SOBRE A REDUNDÂNCIA DE GRAU β DE CÓDIGOS BINÁRIOS

ORIENTADOR: INDER JEET TANEJA

CÉLIA JANTSCH FIUZA

AGOSTO DE 1983

SOBRE A REDUNDÂNCIA DE GRAU β DE CÓDIGOS BINÁRIOS

por

CÉLIA JANTSCH FIUZA

Esta dissertação foi julgada adequada para a obtenção do Título de

" MESTRE EM CIÊNCIAS "

Especialidade em Matemática, e aprovada em sua forma fianl pelo

Curso de Pós-Graduação em Matemática da
UNIVERSIDADE FEDERAL DE SANTA CATARINA

Coordenador em Exercício - PGMTM

Prof. WILLIAN GLENN WHITLEY, PhD.
Coordenador

BANCA EXAMINADORA:

Prof. INDER JEET TANEJA, PhD
Orientador

Prof. GUR DIAL, PhD.

Prof. AUGUSTO C.GADELHA VIEIRA, PhD.

AO

ARMANDO

PATRÍCIA

LEONARDO

ARMANDO JR.

RESUMO

Shannon [1948] provou que é sempre possível construir um código unicamente decifrável tal que $0 \leq r < 1$, onde r é a redundância.

Em[1978], Gallager obteve o seguinte limite superior,

$$r \leq P_1 + \sigma$$

onde P_1 é a probabilidade da palavra mais provável, e

$$\sigma = 1 - \log_2 e + \log_2(\log_2 e) \cong 0{,}086$$

Gallager também mostrou que para $P_1 \geq 0{,}5$ a redundância é limitada superiormente por

$$r \leq 2 - H(P_1, 1-P_1) - P_1$$

onde

$$H(P_1, 1-P_1) = -P_1 \log_2 P_1 - (1-P_1) \log_2 (1-P_1).$$

Em [1980] , Jonhsen melhorou estes resultados de Shannon e Gallager apresentando novos limites superiores e inferiores para a redundância.

No capítulo I apresentamos algumas definições e resultados básicos. os conceitos de entropia e redundância de grau β são introduzidos, definidos respectivamente por

$$H^{\beta}(P) = (2^{1-\beta} - 1)^{-1} \left(\sum_{i=1}^{N} P^{\beta} - 1 \right), \quad \beta \neq 1 \text{ e } \beta > 0$$

e

$$r^{\beta} = L(\beta) - H^{\beta}(P)$$

onde

$$L(\beta) = \sum_{i=1}^{N} P_i n_i(\beta)$$

é o comprimento médio de grau β.

Conforme Bouchon [1978] o comprimento médio de grau β é limitado por

$$H^{\beta}(P) \leq L(\beta) < H^{\beta}(P) + 1 \ , \beta > 1 \ .$$

o que constitui uma generalização do teorema de codificação para canais sem ruído.

No capítulo II analisamos os resultados de Johnsen e retificamos alguns pequenos erros presentes em Johnsen [1980].

No capítulo III apresentamos limites superiores e inferiores para a redundância de grau β, que constituem a contribuição principal desta tese.

ABSTRACT

Shannon [1948] proved that it is always possible to construct a uniquely decipherable code such that $0 \leq r < 1$, where r is the redundancy obtained.

In [1978] Gallager obtained the following upper bound:

$$r \leq P_1 + \sigma$$

where P_1 is the probability of the most probable word, and

$$\sigma = 1 - \log_2 e + \log_2 (\log_2 2) \cong 0{,}086$$

Gallager also showed that for $P_1 \geq 0{,}5$, the upper bound on redundancy is given by

$$r \leq 2 - H(P_1, 1-P_1) - P_1$$

where

$$H(P_1, 1-P_1) = - P_1 \log_2 P_1 - (1-P_1)\log_2(1-P_1).$$

In [1980], Johnsen improved the results of Shannon and Gallager, by giving new upper and lower bounds to redundancy.

In Chapter I we present some basic results and definitions. The concepts of entropy and redundancy of degree β are defined, respectively by:

$$H^{\beta}(P) = (2^{1-\beta} - 1)^{-1} (\sum_{i=1}^{N} P^{\beta} - 1), \quad \beta \neq 1, \ \beta > 0$$

and

$$r^{\beta} = L(\beta) - H^{\beta}(P),$$

where

$$L(\beta) = \sum_{i=1}^{N} P_i \, n_i(\beta),$$

is the mean length of degree β.

Bouchon [1978] showed that the mean length of degree β is limited by

$$H^{\beta}(\beta) \leq L(\beta) < H^{\beta}(P) + 1 \ , \ \beta > 1$$

which generalizes the coding theorem for noiseless channels.

In chapter II we present Johnsen's results [1980] and correct some minor errors that appear in his paper.

In chapter III we give new upper and lower bounds for the redundancy of degree β that constitute the main contribution of this thesis.

I N D I C E

INTRODUÇÃO

O objetivo desse trabalho é apresentar novos limites superior e inferior de redundância de grau β respectivamente para os intervalos $P_1 \geq 0,4$ e $0,4 \leq P_1 < 0,5$.

Dividimos o trabalho em três capítulos:

No capítulo I , apresentamos entropia de uma distribuição de probabilidades (Shannon), suas propriedades, codificação sem ruído, redundância e sua generalização para entropia de grau β.

No capítulo II, analisamos o trabalho realizado por Johnsen [1980] sobre limites inferior e superior de redundância, o qual serviu como base para mostrar limites superior e inferior de redundância de grau β.

No capítulo III, generalizamos limites superior e inferior de redundância de grau β, o qual apresentamos através de dois teoremas para os intervalos $P_1 \geq 0,4$ e $0,4 \leq P_1 < 0,5$, respectivamente.

CAPÍTULO I

REDUNDÂNCIA E SUA GENERALIZAÇÃO PARA ENTROPIA DE GRAU β

INTRODUÇÃO

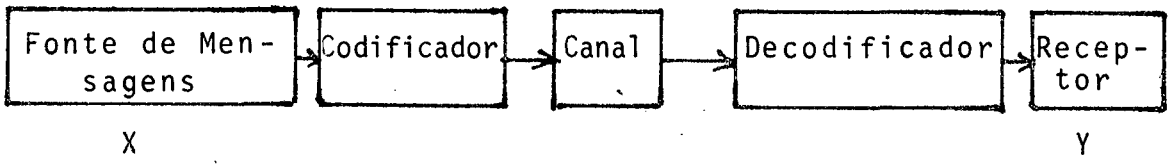
Shannon [1948] desenvolveu a Teoria da Informação, que inicialmente só tratava de problemas relacionados a transmissão de informações, mas hoje devido ao grande interesse por ela despertado, possue grandes aplicações em vários campos do conhecimento, tais como: ciências exatas, linguística, nos sistemas de comunicações, etc.

A idéia de Shannon era o de associar quantitativamente informação à probabilidade de um evento.

Se temos certeza de um evento e este ocorre, o que se obtem em informação é pouca coisa. Por outro lado,se um evento altamente improvável tem lugar, a quantidade de informação assim adquirida é muito grande.

O problema fundamental da Teoria de Informação é o de quantificar o conteúdo informativo de um conjunto de mensagens.

Vamos apresentar um esquema de um sistema da comunicação:



Fonte de Mensagens:

É o componente do sistema que é capaz de produzir mensagens. Geralmente a fonte é representada por uma variável aleatória X.

Codificador:

Transforma a linguagem da fonte para a linguagem do canal, mantendo inalterado seu conteúdo.

Canal:

É o meio através do qual a mensagem é propagada.

Decodificador:

Decifra a mensagem recebida de modo que a mesma seja inteligível pelo receptor.

Receptor:

É o ponto de destino da mensagem é representado pela variável aleatória Y.

1.1 - Entropia de uma distribuição de probabilidade (Shannon).

Seja $X = \{ x_1, x_2, \ldots x_N \}$ uma variável aleatória discreta com distribuição de probabilidade $P \in \Delta_N$, onde $P = (P_1, P_2, \ldots, P_N) \in \Delta_N$.

$$\Delta_N = \{ P = (P_1, P_2, \ldots, P_N) : P_i \geq 0 \ , \ \sum_{i=1}^{N} P_i = 1 \ ,$$

$$i = 1,2, \ldots , N\} \qquad (1.1)$$

A entropia (Shannon) associada à distribuição de probabilidade P é definida por:

$$H(P) = H(P_1, P_2, \ldots, P_N) = - \sum_{i=1}^{N} P_i \log P_i \qquad (1.2)$$

Observação:

Quando não indicarmos a base do logarítmo, esta será sempre na base 2.

Apresentamos agora algumas propriedades da entropia dada em (1.2).

Para demonstração veja Aczél e Daróczy [1975].

1.2 - Propriedades da entropia.

1.2.1 - Não Negativa.

$H(P_1,P_2,\ldots, P_N) \geq 0$ com a igualdade se e somente se, $P_i = 1$ para algum i e $P_j = 0$, $j = 1, 2, \ldots, N$, $i \neq j$.

1.2.2 - Simétrica:

$H_N(P_1,P_2, \ldots, P_N) = H_N (P_{i(1)}, P_{i(2)}, \ldots, P_{i(N)})$ para todo $(P_1,P_2, \ldots, P_N) \in \Delta_N$, onde i é uma permutação arbitrária em $\{1,2, \ldots, N\}$.

1.2.3 - Normalidade:

$H_2(1/2,1/2)=1$

1.2.4 - Expansibilidade:

$H_{N+1}(P_1,P_2, \ldots , P_N,0) = H_N(P_1,P_2, \ldots, P_N)$.

1.2.5 - Decisividade:

$H_2(1,0) = H_2(0,1) = 0$

1.2.6 - Aditividade:

$H_{MN}(PQ) = H_N(P) + H_M(Q)$, onde $P=(P_1,P_2, \ldots,P_N)\in\Delta_N$, $Q=(Q_1,Q_2, \ldots, Q_N) \in \Delta_M$

$PQ=(P_1Q_1 ,P_2Q_2,\ldots,P_1Q_M,P_2Q_1,\ldots,P_2Q_M,\ldots,P_NQ_1,\ldots,P_NQ_M) \in \Delta_{NM}$.

1.2.7 Recursividade:

$$H_N(P_1,P_2,\ldots,P_N) = H_{N-1}(P_1+P_2,P_3,\ldots,P_N)+(P_1+P_2)\cdot$$

$$\cdot H_2\left(\frac{P_1}{P_1+P_2},\frac{P_2}{P_1+P_2}\right), \text{ com } P_1+P_2>0.$$

1.2.8 - Aditividade Forte:

$$H_{MN}(P_1Q_{11},P_1Q_{12},\ldots,P_1Q_{1N},P_2Q_{21},\ldots,P_MQ_{1M},\ldots,P_MQ_{MN}) =$$

$$H_M(P_1,P_2,\ldots,P_M) + \sum_{j=1}^{M} P_j H_N(Q_{j1},Q_{j2},\ldots,Q_{jN}),$$

para todo $(P_1,P_2,\ldots,P_M)\in\Delta_M$, $(Q_{j1},Q_{j2},\ldots,Q_{jN})\in\Delta_N$;

$j = 1,2,\ldots,M$.

1.2.9 - Continuidade:

$H_N(P_1,P_2,\ldots,P_N)$ é uma função contínua das n-variáveis $P_1,P_2,\ldots,P_N$.

1.2.10 - Propriedade de Soma:

$$H_N(P_1,P_2,\ldots,P_N) = \sum_{i=1}^{N} h(P_i), \text{ onde } h(P_i)=-P_i\log P_i,$$

$P_i\in[0,1]$, com $0 \log 0 = 0$

1.2.11 - Maximalidade:

$$H_N(P_1,P_2, \ldots , P_N) \leq \log N = H_N(1/N,1/N,\ldots,1/N)$$

isto é, a entropia é máxima quando todas as probabilidades são iguais.

1.2.12 - Monotonicidade:

$H_N(1/N,1/N,\ldots,1/N)$ é uma função monótona crescente de N, isto é: se $\phi(N)=H_N(1/N, \ldots ,1/N)$ então $\phi(N)\leq\phi(N+1)$.

1.2.13 - Desigualdade:

Para qualquer $(P_1,P_2, \ldots, P_M)\in \Delta_M$ e

$(Q_{j1},Q_{j2}, \ldots, Q_{jN}) \in \Delta_N$, j= 1,2, ..., M , temos:

$$\sum_{j=1}^{M} P_j H_N(Q_{j1},Q_{j2}, \ldots, Q_{jN}) \leq H_N \left(\sum_{j=1}^{M} P_j Q_{j1}, \sum_{j=1}^{M} P_j Q_{j2}, \ldots, \sum_{j=1}^{M} P_j Q_{jN}\right).$$

1.2.14 - Desigualdade de Shannon:

Para $(P_1,P_2, \ldots, P_M)\in \Delta_M$ e $(Q_1,Q_2,\ldots,Q_N)\in \Delta_N$

com $Q_i>0$ temos:

$$-\sum_{i=1}^{N} P_i \log P_i \leq -\sum_{i=1}^{N} P_i \log Q_i$$

1.2.15 - Sub-aditividade Forte:

$$H_{MN}(P_{11},P_{12},\ldots,P_{1N},P_{21},\ldots,P_{2N},\ldots,P_{M1}\ldots,P_{MN})$$

$$\leq H_M\left(\sum_{i=1}^{N}P_{1i},\ \sum_{i=1}^{N}P_{2i},\ \ldots,\sum_{i=i}^{N}P_{Ni}\right) + H_N\left(\sum_{j=1}^{M}P_{j1},\sum_{j=1}^{M}P_{j2},\ \ldots,\ \sum_{j=1}^{M}P_{jN}\right),$$

para todo $(P_{11},P_{12},\ \ldots,\ P_{1N},\ldots\ P_{M1},\ \ldots,\ P_{MN}) \in \Delta_{MN}$.

A seguir apresentamos as definições:

Palavra Código:

Cada símbolo x_i associado com uma sequência finita do alfabeto código é chamado palavra código.

Exemplo:

$$x_1 \longrightarrow a_1a_2a_3 = W_1$$

$$x_2 \longrightarrow a_2a_3a_4a_5 = W_2$$

$$\vdots \qquad\qquad \vdots$$

$$x_N \longrightarrow a_1a_2a_3a_4a_5 = W_N$$

Onde $X = \{x_1,x_2,\ \ldots,\ x_N\}$ é o conjunto de mensagens e $A = \{a_1,a_2,a_3,\ \ldots,\ a_D\}$ é o conjunto alfabeto código com dimensão D.

Comprimento da palavra código

É o número de elementos do alfabeto código em uma palavra código.

Comprimento médio da palavra código

$L = \sum_{i=1}^{N}P_in_i$, onde n_i é o comprimento da palavra

associada ao evento x_i, $i = 1,2, \ldots, N$.

Código:

É a coleção de todas as palavras códigos.

Exemplo:

$W = \{W_1, W_2, \ldots, W_N\}$

1.3 - Codificação sem ruído

Um canal é sem ruído se ele permite uma transmissão perfeita da entrada à saída. Isto quer dizer que não precisamos nos preocupar com o problema de correção de erro.

Se a palavra código associada com x_i, de comprimento n_i para todo $i = 1,2, \ldots, N$, tem probabilidade P_i, tentamos escolher códigos nos quais $\sum_{i=1}^{N} P_i n_i$ é mínimo. Este é o objetivo da codificação sem ruído.

1.3.1 - Código Decifrável unicamente

O código é decifrável unicamente se cada seqüência do alfabeto código corresponde, no máximo, a uma mensagem.

1.3.2 - Código Instantâneo

Se um código tem a propriedade que nenhuma palavra código é prefixo de outra então o código é chamado "código instantâneo".

Cada código instantâneo é decifrável unicamente, mas a recíproca é falsa.

Exemplo:

{0,01,011} é um código decifrável unicamente, mas não é instantâneo.

1.3.3 - Teorema - (Codificação sem ruído)-(Ash[1965])

Dada uma variável aleatória $X = \{X_1, \ldots, X_N\}$ com entropia H(P), existe um código instantâneo de base D (ou dimensão D) cujo comprimento médio L das palavras códigos satisfaz

$$\frac{H(P)}{\text{Log } D} \leq L < \frac{H(P)}{\log D} + 1 \qquad (1.3)$$

Observação:

Quando D=2, isto é, $\log_2 D = \log_2 2 = 1$, neste caso o código é chamado binário e (1.3) reduz-se para:

$$H(P) \leq L < H(P) + 1 \qquad (1.4)$$

1.4 - Redundância

1.4.1 - Definição

A redundância r do código da fonte é definida como sendo o comprimento médio das palavras códigos menos a entropia $H (P_1, P_2, \ldots , P_N)$, isto é:

$$r = \sum_{i=1}^{N} P_i n_i - H(P_1, P_2, \ldots , P_N) \qquad (1.5)$$

Conhecidos os valores $P_1, \ldots , P_N$, Huffman 1952 descreveu um procedimento para a construção de um código ótimo, isto é, com redundância mínima satisfazendo obviamente os limites estabelecidos por Shannon, isto é:

$$0 \leq r < 1 \qquad (1.6)$$

1.4.2 - Limite superior da redundância

Vamos citar um teorema que nos dá o limite superior da redundância quando é conhecido apenas a probabilidade P_1 da palavra mais provável. No Capítulo II apresentamos os resultados de Johnsen [1980] que dão novos limites superiores para $0,4 \leq P_1 < 0,5$ e limites inferiores para $P_1 \geq 0,4$, que são melhores que o apresentado por Gallager [1978] conforme o Teorema seguinte.

1. 4 . 3 - Teorema: (Gallager [1978]).

Seja P_1 a probabilidade da palavra mais provável numa fonte discreta finita. Então a redundância do código binário para a fonte satisfaz:

$$r \leq P_1 + \sigma \tag{1.7}$$

onde $\sigma = 1 - \log_2 e - \log_2 (\log_2 e) \cong 0,086$.

Para $P_1 \geq 1/2$

$$r \leq 2 - H(P_1, 1 - P_1) - P_1 \tag{1.8}$$

1.5 - Entropia de Grau β

Havrda e Charvát [1967] e Daróczy [1970] introduziram o conceito de entropia de grau β para uma distri

buição de probabilidade $(P_1, P_2, \ldots, P_N) \in \Delta_N$ de uma variável aleatória discreta $X = \{x_1, x_2, \ldots, x_N\}$ definindo-a por:

$$H^{\beta}(P) = (2^{1-\beta} - 1)^{-1} \left(\sum_{i=1}^{N} P_i^{\beta} - 1 \right), \beta \neq 1 \quad e \quad \beta > 0. \tag{1.9}$$

É fácil de verificar que:

$\lim_{\beta \to 1} H^{\beta}(P) = H^{1}(P) = -\sum_{i=1}^{N} P_i \log P_i$, que é a entropia de Shannon.

1.6 - Propriedades da entropia de grau β .

A entropia $H_N^{\beta} : \Delta_N \to \mathbb{R}$ (reais) de grau β definida em (1.9) tem as seguintes propriedades: (Para demonstração veja Aczél e Daróczy [1975]).

1.6.1 - Não Negativa.

$H_N^{\beta}(P_1, P_2, \ldots, P_N) \geq 0$ com a igualdade se, e somente se, $P_i = 1$ para algum $i = 1,2,\ldots,N$ e $P_j = 0$, $j \neq i$.

1.6.2 - Simétrica.

$H_N^{\beta}(P_1, P_2, \ldots, P_N) = H_N^{\beta}(P_{i(1)}, \ldots, P_{i(N)})$ para todo $(P_1, P_2, \ldots, P_N) \in \Delta_N$ e i é uma permutação sobre $\{1,2,3,4, \ldots, N\}$.

1.6.3 - Normalidade:

$$H_2^{\beta}\ (1/2,\ 1/2) = 1$$

1.6.4 - Expansbilidade:

$$H_N^{\beta}(P_1,P_2,\ \ldots,\ P_N) = H_{N+1}^{\beta}(P_1,P_2,\ \ldots P_N,0).$$

1.6.5 - Propriedade de soma:

$$H_N^{\beta}(P_1,P_2,\ \ldots\ ,\ P_N) = \sum_{i=1}^{N} f\beta\ (P_i) \text{ onde}$$

$$f_{\beta}(P_i) = (2^{1-\beta} - 1)^{-1}\ [P_i^{\beta} - P_i],\ \beta \neq 1,\ \beta > 0$$
$$\text{e } i = 1,2,\ \ldots,\ N.$$

1.6.6 - Decisividade:

$$H_2^{\beta}\ (1,0) = H_2^{\beta}\ (0,1) = 0$$

1.6.7 - Recursividade:

$$H_N^{\beta}\ (P_1,P_2,\ \ldots,P_N) = H_{N-1}^{\beta}(P_1+P_2,P_3,\ldots,P_N) +$$

$$(P_1+P_2)^{\beta}\ H_2^{\beta}\ \left(\frac{P_1}{P_1 + P_2},\ \frac{P_2}{P_1 + P_2}\right),\quad \text{com } P_1+P_2>0$$

1.6.8 - Aditividade forte:

$$H^{\beta}_{NM}\ (P_1Q_{11},P_1Q_{12},\ldots,P_1Q_{1M},P_2Q_{21},\ldots,P_2Q_{2M},\ldots,$$

$$P_NQ_{N1},\ \ldots,P_NQ_{NM}) = H^{\beta}_{N}(P_1,P_2,\ \ldots,P_N)\ +$$

$$\sum_{i=1}^{N} P_i^{\beta}\ H^{\beta}_{M}\ (Q_{i1},Q_{i2},Q_{i3},\ \ldots,\ Q_{iM}).$$

1.6.9 - Não Aditividade de grau β

$$H^{\beta}_{NM}(P_1Q_1,\ldots,P_1Q_M,P_2Q_1,\ \ldots,P_2Q_M,\ldots,P_NQ_1,\ldots P_NQ_M) =$$

$$= H^{\beta}_{N}\ (P_1,P_2,\ldots,P_N)\ +\ H^{\beta}_{M}(Q_1,Q_2,\ldots,Q_M)+(2^{1-\beta}-1).$$

$$H^{\beta}_{N}\ (P_1,P_2,\ P_3,\ \ldots,P_N)\ H^{\beta}_{M}(Q_1,Q_2,\ \ldots,Q_M).$$

1.6.10 - Continuidade:

$H^{\beta}_{N}(P)$, $\beta > 0$, $\beta \neq 1$ é uma função contínua nas suas N-variáveis P_1, P_2, ..., P_N.

1.6.11 - Maximalidade:

$H^{\beta}_{N}(P)$ é uma função máxima para $\beta > 0$, quando todas as probabilidades são iguais, isto é:

$$H^{\beta}_{N}(P_1,P_2,\ \ldots,P_N) \leq H^{\beta}_{N}\ (1/N,1/N,\ \ldots,1/N),\ \beta > 0$$

1.6.12 - Desigualdade

Para $\beta \geq 1$, temos

$$H_N^{\beta}\left(\sum_{j=1}^{M} P_j Q_{j1}, \sum_{j=1}^{M} P_j Q_{j2}, \ldots, \sum_{j=1}^{M} P_j Q_{jN}\right) \geq \sum_{j=1}^{M} P_j^{\beta}$$

$H_N(Q_{j1}, Q_{j2}, Q_{j3}, \ldots, Q_{jN})$, onde $(P_1, P_2, \ldots, P_M) \in \Delta_M$

e $(Q_1, Q_2, \ldots, Q_N) \in \Delta_N$ $\forall j = 1,2, \ldots, M.$

1.6.13 . Sub-Aditividade:

$$H_{MN}^{\beta}(P_{11}, P_{12}, \ldots, P_{1N}, \ldots P_{M1}, \ldots, P_{MN}) \leq$$

$$H_M^{\beta}\left(\sum_{j=1}^{N} P_{1j}, \ldots, \sum_{j=1}^{N} P_{Mj}\right) + H_N^{\beta}\left(\sum_{i=1}^{M} P_{1j}, \ldots, \sum_{i=1}^{M} P_{Nj}\right),$$

$\cdot \beta > 1.$

A caracterização de entropia de grau β juntamente com a entropia de Shannon pode ser visto em Sharma e Taneja [1975] e Taneja [1975] ,[1977].

1.7 - TEOREMA DE CODIFICAÇÃO

Nesta seção apresentamos uma generalização do Teorema de Codificação (Bouchon [1978]) para entropia de grau β. Iniciamos primeiro com a definição de comprimento mêdio de grau β.

1.7.1 - Definição:

Comprimento mêdio de grau β:

Generalizamos o comprimento n_i da palavra codigo para uma função β dada como

$$n_i(\beta) = \frac{2^{(1-\beta)n_i} - 1}{2^{1-\beta} - 1}, \quad \beta \neq 1, \ \beta > 0 \qquad i=1,2,\ldots,N \ . \quad (1.10)$$

É fácil verificar que

$$\lim_{\beta \to 1} n_i(\beta) = n_i$$

Definimos o comprimento médio de grau β por:

$$L(\beta) = \sum_{i=1}^{N} P_i n_i(\beta)$$

$$= \sum_{i=1}^{N} P_i \left(\frac{2^{(1-\beta)n_i} - 1}{2^{(1-\beta)} - 1} \right),$$

$$= \left(2^{1-\beta} - 1 \right)^{-1} \left\{ \sum_{i=1}^{N} P_i 2^{(1-\beta)n_i} - 1 \right\},$$

$$\beta \neq 1 \ , \ \beta > 0$$

Prova-se agora o teorema que nos dá um limite superior e um limite inferior de comprimento médio de grau β.

1. 7.2 - Teorema (Bouchon [1978])

Dada uma variável aleatória $X= \{x_1, x_2, \ldots, x_N\}$ com entropia de grau β, $H^{\beta}(P)$, existe um código binário instantâneo cujo comprimento médio de grau β das palavras códigos $L(\beta)$ satisfaz

$$H^{\beta}(P) \leq L(\beta) < 2^{1-\beta} H^{\beta}(P)+1, \quad \beta > 0 \qquad (1.12)$$

E para $\beta \geq 1$ temos

$$H^{\beta}(P) \leq L(\beta) < H^{\beta}(P) + 1 . \qquad (1.13)$$

Demonstração:

(1.13) segue-se imediatamente de (1.12), pois para $\beta > 1$, $2^{1-\beta} < 1$ isto é;

$$L(\beta) < 2^{1-\beta} H^{\beta}(P) + 1 < H^{\beta}(P) + 1$$

Para $\beta=1$ em (1.12) ou (1.13), o resultado é válido pela desigualdade (1.4). Provamos (1.12) quando $\beta \neq 1$ e $\beta > 0$.

Sabemos que, é sempre possível construir um código binário (D=2) (Ash [1965]) instantâneo /decifrável unicamente com comprimento das palavras códigos n_i (i= 1,2, ..., n) satisfazendo

$$- \log_2 P_i \leq n_i < - \log_2 P_i + 1 , \quad P_i > 0$$

isto é: $$\frac{1}{P_i} \leq 2^{n_i} < \frac{2}{P_i} \qquad (1.14)$$

Provou-se o resultado (1.12) em dois casos separados isto é, quando $0 < \beta < 1$ e $\beta > 1$.

Caso 1:

Quando $0<\beta<1$ (1.13) implica

$$\left(\frac{1}{P_i}\right)^{1-\beta} - 1 \leq 2^{(1-\beta)n_i} - 1 < 2^{1-\beta} \cdot P_i^{\beta-1} - 1, \tag{1.15}$$

Multiplicando (1.14) por $\frac{1}{2^{1-\beta} - 1}$ (que é maior que zero) temos:

$$\frac{P_i^{\beta-1} - 1}{2^{1-\beta} - 1} \leq \frac{2^{(1-\beta)n_i} - 1}{2^{1-\beta} - 1} < \frac{2^{1-\beta} P_i^{\beta-1} - 1}{2^{1-\beta} - 1}, \tag{1.16}$$

Multiplicando (1.15) por P_i e somando para todos os $i=1,2,\ldots,N$, temos:

$$\frac{\sum_{i=1}^{N} P_i (P_i^{\beta-1} - 1)}{2^{1-\beta} - 1} \leq \frac{\sum_{i=1}^{N} P_i (2^{(1-\beta)n_i} - 1)}{2^{1-\beta} - 1} <$$

$$< \frac{\sum_{i=1}^{N} P_i (2^{1-\beta} P_i^{\beta-1} - 1)}{2^{1-\beta} - 1},$$

$$\therefore \quad \frac{\sum_{i=1}^{N} P_i^{\beta} - 1}{2^{1-\beta} - 1} \leq \frac{\sum_{i=1}^{N} P_i (2^{(1-\beta)n_i} - 1)}{2^{1-\beta} - 1} <$$

$$< \frac{2^{1-\beta} \sum_{i=1}^{N} P_i^{\beta} - 1}{2^{1-\beta} - 1},$$

$$\therefore \quad \frac{\sum_{i=1}^{N} P_i^{\beta} - 1}{2^{1-\beta} - 1} \leq \frac{\sum_{i=1}^{N} P_i (2^{(1-\beta)n_i} - 1)}{2^{1-\beta} - 1} <$$

$$< \frac{2^{1-\beta} \sum_{i=1}^{N} P_i^{\beta} - 1 + 2^{1-\beta} - 2^{1-\beta}}{2^{1-\beta} - 1},$$

$$\therefore \quad \frac{\sum_{i=1}^{N} P_i^{\beta} - 1}{2^{1-\beta} - 1} \leq \frac{\sum_{i=1}^{N} P_i (2^{(1-\beta)n_i} - 1)}{2^{1-\beta} - 1} <$$

$$< \frac{2^{1-\beta} \left(\sum_{i=1}^{N} P_i^{\beta} - 1 \right)}{2^{1-\beta} - 1} + \frac{2^{1-\beta} - 1}{2^{1-\beta} - 1},$$

isto é:

$$H^{\beta}(P) \leq L(\beta) < 2^{1-\beta} \; H^{\beta}(P) + 1 \;, \; 0 < \beta < 1 \tag{1.17}$$

Caso II:

Quando $\beta > 1$

Neste caso a desigualdade (1.14) foi invertida, mas como para $\beta > 1$, $2^{1-\beta} - 1 < 0$, isto é $\frac{1}{2^{1-\beta} - 1} < 0$ isto implica que a desigualdade (1.16) fica o mesmo, isto é:

$$H^{\beta}(P) \leq L(\beta) < 2^{1-\beta} \; H^{\beta}(P) + 1,$$

para todo $\beta > 0$.

1.8 - Redundância de Grau β

Definimos a redundância de grau β como a diferença de comprimento médio de grau β e a entropia de grau β, isto é:

$$r(\beta) = L(\beta) - H^{\beta}(P)$$

$$r(\beta) = \frac{\sum_{i=1}^{N} P \left(2^{(1-\beta)n_i} - 1 \right)}{2^{1-\beta} - 1} - \frac{\sum_{i=1}^{N} P_i^{\beta} - 1}{2^{1-\beta} - 1} \;,$$

$$r(\beta) = (2^{1-\beta} - 1)^{-1} \left\{ \sum_{i=1}^{N} (P_i 2^{(1-\beta)n_i} - P_i^{\beta}) \right\} \quad (1.1.7)$$

$$\beta \neq 1 \quad e \quad \beta > 0.$$

No capítulo II apresentaremos os limites superior e inferior de redundância desenvolvidos por Johnsen [1980], e no Capítulo III faremos uma generalização desses limites superior e inferior de redundância de grau β.

CAPÍTULO II

LIMITES SUPERIOR E INFERIOR DE REDUNDÂNCIA.

Neste capítulo analisamos os limites para a redundância do código binário do trabalho realizado por Johnsen [1980].

Em particular, Johnsen estabeleceu novos limites superior para o intervalo $0,4 \leq P_1 < 0,5$ e inferior para $P_1 \geq 0,4$, onde P_1 é a redundância da palavra mais provável da fonte.

2.1 - Limite Superior e Inferior de Redundância

Suponhamos que $A = \{a_1, a_2, \ldots, a_N\}$ é uma fonte discreta de N palavras independentes $2 \leq N < \infty$, com P_i representando a probabilidade da palavra a_i, $i = 1,2,\ldots, N$.

Sem perda de generalidade, seja a_1 a palavra mais provável e n_i o comprimento da palavra binária a_i.

A redundância r de tal código é definida por (1.5).

Apresentamos agora os teoremas demonstrados por Johnsen 1980 .

2.1 - Teorema (Johnsen [1980]).

A palavra mais provável a_1 de uma fonte pode ser codificada por um único "bit" com o código de Huffman quan

do sua probabilidade de ocorrência P_1 satisfizer:

$$P_1 \geq 0,4 \qquad (2.1)$$

Demonstração:

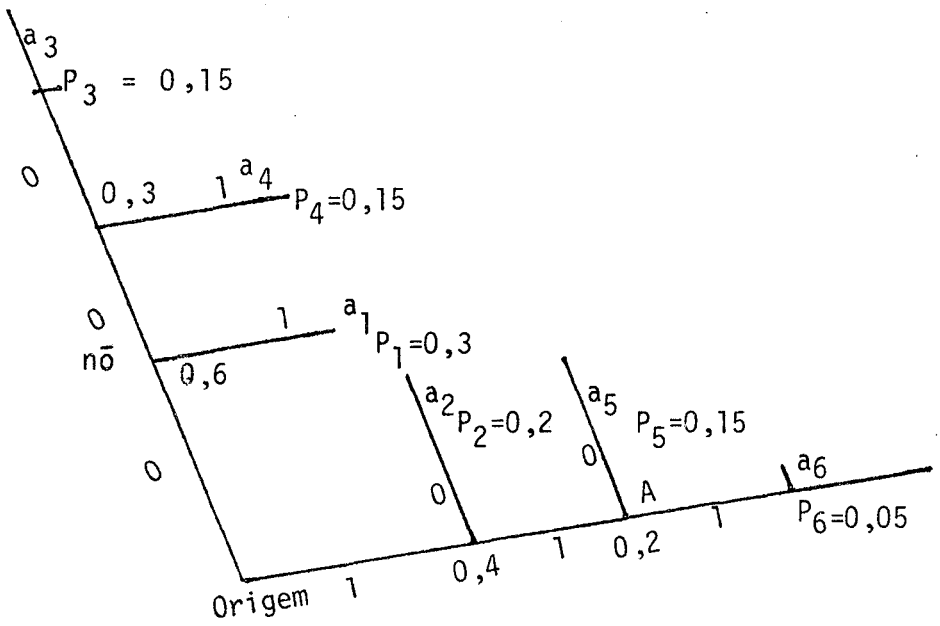


Figura 1 : Exemplo de uma árvore binária de Huffman.

Consideremos um código binário de Huffman e sua árvore correspondente (ver figura 1). A árvore apresenta uma série de ramos que identificam os símbolos códigos (0 ou 1). Cada ramo é terminado por um nó ao qual é associado uma probabilidade. Os nós terminais da árvore caracterizam as letras $\{a_i\}$ do alfabeto original da fonte com suas respectivas probabilidades $\{P_i\}$. A probabilidade de um nó interno é determinada somando-se as probabilidades dos nós terminais relativos aos ramos emergentes do nó em questão. Assim, a probabilidade associada a origem será necessaria-

mente 1.

Pela regra de Huffman devemos construir a árvore associando inicialmente as duas menores probabilidades (no exemplo 0,05 e 0,15 dando 0,20 para o nó A) obtendo-se um novo conjunto de probabilidades (isto é: 0,15;0,15;0,2;0,2; 0,3). Repete-se o processo associando-se as duas menores probabilidades (0,15;0,15) até chegar-se à origem.

Para demonstrar esse teorema foi considerado a etapa quando o conjunto de combinação é reduzido a três probabilidades.

Para que o comprimento da palavra código da palavra mais provável seja $n_1=1$, é necessário que a probabilidade P_1 pertença a este conjunto conforme representado na Figura 2.

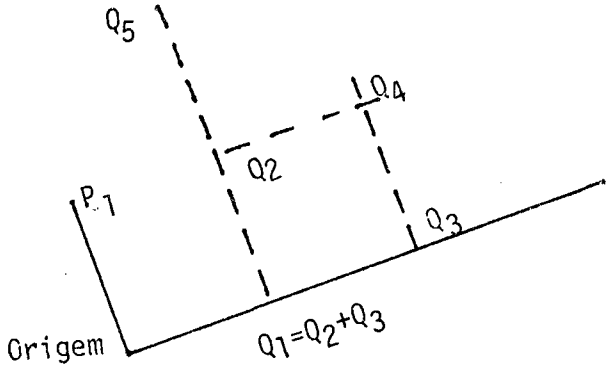


Figura 2: A parte final de uma árvore de Huffman.

Indicou-se as duas probabilidades restantes por Q_2 e Q_3, com $Q_2 \geq Q_3$. Se a palavra fonte é codificada por uma única palavra devemos ter

$$P_1 \geq Q_2 \geq Q_3 \qquad (2.2)$$

(do contrário P_1, seria associado com Q_3 e n_1 seria 2).

Visto que $P_1 + Q_2 + Q_3 = 1$ então $P_1 = 1-Q_2-Q_3$

(2.3)

As relações (2.2) e (2.3) implicam que

$$P_1 \geq 1/3 \qquad (2.4)$$

que é necessário, mas não é condição suficiente.

Supondo Q_2 formado pela adição das probabilidades Q_4 e Q_5 dos dois nós precedentes, então

$$Q_5 \leq Q_3 \qquad Q_4 \leq Q_3 \tag{2.5}$$

$$Q_4 + Q_5 \geq Q_3$$

Desse modo temos :

$$Q_2 = Q_4 + Q_5 \leq 2Q_3 \tag{2.6}$$

e (2.2) pode ser reescrito como:

$$P_1 \geq Q_2 \geq Q_3 \geq Q_2/2 \tag{2.7}$$

A relação (2.7) pode ser satisfeita por todos os valores possíveis de Q_2 e consequentemente também para o máximo valor de Q_2, o qual é obtido para um dado valor de P_1 quando $Q_2 = 2Q_3$.

Então (2.4) torna-se:

$$P_1 = 1 - 3Q_2/2 \tag{2.8}$$

A condição em que $P_1 \geq Q_2$ torna-se:

$$P_1 \geq 0{,}4 \tag{2.9}$$

Mas como $P_1 \geq Q_2$ pela condição anterior e a partir de (2.8) obtem-se, $P_1 \geq 0{,}4$ que é condição suficiente.

Mas se Q_2 não é formada pela adição de duas probabilidades, isto é, ele é a probabilidade terminal, neste caso ocorre que (2.2) é satisfeita visto que P_1 é a palavra mais provável.

No teorema seguinte mostra-se que a redundância é mínima para $1 - H(P_1, 1-P_1)$ e é máxima para $2 - P_1 - H(P_1,1-P_1)$, isto para $P_1 \geq 0,4$.

2. 2 - Teorema: (Johnsen [1980])

Quando a palavra mais provável de uma fonte tem a probabilidade de ocorrer $P_1 \geq 0,4$, um limite inferior e superior para a redundância é dado por:

$$1 - H(P_1, 1-P_1) \leq r < 2 - P_1 - H(P_1,1-P_1) \qquad (2.10)$$

onde

$$H(P_1,1-P_1) = -P_1 \log P_1 -(1-P_1)\log (1-P_1) \qquad (2.11)$$

Demonstração:

Para demonstrar esse teorema considera-se uma fonte produzindo N palavras a_i, $1 \leq i \leq N$ e um subconjunto representado por $B = \{a_2,a_3, \ldots , a_N\}$ com entropia

$$H_B = H\left[\frac{P_2}{1-P_1}, \ldots , \frac{P_N}{1-P_1}\right] = -\sum_{i=1}^{N} \frac{P_i}{1-P_1} \log \frac{P_i}{1-P_1}, \qquad (2.12)$$

Pela propriedade recursiva 1.2.7 tem-se:

$$H = H (P_1, 1-P_1) + (1 - P_1) H_B , \qquad (2.13)$$

isto é:

$$H_B = \frac{H - H(P_1, 1-P_1)}{1 - P_1}$$

Conforme Capítulo I, como L_B é o comprimento médio necessário para a codificação do subconjunto B tem-se:

$$H_B \leq L_B < H_B + 1 \qquad (2.14)$$

Para $P_1 \geq 0{,}4$, onde a_1 é codificada como uma única palavra (Teorema 2.1), o comprimento médio da palavra código para a fonte é relacionada a L_B por:

$$L = 1 + (1-P_1)\, L_B$$

De fato:

$$L = \sum_{i=1}^{N} P_i n_i$$

$$= P_1 n_1 + P_2 n_2 + \ldots + P_N n_N ,$$

e $$L_B = \frac{P_2}{1-P_1} n_2 + \frac{P_3}{1-P_1} n_3 + \ldots, + \frac{P_N}{1-P_1} n_N$$

$$(1-P_1)L_B = P_2n_2 + P_3n_3 + \ldots + P_Nn_N$$

$$(1-P_1)L_B = \sum_{i=2}^{N} P_in_i$$

Para o conjunto $\{a_1\}$ temos $P_1=1$ e $n_1=1$ então $P_1n_1=1$

$$P_1n_1 + (1-P_1)L_B = \sum_{i=2}^{N} P_in_i + P_1n_1$$

$$1 + (1-P_1)L_B = L \qquad (2.15)$$

De (2.13) e (2.15) obtemos:

$$\frac{H-H(P_1,1-P_1)}{1-P_1} \leq \frac{L-1}{1-P_1} < \frac{H-H(P_1,1-P_1)}{1-P_1} + 1 \qquad (2.16)$$

$$H-H(P_1,1-P_1) \leq L < H - H(P_1,1-P_1) + 2 - P_1 \qquad (2.17)$$

e finalmente obtem-se

$$1-H(P_1,1-P_1) \leq r < 2-P_1 - H(P_1,1-P_1) \qquad (2.18)$$

Comentários:

Vemos que Johnsen extendeu o limite superior determinado por Gallager no intervalo $P_1 \geq 0,5$ (conforme (1.8)) para o intervalo $P_1 \geq 0,4$.

Johnsen estabeleceu ainda um novo limite inferior neste intervalo. De acordo com estes resultados temos os limites superior em função de P_1 como mostrados no gráfico I, onde

$$r_{max} = P_1 + \sigma \quad , \quad 0 < P_1 \leq 0,4594$$

$$r_{max} = 2 - P_1 - H(P_1, 1-P_1) \quad , \quad P_1 > 0,4594$$

$$r_{min} = 0 \qquad 0 < P_1 < 0,4$$

$$r_{min} = 1 - H(P_1, 1-P_1) \quad , \quad P_1 \geq 0,4 \ .$$

Note que para $0,4 \leq P_1 \leq 0,4594$,

$$P_1 + \sigma < 2 - P_1 - H(P_1, 1 - P_1).$$

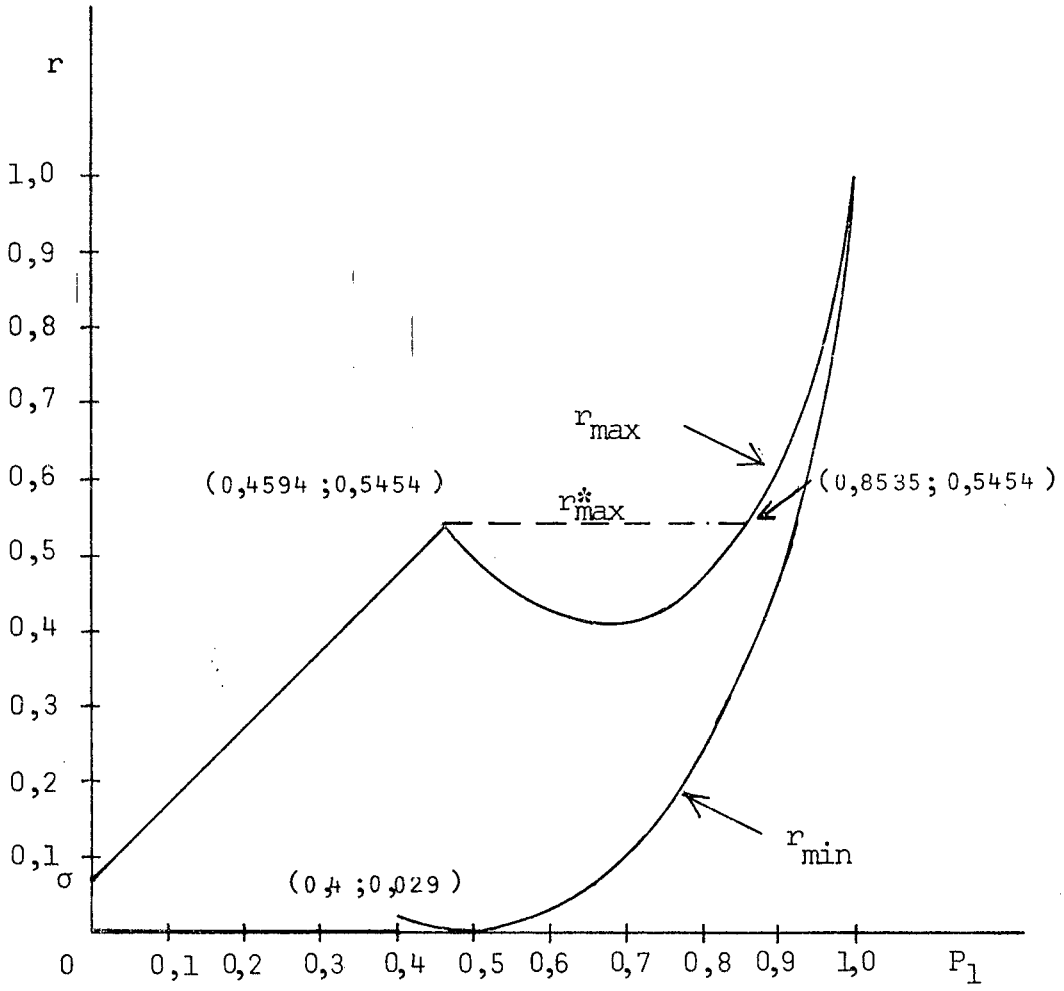


GRÁFICO I

Limites inferior e superior, $r_{mín}$ e $r_{máx}$ de redundância do código binário de Huffman versus probobilidade P_1 para a palavra fonte mais provável como dadas pelo Teorema 2.2, onde $\sigma \cong 0{,}086$.

Quando o valor de P_1 não é conhecido estima-se um valor máximo, seja $P_{1máx}$, que P_1 pode assumir. Seja $r^*_{máx}$ e r^*_{min} os limites superior e inferior, respectivamente, obtidos para tal caso. Como P_1 é um valor (desconhecido) menor que $P_{1máx}$, devemos ter $r^*_{max} \geq r_{máx}$ para todo $r_{máx}$ correspondendo a $P_1 \leq P_{1máx}$. Logo $r^*_{máx}$ (que é uma função de $P_{1máx}$) deve ser função não decrescente. Vemos então que $r^*_{máx}$ não pode acompanhar (ser igual a) $r_{máx}$ para $0,4594 < P_1 \leq 0,8535$ pois neste intervalo devemos ter $r^*_{máx} \geq r_{máx}\ (0,4594) = 0,5454$. $r^*_{máx}$ coincide no entanto com $r_{máx}$ nos outros intervalos.

Similarmente r^*_{min} tem que ser uma função não crescente de P_1 e portanto deve ser igual a zero.

No teorema a seguir apresenta-se um novo limite superior para $0,4 \leq P_1 < 0,5$.

2.3 - Teorema: (Johnsen [1980]).

Quando a palavra mais provável da fonte tem a probabilidade de ocorrer $0,4 \leq P_1 < 0,5$, existe um melhor limite superior da redundância, dado por:

$$r < 1 + 0,5\ (1-P_1) = H(P_1,\ 1-P_1), \quad \text{quando} \quad 0,4 \leq P_1 < \delta \tag{2.19}$$

e

$$r < 3 - 5P_1 - H(2P_1,\ 1-2P_1), \quad \text{quando} \quad \delta \leq P_1 < 0,5. \tag{2.20}$$

onde $\delta \cong 0,4505$.

Demonstração:

Como no Teorema 2.2, toma-se novamente o subconjunto $B = \{a_2, a_3, \ldots, a_N\}$, com probabilidades:

$$\left\{ \frac{P_2}{1-P_1}, \frac{P_3}{1-P_1}, \ldots, \frac{P_N}{1-P_1} \right\}.$$

A probabilidade P_1' da palavra mais provável da fonte B é neste caso:

$$P_1' = \frac{P_2}{1-P_1} \leq \frac{P_1}{1-P_1}, \qquad (2.21)$$

pois $P_1 \geq P_2$.

Se $P_1 < 0,5$ então $\frac{P_1}{1-P_1} < 1$.

Isto quer dizer que se P_1 é conhecido obtém se um limite superior sobre a probabilidade de P_1' e em consequência um limite superior na redundância do subconjunto B como dado por $r^*_{máx}$ do gráfico I. Pode-se então substituir (2.18) por uma desigualdade mais restritiva quando $0,4 \leq P_1 < 0,5$.

Seja $P'_{1máx} = \frac{P_1}{1-P_1}$. Quando $0,4 \leq P_1 < 0,5$,

$P'_{1máx}$ assume os valores $2/3 \leq P'_{1máx} < 1$, e (2.18) torna-se

$$H_B \leq L_B < H_B + 2 - H\left(P'_{1máx}, 1-P'_{1máx}\right) - P'_{1máx} \qquad (2.22)$$

quando $P'_{1máx} \geq 0,8535$, e

$$H_B \leq L_B < H_B + 0,5454 , \qquad (2.23)$$

quando $2/3 \leq P'_{1máx} \leq 0,8535$.

Quando $P'_{1máx} = 0,8535$, então $P_1 \cong 0,4605$.

Substituindo $P'_{1máx}$ por $\frac{P_1}{1 - P_1}$ obtém-se

$$H_B \leq L_B < H_B + 0,5454 , \qquad (2.24)$$

quando $0,4 \leq P_1 < 0,4605$ e

$$H_B \leq L_B < H_B + 2\text{-}H\left[\frac{P_1}{1-P_1}, 1 - \frac{P_1}{1-P_1}\right] - \frac{P_1}{1-P_1} , \qquad (2.25)$$

quando $0,4605 \leq P_1 < 0,5$.

Com passos semelhantes ao Teorema 2.2 a partir de (2.17) e usando (2.24) finalmente obtém-se:

$$1\text{-}H(P_1,1-P_1) \leq r < 0,5454\ (1-P_1) + 1\text{-}H(P_1,1-P_1), \qquad (2.26)$$

quando $0,4 \leq P_1 < 0,4605$.

Usando (2.13) e (2.15) e substituindo em (2.24) obtém-se.

$$\frac{H\text{-}H(P_1,1-P_1)}{1 - P_1} \leq \frac{L - 1}{1 - P_1} < \frac{H\text{-}H(P_1,1-P_1)}{1 - P_1} + 0,5454 ,$$

$$\text{Como } H_\beta = \frac{H - H(P_1,\ 1 - P_1)}{1 - P_1} \quad \text{e} \quad L_\beta = \frac{L - 1}{1 - P_1},$$

substituindo em (2.25) obtem-se:

$$\frac{H - H(P_1,\ 1-P_1)}{1 - P_1} \leq \frac{L - 1}{1 - P_1} < \frac{H - H(P_1,\ 1-P_1)}{1 - P_1} + 2$$

$$- H\left[\frac{P_1}{1 - P_1},\ 1 - \frac{P_1}{1 - P_1}\right] - \frac{P_1}{1 - P_1},$$

$$\therefore \quad 1 - H(P_1,\ 1-P_1) \leq L - H < 3 - H(P_1,\ 1-P_1)$$

$$- 3P_1 - (1-P_1).H\left[\frac{P_1}{1 - P_1},\ 1 - \frac{P_1}{1 - P_1}\right].$$

como $r = L - H$ tem-se:

$$1 - H(P_1,\ 1 - P_1) \leq r < 3 + P_1 \log P_1 + (1-P_1)$$

$$\log (1-P_1) - 3P_1 + (1-P_1)\left[\frac{P_1}{1 - P_1} \log \frac{P_1}{1 - P_1}\right.$$

$$+ \frac{1 - 2P_1}{1 - P_1} \log \frac{1 - 2P_1}{1 - P_1} \Bigg]$$

$$\therefore \quad 1-H(P_1,1-P_1) \leq r < 3-3P_1+P_1 \log P_1 + (1-P_1)\log(1-P_1) +$$

$$P_1 \log P_1 - P_1 \log(1-P_1)+(1-2P_1)\log(1-2P_1)-(1-2P_1)\cdot$$

$$\cdot \log (1-P_1).$$

$$\therefore \quad 1-H(P_1,1-P_1) \leq r < 3-3P_1+2P_1 \log P_1+(1-2P_1)\log(1-2P_1),$$

temos que:

$$2P_1 \log 2P_1 = 2P_1 \log_2 2 + 2 P_1 \log P_1$$

$$2P_1 \log P_1 = 2P_1 \log 2P_1 - 2P_1$$

então

$$1-H(P_1,1-P_1) \leq r < 3-5P_1+2P_1 \log 2P_1+(1-2P_1)\log(1-2P_1).$$

$$\therefore \quad 1-H(P_1,1-P_1) \leq r < 3-5P_1 - H(2P_1,1-2P_1), \qquad (2.27)$$

quando $0,4605 \leq P_1 < 0,5$.

Essas duas desigualdades dão um melhor limite superior da redundância para $0,4 \leq P_1 < 0,5$. A redundância é agora menor ou igual a 0,5 para $P'_{1máx} \leq 0,821$. Usando esse novo limite superior na redundância, (2.24) e (2.25) tornam-se:

$$H_B \leq L_B < H_B + 0,5 , \qquad (2.28)$$

quando $0,4 \leq P_1 \leq \delta$, e

$$H_B \leq L_B < H_B + 2\text{-}H\left[\frac{P_1}{1-P_1} , 1-\frac{P_1}{1-P_1}\right] - \frac{P_1}{1-P_1} , \qquad (2.29)$$

quando $\delta \leq P_1 < 0,5$

onde δ é a probabilidade P_1 correspondendo a $P'_{1máx} = 0,821$:

$\delta = 0,4505$.

Se obedecermos os mesmos argumentos do Teorema 2.2 obtem-se (2.28) e chegamos a desigualdade :

$$1-H(P_1,1-P_1)\leq r<1+0,5(1-P_1)-H(P_1,1-P_1)\,, \qquad (2.30)$$

quando $0,4\leq P_1<\delta$,

e $\qquad 1-H(P_1,1-P_1)\leq r < 3-5P_1-H(2P_1, 1-2P_1)$,

quando $\delta \leq P_1< 0,5$, com $\delta \cong 0,4505$.

Comentários:

Esse novo limite $r_{máx}$ e $r_{mín}$ obtidos por Johnsen [1980] na redundância dados por (1.6), (2.18). (2.30) são representados no gráfico II. Quando somente um valor máximo $P_{1máx}$ de P_1 pode ser estimado, o limite superior é parcialmente modificado pois é necessário uma função não crescente de P_1. Essas modificações denotadas por $r^*_{máx}$, estão também indicadas no gráfico II.

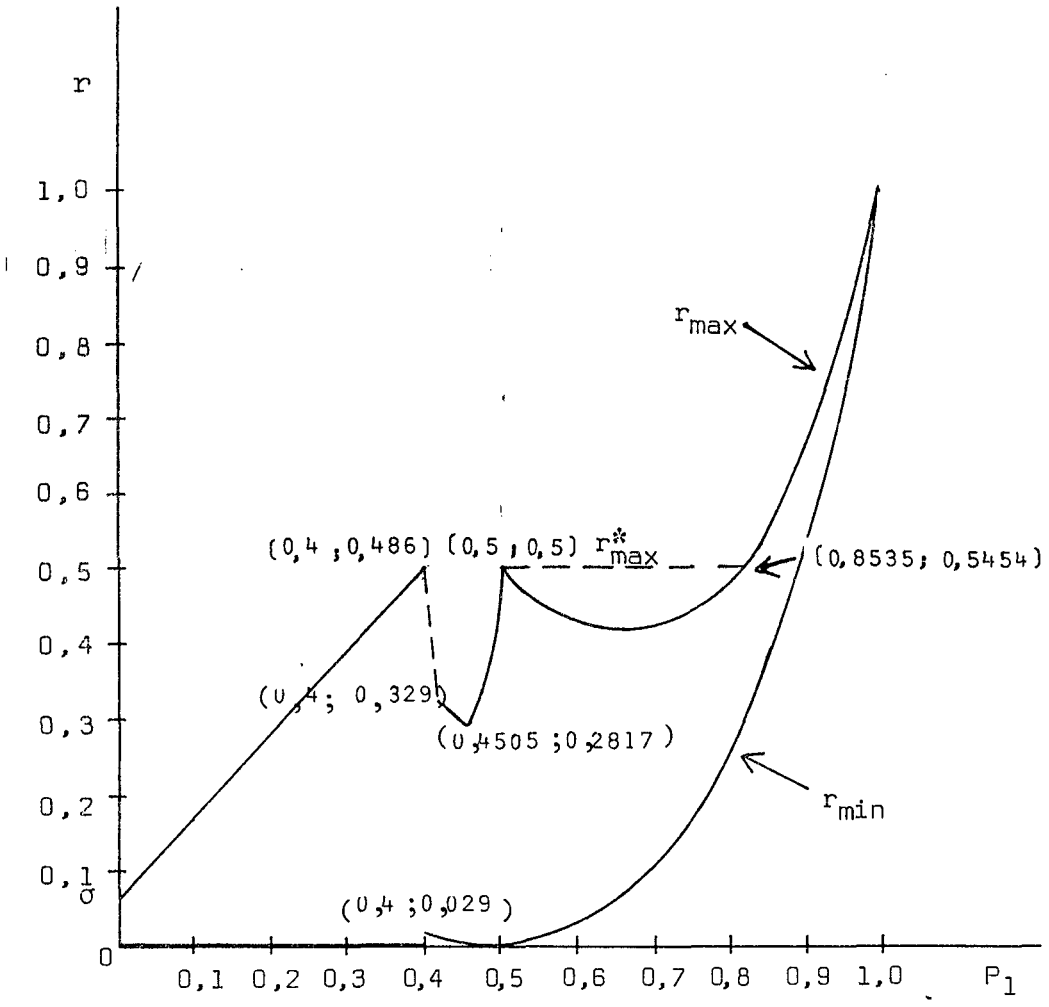


GRÁFICO II

Limites inferior e superior, $r_{mín}$ e $r_{máx}$ de redundância do código binário de Huffman versus probabilidade P_1 para a palavra fonte mais provável como dadas pelo Teorema 2.3, onde $\sigma \cong 0{,}086$.

CAPÍTULO III

LIMITES SUPERIOR E INFERIOR DE REDUNDÂNCIA DE GRAU β

INTRODUÇÃO

Neste capítulo generalizamos os limites superior e inferior de redundância dado por Johnsen [1980], para redundância de grau β, que constituem os resultados principais desta tese.

Pelo Teorema 1.7.2., vemos que, é sempre possível construir um código unicamente decifrável onde a redundância de grau β (β > 0) é limitado entre 0 e 1. Usando argumento similar ao Johnsen, obtemos os limites superior e inferior da redundância quando a palavra mais provável da fonte é $P_1 \geq 0{,}4$. Quando $0{,}4 \leq P_1 < 0{,}5$, o limite superior foi superado.

O conceito de entropia de grau β e o comprimento médio de grau β, encontram-se no Capítulo I em seções 1.5 e 1.7.1 respectivamente.

Definimos a redundância de grau β como sendo:

$$r^{\beta} = L(\beta) - H^{\beta}(P).$$

No Teorema a seguir apresentamos os limites superiore inferior da redundância de grau β quando $P_1 \geq 0{,}4$.

3.1 - Teorema:

Quando a palavra maos provável de uma fonte tem a probabilidade $P_1 \geq 0{,}4$, o limite superior e inferior para redundância de grau β é dado por:

$$\frac{H^{\beta} - H^{\beta}(P_1, 1-P_1)}{(1-P_1)^{\beta-1}} + 1 - H^{\beta} \leq r(\beta) < \frac{H^{\beta} - H^{\beta}(P_1, 1-P_1)}{(1-P_1)^{\beta-1}}$$

$$+ 2 - P_1 - H^{\beta} \ , \ \beta > 1, \quad (3.1)$$

onde $H^{\beta}(P_1 , 1-P_1)$ é a entropia binária de grau β.

Demonstração:

Considerando uma fonte produzindo N palavras a_i com probabilidade P_i, $1 \leq i \leq N$ e um subconjunto B formado pelas palavras $B = \{a_2, a_3, \ldots , a_N\}$ com entropia de grau β,

$$H_B^{\beta} = H_B^{\beta}\left[\frac{P_2}{1 - P_1} , \frac{P_3}{1 - P_1}, \ \ldots \ , \frac{P_N}{1 - P_1}\right], \quad (3.2)$$

Pela propriedade recùrsiva (ref. 1.6.7) sabemos que:

$$H^{\beta} = H^{\beta}(P_1, 1-P_1) + (1-P_1)^{\beta} \ H_B^{\beta} \quad (3.3)$$

Seja $L_B(\beta)$ o comprimento mēdio da palavra cōdi-

go para codificação do subconjunto $B = \{a_2, a_3, \ldots, a_N\}$.

Então pelo Teorema 1.7.2, temos:

$$H_B^{\beta} \leq L_B(\beta) < H_B^{\beta} + 1 \quad , \quad \beta > 1 \qquad (3.4)$$

Para $P_1 \geq 0{,}4$, onde a_1 é codificado por um único "bit" (Teorema 2.1), o comprimento médio de grau β das palavras códigos para toda a fonte é relacionado com $L_B(\beta)$ como:

$$L(\beta) = \sum_{i=1}^{N} P_i n_i(\beta) = P_1 n_1(\beta) + (1-P_1) L_B(\beta) \ ,$$

onde

$$L_B(\beta) = \frac{P_2}{1-P_1} n_2(\beta) + \frac{P_3}{1-P_1} n_3(\beta) + \ldots + \frac{P_N}{1-P_1} \cdot n_N(\beta) .$$

Seja $n_i(\beta) = \dfrac{2^{(1-\beta)n_i} - 1}{2^{1-\beta} - 1}$, $\beta \neq 1$, $\beta > 0$ (3.5)

$i = 1, 2, \ldots, N$.

Agora, quando $n_1 = 1$, $n_1(\beta) = 1$ e

$$\lim_{\beta \to 1} n_i(\beta) = n_i \ , \quad i = 1,2, \ldots, N.$$

Então:

$$(1-P_1)L_B(\beta) = \sum_{i=2}^{N} P_i \frac{2^{(1-\beta)n_i} - 1}{2^{1-\beta} - 1}$$

$$= \frac{\sum_{i=2}^{N} P_i (2^{(1-\beta)n_i} - 1)}{2^{1-\beta} - 1}$$

$$\text{ou, } P_1 n_1(\beta) + (1-P_1)L_B(\beta) = P_1 n_1(\beta) + \frac{\sum_{i=2}^{N} P_i(2^{(1-\beta)n_i} - 1)}{2^{1-\beta} - 1}.$$

Para o conjunto $\{ a_1 \}$ temos $P_1=1$ e $n_1=1$, então

$$n_1(\beta) = \frac{2^{(1-\beta)1} - 1}{2^{1-\beta} - 1} = 1 \quad , \text{ (por (3.5)) isto é:}$$

$$1 + (1-P_1) L_B(\beta) = L(\beta) . \qquad (3.6)$$

De (3.3), (3.4) e (3.6) obtemos:

$$\frac{H^{\beta} - H^{\beta}(P_1,1-P_1)}{(1-P_1)^{\beta}} \leq \frac{L(\beta)-1}{1-P_1} < \frac{H^{\beta} - H^{\beta}(P_1,1-P_1)}{(1-P_1)^{\beta}} + 1 , \qquad \beta > 1 \qquad (3.7)$$

$$(1-P_1)[H^{\beta} - H^{\beta}(P_1,1-P_1)] \leq (L(\beta) - 1)(1-P_1)^{\beta} < (1-P_1)[H^{\beta} - H^{\beta}(P_1,1-P_1)] + (1-P_1)(1-P_1)^{\beta} .$$

Como

$$(1-P_1)^{\beta}[L(\beta) - H^{\beta}] = (1-P_1)^{\beta}[L(\beta) + 1 - 1 - H^{\beta}],$$

i.ē., $(1-P_1)^{\beta} r(\beta) = (1-P_1)^{\beta}(L(\beta) - 1) + (1-P_1)^{\beta}(1-H^{\beta})$.

Então substituindo na desigualdade acima temos :

$$(1-P_1)[H^{\beta} - H^{\beta}(P_1,1-P_1)] \leq (1-P_1)^{\beta} r(\beta) - (1-P_1)^{\beta}(1-H^{\beta}) < (1-P_1)[H^{\beta} - H^{\beta}(P_1,1-P_1)] + (1-P_1)(1-P_1)^{\beta}.$$

Somando $(1-P_1)^{\beta}\ (1-H^{\beta})$ em ambos os lados e simplificando obtemos:

$$\frac{H^{\beta} - H^{\beta}(P_1,1-P_1)}{(1-P_1)^{\beta-1}} + 1-H^{\beta} \leq r(\beta) < \frac{H^{\beta} - H^{\beta}(P_1,1-P_1)}{(1-P_1)^{\beta}} +$$

$$2 - P_1 - H^{\beta} \quad \text{para} \quad \beta > 1 \qquad (3.8)$$

Observação:

Quando fazemos $\beta \to 1$, a equação (3.8) reduz-se para (2.10).

No Teorema a seguir damos um limite superior para a redundância de grau β, quando $0,4 \leq P_1 < 0,5$.

3.2 - Teorema:

Quando a palavra mais provável da fonte tem a probabilidade $0,4 \leq P_1 < 0,5$, então existe um limite superior na redundância dada por:

$$r(\beta) < \frac{[1 - (1-2P_1)^{\beta-1}]H^{\beta}}{(1-2P_1)^{\beta-1}} - \frac{(2P_1)^{\beta} + H^{\beta}(2P_1,1-2P_1)}{(1-2P_1)^{\beta-1}} +$$

$$+ 3 - 3P_1. \qquad (3.9)$$

quando $0,4 \leq P_1 < 0,5$ para $\beta>1$.

Demonstração:

Considere o subconjunto $B = \{a_2, \ldots, a_N\}$ com probabilidades $\{\frac{P_2}{1-P_1}, \frac{P_3}{1-P_1}, \ldots, \frac{P_N}{1-P_1}\}$, seja P_1 a probabilidade da palavra mais provável da fonte B, dado por $P_1' = \frac{P_2}{1-P_1} \leq \frac{P_1}{1-P_1}$, então quando $0,4 \leq P_1 < 0,5$, $P_{1máx}'$ toma valores $2/3 < P_{1máx}' < 1$ e (3.8) torna-se:

$$r_B(\beta) < \frac{H_B^\beta - H^\beta(P_{1máx}', 1-P_{1máx}')}{(1 - P_{1máx})^{\beta-1}} + 2 - P_{1máx}' - H_B^\beta \text{, ou}$$

$$L_B(\beta) < \frac{H_B^\beta - H^\beta(\frac{P_1}{1-P_1}, 1 - \frac{P_1}{1-P_1})}{(1 - \frac{P_1}{1-P_1})^{\beta-1}} + 2 - \frac{P_1}{1-P_1} \qquad (3.10)$$

De (3.3), (3.4) e (3.10) obtemos

$$\frac{L(\beta)-1}{1-P_1} < \frac{\dfrac{H^{\beta} - H^{\beta}(P_1,1-P_1)}{(1-P_1)^{\beta}} - \dfrac{H^{\beta}\left(\frac{P_1}{1-P_1},\ 1-\frac{P_1}{1-P_1}\right)}{}}{\left(\dfrac{1-2P_1}{1-P_1}\right)^{\beta-1}} +$$

$$+ 2 - \frac{P_1}{1-P_1} \qquad (3.11)$$

Agora simplificando (3.11) obtemos o resultado desejado; isto é,

$$r(\beta) < \frac{\left[(1 - (1-2P_1)^{\beta-1})\, H^{\beta} - H^{\beta}(2P_1,1-2P_1) - (2P_1)^{\beta}\right]}{(1-2P_1)^{\beta-1}} +$$

$$+ 3-3P_1$$

para $\beta>1$ e $0,4 \leq P_1 < 0,5$. //

Observação: Quando $\beta \downarrow 1$, a equação (3.9) reduz-se para (2.20).

CONCLUSÕES

Entropias generalizadas tais como entropia de ordem α, entropia de grau β, entropia de ordem α e grau β estão sendo aplicadosem diversas áreas tais como em computação, estatística, reconhecimento de padrões (modelos), conjuntos difusos, teoremas de codificação (teoria da informação), etc.

Blumer [1982] apresentou novos limites superior e inferior de redundância de ordem α de um código binário definidos em função de entropia de ordem α e comprimento médio de ordem α (Campbell [1965]). Os resultados obtidos por Blumer [1982] são baseados nas técnicas de programação linear e generalizam os resultados de Gallager [1978] .

Neste trabalho, generalizamos os resultados de Johnsen[1980] e obtemos os limites superior e inferior de redundância de grau β de um código binário. Os resultados obtidos são condicionados sobre a palavra mais provável da fonte, e principalmente são baseados na propriedade recessiva de entropia de grau β.Os novos resultados obtidos são válidos para $\beta > 1$.

Para mais aplicações de entropias generalizadas em Códigos de Huffman veja Parker [1980] .

BIBLIOGRAFIA

1 - ACZÉL, J. and DARÓCZY Z. (1975): On Measures of Information and Their Characterizations -Academic Press, New York.

2 - ASH, R. (1965): "Information Theory", Interscience Pub. New York.

3 - BLUMER, A. (1982): Bounds on the Redundancy of Noiseless Source Coding - Ph. D. Thesis,University of Illinois at Urbana - Champaign.

4 - BOUCHON, B. (1978) : Sur La Réalisation de Questionnaires - Thèse de Doctorat d'Etat,Université Pierre et Marie Curie, Paris.

5 - CAMPBELL, L. L. (1965): A Coding Theorem and Rényi's Entropy - Information and Control, Vol. 8, 423-429.

6 - DARÓCZY, Z. (1970): Generalized Information Functions - Information and Control, Vol. 16 , 36-51.

7 - GALLAGER, R.G. (1978): Variotions on Theme by Huffman - IEEE Trans. Inform. Theory, Vol. II-24, nº 6,668-674

8 - HÁVRDA, J. and F. CHÁRVAT (1967): Quantification Method of Classification Processes. Concept of Structure a-Entropy-Kybernetika (Prague), Vol. 3 , 30-35.

9 - HUFFMAN, D.A. (1952): A Method for the Constrution of Mininum Redundancy Codes - Proc. IRE, Vol.40 1098-1101.

10 - JOHNSEN, O. (1980): On the Redundancy of Binary Codes - IEEE Trans. Inform. Theory,Vol. II-26, 220-222.

11 - PARKER, D.S. Jr. (1980): Conditions for Optimality of the Huffman Algorithms - SIAM J. Comput., Vol.9, 470-489.

12 - SHANNON, C.E. (1948): The Mathematical Theory of Communication - B.S.T.J. , Vol. 27, 379-423, 623-656.

14 - SHARMA, B.D. and TANEJA I.J. (1975): Entropy of Type (α ,β) and Other Generalized Measures in Information Theory - METRIKA, Vol. 22, 205-215.

14 - TANEJA, I. J. (1975): A Joint Characterization of Shannon's Entropy and Entropy of type β Through a Functional Equation - J.of Math. Sci. , Vol. 10 69-74.

15 - TANEJA, I.J. (1977): On the Branching Property of Entropy - Ann. Polonici Math, Vol XXXV, 67-75.