i-1}) P_{X_1 \dots X_{i-1}}(d_1 \dots d_{i-1}) \quad (4.4)$$

O máximo valor de  $H(X'_i | X'_1 \dots X'_{i-1})$  é atingido quando, no  $i$ -ésimo passo da construção do algoritmo, escolhermos testes que maximizem todos  $H(X'_i | X'_1 \dots X'_{i-1} = x_1 \dots x_{i-1})$ ,  $x_1 \dots x_{i-1} \in Q_{i-1}$ . Uma vez que o algoritmo está construído até o  $(i-1)$ -ésimo passo,  $\bar{W}_i$  será o mesmo para qualquer escolha de testes no  $i$ -ésimo passo que não produza

$$H(X'_i | X'_1 \dots X'_{i-1} = x_1 \dots x_{i-1}) = 0$$

Entretanto se isso acontecer, implica que  $x_1 \dots x_{i-1}$  já identificou uma variável e não é necessário aumentarmos o seu comprimento. Logo, os testes do  $i$ -ésimo passo na construção de um algoritmo que maximizam

$$\frac{H(X'_i | X'_1 \dots X'_{i-1})}{\bar{w}_{X_i}}$$

também maximizam

$$H(X'_i | X'_1 \dots X'_{i-1} = d_1 \dots d_{i-1}), \quad (4.5)$$

onde  $d_1 \dots d_{i-1}$  é o prefixo de, pelo menos, duas variáveis aleatórias. Entretanto, maximizar a Equação (4.5) a cada passo é o critério usado na construção do algoritmo TOPO de Massey.

C.q.d.

Exemplo 4.3: Apresentemos mais uma ilustração relativa ao Exemplo 1.1. Para tanto consideremos sua Tabela de Decisão, reproduzida na Fig. IV.1 do Exemplo 4.1. Construiremos o Algoritmo TOO (1,...,1) esperando obtermos  $\bar{w}_{M2} = 2.70$  t/i, valor já obtido para o  $\bar{w}_{M1}$  do Algoritmo TOPO no Exemplo 2.8 do Capítulo II.

A Equação (4.1) pode ser apresentada sob a forma:

$$H(U) = \left( \frac{H(X'_1)}{\bar{w}_{X_1}} \right) \bar{w}_{X_1} + \left( \frac{H(X'_2 | X'_1)}{\bar{w}_{X_2}} \right) \bar{w}_{X_2} + \left( \frac{H(X'_3 | X'_1 X'_2)}{\bar{w}_{X_3}} \right) \bar{w}_{X_3} + \dots +$$

$$\dots + \left( \frac{H(X'_N | X'_1 \dots X'_{N-1})}{\bar{w}_{X_N}} \right) \bar{w}_{X_N} \quad (4.6)$$

A construção do Algoritmo TOO(1,...,1) então procurará maximizar os termos acima entre parênteses. Assim, o primeiro este escolhido será aquele que maximizar o primeiro termo .....

$\frac{H(X'_1)}{\bar{W}_{X_1}}$  . Como toda palavra código tem comprimento maior ou igual a um, então maximizar o primeiro termo entre parenteses da Equação (4.6) é maximizar  $H(X'_1)$ . Assim:

| 1º TESTE | $\frac{H(X_1)}{\bar{W}_{X_1}}$ |
|----------|--------------------------------|
| $T_1$    | $h(0,50)$                      |
| $T_2$    | $h(0,40)$                      |
| $T_3$    | $h(0,20)$                      |
| $T_4$    | $h(0,40)$                      |
| $T_5$    | $h(0,30)$                      |

Então o teste escolhido para iniciar o Algoritmo (1,...,1) será o  $T_1$ , e o seu início será idêntico à Fig. II.15 (Exemplo 2.8). Analogamente, os segundos testes serão aqueles que maximizem

$$\frac{H(X'_2|X'_1)}{\bar{W}_{X_2}}$$

Mas podemos desdobrar o termo acima em:

$$\frac{H(X'_2|X'_1)}{\bar{W}_{X_2}} = \left( \frac{H(X'_2|X'_1=0)}{\bar{W}_{X_2}} \right) P_{X_1}(0) + \left( \frac{H(X'_2|X'_1=1)}{\bar{W}_{X_2}} \right) P_{X_1}(1), \quad (4.7)$$

de onde podemos concluir que, dado o teste inicial, maximizar o primeiro termo da Equação (4.7) é maximizar  $H(X_2|X_1=0)$  e  $H(X_2|X_1=1)$ . Assim, consultando a tabela do Exemplo 4.1 que tem  $T_1$  como teste inicial podemos ver que o  $T_5$  maximizará ambas entropias acima.

Então o segundo passo da construção do algoritmo é idêntico à Fig.II.17 (Exemplo 2.8). Procedendo analogamente pa-

$H(X'_3 | X'_1 X'_2)$  podemos construir a árvore binária do nosso Algoritmo TOO (1,...,1), totalmente semelhante a da Fig. II.18. E o comprimento do código a esse algoritmo associado será (como também já calculado no Exemplo 2.8),

$$\bar{w}_{M_1} = 2,70 \text{ t/i}$$

#### IV.2 - ESTUDO DA COMPLEXIDADE DE CONSTRUÇÃO DO NOVO ALGORITMO:

Estudemos agora a complexidade do algoritmo TOO ( $L_1, \dots, L_K$ ). O nosso critério de medida será pelo número de entropias a serem calculadas na construção desse algoritmo. Primeiramente vejamos o número máximo de entropias a serem calculadas para que se determine

$$\max \left( \frac{H(X'_1 \dots X'_{L_1})}{\bar{w}_{X_1 \dots X_{L_1}}} \right) .$$

Sabemos que,

$$H(X'_1 \dots X'_{L_1}) = H(X'_1 \dots X'_{L_1-1}) + \sum_{d_1=0}^{D-1} \sum_{d_2=0}^{D-1} \sum_{d_{L_1-1}=0}^{D-1} H(X'_{L_1} | X'_1 \dots X'_{L_1-1} = d_1 \dots d_{L_1-1}) \dots$$

$$\dots P_{X_1 \dots X_{L_1-1}} (d_1 \dots d_{L_1-1}) .$$

Esta complexidade será máxima quando a sub-árvore associada ao Algoritmo TOO ( $L_1, \dots, L_K$ ) no passo  $L_1$ -ésimo for uma árvore completa. Analogamente ao argumento usado na demonstração do teorema 4.1, uma vez dado o valor de  $H(X'_1 \dots X'_{L_1-1})$ , os testes do  $L_1$ -ésimo passo na construção do algoritmo que maximizam

$$\frac{H(X'_1 \dots X'_{L_1})}{\bar{w}_{X_1 \dots X_{L_1}}}$$

são aqueles que também maximizam

$$H(X'_{L_1} | X'_1 \dots X'_{L_1-1} = d_1 \dots d_{L_1-1}),$$

onde  $d_i \in \{0, 1, \dots, D-1\}$  para  $i=1, \dots, L_1-1$ .

Seja NT o número total de testes disponíveis. O número total de diferentes algoritmos até o  $(L_1-1)$ -ésimo passo é dado por:

$$(NT) (NT-1)^D (NT-2)^{D^2} \dots (NT - (L_1-2))^{D(L_1-2)}$$

Assim, o número total de diferentes algoritmos até o  $L_1$ -ésimo passo será:

$$D^{L_1-1} (NT - L_1 + 1) \prod_{i=2}^{L_1} (NT - L_1 + i)^{D(L_1-i)}$$

Agora vejamos essa complexidade para acharmos

$$\max \left( \frac{H(X'_{L_1+1} \dots X'_{L_1+L_2} | X'_1 \dots X'_{L_1})}{W_{X'_{L_1+1} \dots X'_{L_1+L_2}}} \right)$$

Sabemos que

$$H(X'_{L_1+1} \dots X'_{L_1+L_2} | X'_1 \dots X'_{L_1}) = H(X'_{L_1+1} \dots X'_{L_1+L_2-1} | X'_1 \dots X'_{L_1}) +$$

$$+ \sum_{d_1=0}^{D-1} \sum_{d_2=0}^{D-1} \sum_{d_{L_1+L_2-1}=0}^{D-1} H(X'_{L_1+L_2} | X'_1 \dots X'_{L_1+L_2-1} = d_1 \dots d_{L_1+L_2-1}) \dots$$

$$\dots P_{X'_1 \dots X'_{L_1+L_2-1}} (d_1 \dots d_{L_1+L_2-1}) \cdot$$

Novamente usaremos o fato de que os testes do  $(L_1+L_2)$  -ésimo passo, na construção do algoritmo, que maximizam

$$\left( \frac{H(X'_1 \dots X'_{L_1+L_2})}{\bar{W}_{X_1 \dots X_{L_1+L_2}}} \right)$$

São os mesmos que também maximizam

$$H(X'_{L_1+L_2} | X'_1 \dots X'_{L_1+L_2-1} = d_1 \dots d_{L_1+L_2-1}),$$

onde  $d_i \in \{0, 1, \dots, D-1\}$ , para  $i = 1, \dots, L_1+L_2 - 1$ .

O número total de diferentes algoritmos até o  $(L_1+L_2-1)$  -ésimo passo a serem inspecionados, é dado por:

$$\left[ (NT-L_1) (NT-L_1 - 1)^D (NT-L_1-2)^{D^2} \dots \left( (NT-L_1)-(L_2-2) \right)^{D^{(L_2-2)}} \right]^{D^{L_1}}$$

Então esse número até o  $(L_1+L_2)$  -ésimo passo é dado por:

$$D^{(L_1+L_2-i)} (NT-L_1-L_2+1) \prod_{i=2}^{L_2} (NT-L_1-L_2+i)^{D^{(L_1+L_2-i)}}$$

Como a análise foi feita para o caso mais complexo possível, a Complexidade total, CT, para qualquer Algoritmo TOO ...  $(L_1, \dots, L_k)$  será:

$$CT \leq \sum_{i=1}^K \left( D^{(s_i-1)} (NT-s_i+1) \prod_{j=2}^{L_i} (NT-s_i+j)^{D^{(s_i-j)}} \right) \quad (4.8)$$

No caso de  $L_i=1$ , então

$$\prod_{j=2}^{L_i} \left( (NT-s_i+j)^{D^{(s_i-j)}} \right) = 1..$$

O exemplo abaixo dá uma ilustração da Eq.(4.8) acima:

Exemplo 4.3 : Calculando a complexidade do Exemplo 4.1 pelo Eq. (4.8) teremos:

$$NT = 5$$

$$L_1 = 2$$

$$L_2 = 1$$

$$K = 2$$

$$D = 2$$

$$CT \leq \sum \left( 2^{s_i-1} (5 - s_i + 1) \prod_{j=2}^{L_i} (5 - s_i + j) 2^{(s_i-j)} \right)$$

$$\leq \left[ 2^{2-1} (5-2+1) (5-2+2) 2^{(2-2)} \right] + 2^2 (5-3+1) = (2 \times 4 \times 5) + (4 \times 3) =$$

$$CT \leq 52$$

Ao construirmos o algoritmo T00 (2.1) no exemplo 4.2 tivemos uma complexidade igual a 40, concordando com o valor acima obtido. Dessas 40 entropias tivemos várias delas iguais, tendo sido, portanto necessário apenas nos preocupar com 11 tipos (cálculos) diferentes. Além disso, na maioria dos casos (caso binário) esses cálculos se resumem a simples consultas de tabelas, como a apresentada no Apêndice A.

## CAPÍTULO V

## CONCLUSÕES

## CONCLUSÕES:

### V.1 - COMENTÁRIOS FINAIS:

Neste trabalho nos generalizamos o Algoritmo de Teste Ótimo de Primeira Ordem (TOPO) introduzido por Massey [4]. Devemos ressaltar aqui que o Algoritmo Generalizado, isto é, o Algoritmo TOO ( $L_1, \dots, L_K$ ) tem o limite superior tal que,

$$\frac{H(U)}{H_{\min}(L_1, \dots, L_K)} \leq \frac{H(U)}{H_{\min}}$$

Intuitivamente se espera que o Algoritmo TOO ( $L_1, \dots, L_K$ ) seja, pelo menos, tão eficiente quanto o Algoritmo TOPO introduzido por Massey, e isto é confirmado pelo Exemplo 4.1 do capítulo anterior. Entretanto ele apresenta uma complexidade superior ao TOPO, conforme mostra a Equação (4.8), a não ser para o caso em que  $L_1 = L_2 = \dots = L_K = 1$ . Nesse caso essa complexidade de construção será igual para os dois algoritmos, mas sua eficiência também.

Como salientado por Shwayder [8], a solução ótima requer uma procura exaustiva, com exceção de alguns casos especiais em que dispomos dos testes necessários para a construção de um algoritmo, cujo código correspondente seja um Código de HUFFMAN. Essa procura exaustiva, na quase totalidade dos casos, é impraticável. O nosso trabalho proporciona um compromisso entre a complexidade de construção do algoritmo e o valor do limite superior do algoritmo implementado.

### V.2 - SUGESTÃO PARA CONSTRUÇÃO DE ALGORITMOS EFICIENTES

Sugerimos a seguir um procedimento para a construção de algoritmos eficientes:

- 1) Construa todos os Códigos de HUFFMAN associados ao espaço de probabilidades do problema a ser solucionado, e veja se os testes disponíveis são suficientes para a implementação de um deles. Em caso afirmativo esse será o Algoritmo Ótimo.
- 2) Em caso negativo construa o Algoritmo TOPO e compare o seu comprimento,  $\bar{W}_{M1}$ , com o  $\bar{W}_{HUF}$ . Se essa diferença foi aceitável esta será uma das soluções satisfatórias.
- 3) Caso o  $\bar{W}_{M1}$  difira em muito do  $\bar{W}_{HUF}$ , construa o Algoritmo T00 (2,1,...,1) e compare seu comprimento médio,  $\bar{W}_{M2}$ , com o  $\bar{W}_{M1}$ . Caso estes dois valores sejam muito próximos, provavelmente indicará que tanto  $\bar{W}_{M1}$  quanto  $\bar{W}_{M2}$  são valores próximos do  $\bar{W}$  mínimo, e qualquer um dos dois algoritmos podem ser tomados como a solução desejada.
- 4) Caso  $\bar{W}_{M2}$  difira muito do  $\bar{W}_{M1}$  e também do  $\bar{W}_{HUF}$ , construa o Algoritmo TOPO (2,2,1,...,1), e assim sucessivamente até que a solução desejada seja encontrada.

Os exemplos por nós apresentados no desenvolvimento do Trabalho são de caráter ilustrativo. Na prática podemos ter, por exemplo, listas com algo em torno de cem objetos (ou mais) a serem unicamente identificados, com probabilidades de ocorrência que variem, desde 0,30 até  $10^{-5}$ . Nesse caso então a utilização do Algoritmo T00 ( $L_1, \dots, L_K$ ) poderia nos levar a reduções substanciais no tempo de operação de equipamentos sofisticados, em relação ao de Massey.

Finalmente devemos salientar que o trabalho adicional requerido na construção de um Algoritmo T00 ( $L_1, \dots, L_K$ ) será executado apenas uma única vez, enquanto que os reflexos da economia por ele proporcionado serão sentidos a cada vez que for utilizado.

### SUGESTÕES PARA PRÓXIMAS PESQUISAS:

- 1) Provar que o comprimento médio do Algoritmo  $(L'_1, L'_2, \dots, L'_j)$  é menor ou igual ao do Algoritmo  $(L_1, L_2, \dots, L_K)$ , onde  $L'_j = L'_{t_{j-1}+1} \dots L_{t_j}$ ,  
 $j = 1, \dots, J$  e onde  $t_0 = 0$
- 2) Estabelecer um limiar entre as entropias associadas aos testes de distinção, limiar esse que eliminará, a priori, a análise de todos aqueles testes que tiverem entropias abaixo desse valor pré-fixado. Isso iria simplificar e diminuir a complexidade de implementação do Algoritmo TOO  $(L_1, \dots, L_K)$ .
- 3) Desenvolver um programa computacional que, a partir do espaço de probabilidades e da Tabela de Decisão, construa o Algoritmo  $(L_1, \dots, L_K)$  pré-fixado.

APÊNDICE A

VALORES DA FUNÇÃO ENTROPIA BINÁRIA,  $h(p)$

APÊNDICE A: Valores da função Entropia Binária,  $h(p)$ :

| p    | $-p \log p$ | $-(1-p)\log(1-p)$ | $h(p)$ |
|------|-------------|-------------------|--------|
| 0    | 0           | 0                 | 0      |
| 0,01 | 0,0664      | 0,0144            | 0,0808 |
| 0,02 | 0,1129      | 0,0285            | 0,1414 |
| 0,04 | 0,1858      | 0,0565            | 0,2423 |
| 0,06 | 0,2435      | 0,0839            | 0,3274 |
| 0,08 | 0,2915      | 0,1107            | 0,4022 |
| 0,10 | 0,3322      | 0,1368            | 0,4690 |
| 0,12 | 0,3671      | 0,1623            | 0,5294 |
| 0,14 | 0,3971      | 0,1871            | 0,5842 |
| 0,16 | 0,4230      | 0,2113            | 0,6343 |
| 0,18 | 0,4453      | 0,2348            | 0,6801 |
| 0,20 | 0,4644      | 0,2575            | 0,7219 |
| 0,22 | 0,4806      | 0,2796            | 0,7602 |
| 0,24 | 0,4941      | 0,3009            | 0,7950 |
| 0,26 | 0,5053      | 0,3214            | 0,8267 |
| 0,28 | 0,5142      | 0,3412            | 0,8554 |
| 0,30 | 0,5211      | 0,3602            | 0,8813 |
| 0,32 | 0,5260      | 0,3784            | 0,9044 |
| 0,34 | 0,5292      | 0,3956            | 0,9248 |
| 0,36 | 0,5306      | 0,4121            | 0,9427 |
| 0,38 | 0,5304      | 0,4276            | 0,9580 |
| 0,40 | 0,5288      | 0,4422            | 0,9710 |
| 0,42 | 0,5257      | 0,4558            | 0,9815 |
| 0,44 | 0,5212      | 0,4684            | 0,9896 |
| 0,46 | 0,5154      | 0,4800            | 0,9954 |
| 0,48 | 0,5083      | 0,4905            | 0,9988 |
| 0,50 | 0,5000      | 0,5000            | 1,000  |

onde  $h(p) = -p \log p - (1 - p) \log (1 - p)$

## BIBLIOGRAFIA

- [1]- PAPOULIS, A. "PROBABILITY RANDOM VARIABLES AND STOCHASTIC PROCESSES", McGraw-Hill, 1965.
- [2]- ASH, R., "INFORMATION THEORY", John Wiley, New York, 1965.
- [3]- GALLAGER, R.G., "INFORMATION THEORY AND RELIABLE COMMUNICATION", John Wiley, 1968.
- [4]- MASSEY, J.L., "TOPICS INDISCRETE INFORMATION PROCESSING", Unpublished Manuscript, University of Notre Dame, Indiana, 1976.
- [5]- GOEL, D.K., "RANDOM TEST GENERATION FOR FAULT DETECTION AND DIAGNOSIS", Dissertation for Ph.D. Degree, Syracuse University, New York, 1974.
- [6]- HUFFMAN, D.A., "A METHOD FOR CONSTRUCTION OF MINIMUM REDUNDANCY CODES", Proc. IRE 40,10 (Oct. 1952), 1098-1101.
- [7]- SHWAYDER, K., "EXTENDING THE INFORMATION THEORY APPROACH TO CONVERTING LIMITED-ENTRY DECISION TABLES TO COMPUTER PROGRAMS", Communications of the ACM 17,9 (September 1974), 532-537.
- [8]- SHWAYDER, K. "CONVERSION OF LIMITED-ENTRY DECISION TABLES TO COMPUTER PROGRAMS - A PROPOSED MODIFICATION TO POLLACK'S ALGORITHM", Communication of the ACM 14, 2 (February 1971) ; 69-73.
- [9]- VERHELST, M., "THE CONVERSION OF LIMITED-ENTRY DECISION TABLES TO OPTIMAL AND NEAR-OPTIMAL FLOWCHARTS: TWO NEW ALGORITHMS", Communication of the ACM 15,11 (November 1972), 974 - 980.